

COLLECTION
bioéthique
critique

Sous la direction de
Marie-Hélène Parizeau
Soheil Kash

La société robotisée

Enjeux éthiques et politiques



La société robotisée

Enjeux éthiques et politiques

Sous la direction de
MARIE-HÉLÈNE PARIZEAU
ET SOHEIL KASH

COLLECTION
bioéthique
critique

Sous la direction de Marie-Hélène PARIZEAU

L'objectif principal de cette collection est la diffusion en langue française de résultats de recherche originaux dans le champ de la bioéthique et de l'éthique de l'environnement.

La bioéthique est entendue dans un sens large qui inclut son orientation biomédicale traditionnelle, mais aussi ce qu'on appelle l'éthique de l'environnement.

«Bioéthique critique» est donc une collection conçue comme un carrefour de réflexion multidisciplinaire et interdisciplinaire, d'abord éthique, mais à l'heure de la mondialisation, également en lien avec le politique. Ce carrefour permet que l'on s'interroge collectivement, d'une part sur les conséquences des nouvelles technologies en biomédecine sur l'être humain et, d'autre part, sur la relation de l'être humain à la nature-environnement. Cette collection constitue un lieu de discussion ouvert sur le monde et sa diversité culturelle.

La collection veut promouvoir les résultats de recherche originaux, les ouvrages collectifs, les essais et, à l'occasion, la traduction d'ouvrages significatifs de bioéthique en langue française.

Une liste des titres parus dans la collection est disponible à la fin du volume.

La société robotisée

Enjeux éthiques et politiques

Sous la direction de
MARIE-HÉLÈNE PARIZEAU
ET SOHEIL KASH



Presses de
l'Université Laval

Financé par le gouvernement du Canada
Funded by the Government of Canada

Canada

Nous remercions le Conseil des arts du Canada de son soutien.
We acknowledge the support of the Canada Council for the Arts.



Conseil des arts Canada Council
du Canada for the Arts

Les Presses de l'Université Laval reçoivent chaque année du Conseil des Arts du Canada et de la Société de développement des entreprises culturelles du Québec une aide financière pour l'ensemble de leur programme de publication.

SODEC

Québec 

Maquette de couverture: Laurie Patry

Mise en pages: Danielle Motard

ISBN : 978-2-7637-3993-9

ISBN PDF : 9782763739946

© Les Presses de l'Université Laval

Tous droits réservés.

Imprimé au Canada

Dépôt légal 3^e trimestre 2019

Les Presses de l'Université Laval

www.pulaval.com

Toute reproduction ou diffusion en tout ou en partie de ce livre par quelque moyen que ce soit est interdite sans l'autorisation écrite des Presses de l'Université Laval.

Table des matières

Présentation des auteurs	XV
Introduction	1
Marie-Hélène PARIZEAU et Soheil KASH	

PREMIÈRE PARTIE LE ROBOT COMME OBJET TECHNIQUE : ENJEUX PHILOSOPHIQUES ET ÉTHIQUES

CHAPITRE 1

Le robot, alibi d'une humanité fatiguée d'elle-même	23
Jean-Michel BESNIER	

CHAPITRE 2

Généalogie de la morale automatisée	31
Vanessa NUROCK	
1. Un sens de la justice?	32
2. Piaget et les règles morales	36
3. Dilemmes et cognition morale	38
4. Dilemmes en série: la tramwaylogie et ses limites	43

CHAPITRE 3

À propos du statut moral des robots	51
Tomislav BRACANOVIĆ	
Introduction	51
1. Le statut moral et ses théories	52
2. Les robots?	55

3. Les robots cognitifs.	57
4. Biorobots	61
Conclusion	65
CHAPITRE 4	
Quel(s) droit(s) pour la robotique?	69
Pistes de réflexion sur la gouvernance par le droit du développement de la robotique interactive.	69
Charles-Étienne DANIEL	
Introduction	69
1. Les limites du droit dans l'encadrement du développement de robots interactifs et les solutions de remplacement proposées	72
1.1 Les limites du droit face au développement de la robotique ...	73
1.2 Les solutions de remplacement proposées pour encadrer la robotique interactive.	77
2. Le recours aux normes alternatives pour gouverner le développement de la robotique interactive	79
2.1 La nature et les caractéristiques des différentes formes de droit retrouvées dans la gouvernance par le droit	79
2.2 La considération des propositions dans l'encadrement normatif de la robotique.	83
Conclusion	88
CHAPITRE 5	
Du robot interprète de la volonté de l'homme au robot réducteur de l'humain	95
Hoda NEHMÉ	
1. Innovations et mutations.	95
2. État des lieux : Confrontations entre partisans et opposants	96
L'incontournable développement robotique.	99
3. L'incontournable nécessité d'une identité juridique de « la personne électronique ».	100
4. Genèse de la robotique	101
5. Vers quel homme s'achemine-t-on?	106
Conclusion	107

CHAPITRE 6

**Ingénierie neuromorphique et intelligence artificielle:
construire un oiseau qui ne sait possiblement pas voler 111**
Frédéric DUBOIS

- 1. Le battement des ailes:
 - un isomorphisme fonctionnel 115
 - 1.1 Vers une théorie fonctionnaliste du vol 116
 - 1.2 Isomorphisme et multiréalisabilité 119
- 2. Le fonctionnalisme et l'ingénierie neuromorphique 121
 - 2.1 Le cerveau-ordinateur 121
 - 2.2 Les modèles du cerveau-ordinateur. 123
- 3. Qu'est-ce que l'objet technique? 126
 - 3.1 Multiréalisabilité et causalité instrumentale 126
 - 3.2 De l'abstraction à la réalisation concrète de l'objet technique . . . 127
- Conclusion 131

CHAPITRE 7

**Robotique développementale et intelligence artificielle «incorporée»:
questions éthiques du point de vue du robot 135**
Alexandre PITTI

- Introduction 135
- 1. De la corporéité. 137
- 2. Développement et apprentissage 140
- 3. Inférence active 142
- 4. Les régions soi-autrui du cerveau et les neurones-miroir. 143
- 5. Algorithmes d'apprentissage explicatifs 144
- 6. Une science d'apprentissage 146

CHAPITRE 8

**Le vivant et la machine en biologie de synthèse:
entre ingénierie et bricolage 151**
Guillaume PELLETIER

- 1. Qu'est-ce qu'une bonne machine? 152
- 2. De l'organisme à la machine 157
- 3. Le bricolage et l'exploration technique
en biologie de synthèse 160
- Conclusion 164

CHAPITRE 9

Robotisation et aliénation: de l'objectivation de l'homme à la valorisation de la machine	167
Chantale PILON	
1. L'idéologie de la domination et déshumanisation.	169
2. Rationalisation, dévalorisation et auto-instrumentalisation.	172
3. L'homme-machine et l'organisation scientifique du travail.	176
4. L'obsolescence de l'homme	180
5. En conclusion.	182

DEUXIÈME PARTIE

LES ROBOTS DANS LA SOCIÉTÉ: ENJEUX ÉTHIQUES ET POLITIQUES

CHAPITRE 10

Les implications éthiques de la perte progressive du contrôle humain	189
John L. FINNEY	
Introduction	189
1. Conventions internationales	190
2. Armes robotisées contrôlées à distance	191
3. Armes autonomes.	195
4. L'autorisation de tuer	198
5. Implications stratégiques.	199
6. Responsabilités personnelles	199
Conclusion	200

CHAPITRE 11

Le Kamikaze et le drone Éthique de l'assassinat	201
Soheil KASH	

CHAPITRE 12

Les robots: une expérimentation sociale en mal d'encadrement éthique	211
Marie-Hélène PARIZEAU	
1. Mise en situation.	211

2. Les robots industriels: une logique de productivité économique	213
3. Les robots sociaux à l'hôpital, à l'école et à la maison: une expérimentation sociale à large échelle	218
4. Les robots, une réponse technique à des problèmes sociaux?	221
5. Critères et principes éthiques d'une expérimentation sociale Avec les robots sociaux	224
 CHAPITRE 13	
Les robots humanoïdes peuvent-ils venir en aide aux enfants atteints du trouble du spectre de l'autisme?	229
Josée Anne GAGNON	
 CHAPITRE 14	
Robots soignants en gériatrie: réalité ou illusion?	237
Félix PAGEAU	
Introduction	237
1. Des robots pour le soin des personnes âgées: mise en contexte.	238
1.1 Les données sociodémographiques	238
1.2 <i>Hype</i> , travail et transhumanisme	239
1.3 Les raisons médicales en faveur de l'emploi du robot	240
2. Avantages et inconvénients des robots dans les soins aux personnes âgées	241
2.1 La sécurité et l'autonomie fonctionnelle (AVD, AVQ).	241
2.2 Sécurité et liberté	243
2.3 La vie privée	244
2.4 Les relations sociales.	245
2.5 La fiabilité	246
2.6 Adaptation aux personnes avec démence	247
2.7 Les coûts	248
3. Reprise des critiques des robots dits «soignants».	248
3.1 Le jugement clinique: aspect critique.	251
3.2 Un argument majeur: l'analyse sociale	252
3.3 L'environnement	253
Conclusion	254

CHAPITRE 15

Un amour de robot: robot émotionnel et travail d'aidant	259
Monique LANOIX	
1. Pourquoi PARO?	260
2. L'impossibilité de se décentrer	265
3. Le travail d'aidant	269
4. L'aidant comme outil	271
5. L'effet PARO	274
Conclusion	276

CHAPITRE 16

Enjeux et perspectives de l'apprentissage avec des robots pédagogiques	281
Raoul KAMGA et Margarida ROMERO	
Introduction	281
1. Usage pédagogique des robots et perspectives d'apprentissage	282
2. Défis de robotique pédagogique et rapport à l'erreur	284
3. Les défis liés à l'intégration pédagogique des robots à l'école	285

CHAPITRE 17

Les animaux robots: quelles conséquences pour les enfants?	289
Sophie CLOUTIER	
1. Remarques préliminaires sur l'aliénation au monde	290
2. Les animaux robots et l'attachement narcissique	293
3. L'apprentissage moral par les relations avec les animaux biologiques	299

CHAPITRE 18

Ferme laitière et robot de traite: Que reste-t-il de la relation entre l'humain et l'animal?	311
Sandra BLOUIN	
1. Pourquoi se procurer un robot de traite?	312
2. Les inconvénients du robot de traite	314
3. Comprendre les relations entre les humains et les animaux	317
4. L'importance de la relation pour l'animal	323
Conclusion	326

CHAPITRE 19

Face à la culture: animaux, robots et êtres humains 329

Georges CHAPOUTHIER

- 1. L'animal, être de culture 329
- 2. Langage, morale, esthétique chez les animaux. 331
- 3. L'intelligence au-delà de la culture 333
- 4. Les robots et la culture 334
- 5. Les limites des robots actuels 335
- 6. La place de l'être humain 336

CHAPITRE 20

**La robotisation des environnements: de l'émancipation des corps
à la domination de la nature 339**

Louis-Étienne PIGEON

- Introduction 339
- 1. La technosphère, contexte et horizon
de la robotisation 340
- 2. Qu'entend-on par robotisation? 344
- 3. La robotisation des corps 345
- 4. La robotisation des environnements 347
- 5. Problématique ontologique et éthique
des environnements robotisés. 350
- Conclusion 352

CHAPITRE 21

**Quelle place les robots peuvent-ils avoir
dans la transition écologique? 355**

Catherine LARRÈRE

- 1. Quel impact les robots et l'automatisation croissante ont-ils
sur l'engagement de la transition écologique? 356
- 2. Quelle place pour les robots dans le projet écologique? 361
- 3. L'écologie des robots. 364

Présentation des auteurs

Jean-Michel BESNIER, philosophe, professeur émérite à l'Université Paris-Sorbonne, France.

Sandra BLOUIN, étudiante à la maîtrise à la faculté de philosophie de l'Université Laval, Québec.

Tomislav BRACANOVIĆ, philosophe à l'Institut de philosophie de Zagreb, vice-président de la COMEST/UNESCO, Croatie.

Georges CHAPOUTHIER, neurobiologiste et philosophe, directeur de recherche émérite au CNRS, France.

Sophie CLOUTIER, philosophe, professeur agrégée à la faculté de philosophie et à l'école d'éthique publique de l'Université Saint-Paul, Ottawa, Ontario.

Charles-Étienne DANIEL, juriste, professeur à la faculté de droit de l'Université de Sherbrooke, Québec.

Frédéric DUBOIS, doctorant à la faculté de philosophie de l'Université Laval, Québec.

John FINNEY, physicien, professeur émérite au London University college, membre du Pugwash Conference on Science and World Affairs, membre de la COMEST/UNESCO, Grande-Bretagne.

Josée-Anne GAGNON, pédiatre intensiviste, chef de service de pédiatrie au CHUL, Québec.

Raoul KAMGA, doctorant au département d'études sur l'enseignement et l'apprentissage à la faculté d'éducation, Université Laval, Québec.

Soheil KASH, philosophe, professeur associé à la faculté de philosophie de l'Université Laval, Québec.

Monique LANOIX, philosophe, professeur agrégée à la faculté de philosophie et à l'école d'éthique publique de l'Université Saint-Paul, Ottawa, Ontario.

Catherine LARRÈRE, philosophe, professeur émérite de l'Université Paris 1 Panthéon-Sorbonne, France.

Hoda NEHMÉ, vice-rectrice aux affaires sociales, Université Saint-Esprit de Kaslik-Liban.

Vanessa NUROCK, philosophe, maîtresse de conférence au département de sciences politiques à l'Université Paris 8, France.

Félix PAGEAU, gériatre, Université Laval, Québec.

Marie-Hélène PARIZEAU, philosophe, professeur titulaire à la faculté de philosophie de l'Université Laval, présidente de la COMEST/UNESCO 2016-2019, Québec.

Guillaume PELLETIER, doctorant à la faculté de philosophie de l'Université Laval, Québec.

Louis-Étienne PIGEON, philosophe, chargé d'enseignement à la faculté de philosophie de l'Université Laval, Québec.

Chantale PILON, doctorante, faculté de philosophie à l'Université Laval, Québec.

Alexandre PITTI, chercheur informaticien au laboratoire ETIS à l'Université de Cergy-Pontoise, UMR 8051 CNRS, ENSEA.

Margarida ROMERO, directrice du Laboratoire d'innovation et numérique pour l'éducation de l'ESPE de Nice, Université Côte d'Azur, France.

Introduction

Marie-Hélène PARIZEAU
et Soheil KASH

La modernité a sonné le glas des « Gens de l'épée¹ » qui, avec l'apparition des fusils et de la poudre à canon, ont dû abandonner leurs armes, et assister impuissants à l'écroulement de leur système de valeurs chevaleresque fondé sur le courage, l'honneur et les qualités guerrières individuelles. C'est la disparition de ce monde prémoderne que déplore le *Don Quichotte* de Cervantès. Les progrès de la technique guerrière moderne sont venus remplacer le courage du chevalier par un homme « sans qualités » muni d'un pistolet, d'un canon ou d'un drone pour battre le chevalier le plus courageux et le plus aguerri.

D'homme d'actes et de paroles qu'il était, l'être humain se transforme avec la modernité en « inventeur de machines » qui tendent à le remplacer dans ce qu'il fait et ce qu'il dit. Il a même inventé un verbe – robotiser – pour désigner l'ère de ces objets capables de tout faire à sa place : robots industriels, robots de soins, robots de service, robots tueurs, etc. Ces machines sont si « intelligentes » qu'elles nous promettent de dépasser par leur artifice l'intelligence humaine qui

1. Ibn Khaldoun, dans un court chapitre de la Muqaddima s'intitulant « L'importance relative des fonctions de l'épée et de la plume dans l'histoire des États » (chapitre XXXIII dans la traduction de la Pléiade, chapitre 35 dans la version arabe), distingue les « Gens de l'épée » des « Gens de la plume ». Cette division est fonction des besoins du pouvoir d'une dynastie entre le début guerrier de son installation et sa stabilité de luxe qui suit sa domination. Ibn Khaldoun cite une maxime perse : « Ce que craignent le plus les gens de l'épée, c'est quand le peuple retrouve la paix » (p. 570).

les a créées. Il ne leur manquerait que l'âme pour se mesurer à l'être humain de plus en plus étranger à sa création robotique.

En effet, les robots contemporains sont caractérisés, selon le rapport *L'éthique de la robotique* de la Commission mondiale pour l'éthique des connaissances scientifiques et des technologies de l'UNESCO (COMEST), par quatre attributs essentiels: la mobilité, l'interactivité qui permet aux robots d'agir sur leur environnement, la communication et l'autonomie ou la capacité à « penser » par eux-mêmes et à prendre leurs propres décisions pour agir. Selon ce rapport, « la robotique contemporaine repose de plus en plus sur les technologies de l'intelligence artificielle (IA), qui permettent le développement de capacités de type humain telles que la perception, l'utilisation du langage, l'interaction, la résolution de problèmes, l'apprentissage et même la créativité » (COMEST, Rapport, p. 4).

Le fait que le robot puisse agir de façon plus ou moins autonome a donné naissance à tout un débat éthique sur son libre-arbitre et la responsabilité de ses actes en tant qu'agent moral alors que l'on n'attribuait jusqu'ici cette capacité morale qu'au seul sujet humain. La perspective que l'intelligence humaine puisse un jour fabriquer des machines capables de « penser » par elles-mêmes ou d'éprouver des émotions et des sentiments vient faire basculer les limites entre la pensée humaine et ses objets. Une telle perspective a toutes les qualifications d'un enjeu philosophique qui a toujours accompagné le progrès scientifique et technologique. Une telle aventure vient nous rappeler le mythe biblique de la révolte de l'être humain contre Dieu, révolte qui lui a valu d'être chassé du paradis céleste pour avoir transgressé les limites de la connaissance tracées par son créateur.

Cette prétention robotique de remplacer l'action humaine, et de pouvoir un jour se mettre « à penser » à la place de l'être humain, crée en fait un grand amalgame entre les trois instances de la *Vita activa* telles que délimitées par Hannah Arendt dans *Condition de l'homme moderne*, à savoir: le travail, l'œuvre et l'action. En parlant de l'agir autonome du robot, ne risque-t-on pas de confondre son activité technique avec le concept arendtien d'action, qui est le propre de l'être humain, au sens d'agir collectivement pour créer un monde et improviser un avenir inédit? Ne risque-t-on pas de confondre également le travail que viendrait remplacer le robot avec l'œuvre telle que définie par Arendt dans la lignée de Heidegger?

En effet, en opposant l'œuvre de nos mains (*homo faber*) au travail de nos corps (*homo laborans*), Hannah Arendt explique la durabilité des objets d'usage qui meublent le monde de l'artifice humain. L'usage auquel ces objets se prêtent donne à l'artifice humain cette durabilité qui héberge l'être humain et lui donne la conviction de se dresser contre le temps de sa finitude.

En attribuant aux objets produits par l'artifice humain la durabilité qui défie le temps, Hannah Arendt reprend l'analyse de Heidegger réservée à l'œuvre d'art qui, en advenant, ouvre un Monde dont elle maintient à demeure le règne.

Un monde, ce n'est pas le simple assemblage des choses données, dénombrables et non dénombrables, connues ou inconnues. Un monde, ce n'est pas non plus un cadre figuré qu'on ajouterait à la somme des étants donnés. [...]

Là où se décident les options essentielles de notre Histoire, que nous recueillons ou délaissions, que nous méconnaissions ou mettons à nouveau en question, là s'ordonne un monde (Heidegger, 1962, p. 47).

Entre Hannah Arendt et Heidegger émerge le Monde qui est le propre de l'être humain. La distinction de l'être humain n'est pas ici synonyme d'une supériorité quelconque vis-à-vis des autres étants, elle est plutôt le signe d'une inquiétude, angoisse ou souci de l'être humain devant la mort inéluctable. C'est ce souci d'écarter la mort qui est à la base de toutes les œuvres artistiques aussi éternelles que gratuites et ne visant point l'utile. C'est dans ce sens que les robots les plus sophistiqués et les plus utiles n'ont pas de monde. Hannah Arendt disait :

On déplore souvent la perversion des fins et des moyens dans la société moderne, où les hommes deviennent les esclaves des machines qu'ils ont inventées et s'adaptent aux exigences de ces machines au lieu de les mettre au service des besoins humains : c'est se plaindre de la situation de fait de l'activité de travail. Dans cette situation, où la production consiste avant tout en une préparation à la consommation, la distinction même de la fin et des moyens, si nettement caractéristique des activités de l'*homo faber*, n'a tout simplement aucun sens (Arendt, 1961, p.197).

En effet, quand la société moderne remplace l'œuvre (domaine de la liberté) par le travail (domaine de la nécessité vitale), il devient vain de soulever des questions qui supposent la catégorie de la fin et des

moyens, comme de savoir si les êtres humains vivent et se reproduisent afin de pouvoir recommencer le travail ou s'ils travaillent afin de survivre et de se reproduire. Ceux qui persistent à dire que les robots viennent faire tout à la place de l'être humain sont tenus de préciser ce qu'ils entendent par «faire» et par «penser».

Face à ce cycle de reproduction biologique du travail, l'œuvre et l'action s'inscrivent plutôt dans un jeu avec l'éternité, jeu qui protège l'être humain de ne pas devenir un robot lui-même.

Pour désigner et nuancer l'éternité, la langue arabe nous a fourni le concept de *Sarmadi* pour dire ce qui n'a ni commencement (*Azali*) ni fin (*Abadi*) et qui est réservé en métaphysique à Dieu seul, qui est toujours déjà là au centre de l'univers. Il est à la base de tout questionnement philosophique jusqu'à l'avènement de la modernité et de son avatar, l'humanisme. En comparaison avec Dieu qui n'a ni commencement ni fin, l'homme n'a osé prendre la parole que lorsqu'il a marginalisé Dieu et l'a condamné au silence grâce à «son progrès scientifique» qui mettait la raison humaine au centre de l'univers. Ce retrait de Dieu est décrit par Nietzsche comme étant un meurtre commis par l'homme moderne qui essaie de prendre sa place et ses attributs (Nietzsche, 1971).

C'est dans ce sens que, à la différence de Nietzsche et de Heidegger qui font remonter l'apparition de l'Homme soit à Socrate (Nietzsche, 1971), soit à la république romaine (Heidegger, 1966), Foucault se permet de constater que l'homme n'est pas le plus vieux problème ni le plus constant qui se soit posé au savoir humain. Foucault conclut :

En prenant une chronologie relativement courte et un découpage géographique restreint – la culture européenne depuis la fin du XVI^e siècle – on peut être sûr que l'homme y est une invention récente... L'homme est donc une invention dont l'archéologie de notre pensée montre aisément la date récente, et peut-être la fin prochaine» (Foucault, 1966, p. 298).

Si le Dieu créateur de l'être humain a décidé de le chasser de son paradis céleste parce qu'il voulait savoir plus en se mesurant à lui, il ne savait peut-être pas que sa créature finirait un jour par le chasser à son tour de son paradis terrestre sans arriver, avec tous les robots sans âme qu'il a pu inventer, à échapper lui-même à la mortalité et devenir aussi éternel que Dieu.

Quoi qu'il en soit du bien-fondé de ce mythe, la signification philosophique à en tirer est que l'être humain a toujours flirté avec les limites entre le bien et le mal, et il a toujours été le seul étant, se sachant mortel, qui a trouvé dans l'œuvre et l'action son mode d'être pour surmonter sa mortalité en créant sa propre éternité contre le temps qui le rapproche de sa finitude. Cela a toujours été son champ réservé de jouer avec l'Éternel.

En effet, longtemps avant l'avènement de la modernité occidentale, l'affrontement de la mort a toujours hanté les différentes civilisations avant de devenir la préoccupation de toutes les religions: «*l'épopée de Gilgamesh, ou le grand homme qui ne voulait pas mourir*» des Assyriens ne faisait que rappeler l'épopée grecque qui était destinée à perpétuer l'immortalité du héros qui, en acceptant la mort, ne faisait que la déjouer par ses «actes et paroles» comme disait Hannah Arendt (1961). La parenté entre la parole et la mort n'est que trop manifeste dans le récit arabe des *Mille et Une Nuits*. Foucault avait déjà fait le lien entre le conte, l'écriture et la mort: «on parlait, on racontait jusqu'au petit matin pour écarter la mort, pour repousser cette échéance qui devait fermer la boucle du narrateur. Le récit de Shéhérazade, c'est l'envers acharné du meurtre, c'est l'effort de toutes les nuits pour arriver à maintenir la mort hors du cercle de l'existence.» (Foucault, 1994, p. 793)

On peut conclure en disant que la parole, l'écriture, l'œuvre et l'action sont les derniers subterfuges que l'être humain a découverts pour ne pas mourir, et échapper peut-être à la punition qui lui a été infligée de passer sa vie à travailler comme un robot.

* * *

C'est bien dans le travail et particulièrement à l'usine que les robots se sont multipliés au cours du XX^e siècle, puis au XXI^e siècle dans la logique capitaliste de la mécanisation du travail et de la production de masse. Rapide, fiable, précis, efficace dans des tâches précises et permettant d'augmenter la productivité et les profits d'une entreprise, les robots ont changé l'organisation des usines au point que plusieurs d'entre elles sont entièrement robotisées.

Les robots contrairement aux ouvriers ne se fatiguent pas (ils tombent cependant en panne): ils ne se plaignent pas; ils ne sont jamais distraits dans leur travail; ils ne font pas la grève; ils n'ont pas la gueule de bois le lundi matin... Les robots coûtent moins

cher. Ils sont souvent plus efficaces et précis que les travailleurs humains. Ils n'ont pas de retraite, ni assurance santé, ni droits légaux. Nous voulons que les robots aient toutes les qualités que les maîtres recherchent dans leurs esclaves, les patrons dans leurs employés, les commandants dans leurs soldats, et nous voulons qu'ils n'aient ni leurs faiblesses, ni leurs manques, ni surtout cette tendance irrépressible à l'insubordination, à l'indépendance et à n'en faire qu'à leur tête qui caractérise les travailleurs humains. (Dumouchel, Damiano, 2016, p. 10).

Dans ces usines, les ouvriers ont donc peu à peu disparu, remplacés par des techniciens et des ingénieurs qui veillent au bon fonctionnement et à la réparation des robots. Le travail répétitif et pénible est à la charge des robots, dont croissent les capacités de s'adapter à ce qui les entoure par la mobilité, les capteurs ou la transformation de leurs tâches. Les ouvriers ont appris à se reconvertir dans d'autres activités de production. Or voilà, la mécanisation du travail humain entraîne les robots à sortir de l'usine et à remplacer les travailleurs dans d'autres lieux et pour une variété grandissante de fonctions. Les robots sont devenus, en effet, des machines plus complexes et capables d'interagir avec les gens dans certaines tâches, certes définies, mais plus polyvalentes. Les robots sont donc programmés pour avoir un certain degré d'autonomie avec la capacité de faire des choix, des interactions et des adaptations au contexte extérieur. Les capacités d'apprentissage des robots développées par l'intelligence artificielle, à partir de l'apprentissage profond (*deep learning*), permettent une exploration des possibles avec une efficacité croissante dans la réponse au fil des expériences répétées².

Ce mouvement de robotisation au sein de la société se manifeste dans de nombreux pays du monde, des États-Unis en passant par l'Allemagne, le Japon, et la Chine, et induit une vaste expérimentation sociale à l'échelle mondiale. Depuis environ une dizaine d'années, le public a découvert comment des robots pouvaient être utilisés pour remplacer l'être humain dans un nombre croissant d'activités économiques et politiques. Les robots-drones sont utilisés dans les conflits armés ou encore dans des contextes non armés pour la surveillance

2. Cet apprentissage n'est rendu possible que par les millions de travailleurs « du clic » qui entraînent les algorithmes à travers le monde. Cette face cachée du micro-travail derrière l'intelligence artificielle a été très bien décrite par Antonio Cassili dans son livre « En attendant les robots » (2019).

ou l'assassinat ciblé. Des robots de forme androïde ont fait leur apparition dans le domaine de la santé et du bien-être. Dans les hôpitaux, des robots opèrent sous la direction du chirurgien, d'autres robots aux formes plus androïdes jouent un rôle complémentaire à celui de soignant ou de compagnon pour les personnes âgées. Dans les écoles, certains robots sont utilisés par les enseignants pour l'apprentissage des langues ou des matières scientifiques. Des robots androïdes comme compagne sexuelle ont aussi fait leur apparition. Dans les maisons, des robots de service aspirent la poussière des tapis tandis que des jouets robots s'occupent des enfants. Dans le domaine de l'agriculture, les robots sont utilisés pour traire les vaches et nettoyer l'étable, transformant ainsi les relations entre l'agriculteur et ses animaux.

Un autre des exemples très connus du public est celui de la voiture « autonome » sans chauffeur conçue, entre autres, pour diminuer l'encombrement des voitures en ville. Si elles sont nombreuses et coordonnées, les voitures autonomes permettraient en théorie de régulariser la circulation routière. Introduire des robots-voiture ou robot-autobus dans une ville peut ainsi changer les comportements de transport des citoyens par rapport à l'utilisation de la voiture individuelle, du transport en commun ou du transport actif. Le choix des voitures autonomes a nécessairement un impact sur l'organisation et l'urbanisme des villes – surtout si celles-ci deviennent « intelligentes ou connectées » -, c'est-à-dire si elles multiplient les services et les applications par le biais de capteurs et par le recueil de données massives. Les robots comme les voitures autonomes se conçoivent dans un environnement numérique généralisé qui fournit les données massives, les algorithmes et les machines. Les robots font partie de ce système technique numérique qui tend à nous immerger, à médiatiser nos émotions, nos agissements et nos comportements sociaux, à transformer l'économie, le travail et nos institutions. Le téléphone portable représente une des interfaces essentielles de ce système technique numérique.

Le développement et l'utilisation des robots se sont étendus au point que, dans certains métiers, « il faut vivre avec les robots ». L'ouvrier ou l'employé doit s'adapter et modifier sa manière de faire en interagissant avec les robots. Changeons-nous pour autant notre manière de nous comporter entre humains ? Et si des robots se retrouvent à la maison dans notre intimité ? Si ce sont des robots qui nous soignent ? Si ce sont des robots qui éduquent nos enfants ? Est-ce cela une société robotisée ? Est-ce que ces robots que nous créons entraînent des changements

sociaux et culturels majeurs? Modifient-ils en profondeur la qualité de nos relations humaines? Menacent-ils nos valeurs fondamentales comme la liberté et l'égalité? Qu'en est-il de notre responsabilité dans l'acte de tuer un être humain quand sont utilisés des drones tueurs? Les robots génèrent-ils davantage encore d'exclusion sociale ou de puissance de domination? Et puis, quels impacts la production de robots à grande échelle a-t-elle sur l'environnement et la consommation d'énergie émettrice de CO₂ et cause des changements climatiques?

L'ensemble de ces préoccupations d'ordre philosophique, éthique et politique traverse cet ouvrage collectif. Ce livre est le fruit d'un colloque international et interdisciplinaire³, intitulé « Robots et sociétés: quelles transformations? Quelles régulations? » qui s'est déroulé à l'Université Laval à Québec, les 28 et 29 mars 2017. En marge de ce colloque a eu lieu la réunion du groupe de travail de l'éthique de la robotique de la COMEST qui a déposé son rapport en septembre 2017 lors de sa réunion annuelle à l'UNESCO à Paris. Ce colloque a été une source d'inspiration pour le groupe de travail par la richesse de ses points de vue et le dynamisme des discussions.

* * *

Cet ouvrage collectif est divisé en deux parties. La première partie est intitulée « Le robot comme objet technique: enjeux philosophiques et éthiques ». Elle regroupe neuf chapitres qui analysent différentes dimensions plus théoriques liées à la conception des robots. La deuxième partie dont le titre est « Les robots dans la société: enjeux éthiques et politiques » s'intéresse à l'utilisation des robots dans leurs interactions sociales et contient 12 chapitres.

La première partie du livre s'ouvre sur le chapitre de **Jean-Michel Besnier**. Dans ce *premier chapitre*, J.-M. Besnier soulève la question du robot qui est conçu au départ comme outil devant prolonger les capacités de l'humain et être à son service. Lorsque le robot cesse d'être un simple instrument, il pose alors problème. Le robot cesse d'être un instrument lorsqu'il rend un service qui paraît relever d'une fonction symbolique, c'est-à-dire qui ne se réduit pas à une stricte finalité pragmatique. Le robot pose alors problème parce qu'il semble pouvoir

3. Ce colloque a été financé par une subvention CRSH connexion avec des contributions de l'UNESCO, de l'Université Laval (la faculté de philosophie, La Chaire dans le monde actuel, l'Institut Environnement, Développement et Société, l'Institut d'éthique appliquée) et le Centre de recherche d'éthique publique et de gouvernance de l'Université Saint-Paul.

se substituer à l'humain et non plus seulement l'épauler. Cette façon de poser la question du rapport de l'homme à sa création artificielle permet à J.-M. Besnier de cerner les contours du débat philosophique et éthique que soulève la robotisation de nos sociétés. Dans les utopies posthumaines, conclut J.-M. Besnier, le robot figure l'anticipation d'une humanité qui se serait débarrassée de ses limites en ayant fusionné avec la machine et aboli les oppositions ontologiques ou de genres qui sont censées l'aliéner.

Dans un *deuxième chapitre*, **Vanessa Nurock** pose la question de la morale des machines qui peut s'entendre au sens de notre relation morale aux machines, d'une part, et au sens de la morale que les machines peuvent constituer ou manifester, d'autre part. De ces deux aspects de la morale mécanisée, V. Nurock pose la question du monde commun que nous pouvons souhaiter construire (ou non) avec les machines – ou certaines d'entre elles. Son hypothèse de travail est que la mécanisation de la morale aujourd'hui à l'œuvre trouve ses racines dans un double mouvement : celui engagé par la psychologie morale de Lawrence Kohlberg, d'une part, et celui engagé par les sciences cognitives, d'autre part. Pour démontrer son hypothèse de travail, V. Nurock revisite la généalogie de la morale classique à travers les travaux de John Rawls, Jean Piaget et Lawrence Kohlberg.

Dans un *troisième chapitre*, **Tomislav Bracanović** examine de façon systématique les différentes théories et arguments éthiques permettant ou non d'inclure les robots dans la communauté morale et de les reconnaître comme agents moraux. Reprenant les arguments kantien sur le libre arbitre, utilitaristes sur la capacité de ressentir le plaisir et la douleur, du *care* sur l'inclusion au sein d'une relation ou de l'éthique environnementale sur la valeur intrinsèque des vivants, il analyse comment les différents types de robots (robot programmé, robot cognitif, biorobot) peuvent ou non se voir attribuer un statut moral. Le robot déterministe ne peut pas avoir de statut moral, puisqu'il n'est qu'une machine programmée par des humains pour des tâches précises et répétitives. Le robot cognitif présente par contre des capacités d'apprentissage continu et il peut s'adapter à un nouvel environnement et prendre des décisions de plus en plus complexes et opaques. T. Bracanović conclut que si l'on considère que la rationalité est un critère moral déterminant, il faudrait évaluer dans quelle mesure le robot cognitif pourrait être un candidat sérieux pour obtenir un statut moral. Concernant les biorobots, cette fois-ci, si on utilise le critère du vivant pour attribuer un statut moral, alors la même conclusion

s'impose. Ces considérations métaéthiques, explique T. Bracanović, permettent de voir les contradictions dans lesquelles nous sommes lorsqu'il s'agit d'attribuer au robot un statut moral. À mesure que les robots deviennent omniprésents dans notre vie quotidienne, les débats risquent de se polariser davantage.

Dans un *quatrième chapitre*, **Charles-Étienne Daniel** se demande comment l'on peut baliser le développement de la robotique afin que l'être humain puisse profiter des avantages engendrés par l'introduction de machines de plus en plus sophistiquées, tout en évitant la réalisation de conséquences adverses et indésirables. Or, par un survol des principales critiques adressées au droit positif, C.-É. Daniel aboutit à la conclusion que la rigidité du droit positif face à l'essor de la robotique interactive justifie l'exploration de solutions juridiques alternatives du côté du droit souple (*soft law*). En examinant la nature des autres formes de droit retrouvées dans le concept général de *soft law* et mobilisées par ces réponses normatives alternatives, C.-É. Daniel effectue une brève analyse de la place qui pourrait être accordée aux différents instruments normatifs alternatifs proposés pour encadrer le développement de la robotique.

Dans un *cinquième chapitre*, et après avoir affiché son inquiétude de voir l'intelligence artificielle se substituer à l'intelligence humaine, **Hoda Nehmé** affirme que l'homme postmoderne plonge le monde dans le chaos. L'homme scientifique s'approprie les ingrédients nécessaires à une révolution contre l'homme. Cette révolution prépare la disparition de l'homme, livrant ainsi la terre et la vie au gré d'un enjeu technologique susceptible de jeter dans l'oubli le jardin d'Eden et de rendre caduc le contenu des bibliothèques. Son argumentation aboutit à la conclusion que la robotisation sans limites vient occulter l'extrême pauvreté de l'automatisme.

Au *sixième chapitre*, **Frédéric Dubois** met en question les théories computationnelles de l'esprit d'Hilary Putnam et de Jerry Fodor qui se sont imposées depuis la deuxième moitié du XX^e siècle avec l'idée qu'il serait envisageable dans le futur de reproduire les facultés cognitives humaines en algorithme calculable par ordinateur. Suivant leur leitmotiv, l'ordinateur serait au cerveau ce que l'esprit serait à un programme informatique. Cette hypothèse se laisse facilement interpréter comme une relation d'identité entre le cerveau et l'ordinateur où le cerveau ne serait rien de plus qu'une machine sophistiquée de traitement symbolique. Or, la cybernétique a déjà depuis longtemps défini le

cerveau et les organes sensoriels biologiques comme des outils autorégulés de traitement d'informations. Le fonctionnalisme computationnel n'est finalement qu'un pas de plus vers l'affirmation de cette idée que le corps humain est une machine complexe. C'est à travers l'exemple de l'avion, oiseau d'artifice et argument fort du fonctionnalisme contre le behaviorisme, que F. Dubois retrace les enjeux philosophiques et éthiques soulevés par la prétention du robot à remplacer l'homme sans âme comme l'oiseau d'artifice qui n'arrive pas à voler.

Au septième chapitre, Alexandre Pitti essaie, dans son étude du développement futur des robots humanoïdes, de trouver un antécédent à cet artifice dans les technologies numériques (ordinateurs, cellulaires et Internet), modèle qui peut servir de représentation pour l'évolution future du monde des robots, ainsi que de ses normes et règlements. Internet a déjà changé la façon dont nous organisons notre vie sociale, ainsi que les limites de notre sphère privée. Comme cela s'est déjà produit avec les technologies précédentes, il est tout à fait probable que l'on produise une réglementation définie par la société pour les robots dans les maisons. En adoptant la théorie du Désir mimétique de René Girard, A. Pitti bâtit son raisonnement sur le mimétisme qui renferme les êtres humains qui se sentent en même temps attirés et rejetés par ce qui leur ressemble et les imite. Dans cet état de mimétisme réciproque, il est difficile de comprendre qui est l'original et qui l'imitateur, car la distance entre eux disparaît. Selon cette logique, les robots et l'intelligence artificielle sont nos miroirs modernes, qui nous renvoient nos fantaisies et nos désirs.

Au huitième chapitre, Guillaume Pelletier aborde le sujet du rapport entre artifice et vivant à partir des travaux de Georges Canguilhem. Si on cherche à évaluer comment un être artificiel pourrait éventuellement se rapprocher du vivant, on ne fait qu'aborder cette réflexion à partir du point de vue de la machine. Il semble pertinent, selon G. Pelletier, de considérer le chemin inverse, en examinant le point de vue de certains ingénieurs qui, dans le domaine de la biologie de synthèse, font des efforts considérables pour rapprocher le vivant des machines qui nous sont les plus familières. Comment pourrait-on réduire la complexité du vivant pour éviter ou surmonter des problèmes liés à l'émergence, à l'autoréplication ou aux aléas du contexte cellulaire, dans le but souvent explicite de concevoir des vivants analogues à nos machines électroniques. La mise en perspective de ces deux attitudes semble être un moyen efficace de susciter une réflexion philosophique sur les relations entre l'organisme et la machine. Bien que toutes

deux reconnaissent à la vie une certaine imprévisibilité, l'une y voit un modèle de conception technique alors que l'autre y voit un obstacle à la construction de machines. En s'appuyant sur une réflexion amorcée par Canguilhem, G. Pelletier propose de voir entre ces deux conceptions respectives de l'objet technique, non pas une opposition, mais une continuité biologique.

Au neuvième chapitre, **Chantale Pilon** fait le diagnostic du processus de «mécanisation» de nos sociétés qui semble s'accroître de manière inédite. L'humain ne serait plus seulement comme une machine, il pourrait maintenant «devenir un robot». Le fantasme du progrès scientifique nous promet un monde dominé par les robots et les cyborgs, où l'homme aura quitté son enveloppe charnelle pour devenir un être hybride, augmenté de puces et de prothèses, à mi-chemin entre la machine et le réseau. C. Pilon souligne d'abord le renversement de valeurs que trahit cet engouement pour la robotisation. Cet intérêt suppose, d'un côté, que l'identification à la machine, traditionnellement entendue de façon péjorative et réductrice, soit aujourd'hui perçue positivement, jusqu'à devenir un idéal à atteindre. De l'autre, il implique que l'humanité ait développé un degré suffisamment élevé d'insatisfaction face à sa nature pour souhaiter sa propre transformation et développer des moyens technologiques afin de s'émanciper des limites de sa condition. Autrement dit, la démesure du pouvoir technique ne serait pas la cause de cette volonté de transformer la condition humaine par la robotisation, mais s'inscrirait plutôt, selon C. Pilon, dans l'extension de la logique capitaliste, comme son corolaire.

John Finney inaugure la deuxième partie du livre intitulée «Les robots dans la société: enjeux éthiques et politiques» en analysant, *au dixième chapitre*, les implications éthiques de la perte du contrôle humain sur sa création technique. Il part de la prémisse simple que nous sommes chacun responsable en matière d'éthique et devons chacun pouvoir justifier nos actions. Nos responsabilités sont assez faciles à comprendre, lorsque la chose ou la personne en question se trouve à proximité. Mais qu'advient-il lorsque la relation éthique est interprétée par un dispositif technologique qui pourrait ne pas être complètement contrôlé par un humain? Prenant les exemples des systèmes télécommandés puis celui des systèmes autonomes dans le contexte militaire, J. Finney s'interroge sur la capacité des conventions que sont le droit international humanitaire (DIH) et le droit international des droits de l'homme, à réguler l'utilisation des armes robotisées. Dans son analyse, il utilise les trois principes canoniques

de *distinction*, *proportionnalité* et *imputabilité* et les applique aux armes robotisées contrôlées à distance puis aux armes autonomes afin d'identifier les problèmes éthiques liés à leur application. La question de la responsabilité est centrale et J. Finney rejette la possibilité de déléguer à une machine l'autorisation de tuer. Il montre également comment les armes robotisées ont aussi des implications stratégiques dans la façon de faire la guerre, avec la possibilité d'imaginer un monde où la guerre de faible intensité basée sur des exécutions ciblées et continues serait la règle. J. Finney en appelle à des règles éthiques et juridiques ainsi qu'à notre volonté de réaffirmer le jugement humain dans l'art de la guerre.

Au *onzième chapitre*, et en accord avec la critique adressée par Carl Schmitt et Michel Foucault à l'idée que la guerre est le prolongement de la politique, **Soheil Kash** retrace la logique des guerres irrégulières (kamikaze) pour contourner la sophistication de la guerre postmoderne consistant à tuer à distance (drone militaire) sans faire la guerre. À la lumière de la dialectique hégélienne du maître et de l'esclave fondée sur le jeu avec la mort, S. Kash analyse que c'est toute la logique de Hegel qui se présente à l'envers. Du côté kamikaze, c'est l'engagement intégral, alors que du côté de l'opérateur du drone, c'est le désengagement total; alors que le kamikaze implique la fusion complète du corps du combattant avec son arme, le drone assure leur séparation radicale. L'attentat-suicide implique la mort du kamikaze, alors que le drone l'exclut de façon absolue, et l'opérateur du drone n'a plus besoin d'avoir le profil d'un combattant prêt à risquer sa vie sur le terrain, car il n'a qu'à appuyer sur des boutons dans une salle climatisée du Nevada pour faire disparaître sa cible en Afghanistan ou au Yémen. Sa philosophie est devenue le contraire du maître, qui, chez Hegel, cherche la mort pour imposer sa volonté. Plus que cela, il ne cherche plus, comme dans la dialectique de Hegel, à sauvegarder la vie de l'esclave parce qu'il a besoin de sa reconnaissance, il cherche plutôt à l'anéantir. Entre le kamikaze et le drone, les définitions mêmes de maître et d'esclave sont piégées: le maître de la haute technologie fuit la mort pour faire au kamikaze une chasse à l'homme à distance, en s'épargnant l'affrontement guerrier comme au beau vieux temps des guerres classiques.

Dans le *douzième chapitre*, **Marie-Hélène Parizeau** s'interroge sur ce phénomène en cours où les robots sont « lâchés » dans la société et sont expérimentés par les travailleurs, les consommateurs, les patients, les élèves, les enfants et les gens de tout âge. Elle montre tout d'abord

comment les logiques économiques d'efficacité et de productivité ont multiplié les robots dans les usines depuis 30 ans. S'appuyant sur Gunther Anders, elle illustre comment s'est mis en place une ingénierie humaine du travail, c'est-à-dire, une transformation expérimentale du corps et de l'esprit du travailleur le soumettant à ces nouvelles contraintes de la robotisation qui le conduit à la perte de sa liberté. On peut alors se demander si cette extension de la robotisation industrielle vers les robots sociaux aura aussi des effets sur la transformation des institutions, des rapports sociaux et moraux. M-H Parizeau en prenant les exemples des robots soignants et des robots éducatifs, examine comment ceux-ci s'immiscent dans les lieux (maison, hôpital, école) où s'exprime la relation de la sollicitude, du *care*. Les questions éthiques concernant la finalité, l'utilité, les limites de l'utilisation des robots sociaux questionnent directement la responsabilité humaine dans le développement de la robotique. Dans la foulée du rapport de la COMEST, M-H Parizeau en appelle à un effort normatif similaire à celui qui a prévalu, dans les années 1970, en bioéthique pour l'expérimentation humaine, afin d'encadrer, voire de régler le développement des robots dans la société.

Dans le *treizième chapitre*, Josée-Anne Gagnon examine l'utilisation des robots dans la thérapie comportementale du spectre de l'autisme. La socialisation de chacun de ces enfants atteints du trouble du spectre de l'autisme (TSA) représente une difficulté puisqu'ils ont des relations avec les pairs qui semblent absentes, inadéquates ou étranges. L'utilisation des robots dans la thérapie des enfants qui présentent un TSA étant donné leur attrait pour les objets semble logique. Avec un comportement programmé pour être d'humeur égale et prévisible, le robot constitue un atout indéniable pour capter l'attention de ces enfants ce qui aiderait à établir des interactions entre l'enfant et le robot. Encore faut-il cependant que cet effet bénéfique soit reproductible à plus ou moins long terme dans les interrelations de l'enfant présentant un TSA avec de vrais humains. J.-A. Gagnon distingue ensuite la communication de la création d'une relation vivante et dynamique. La communication est l'échange d'un message entre des interlocuteurs, ce que peuvent faire les robots par programmation de phrases circonstanciées joyeuses ou neutres. Cependant, la relation est un lien qui se crée entre deux personnes qui communiquent de manière signifiante l'une pour l'autre, impliquant un référent émotif et une transmission d'émotions par rapport à ces référents. La difficulté relationnelle de l'enfant qui présente un TSA

et le but de la thérapie portent sur cette composante émotive de la relation humain-humain. Or, le robot n'entre pas en relation au sens humain du terme. Il est un objet animé programmé qui parle à un enfant et cet enfant lui parle justement parce qu'il n'est pas humain. Est-il éthique, conclut-elle, d'utiliser des robots comme figure de l'être humain relationnel ?

Dans le *quatorzième chapitre*, **Félix-Antoine Pageau** s'interroge, à l'autre bout du spectre, sur l'utilisation des robots soignants auprès des personnes âgées qui ont des problèmes de mobilité ou de démence. La sécurité est un élément primordial en faveur de l'intégration des robots dans les soins en gériatrie. F.-A. Pageau, dans une analyse de la littérature, met en évidence comment l'utilisation des robots relance la tension entre la sécurité et l'autonomie de la personne âgée, pose des problèmes relatifs à la surveillance et à la vie privée, et, enfin, interroge les formes de sociabilité et d'entraide familiale. « Ne t'inquiète pas pour Mamie, elle a le robot pour lui tenir compagnie » est une phrase indicatrice des phénomènes de dérives que F.-A. Pageau examine. Il s'interroge sur le type de société que nous avons bâti, sur la qualité de nos relations familiales, sur la dévalorisation de l'entraide sociale et étatique dans les soins aux personnes âgées. Il conclut que seul le robot-outil demeure pertinent pour réduire la charge de travail des professionnels de la santé et exclut pour des raisons éthiques l'utilisation du robot-soignant.

Le *quinzième chapitre* porte également sur l'utilisation des robots chez les personnes âgées. **Monique Lanoix** y examine dans ce contexte l'impact du robot PARO sur les relations humaines et de soins. Elle met en évidence ce paradoxe : le robot compagnon tend à isoler davantage la personne en milieu institutionnel, en offrant un semblant d'attachement et surtout en procurant le calme. Or, les lieux d'hébergements spécialisés mettent déjà les personnes âgées en marge de la société, dans un lieu à l'écart. En s'appuyant sur les analyses d'Helmut Plessner, M. Lanoix montre que les robots de compagnie peuvent faire obstacle à la relation avec les autres humains par l'appauvrissement de la vie intérieure, et par la perte progressive de la capacité de décentrement et d'appréhension de la diversité du monde. Elle constate également que les robots de compagnie ont un impact sur le travail d'aidant. Puisque le travail de préposé aux bénéficiaires est conçu selon le modèle du travail à la chaîne, les dimensions relationnelle et émotionnelle des soins sont de plus en plus réduites. Introduire des robots compagnons, comme PARO, ne fera qu'aliéner davantage le travail du préposé,

explique M. Lanoix. En s'appuyant sur Giorgio Agamben, elle explique que le travail d'aidant est traditionnellement considéré comme non productif et qu'en devenant salarié, il s'est inscrit dans une logique de productivité où l'aidant devient un outil. Elle conclut que les soins qui ont trait aux émotions sont aliénés du travailleur et resitués dans le personnage du robot. Le travail de l'aidant devient davantage un travail instrumental, car toute communication et tout échange seront ciblés sur les besoins corporels. C'est l'aidant qui devient un robot.

Le *seizième chapitre*, de **Raoul Kamga** et **Margarida Romero**, aborde l'utilisation des robots dans l'éducation. Les auteurs montrent que l'usage des robots dans l'éducation s'est diversifié, par exemple avec l'usage des robots pour l'apprentissage des sciences ou des langues, la résolution de problèmes en mathématique. Il existe également des activités de robotique pédagogique qui engagent l'apprenant dans un défi d'une certaine complexité, puisque ce dernier doit alors faire des démarches de résolution de problèmes. Plusieurs problèmes se posent dans l'utilisation des robots dans l'éducation : la formation des enseignants, le coût des technologies robotiques dans un secteur où l'argent de l'État se fait rare. La question éthique de l'égalité des chances liées à l'accès à une éducation de qualité pour tous les apprenants est posée. Est-ce une pratique réservée ou des outils destinés aux pays « riches » par rapport aux pays « pauvres » ? Précisément aux écoles « riches » et privées ?

Au *dix-septième chapitre*, **Sophie Cloutier** aborde la question des robots destinés aux enfants en prenant comme point de départ la critique d'Hannah Arendt de la science moderne, qui contribue selon elle, à l'aliénation par rapport au monde. En effet, S. Cloutier analyse les animaux robots comme une forme de technologie symptomatique de cette aliénation au monde. En s'appuyant notamment sur les recherches de Sherry Turkle, elle s'interroge sur l'authenticité, la réciprocité et les interactions émotionnelles entre les humains et les robots. De cette analyse, il ressort un individu moderne aliéné au monde, rejeté en lui-même, ayant perdu la capacité de supporter l'étrangeté, et qui ne trouve de réconfort émotionnel que dans la machine qu'il peut contrôler. Il se dégage en filigrane une opposition entre une relation avec un être vivant, où il faut s'ouvrir à l'autre pour établir une réciprocité, et une relation avec un objet qui n'est qu'un miroir de soi-même. S. Cloutier insiste sur le fait que les relations de qualité avec les animaux contribuent au développement et à l'actualisation du potentiel éthique des enfants et pourraient ainsi servir de rempart

à l'aliénation, contrairement à ce que pourraient offrir les robots, qui risquent de reconduire et d'alimenter l'aliénation et le mépris pour notre condition humaine vulnérable et mortelle.

Dans le *dix-huitième chapitre*, **Sandra Blouin** analyse la relation humain-animal à partir de l'utilisation des robots de traite dans les fermes laitières. Ces robots participent à une mécanisation et à une automatisation de plus en plus poussée dans l'agriculture et la production animale. Si ces équipements peuvent faire miroiter un travail plus facile pour les fermiers, qui se font remplacer par des robots, qu'en est-il de la relation entre l'agriculteur et ses vaches lorsqu'il introduit un robot de traite ? S. Blouin analyse donc les avantages et les inconvénients des robots de traite et les adaptations nécessaires pour les fermiers et les vaches. Les robots de traite imposent finalement un autre modèle de gestion de ferme qui vise à optimiser la production et aider les agriculteurs à prendre les meilleures décisions à partir des données fournies par le robot sur les vaches. Cette médiation par les chiffres éloigne l'agriculteur de sa connaissance individuelle de chacune de ses vaches, fondée sur son expérience. S'appuyant sur Dominique Lestel, S. Blouin montre que les rapports coopératifs entre humains et animaux induisent des aspects relationnels de part et d'autre. Or les robots de traite déshabituent les animaux à l'humain, puisque le contact est réduit et médiatisé par le robot. La traite et l'alimentation des vaches ne sont plus le moment d'échanges agréables entre l'agriculteur et ses vaches. Qu'en est-il alors du bien-être animal ? S. Blouin conclut que si toutes les tâches de la ferme ne sont pas encore robotisées, il devient essentiel de préserver une bonne relation avec l'animal pour qu'il se sente en confiance et collabore bien.

Le *dix-neuvième chapitre* s'intéresse à la comparaison entre les animaux et les robots. Dans ce chapitre, **Georges Chapouthier** soulève la différence radicale qui existe entre les deux et ce, malgré le fait qu'ils possèdent tous les deux une certaine intelligence. Par rapport aux animaux, les robots actuels, malgré leur complexité, n'atteignent pas un niveau de conscience, probablement lié à la possibilité qu'ont les animaux de dialoguer en permanence avec leur environnement. Étant donné que les robots actuels sont dépourvus de la conscience de la douleur et d'un vécu émotionnel qui lui est lié, ils sont loin de partager avec l'être humain les bases de la morale. L'être humain est un être moral dans la mesure où il partage les traits de ses cousins animaux, et échappe ainsi à la robotisation

Au *vingtième chapitre*, **Louis-Étienne Pigeon** pose la problématique à la fois ontologique et éthique du rapport de l'être humain à l'artefact qu'il a créé entre la biosphère, la partie vivante de la planète, et la technosphère, l'ensemble de la nature transformée par l'être humain. Ces deux concepts ont été élaborés par le savant russe Vladimir Ivanovitch Vernadski. L.-E. Pigeon considère que cette catégorisation permet de bien décrire la problématique de la robotisation des environnements. Il explique le phénomène de la robotisation des environnements dans le contexte de l'émergence de la technosphère, pour spécifier ensuite ce qu'il entend par la robotisation des corps et des environnements. Il conclut que le processus de robotisation des milieux, propre à l'émergence de la technosphère comme environnement global, doit se juger à l'aune de deux problématiques, l'une ontologique et l'autre éthique ou normative. La société dans laquelle nous vivons est de plus en plus déterminée par les processus automatisés et les machines qui les rendent opérationnels. La construction de la technosphère ne relève pas d'un « plan » préétabli, mais d'un foisonnement culturel propre à une société désormais qualifiée de « technoscientifique ». La question devient donc : savons-nous réellement ce que nous produisons ? Ou alors, ne sommes-nous pas simplement balancés par les rêves de la technoscience ?

Au *vingt-unième chapitre*, **Catherine Larrère** soulève des questions pertinentes concernant la place que peut avoir l'automatisation robotique en marche dans la transition écologique. En examinant ces questions, elle cherche à montrer les limites du débat sur le revenu universel, envisagé comme la médiation entre l'automatisation et la transition écologique. Ceci la conduit à la question centrale, rarement posée, de l'empreinte écologique des robots ainsi qu'à la réflexion éthique et philosophique suscitée par la robotisation de la société. C. Larrère constate au départ que le capitalisme industriel a fait place au capitalisme cognitif, ou à une économie de la connaissance, apte à faire des connaissances, comme plus largement du vivant, une production rentable, insérée dans les circuits économiques. Cela a été rendu possible par l'innovation technologique du numérique, qui s'est imposée, dès les années 1980, avec les ordinateurs personnels et Internet. L'intelligence artificielle et les robots peuvent être ainsi vus comme participant à la fois « d'une deuxième vague de déploiement du capitalisme cognitif », et « d'un deuxième âge des machines », celui de la mécanisation de l'esprit. C'est ainsi que la robotisation tend à se généraliser pour couvrir toutes les tâches du travail des sociétés

modernes, généralisation accompagnée d'une diminution importante des emplois qui touche même les classes moyennes. S'il apparaît ainsi que la transition écologique peut tirer indirectement avantage, grâce au revenu universel, de l'automatisation croissante et des pertes d'emplois qu'elle entraîne, il ne s'agira plus alors de savoir dans quel type de société les robots peuvent trouver une place, mais quelle place une société écologique peut-elle offrir aux robots ? Il s'agit donc de savoir comment nous devons nous comporter avec les robots que nous avons installés parmi nous. En éthique animale comme en éthique environnementale, la question s'est posée de l'extension aux non-humains de la communauté morale, jusque-là réservée aux seuls humains, et on a recherché les qualités qui permettent de se voir attribuer une place dans la communauté. Or, nous reconnaissons si peu aux robots un être qui leur serait propre que la seule possibilité que nous leur laissons est celle de se faire passer pour ce qu'ils ne sont pas, mais que nous les poussons à être ou à prétendre être : nous-mêmes. Ce sont des imposteurs par destination. Il faut qu'ils nous ressemblent le plus possible ; ce faisant, ils menacent de prendre notre place. Il faut donc trouver la bonne distance. Quelle place réserver aux robots dans la transition écologique ? Catherine Larrère laisse la question ouverte en concluant : Ni animaux ni humains, des robots : une relation à étudier. Une chance à leur donner.

* * *

Peut-être le lecteur trouvera-t-il cet ouvrage collectif très interrogatif sur l'utilisation actuelle et future des robots dans nos sociétés et au sein de notre oïkos (maisonnée). Dans cette quatrième révolution que promettent l'intelligence artificielle et le numérique, il convient, nous semble-t-il, de prendre une distance critique face à cette fascination technologique anthropocentrique afin de créer des conditions de choix individuel et collectif sur le genre de société ou de milieux dans lesquels nous voulons vivre, avec ou à bonne distance des robots !

BIBLIOGRAPHIE

- Arendt, Hannah (1961). *Condition de l'homme moderne*, Paris, Calmann-Lévy.
- Cervantès, Miguel de (2001). *L'ingénieur Hidalgo Don Quichotte de la Manche*, Paris, Seuil, tomes 1 et 2.
- Casilli, Antonio, (2019). *En attendant les robots. Enquête sur le travail du clic*, Paris, Seuil.

- COMEST/UNESCO (2017). *Rapport sur l'Éthique de la robotique*. Consulté en ligne.
- Dumouchel, Paul, Damiano, Luisa (2016). *Vivre avec les robots. Essai sur l'empathie artificielle*, Paris, Seuil.
- Foucault, Michel (1966). *Les mots et les choses*, Paris, NRF.
- Foucault, Michel (1994). *Dits et écrits vol I (1954-1969)*. Paris, Gallimard, chapitre: Qu'est-ce qu'un auteur?».
- Heidegger, Martin (1962). *Chemins qui ne mènent nulle part*, chapitre «L'origine de l'œuvre d'art». Paris, Gallimard.
- Heidegger, Martin (1966). *Lettre sur l'humanisme*. Paris, NRF.
- Ibn Khaldoun (2002). *Al-Muqaddima, Le Livre des exemples*, Paris, Paris, Gallimard collection La Pléiade. Tome 1.
- L'épopée de Gilgamesh, Le grand homme qui ne voulait pas mourir* (1992). Traduit et adapté par Jean Bottéro, Paris NRF.
- Musil, Robert (1995). *L'homme sans qualités*. Paris, Seuil, tomes 1 et 2
- Nietzsche, Friedrich (1971). *Ainsi parlait Zarathoustra*. Paris, NRF.
- Nietzsche, Friedrich (1974). *L'Antéchrist*. Paris, NRF.

PREMIÈRE PARTIE

**Le robot comme objet technique :
enjeux philosophiques et éthiques**

CHAPITRE 1

Le robot, alibi d'une humanité fatiguée d'elle-même

Jean-Michel BESNIER

Le robot a aujourd'hui de multiples visages. Pour peu que nous ayons affaire à une machine qui paraît jouir d'une certaine mobilité ainsi que d'une apparente autonomie, on la classe parmi les robots. Ainsi, un aspirateur, un drone mais aussi un moteur de recherche sur Internet sont-ils en ce sens des robots.

Il n'empêche que notre imaginaire privilégie le robot androïde (Nao, Affetto, Takéo...). On demande parfois aux roboticiens pourquoi ils s'obstinent à vouloir donner forme humaine à leurs machines, alors qu'elles rendraient d'aussi bons services si elles ressemblaient à des lessiveuses sur roulettes. On se trompe quand on croit qu'il y a là-dedans une ambition démiurgique. La réponse est souvent d'une grande platitude et elle témoigne d'un finalisme archaïque : sélectionnée par l'évolution biologique, notre morphologie serait adaptée aux exigences de notre environnement, de sorte que pour être aussi efficaces que nous le sommes dans ce même environnement, nos robots ne pourraient faire moins que comporter une tête avec des yeux, ainsi que des bras et des jambes. Tant pis pour la poésie ! Le robot humanoïde restera avant tout un instrument de travail, au même titre que l'humain accomplit les tâches que son évolution morphodynamique lui a rendues accessibles.

S'il est un simple outil, cela veut dire que le robot doit prolonger les capacités de l'humain. On peut se réjouir que des exosquelettes permettent une mobilité accrue des soldats sur le terrain des combats ou le déplacement de malades grabataires dans les services hospitaliers. Le robot autorise aussi des activités qui seraient dangereuses pour les humains et, en ce sens, il rend le service qu'on attend de n'importe quel outil : il intervient dans le cœur d'un réacteur nucléaire, sur la scène de catastrophes naturelles ou, plus simplement, sur les chaînes de montage automobile. C'est quand le robot cesse d'être seulement un instrument qu'il pose un problème et il cesse de l'être au moment où il rend un service qui paraît relever d'une fonction symbolique, c'est-à-dire qui ne se réduit pas à une stricte finalité pragmatique. Il pose alors problème parce qu'il semble pouvoir se substituer à l'humain et non plus seulement l'épauler. On dit alors qu'il devient un compagnon, et en tant que tel, il peut devenir « abusif » : une unité de soins palliatifs, un service psychiatrique ou simplement une école maternelle qui seraient administrés par des machines, aussi empathiques fussent-elles, annonceraient en effet un monde dans lequel l'humanité deviendrait superflue. Le remplacement de l'humain par les robots devient dès lors un problème qui relève autant du niveau sociopolitique que de la métaphysique.

La comparaison avec l'animal a tenté quelques auteurs de science-fiction comme Simak dans *Demain les chiens* (1970 [1952]). Personne ne doute, en effet, que l'animal a contribué à domestiquer l'humain et on a pu imaginer qu'il pourrait lui survivre. Le chien impose des rituels de promenades à son maître, quand il ne lui dicte pas une manière de parler sinon de penser. Le robot de service s'annonce déjà comme un être avec lequel un dialogue sera nécessaire. C'est comme s'il constituait à nos yeux un point de vue sur le monde, au même titre qu'un chat ou un chien, et qu'il appelait à ce titre au dialogue. C'est en quoi il n'est pas un pur outil mais potentiellement un partenaire devant partager avec nous une culture qui, comme toute culture, nous transforme. Comment éviterons-nous de changer quand nous serons entourés de machines capables de mobiliser un langage et de répondre à nos attentes ? Mais, de même que le propriétaire d'un chien cède à des comportements qui paraissent plutôt élémentaires au regard des raffinements qu'exige la communication humaine, de même il pourrait se faire que la fréquentation permanente de nos robots se solde par une simplification de nos gestes, de notre langage et de notre façon d'envisager le monde. Les neurosciences diagnostiqueront peut-être

un amoindrissement de la plasticité cérébrale chez ceux qui auront écarté leurs semblables au profit des machines.

Déjà, le raisonnement sommaire des informaticiens surprend dans les situations compliquées que ne manquent pas de produire les relations humaines. Voyez comment ils s'expriment au sujet de la conscience dont ils pourraient vouloir doter leurs machines : une simple heuristique réflexive mise en acte par une expérience ponctuée d'essais et d'erreurs. Rien de plus. Le comportementalisme issu du positivisme nous avait habitués à cette réduction de la vie mentale au schéma stimulus-réponse, la robotique s'en satisfait parfaitement et prétend équiper les robots de l'avenir de l'équivalent de cette conscience réduite au minimum.

Comment en est-on venu à construire une représentation du robot comme cet équivalent d'un être conscient et, par suite, comme l'équivalent d'un être social ? Un être digne d'apparaître comme partenaire dans un dialogue interpersonnel, susceptible par là-même de s'attirer des sentiments d'empathie – des sentiments qui peuvent être vécus comme dérangeants ainsi que le démontre Masahiro Mori dans la « La vallée de l'étrange » (2012). Un être qui n'a pas besoin d'être support de conscience pour mériter qu'on le considère comme accessible au droit, à l'éthique et à la morale. Il faut bien considérer que devenant cet être, il n'est pas déraisonnable d'accepter que le robot soit appelé à remplacer l'humain.

Il y a donc de bonnes raisons pour que le robot constitue désormais un thème de réflexion philosophique pris au sérieux. Ce qui n'était qu'affaire de science-fiction touche à présent au plus sensible de nos vies quotidiennes. Ainsi, on s'interroge sur l'avenir du travail dans l'éventualité que les robots prennent en charge les activités professionnelles que nous exerçons encore (pas seulement les activités pénibles ou répétitives mais aussi celles qui demandent du jugement). Un rapport de l'OCDE, en octobre 2017, établit que 13 % seulement de ces activités professionnelles seraient mieux réalisées par des humains que par des robots. Ainsi, convaincue par les promesses d'une robotique de service, on voit des communautés d'agglomération (comme Lyon) mettre en place les conditions d'un dialogue sur la robotique avec la société urbaine. Autre indice, la Direction générale des Armées (DGA) a commandité un rapport sur l'éthique des robots (le rapport Crémieux) qui paraît fort surprenant mais qui ne fait qu'affronter culturellement les défis portés par la multiplication des machines

dans les conflits armés : le récit de certains démineurs en Irak ayant confessé la sympathie qu'ils avaient fini par éprouver pour leurs robots a même accredité l'animisme dont les Modernes que nous sommes ne se sont jamais totalement débarrassés (Latour, 1991). Enfin, les crises financières qui menacent l'équilibre du monde sont désormais largement imputées à quelques robots traders dont la réactivité interdit le contrôle des marchés. L'avenir de la civilisation n'aura jamais paru autant suspendu aux performances de ses machines.

Tout se passe comme si le robot devenait le support de toutes les anticipations et de l'imaginaire sociétal que les politiques sont incapables de polariser. Le robot est devenu le symbole des sociétés qui subordonnent leur prospérité au développement des automatismes de toutes sortes. La moindre entreprise de services se glorifie aujourd'hui de la mise en place de serveurs vocaux pour satisfaire ses clients et l'on ne s'inquiète pas de se dire qu'il n'est plus d'autres choix offerts à ces clients que d'ajuster leurs comportements aux exigences de ces machines supposées leur faciliter la communication... Derrière le robot, c'est la mutation opérée par les technologies sur nos représentations, nos croyances et nos idéaux, qui s'affiche sans arrière-pensées. En ce sens, il est un analyseur de l'esprit du temps.

L'association du robot aux utopies posthumaines est à cet égard un signe tout à fait éloquent : qu'il soit androïde ou pas, le robot est présenté comme une invitation à élargir le concept d'humanité au-delà des seuls hommes et femmes qui y prétendaient depuis toujours. Avec le robot, le clone, le cyborg, l'animal sont désormais les candidats à cette humanité élargie dont il faut d'ores et déjà envisager le système de valeurs qui la régulera. Le dialogue avec le robot, cette machine qui suscite l'empathie et nous impose de faire l'épreuve de notre sensibilité à l'égard des autres, fait l'objet de l'interrogation des philosophes dont la vie conceptuelle n'est pas trop frileuse. Le Japonais Ishiguro, dont les machines fascinent autant qu'elles effraient, le disait simplement : la relation aux robots androïdes en dit d'abord long sur la relation que les humains entretiennent entre eux. N'en est-il pas ainsi de la relation que nous établissons avec les animaux ? La chose est tout de même étrange, n'en doutons pas : on oublie que la relation aux robots ne tient pas davantage qu'à l'échange de signaux (même élaborés) et on soutient que cette relation devrait permettre d'éclairer celle que les humains pourraient se proposer. C'était le thème directeur de *Real Humans*, une série suédoise télévisée qui fut accueillie en France avec une certaine stupeur. Ma question est la suivante : Jusqu'où se

laissera-t-on simplifier par les machines dans l'image que nous nous faisons de nous-mêmes ?

Dans les utopies posthumaines, le robot figure l'anticipation d'une humanité qui se serait débarrassée de ses limites en ayant fusionné avec la machine et aboli les oppositions ontologiques ou de genres qui sont censées nous aliéner. On reconnaît là aussi bien l'annonce de Donna Haraway dans son *Manifeste Cyborg* en 1986 que celle de la Singularité exploitée par Ray Kurzweil, dans *L'Humanité 2.0*. Mais cet appel fait au robot pour délivrer l'humanité d'elle-même est évidemment ancien. On en trouve un précédent célèbre avec le futurisme italien et les Manifestes de Marinetti en faveur de l'homme d'acier qui nous permettra de nous débarrasser de la nature en nous, c'est-à-dire de la passivité dont la femme, au passage, est supposée être le désolant symbole.

C'est bien par-là que s'impose le robot dans notre paysage culturel : il incarne, si j'ose dire, l'idéal de non-fragilité que nos sociétés aimeraient tant réaliser. Il est aussi l'emblème d'une absence d'ambiguïté que toutes nos technologies qui exigent la transparence voudraient illustrer, au prix de la simplification extrême de l'humain que je viens de mentionner. Avec le robot, on aimerait s'exempter du « ratage » originel qu'est l'homme, à en croire le mythe de Prométhée qui nous rappelle que nous sommes le fruit de l'étourderie d'Épiméthée et qu'à cet égard nous n'aurons jamais fini de recourir aux expédients destinés à compenser le manque des qualités dont bénéficient par nature les animaux pour leur survie. Nous sommes des « ratés » – d'éternels prématurés – et la technique vise à remédier à la néoténie dont nous souffrons. Au lieu de conclure, comme Roger Caillois jadis, que ce ratage originel est l'indice de la dignité de notre humanité, puisqu'il explique que nous soyons des êtres offerts à l'histoire et à la construction volontaire de soi, les thuriféraires du posthumanisme nous expliquent qu'en fusionnant avec la machine, nous allons le dépasser – et cesser donc d'être ces pauvres humains en manque de perfection (animale ou divine). Le robot endosse donc la responsabilité de nous arracher à nous-mêmes, non pas comme y viserait l'éducation du genre humain telle qu'on la pense au XVIII^e siècle mais comme une transgression susceptible de nous faire basculer dans l'autre de l'humain, dans la « Singularité ».

Qu'on ne s'étonne donc pas si le robot est si communément invoqué comme la manifestation du triomphe de l'automatisme sur la réflexion et la conscience délibérative, lesquelles freinent l'action et la prise

de décision. Le robot, c'est l'expert par excellence. On aime chez les experts qu'ils soient intuitifs, que leur cerveau sache mobiliser les algorithmes et les routines cognitives qui sont le gage de l'efficacité. On a rêvé très tôt dans l'histoire de l'IA de « systèmes experts » qui rempliraient le programme attendu d'experts sans faille. On ne s'étonne même plus que les Coréens du Sud se disposent à équiper leurs écoles maternelles de ces robots-enseignants qui sauront administrer un savoir sans réplique et sans hésitation. Les tâtonnements et les humeurs des maîtresses d'école ne sauraient faire le poids devant ces machines à enseigner et à déjouer les problèmes relationnels qui conduisent tant d'élèves à l'échec. On ne se demande pas encore ce qu'il pourrait advenir de ces enfants, élevés et éduqués par des robots au comportement efficace et neutre, lorsqu'ils devront pénétrer le monde chaotique et capricieux des adultes... C'est peut-être parce qu'on se dit que ce monde va disparaître sous la pression des technologies.

Le robot est un être sans intériorité, voilà sa force pour qui souhaite optimiser la synergie des humains. Skinner, le psychologue du comportement, a décrit dans une utopie intitulée *Walden II* le fonctionnement d'une société qui en aurait fini avec les états d'âme de ses membres et qui n'obéiraient plus qu'aux lois béhavioristes. Les spéculations contemporaines sur la robotique, même si elles sont plus fines, participent de la même inspiration : permettre de manager les relations interindividuelles en faisant comme si chacun était une boîte noire offerte aux jeux des stimuli et des réponses dans lesquels le robot excelle. Les sciences cognitives prétendent avoir résolu le « mind-body problem » et accrédi-ter le bien-fondé du test de Turing : il n'y a pas de « fantôme dans la machine », comme dit Gilbert Ryle ; se comporter adéquatement dans une situation problématique, c'est être intelligent sans abus de langage, comme le soutint Alan Turing ; éprouver des émotions, c'est seulement émettre des signaux susceptibles d'être décodés et imités par des robots. Il n'en faut pas plus pour étayer les prétentions épistémologiques des roboticiens et faire de leur laboratoire le lieu où l'on expérimente le caractère superflu de l'hypothèse de l'âme (ou de l'esprit). Les post-humanistes en tirent désormais tous les bénéfices : l'idéal d'une fusion avec les machines a cessé d'être sulfureux, si l'âme n'est plus qu'une illusion archaïque réfutée par les sciences. Un matérialisme sommaire a triomphé, qui arrange bien notre société productiviste et concurrentielle.

Paraîtra-t-il oiseux d'avancer une thèse – que mon livre *Demain les posthumains* (Besnier, 2010) a croisée ? L'intérêt pour le robot, dans nos

sociétés modernes, s'étaye sur le constat que l'humanité est fatiguée d'elle-même, qu'elle éprouve un sentiment croissant d'impuissance devant les performances acquises par ses machines – sentiment que le philosophe G. Anders baptisa, en 1956, «honte prométhéenne d'être soi». Le témoignage de nombreux discours tenus par les propagandistes du transhumanisme apporterait de l'eau au moulin de cette thèse: le robot est le dernier idéal du moi aux yeux d'une humanité qui éprouve comme une faiblesse le fait de naître par hasard, au lieu d'être fabriquée avec la rigueur déployée par les ingénieurs high tech d'aujourd'hui. Tout ce dont rêve l'être débile et fatigué d'être soi se trouve ainsi projeté sur le robot tout-puissant. Jean-Michel Truong le disait avec une certaine brutalité: après les désastres provoqués dans le monde et l'histoire par l'humanité – après Auchwitz et Hisroshima, en particulier –, on ne saurait vouloir que l'avenir ait le visage de l'homme. La sagesse est de vouloir ménager la place au Successeur de l'homme, sachant que c'est la technologie qui le prépare. Comble de cette lassitude d'être humain, il n'est plus temps de vouloir aller plus loin. Le moment est venu de souhaiter la rupture qui livrera l'inédit grâce à quoi autre chose commencera.

En ce sens, s'est-on avisé des représentations volontiers données du robot dans la littérature ou le cinéma? Un être serein comme un Bouddha, dépourvu de passions autant que de liberté. Le contraire même des agités et des déprimés que nous sommes. «Plus je connais les hommes, plus j'aime les machines»: la formule connaît un espace d'applications appelé à croître, surtout si le robot perfectionne une apparence susceptible de s'attirer l'empathie que nos semblables éveilleraient de moins en moins...

Il y a un petit nombre de philosophes pour s'inquiéter de cette tendance: le robot signe la mort du symbolique traditionnellement attaché à l'humanité. Il nous contraint, disent certains, à une dangereuse «désublimation»: ainsi en faisant prévaloir un langage-signal sur l'échange des signes porteurs de conversations, il règle son compte à l'idée soutenue par la paléanthropologie selon laquelle l'hominisation se serait faite grâce à la co-évolution de la parole et de l'outil. Désormais, la part accordée à la technologie instrumentalise le langage au point de le réduire à la seule transmission des informations efficaces, parce que numérisables. Le thème est souvent décliné et référé à la novlangue d'Orwell dont on n'aura jamais si bien compris les inquiétudes. Le robot triomphant parlera naturellement la novlangue et déjà,

nous oblige aussi à la parler. Comment obtiendrait-on de lui les services qu'on en attend si on ne parlait pas sa langue ? Son langage sans ambiguïté ni mensonge signera la mort de l'humanité si l'on pense, comme George Steiner, que celle-ci naît avec la faculté découverte par les hommes de dire ce qui n'est pas et de ne pas se laisser imposer la réalité comme seule norme. Ajoutons, pour finir, que le robot dont la mémoire sera évidemment illimitée nous éloignera définitivement de la culture qui a besoin de l'oubli et de la déformation du souvenir qui résulte en général d'échanges humains, trop humains.

BIBLIOGRAPHIE

- Besnier, Jean-Michel (2010). *Demain les posthumains: Le futur a-t-il encore besoin de nous?* Paris, Éditions Fayard.
- Latour, Bruno (1991). *Nous n'avons jamais été modernes: Essai d'anthropologie symétrique*. Paris, Éditions La découverte.
- Mori, Masahiro (2012). « La vallée de l'étrange », dans *Gradhiva*, vol. 15, n° 1, p. 26-33.
- Simak, Clifford D. (1970 [1952]). *Demain les chiens*. Paris, Éditions J'ai lu.

CHAPITRE 2

Généalogie de la morale automatisée

Vanessa NUROCK

La morale peut-elle être automatisée? Cette question, très classique, se pose aujourd’hui avec d’autant plus d’acuité que la présence des machines – restées jusqu’à présent cantonnées à certaines tâches et encore rarement implantées au sein du corps humain – est aujourd’hui en train de s’étendre graduellement jusqu’à des lieux, des relations, des usages susceptibles de mettre en jeu l’éthique, que ce soit dans les prises de décision, les actions ou encore le type de relations instituées. Ces machines répondent également à des usages très divers et s’insinuent dans tous les aspects de notre vie collective, familiale ou personnelle, des véhicules autonomes aux « machines à tuer », en passant par les robots de soin à la personne ou les implants cérébraux.

Une telle situation rend donc aujourd’hui, plus que jamais, nécessaire de poser la question de la morale des machines. Celle-ci peut s’entendre en deux sens : premièrement, celui de notre *relation morale aux machines*; deuxièmement, la question de *la morale que les machines peuvent constituer* ou manifester. Mais ces deux questions débouchent logiquement sur une troisième : celle du *monde commun* que nous pouvons souhaiter construire (ou non) avec les machines – ou certaines d’entre elles –, c’est-à-dire *l’oïkos* que nous souhaitons développer, au sens propre du terme. On parle non seulement des habitants de notre maisonnée, mais également du type de relations, voire de hiérarchies,

que nous désirons ou non mettre en place et entretenir, y compris avec les machines ou par l'entremise des machines.

Pour autant, avant de pouvoir répondre à ces questions, il importe de chercher à analyser le processus à l'œuvre dans la mécanisation de la morale ainsi que la vision du monde qui y préside. En d'autres termes, la question que je souhaite poser est celle de la généalogie de l'automatisation de la morale.

Mon hypothèse de travail est que l'automatisation de la morale trouve ses racines dans un double mouvement : celui engagé par la psychologie morale de Lawrence Kohlberg dans la suite de Piaget, d'une part ; et celui engagé par les sciences cognitives, au sein du complexe NBIC (combinant les sciences cognitives avec les biotechnologies, les nanotechnologies et les sciences de l'information), d'autre part. Ces deux voies distinctes ont, en l'espèce, convergé pour offrir une certaine conception de la morale comme susceptible d'être automatisable, à partir d'un ensemble de règles. Si on les lit ensemble, elles sont porteuses d'une certaine vision du monde que je souhaiterais à la fois expliciter et remettre en question, pour tenter d'explorer d'autres voies possibles.

1. UN SENS DE LA JUSTICE ?

Peut-il y avoir une science de la morale ? Cette question est un *leitmotiv* des sciences humaines en général, et de la philosophie en particulier. Elle est complexe non seulement parce qu'elle a été engagée sous divers aspects, par des auteurs très différents, mais aussi parce que la définition même des termes « sciences » et « morale » est loin d'être unanime au sein de ce débat – j'aurai l'occasion d'y revenir.

La question d'une science de la morale constitue une interrogation classique de la philosophie, et mon objet n'est pas d'en explorer exhaustivement les tours et contours, mais plutôt de montrer comment cette conception a tracé sa route jusqu'à aujourd'hui.

Développée en large partie à l'époque classique par les philosophes sentimentalistes de la tradition du sens moral, de Shaftesbury à Smith ou encore Rousseau, cette hypothèse est reprise à nouveaux frais par John Rawls au XX^e siècle. Sa théorie est intéressante pour notre objet à deux égards au moins. D'une part, l'interrogation fondamentale de Rawls, lorsqu'il développe sa conception du sens de la justice, vise à

articuler ensemble éthique et politique. D'autre part, il élabore cette conception philosophique en s'appuyant sur différents éléments issus de la psychologie développementale et des sciences cognitives.

Pour être juste, selon Rawls, il faut non pas être capable d'appliquer la justice, comme une machine ferait tourner un programme, mais plutôt la comprendre afin d'être motivé à agir justement. Pour la comprendre, il faut à la fois en appréhender librement les principes et le fonctionnement. Il faut également être apte à analyser les procédures ou situations comme étant ou non justes. C'est pourquoi la théorie de la justice doit être non seulement désirable, souhaitable, mais aussi faisable au sens où elle est réaliste et où ses principes « décrivent [...] notre sens de la justice » (Rawls, 1997, p. 73), notamment en s'appuyant sur ce que l'on peut connaître de notre capacité morale. Ainsi, nous serons motivés par la coopération, et c'est sur cette motivation que peut reposer, entre autres, la stabilité du gouvernement. Le fait d'être dotés du sens de la justice nous permet à la fois de coopérer avec autrui et d'avoir suffisamment confiance en autrui pour coopérer sans appréhension. Son approche s'appuie ainsi sur ce qu'il appelle une « psychologie morale raisonnable (Rawls, 2003, p. 265). Rawls pose que le sens de la justice est, avec la capacité d'une conception du bien, l'une des deux facultés morales¹ permettant de définir les personnes comme libres, égales, et déterminées à s'engager dans un schème stable de coopération. Le sens de la justice est compris comme la capacité « de comprendre, d'appliquer, et d'agir selon (et non seulement en conformité avec) les principes politiques qui spécifient les termes équitables de la coopération sociale » (Rawls, 2003, p. 39).

Cette attention portée à la structure et au développement de notre capacité morale est précoce dans l'œuvre de Rawls, puisqu'elle apparaît de manière structurée près de 10 ans avant la publication de sa *Théorie de la Justice*, dans un article de 1963, intitulé « Le sens de la justice ». Alors que l'article de 1963 reprend dans ses grandes lignes la conception piagétienne du développement moral – sur laquelle j'aurai l'occasion de revenir –, la *Théorie de la Justice* propose une approche appuyée sur deux conceptions distinctes : d'une part, la théorie du développement moral proposée par Lawrence Kohlberg ; d'autre part, la conception de la cognition linguistique développée par Noam Chomsky, pour proposer une analogie avec le développement moral.

1. Les facultés morales se distinguent des facultés intellectuelles. Notons cependant que les facultés intellectuelles sont indispensables pour le bon exercice des facultés morales.

Ces deux éléments, que Rawls conjoint plus ou moins pour avancer son hypothèse du sens de la justice, ont un statut différent. Alors qu'il reprend au moins en partie la conception développée par Kohlberg – à d'importantes différences près, comme on le verra –, l'analogie avec la conception chomskyenne est très originale et a contribué à orienter un faisceau de recherches ces 15 dernières années. Le but de Rawls est de construire une psychologie philosophique « purement hypothétique » bien que « raisonnablement plausible » (Rawls, 1999, p. 100), à mi-chemin entre la science de la nature humaine de Hume et la conception kantienne de la personne appuyée sur des concepts et principes. Il a toutefois donné lieu, ainsi qu'on le verra, à une approche expérimentale de la morale qui s'est peu à peu éloignée de cette visée première.

Le point de départ de Rawls est la complexité de notre capacité morale, qu'il compare à la complexité de notre capacité linguistique. En paraphrasant Chomsky, on pourrait en effet s'étonner que le jeune enfant, tout comme il est capable de produire (et comprendre) des phrases qu'il n'a jamais entendues ni énoncées auparavant, soit susceptible de produire (et comprendre) des évaluations morales auxquelles il n'a jamais été exposé auparavant. C'est pourquoi Rawls propose de reprendre le terme technique de « compétence » pour parler de notre capacité morale, tout comme Chomsky emploie ce terme pour désigner notre capacité linguistique. Il s'agit donc de penser ce sens de la justice comme une capacité cognitive innée humaine, universelle et inconsciente qui se distingue de la « performance » (l'emploi effectif dans telle ou telle situation concrète). Elle ne peut donc se confondre avec les règles consciemment explicitées. C'est pourquoi elle doit être étudiée scientifiquement : « De même que les psycholinguistes dépendent des linguistes, de même la théorie de l'apprentissage moral dépend d'une analyse de la nature et des formes de la moralité. Les objectifs de la théorie ne peuvent se contenter sur ces questions de nos idées du sens commun. » (Rawls, 1997, p. 531)

Il s'agirait donc de prendre comme modèle la « Grammaire Universelle » qui vise à analyser les propriétés biologiques universelles de la nature humaine. Cette théorie est « descriptivement adéquate » : elle permet de rendre compte des faits linguistiques en formulant des principes universels, mais aussi en expliquant les mécanismes d'engendrement qui nous permettent de formuler des cas particuliers. De même, Rawls affirme qu'il faudrait construire une théorie permettant de « former un ensemble de principes qui, complétés par nos croyances,

et notre connaissance du contexte, nous conduiraient à formuler ces jugements ainsi que les raisons qui les appuient... » (Rawls, 1997, p. 72).

Le sens de la justice est une capacité développementale qui suit, selon Rawls, le schéma proposé par Piaget (dans l'article de 1963) puis par Kohlberg (dans *Théorie de la Justice*). Des deux psychologues, le philosophe reprend l'idée que le sens de la justice se développe en suivant trois grands stades qui évoluent dans le sens d'un élargissement de la communauté morale (des parents à l'État) et d'une abstraction progressive. Rawls part du cadre le plus élémentaire, celui de la relation réciproque entre parents et enfants, pour en montrer l'extension vers une association équitable. De là, il montre comment il devient possible de développer le sens moral jusqu'à la compréhension des principes proprement moraux. À l'inverse du raisonnement mis en œuvre dans l'élaboration de la position originelle (qui va des principes aux relations interpersonnelles), il opère ainsi des relations interpersonnelles vers les principes pour proposer dans *Théorie de la Justice* trois lois psychologiques (Rawls, 1997, p. 530). Premièrement, l'amour des parents pour leurs enfants dans des conditions familiales justes permet, en retour, l'amour des enfants pour leurs parents. Deuxièmement, sur cette base d'amour parental/filial, une société juste reconnue comme telle permet l'élargissement des relations d'amitié et de confiance aux autres membres du groupe qui s'en montrent dignes. Enfin, sur le socle des deux premières lois, au sein d'institutions justes et reconnues telles, l'élargissement s'opère vers les principes eux-mêmes.

L'une des originalités de la conception rawlsienne, qui repose sans doute en bonne part sur l'héritage recueilli des philosophes des Lumières (notamment Rousseau et Hume), est l'importance accordée, dans ce développement, à l'émotion ainsi qu'au développement de l'empathie et de la confiance au fur et à mesure du processus. Le sens de la justice est inséparable de ce que Rawls nomme des attitudes naturelles, comme l'amitié ou la confiance, qui ne peuvent être confondues avec de simples réactions physiologiques. Elles supposent également que l'on puisse en donner raison, par exemple en mobilisant les notions de bien et de mal – comme lorsque je ne peux plus faire confiance à quelqu'un qui a pris davantage que sa part et a ainsi rompu le schème de coopération.

Le sens de la justice n'est donc ni purement émotionnel ni purement rationnel : c'est « un authentique sentiment du cœur éclairé par la raison » (Rawls, 1999, p. 96 ; Nurock, 2008, p. 95-109). Pour cette raison,

le sens de la justice est à la fois justificatif et motivant : il permet à la fois de distinguer le juste de l'injuste et d'être motivé à agir justement. Or, selon Rawls, ce sens de la justice fait partie de la nature humaine et est lié à une disposition à avoir des sentiments moraux ainsi que des attitudes naturelles. Il est donc ce qui lie le rationnel et le raisonnable, j'aurai l'occasion d'y revenir.

2. PIAGET ET LES RÈGLES MORALES

Afin de montrer comment le sens de la justice proposé par Rawls a indirectement permis le développement d'un faisceau de travaux sur la cognition morale, je souhaiterais revenir aux deux points d'appui de cette conception, d'une part les deux figures majeures de la psychologie morale du vingtième siècle (Piaget et Kohlberg) et, d'autre part, les sciences cognitives, pour mettre en lumière les conceptions du monde (et de la morale) qui leur sont sous-jacentes. Je commencerai, dans cette deuxième partie, par expliciter les éléments de sa psychologie morale.

L'un des points les plus singuliers de la conception du développement moral développée par Piaget est son ancrage dans la question de la règle et, plus précisément, dans celle de la règle du jeu, plus précisément du jeu de billes. Piaget distingue ainsi l'étude de la morale pratique – telle qu'elle se dévoile à travers ces jeux – de l'étude de la morale verbale, dont témoigne la manière dont les enfants sont susceptibles de répondre à des questions sur les règles du jeu ou encore à de petits scénarii mettant en scène des situations simples. Ces scénarii ne retiendront pas mon attention dans les paragraphes qui suivent, mais j'y reviendrai plus tard. Contrairement à ce qui pourrait sembler à première vue, cet ancrage dans le jeu de billes est particulièrement signifiant. C'est pourquoi plutôt que de résumer la conception piagétienne du développement moral, je me concentrerai sur cet aspect de son approche dans les pages qui suivent.

Le premier chapitre de *Le jugement moral chez l'enfant* s'intitule ainsi « Les règles du jeu » : le point de départ n'est pas l'étude des règles morales, mais celle des jeux auxquels s'adonnent les enfants. Piaget considère donc un système de normes que l'on peut envisager comme

définissant une pratique, au sens où elles sont la condition sine qua non de cette dernière².

De manière surprenante, Piaget ne fait porter son enquête ni sur des règles spécifiquement sociales ni sur des règles de droit. Plus encore, il ne cherche pas à comprendre comment les règles peuvent ou non être justes. La focale de son enquête est la manière dont les enfants pratiquent les règles lorsqu'ils jouent à un jeu apparenté à un jeu collectif: le jeu de billes chez les garçons³.

Ce jeu présenterait en effet deux caractéristiques qui le rendent particulièrement pertinent pour comprendre la morale d'après Piaget. D'une part, il se transmet dans l'interaction interpersonnelle, éventuellement de génération en génération, sous la forme d'un droit coutumier possédant tout un code de même qu'une jurisprudence⁴; d'autre part, il repose sur le respect des règles. Or, le respect de la règle constitue pour Piaget le critère distinctif de la morale: même s'il ne nie pas qu'il puisse exister des règles non morales, il ne peut cependant, selon lui, exister de morale sans règle. Ainsi qu'il l'écrit: « Toute morale consiste en un système de règles et l'essence de toute moralité est à chercher dans le respect que l'individu acquiert pour ces règles. » (Piaget, 1932, p. 2)

Porter l'accent sur les règles, c'est mettre en valeur la dimension sociale et interactionnelle de la morale qui est mise en valeur. De manière originale, en se focalisant sur les jeux d'enfants, Piaget vise à resserrer l'analyse sur des règles dont le contenu serait fixé par les enfants

-
2. Il importe peu que celle-ci soit donc ou non sociale, car ces règles ne sont pas nécessairement sociales ni rationnelles, ainsi que le souligne entre autres Ernst Tugendhat. Voir Tugendhat, E. «The Necessity for Cooperation between Philosophical and empirical Research in the Clarification of the Meaning of the Moral "Ought"» (Wren, 1990, p. 3-14).
 3. Ce choix, qui pourrait sembler pour le moins surprenant, est justifié par Piaget en ces termes: « Comme toutes les réalités dites morales, les règles du jeu de billes se transmettent de générations en générations et se maintiennent uniquement grâce au respect que les individus ont pour elles. La seule différence est qu'il ne s'agit que de rapports entre enfants. Les petits qui se mettent à jouer sont peu à peu dressés par les grands au respect de la loi, et, d'ailleurs, ils tendent de tout leur cœur à cette vertu, éminemment caractéristique de la dignité humaine, qui consiste à pratiquer correctement les usages du jeu. Quant aux grands, il est en leur pouvoir de modifier les règles. S'il n'y a pas là de morale – mais où commence la morale? – il y a donc tout au moins un respect de la règle, et c'est par l'étude d'un tel fait qu'une recherche comme la nôtre se doit de débiter » (Piaget, 1932, p. 2).
 4. Ce sont les termes utilisés par Piaget dans son ouvrage *Le jugement moral chez l'enfant* (Piaget, 1932, p. 2).

eux-mêmes⁵ et non pas sur la transmission des règles morales aux enfants par les adultes ou par la société⁶. L'analyse du jeu de billes permet de porter l'accent sur un paradigme social d'élaboration des règles par les enfants dans une relation entre pairs et dans le respect mutuel, la coopération, qui sont fondamentaux dans la conception de Piaget. Le but de Piaget est de tenter d'échapper à une conception transcendante des règles pour les établir dans l'immanence. Le matériau empirique choisi est donc tout à fait révélateur – et symptomatique – du cadre théorique de Piaget. Celui-ci pose en effet avant tout la morale comme une morale sociale nouée non pas – comme chez Durkheim – dans le rapport de l'individu et de la société, mais dans l'expérience des rapports interpersonnels.

En résumé, l'accent porté par Piaget sur les règles du jeu de billes est particulièrement révélateur dans la mesure où il lui permet d'insister sur l'articulation entre la dimension subjective et la dimension intersubjective de la morale ainsi que sur sa dimension à la fois naturelle et conventionnelle⁷. Le choix même du jeu de billes comme paradigmatique pour étudier la morale est ainsi particulièrement important, d'un point de vue non seulement méthodologique mais également métaéthique sur la conception de la morale qui sous-tend l'enquête, que l'on pourrait croire purement descriptive, de Piaget sur la morale.

3. DILEMMES ET COGNITION MORALE

La postérité de Piaget a délaissé l'étude des jeux d'enfants pour se focaliser sur un autre objet : le dilemme moral. Le dilemme moral peut être défini comme une alternative dont les deux membres sont également souhaitables ou non souhaitables d'un point de vue éthique.

L'utilisation des dilemmes moraux pour l'étude du développement moral a été généralisée par Lawrence Kohlberg au tournant de la Seconde Guerre mondiale. La méthodologie de Kohlberg est inséparable de sa vision d'ensemble, c'est-à-dire de sa caractérisation de la

5. «...nous sommes en présence, dans le cas des jeux sociaux les plus simples, de règles élaborées par les enfants seuls» (Piaget, 1932, p. 2).

6. Cette posture s'inscrit en opposition explicite avec l'approche durkheimienne.

7. Ainsi que le souligne Vincent Descombes commentant Wittgenstein, un « nouveau venu » ne peut savoir qu'il suit une règle que s'il ne le fait pas pour la première fois. Sur la difficulté conceptuelle de l'expression « se fixer une règle à soi-même », voir Descombes (2004, p. 451-456).

morale et de sa conception de ce que peut et doit faire la psychologie morale.

Selon lui, il importe de concevoir une théorie morale universelle et idéale opposée au relativisme, qui s'entend sous l'égide d'une vertu englobant toutes les autres: la justice. Sa conception de la morale est constituée, comme c'était déjà le cas chez Piaget, d'une capacité dont le développement progressif est opéré par stades – même s'il accorde à ces derniers une dimension plus stricte que ne le faisait Piaget. Mais, contrairement à Piaget, la question des règles ne saurait être au premier plan de l'enquête, car :

La justice n'est pas une règle ou un ensemble de règles, c'est un principe moral. Par un principe moral, je veux dire un mode de choix qui est universel, une règle de choix que l'on souhaite que tout le monde adopte dans toutes les situations. [...] Il y a des exceptions aux règles mais pas d'exception aux principes. Une obligation morale consiste en une obligation de respecter la revendication de quelqu'un d'autre. Un principe moral est un principe qui permet de résoudre le problème des revendications concurrentes. Toi contre moi, toi contre une tierce personne. Il n'y a qu'une seule base, un seul principe pour résoudre le problème des revendications concurrentes: la justice ou l'égalité. Traiter la revendication de chaque personne impartialement, sans considérer la personne en question. Un principe moral n'est pas seulement une règle d'action, mais une raison d'action. (Kohlberg, 1981, p. 39-40)

À la différence de Piaget, qui propose une définition assez large de la morale et qui entend étudier cette dernière tant dans sa dimension pratique que dans sa dimension verbale, Kohlberg considère, on l'a vu, que la morale se limite à la justice et qu'elle se restreint à la morale verbale, c'est-à-dire aux jugements moraux explicites. L'étude de la morale se limite donc au jugement moral. Il caractérise le jugement moral par trois éléments essentiels: la prescriptivité (qu'il distingue de la dimension descriptive); l'universabilisation⁸ (Kohlberg *et al.*, 1983, p. 17); et la cognition. Ce dernier élément caractérise le jugement moral comme rationnel et conscient.

L'une des principales préoccupations de Piaget, très marqué par la Première Guerre mondiale – ainsi qu'en témoigne notamment son

8. Par cette dimension prescriptive, Kohlberg signifie «l'obligation catégorique d'agir».

essai de jeunesse *Recherche* –, était de dégager les grandes règles de la coopération afin de la renforcer, ce qui explique probablement en partie l'importance qu'il accorde à l'étude des règles du jeu. Kohlberg, quant à lui, très marqué par la Seconde Guerre mondiale cherche à comprendre ce que l'on devrait faire, selon une approche très déontologique. Ainsi que le souligne Habermas⁹, la posture de Kohlberg est non pas descriptive, mais véritablement normative. Il cherche à se démarquer résolument des conceptions relativistes, qui présentent le danger de défendre l'idée que toutes les positions se valent d'un point de vue moral, et promeut une conception de la morale universaliste.

De manière très cohérente, l'universabilité (ou la possibilité d'universalisation) est étroitement liée à l'un des éléments essentiels de la méthodologie kohlbergienne : la notion de réversibilité. Dans une rhétorique qui n'est pas sans similitudes avec celle utilisée par Rawls dans son argumentaire sur la position originelle, Kohlberg souligne que le choix moral doit pouvoir convenir à tous et donc que les positions doivent être réversibles au sens où il doit pouvoir être acceptable moralement d'échanger sa position. À cette condition, on peut garantir un équilibre entre des revendications concurrentes (Kohlberg, 1981, p. 197). On l'a dit, l'une des caractéristiques principales du jugement moral selon Kohlberg est sa dimension cognitive, c'est-à-dire mettant au premier plan la rationalité – si l'affect n'est pas exclu par principe, ce dernier est toujours structuré ou médié par des mécanismes cognitifs (Kohlberg *et al.*, 1983, p. 79).

Au premier rang de ces derniers se trouve le *role taking*, qui est le mécanisme clé de la réversibilité, emprunté au philosophe, sociologue et psychologue béhavioriste George Herbert Mead. Dans *L'esprit, le Soi et la Société*, initialement paru en 1932, Mead insiste sur le rôle de l'imagination interpersonnelle dans la constitution du sujet et de ses jugements¹⁰. Le *role taking* comme processus de constitution se distingue du « soi généralisé » (*generalized other*) constitué par la société¹¹ au sens où il constitue une instance interne et non externe ou intériorisée. Il consiste à se projeter en imagination à la place d'autrui.

9. Habermas le souligne dans *Morale et Communication* (1986) et dans « Justice and solidarity: on the discussion concerning stage 6 » (1990, p. 224 sq.).

10. De la même manière, Piaget, inspiré par les théories de Baldwin sur l'imitation, insiste sur le rôle de l'imaginaire interpersonnel dans la constitution des jugements moraux. Cependant, ce processus ne consistait qu'en une étape préliminaire et en quelque sorte par défaut, l'important résidant avant tout dans les rapports interpersonnels réels.

11. Voir notamment George. H. Mead, (1962 [1934]), p. 154.

Kohlberg restreint l'usage de ce processus à la morale et insiste sur sa dimension rationnelle et intellectuelle, même si les émotions ne sont pas totalement exclues. Elles sont cependant subordonnées à la raison, seule capable de former des justifications morales cohérentes et acceptables : les émotions ne sont pas fiables pour Kohlberg, pis encore, elles sont subjectives et irrationnelles. Le jugement moral trouve donc à s'exprimer rationnellement et explicitement grâce aux justifications morales. Ces dernières sont partageables dans la mesure où il nous est possible de nous mettre à la place d'autrui. Le développement moral est achevé lorsque le sujet est capable de se mettre à la place d'autrui pour embrasser en même temps toutes les positions, en « spectateur impartial ».

Kohlberg considère que la méthode la plus à même de tester le jugement moral défini par sa dimension prescriptive, universalisable et cognitive est de mettre au premier plan les justifications offertes par les sujets. Dans cet ordre d'idées, l'outil privilégié par sa conception est l'utilisation de dilemmes moraux – plutôt que de questions sur le jeu ou de petits scénarii comme chez Piaget, même si Kohlberg ne remet pas en question la conception de la règle qui structure l'approche piagétienne. Pour la même raison, il se concentre sur la morale verbale en laissant de côté l'aspect pratique, que Piaget considérerait comme essentiel.

Lors d'entretiens, les expérimentateurs proposent aux sujets des dilemmes en leur demandant non seulement de donner une réponse, mais surtout de justifier la réponse choisie. Pour aider à cette justification, l'expérimentateur pose aux sujets une batterie d'une dizaine de questions. La raison pour laquelle l'usage des dilemmes est privilégié est parfaitement cohérente avec la conception kohlbergienne. Il ne s'agit pas de déterminer la bonne réponse, puisqu'il n'y a pas de bonne réponse à un dilemme par définition : les deux propositions de l'alternative doivent justement être également souhaitables ou non souhaitables d'un point de vue moral pour que le dilemme « fonctionne », et qu'il soit si difficile de la trancher. Mais ce qui est essentiel aux yeux de Kohlberg, c'est justement non la réponse privilégiée, mais la structure du raisonnement, le cheminement emprunté et la justification choisie¹². Par ailleurs, ces dilemmes peuvent prendre la forme de

12. À cet égard, Kohlberg insiste sur l'idée que sa méthodologie permet de tester la compétence (la capacité à comprendre sinon à former des arguments moraux) des sujets et non pas seulement leur performance (leur capacité à les formuler explicitement) grâce aux ques-

conflits de valeurs¹³ (comme la vie ou l'obéissance à la loi, par exemple à travers le respect de la propriété) en conduisant le sujet à établir une hiérarchie entre les valeurs. La procédure met en œuvre trois formes parallèles d'entretiens, chacun constitué de trois dilemmes, dont le conflit de valeurs est prédéfini. Le premier dilemme porte toujours sur un conflit entre la vie et la loi, le second entre la conscience et la punition et le troisième entre l'autorité et le contrat.

Les dilemmes utilisés par Kohlberg font appel à des situations diverses et sont suffisamment souples pour pouvoir éventuellement être adaptés aux cultures dans lesquelles ils sont testés. Certains mettent notamment en scène – comme c'était déjà le cas dans certains scénarii de Piaget – le vol altruiste (Kohlberg, 1984, p. 640), des questions liées à l'euthanasie et à la dénonciation¹⁴. Le dilemme le plus connu est le dilemme de Heinz, dont voici le scénario.

En Europe, une femme était proche de mourir d'un type de cancer très particulier. Les médecins pensaient qu'un certain médicament pouvait la sauver. Il s'agissait d'une forme de radium qu'un médecin de la même ville venait tout juste de découvrir. Ce médicament était cher à faire, mais le pharmacien facturait ce médicament dix fois ce qu'il lui coûtait à réaliser. Le radium lui coûtait 400 \$ et il vendait une petite dose du médicament 4000 \$. Le mari de la femme malade, Heinz, était allé voir toutes les personnes qu'il connaissait pour leur emprunter de l'argent et avait essayé tous les moyens légaux, mais il n'était parvenu à réunir que 2000 \$, soit la moitié du prix. Il dit au pharmacien que sa femme était en train de mourir et lui demanda de lui faire une réduction sur le prix du médicament ou de lui faire crédit d'une partie du montant. Mais le pharmacien lui répondit : « Non, j'ai découvert le médicament, et je veux qu'il me rapporte de l'argent. »

tions posées par l'expérimentateur, lesquelles visent notamment à déterminer les limites « supérieures » du jugement moral du sujet interrogé. Voir par exemple Kohlberg, Lawrence (1984), p. 245-246. Ce point est important dans la mesure où, on le verra, chaque stade englobe le précédent et un sujet susceptible de comprendre un argument moral typique du stade 4 est capable de comprendre ceux des stades précédents tout en jugeant celui-ci comme moralement plus adéquat.

13. Ce choix est en lui-même significatif. En effet, des dilemmes d'obligation (A a promis à B de l'aider à aller au four et à C de l'emmener au moulin, or les deux lieux ne sont accessibles qu'en même temps et il ne peut, selon l'adage, être à la fois au four et au moulin et donc remplir ses deux obligations) ne permettent pas de rendre compte de la vision morale du sujet de manière aussi évidente.
14. La dernière version du test, publiée en 1987 propose à l'expérimentateur trois séries de trois dilemmes différents.

Ayant ainsi essayé tous les moyens légaux, Heinz devient désespéré et songe à cambrioler la pharmacie pour voler le médicament (Kohlberg, 1984, p. 640).

Les questions posées par l'expérimentateur dans l'entretien semi-structuré visent non seulement à conduire le sujet à proposer une élucidation de l'alternative proposée, mais également à expliciter sa conception de la morale. Ces réponses sont « codées » par l'expérimentateur selon une procédure stricte, souvent révisée. Mais à côté de cet aspect qui peut paraître un peu automatique, Kohlberg insiste sur la dimension herméneutique de la pratique du psychologue : les justifications apportées par le sujet peuvent déjà être considérées comme une forme d'interprétation dans la mesure où elles témoignent de la lecture faite par le sujet de la situation. Non seulement le jugement moral se détermine grâce aux justifications apportées par le sujet, lesquelles peuvent déjà être considérées comme une forme d'interprétation, mais encore ces justifications sont interprétées par l'expérimentateur. Selon Kohlberg, la relation entre le sujet et l'expérimentateur repose sur une forme d'empathie, car ce dernier cherche à comprendre la vision du monde révélée que le sujet transmet dans ses réponses.

Pourtant, au sein même de l'équipe de Kohlberg, des voix dissonantes ont assez rapidement conduit à remettre en cause deux éléments structurant l'histoire de la psychologie morale du XX^e siècle esquissée ci-dessus : d'une part, la conception des règles ; d'autre part, le rôle structurant du dilemme moral. Ces deux éléments ont notamment conduit à décentrer la conception du sujet moral implicite dans le cadre kohlbergien : l'homme blanc, américano-européen, de classe moyenne...

4. DILEMMES EN SÉRIE : LA TRAMWAYLOGIE ET SES LIMITES

La contestation du cadre kohlbergien a pris place au tournant des années 1980 sur arrière-fond de montée de la lutte pour les droits civiques, de contestation de la guerre du Viêt Nam et de progression du féminisme. La remise en question du cadre de référence s'est pour partie construite à partir d'une critique de l'utilisation du dilemme comme pierre de touche du développement moral. Carol Gilligan, l'une des anciennes collaboratrices de Kohlberg, a ainsi mis en avant la manière dont une petite fille, Amy (Gilligan, 2008, p. 48 *sq.*), est conduite à reconfigurer le dilemme de Heinz, montrant par-là que

la manière binaire de poser le cadre limite *a priori* la pensée morale. Alors que Jake, un petit garçon du même âge propose une réponse typique du cadre kohlbergien, à savoir que Heinz devrait voler le médicament puisque la valeur de la vie est plus importante que l'argent, Amy répond :

S'il volait le médicament, il sauverait peut-être la vie de sa femme, mais alors il risquerait d'aller en prison. Si sa femme retombait malade par la suite, il ne serait plus en mesure de lui procurer un nouveau médicament et la vie de sa femme serait de nouveau en danger. Ils devraient discuter à fond du problème et trouver un moyen de réunir l'argent. (Gilligan, 2008, p. 53)

Elle ne configure pas la situation sous la forme d'un face-à-face, mais plutôt comme une relation à trois où le problème est non le dilemme voler/ne pas voler, mais où le problème réside dans le refus du pharmacien d'aider la femme de Heinz. La question centrale à ses yeux est le manque de compréhension et d'empathie du pharmacien, les motivations qu'il peut avoir et la manière de le convaincre en trouvant une solution de consensus par la discussion. Il ne s'agit donc pas de trancher, mais de dénouer. Amy ne parvient pas à répondre « correctement » aux réponses de l'expérimentateur, car elle ne les saisit pas comme il s'y attend. Dès lors, elle « échoue » au test parce qu'elle ne rentre pas dans les cases en appréhendant la situation différemment. Ne trouvant pas ce qu'il cherche dans l'entretien, l'expérimentateur reste sourd à la « voix différente » qui s'y exprime. Mais plus encore, Amy est retorse à répondre directement à l'expérimentateur et emprunte un chemin buissonnier : ce sont les relations entre les protagonistes et le contexte qu'explore Amy sans chercher à se hausser sur un piédestal impartial.

Pourtant, avoir le loisir d'emprunter un tel chemin buissonnier est loin d'aller de soi, ainsi qu'en témoigne la dernière tendance à la mode tant en philosophie qu'en psychologie ou en intelligence artificielle : la trolleylogie ou tramwaylogie. Cette dernière tire sa source des travaux classiques de Philippa Foot et Judith Jarvis Thomson¹⁵, et a notamment été popularisée par John Mikhail, qui affirme explicitement souhaiter

15. Ces questions ont été introduites dans Foot, Philippa (1967). «The Problem of Abortion and the Doctrine of Double Effect», dans *Virtues and Vices and Other Essays in Moral Philosophy*, Oxford Review, n° 5, p. 204-217. Elles sont depuis l'objet d'un débat tel qu'il faudrait une trop longue bibliographie pour en rendre pleinement compte.

développer l'intuition proposée par John Rawls d'une cognition morale inspirée de la cognition linguistique chomskyenne (Mikhail, 2011).

Imaginons que vous soyez en train de faire votre jogging lorsque vous voyez arriver un tramway qui fonce à toute allure sur cinq ouvriers travaillant sur la voie. Le conducteur, qui les a vus, a tant et tant actionné les freins qu'ils ont lâchés et que, horrifié, il a lui-même fini par tomber dans les pommes. Par chance, vous vous trouvez juste à côté d'un levier permettant de faire passer le tramway sur une voie secondaire – qui, dans certaines versions, fait une boucle avant la voie principale. Sur cette voie se trouve un objet suffisamment lourd pour arrêter le train – ou au moins faire gagner suffisamment de temps aux ouvriers pour quitter la voie principale avant que le tramway ne la rejoigne. Est-il vraiment besoin de poser la question? Vous actionnez le levier n'est-ce pas? Mais que faut-il faire si la voie se sépare en deux et que sur la voie secondaire se trouve une personne? Faut-il dévier le tramway pour sauver les cinq ouvriers? Et si vous vous trouvez maintenant non à côté du levier, mais sur un pont où se trouve un homme si gros – ou, dans une version plus politiquement correcte, portant un si gros sac à dos – qu'en le poussant pour qu'il tombe sur la voie, vous pourriez arrêter le tramway ou permettre aux cinq ouvriers de se sauver? Et si les cinq hommes étaient des criminels, et s'ils n'étaient pas très blancs, et s'ils étaient pauvres, malades, et s'ils étaient des femmes et si la personne isolée était votre fiancée, et si...

Cette série de petits scénarii constituent des variations de dilemmes moraux où, comme on le sait, aucune des réponses ne devrait normalement être moralement acceptable. Ils sont un bon exemple de ce que la philosophie appelle traditionnellement le «double effet»: l'action vise une conséquence bonne (sauver les cinq ouvriers) qui a elle-même une conséquence «secondaire» prévue, mais non voulue, qui est négative (la mort de la personne sur l'autre voie ou celle du gros monsieur, etc.). Ils décrivent des situations rares dans la vie courante¹⁶ qui sont pourtant, semblerait-il, devenues paradigmatiques, voire emblématiques de la manière dont certains envisagent la vie morale,

16. Mais moins dans des circonstances extraordinaires ou dans la littérature, ainsi que le montrent les recherches de Frédérique Leichter-Flack (2015).

et qui sont aujourd'hui très largement utilisées par les psychologues pour tester nos intuitions morales¹⁷.

Cet usage des dilemmes par la psychologie cognitive se démarque de la méthodologie kohlbergienne par deux aspects essentiels. Premièrement, les dilemmes sont centraux, car il s'agit de tester les intuitions et non les justifications, ainsi que c'était le cas chez Kohlberg. L'hypothèse, ici, est plutôt que la morale « véritable » n'est guère affaire de justification parce que les justifications sont souvent *post hoc* : elles précèdent la prise de décision, mais ne la déclenchent pas parce que les causes nous conduisant à opérer nos choix dans ces cas sont implicites et inconscientes¹⁸. Cette conception s'oppose à l'idée classique selon laquelle nous formerions nos évaluations morales en faisant (consciemment) appel à des raisons ou à des « principes » *a priori*.

Deuxièmement, la morale est considérée sous son aspect descriptif plutôt que prescriptif. Même si la méthode utilisée peut sembler très similaire en apparence, les enjeux sont donc radicalement différents, pour ne pas dire opposés. Or, ce sont ces exemples et cette méthodologie qui ont été investis dans l'idée de programmer une morale aux machines dites autonomes, par l'intelligence artificielle, et ainsi de créer une morale « automatique », et mécanisable. Les applications en sont multiples, des voitures sans chauffeur aux machines tueuses, en passant par le triage médical. Contrairement à ce que l'on pourrait penser, ces IA ne sont pas de la science-fiction. Par exemple, des voitures sans chauffeur circulent déjà aujourd'hui aux États-Unis ou à Singapour, mais aussi à Toulouse, sur le campus de l'Université Paul-Sabatier.

Les variations sur ces dilemmes moraux ont servi de base à la constitution de deux sites web, considérés comme des supports permettant de tester en ligne notre sens moral : *The Moral Sense Test* (<http://www.moralsensetest.com/>) et *The Moral Machine* (<http://moralmachine.mit.edu/>). Le premier, conçu dans la lignée des travaux doctoraux

17. Même le *New Yorker* sous la plume de Mark Remy, a fait la part belle à ces variations, dans un article du 4 décembre 2017.

18. Il s'agit là d'une démarche classique au sein des sciences cognitives ainsi qu'on le verra. Toutefois, on peut souligner qu'on se trouve ici au croisement de deux types de suggestions : d'une part, l'une des hypothèses de Haidt, selon qui les justifications morales sont le plus souvent *a posteriori* et ne nous renseignent guère sur les mécanismes sous-jacents à notre capacité morale, et d'autre part, un ensemble de travaux en psychologie sociale, selon lequel les raisons de certains de nos choix non moraux sont inconscients, d'où la nécessité d'élaborer une trame narrative ou argumentative pour les justifier (Nisbett et Wilson, 1977).

de Mikhail, par l'équipe de Mark Hauser à Harvard au début des années 2000, est aujourd'hui encore hébergé sur le site dirigé par l'ancien collaborateur de Mark Hauser, Fiery Cushman (Hauser et Cushman, 2000). Le second, plus récent, a été conçu par une équipe du MIT, en collaboration avec une équipe toulousaine, afin de mécaniser la morale, notamment en implémentant des programmes pour les véhicules sans chauffeur.

Les dilemmes, par leur côté à la fois ludique et systématique qui permet une forme d'économie cognitive – ils ne sont pas difficiles à appréhender une fois que l'on a compris le schème de base puisqu'il suffit de faire varier un seul élément à chaque fois – attirent l'attention assez facilement. La vogue des dilemmes moraux a vite connu un succès tel qu'il est difficile de dénombrer combien de variations sont aujourd'hui en circulation. Une page Facebook leur est même consacrée¹⁹ et ils apparaissent dans de nombreuses séries TV – de *Orange is the New Black* à *The Good Place*. L'emploi de ces dilemmes s'est imposé comme « la » méthode idéale pour automatiser la morale et produire ainsi des machines « morales ».

Pour autant, il est loin d'aller de soi que les dilemmes moraux peuvent permettre une automatisation de l'éthique, et donc une « science de la morale » pour différentes raisons éthiques et métaéthiques – je laisse ici de côté les importantes réserves techniques sur ces véhicules, qui peinent à catégoriser correctement les éléments (notamment humains!) du paysage pour me concentrer sur les questions proprement éthiques.

Les dilemmes moraux explorés par la tramwaylogie sont loin d'être des situations couramment rencontrées par les conducteurs, il n'est donc pas évident que ce type de test puisse être d'une quelconque utilité – mais on peut toujours imaginer qu'il propose une matrice permettant d'extrapoler d'autres réponses, bien entendu. En revanche, un véhicule sans chauffeur pourrait tout à fait être programmé pour sauver à tout prix la vie du conducteur et/ou des passagers de la voiture, quel qu'en soit le prix.

Ce véhicule pourrait également être programmé pour sauver l'un ou l'autre sans que nous le sachions, ou encore être piraté, ce qui permettrait à une « boîte noire », à une entreprise, à un pirate informatique, par exemple, de décider qui doit vivre ou mourir. Il pourrait donc proposer une hiérarchie « programmée » entre les êtres humains.

19. www.facebook.com/TrolleyProblemMemes

Toutefois, si ce type de test avait une quelconque réalité et que l'on se trouve dans l'une des situations décrites, alors, pourrait-on se réjouir : «heureusement que ma voiture a pris la bonne décision et écrasé une personne plutôt que cinq»? Il est évident qu'une telle exclamation passerait complètement à côté de ce qui fait la spécificité des dilemmes moraux. Répétons-le, dans un dilemme, les deux propositions de l'alternative doivent être également souhaitables ou non souhaitables d'un point de vue moral pour que le dilemme «fonctionne», et qu'il soit si difficile de le trancher. Bref, aucune réponse n'est acceptable d'un point de vue moral : ce qui fait la spécificité du dilemme, c'est qu'on ne s'en sort pas indemne, qu'il reste toujours des «résidus moraux²⁰». Dans le célèbre roman de William Styron et le film qui en est issu, *Le choix de Sophie*, Sophie est sommée de choisir entre la vie de sa fille ou celle de son fils : on ne peut pas l'imaginer satisfaite ou ne se sentant pas coupable.

Pour le dire autrement, l'utilisation de ces dilemmes schématise la vie morale d'une manière immorale, en nous forçant à faire des choix inacceptables et irréalistes, ce qui risque de reconfigurer la manière dont nous appréhendons ce qu'est la morale. Il s'agit donc d'un déni total de la voix d'Amy et de la richesse de la vie morale. Le problème central de l'automatisation de la morale est que celle-ci dénie la volonté libre comme composante essentielle de la vie morale. Automatiser la morale, c'est non seulement annihiler la responsabilité, mais encore se débarrasser des résidus moraux. Par conséquent, sous l'appellation fallacieuse de machine «autonome» – puisque les règles ne sont pas dictées indépendamment par l'IA, même si elle peut les développer en «apprenant» – c'est l'autonomie humaine qui se trouve en question, à la fois dans sa configuration et dans ses conditions de possibilité.

* * *

La généalogie de l'automatisation de la morale dévoile à quel point c'est une certaine conception de la morale, limitée à un système de règles binaires, qui s'y constitue comme dominante. Or, ainsi que le montrait très bien Rawls, une machine ne peut pas être motivée à agir moralement et on ne peut pas appliquer la morale comme un problème. Sinon, le sens moral ne serait pas autre chose qu'un sens des règles et n'importe quel psychopathe pourrait être moral, du

20. Voir par exemple Greenspan, Patricia (1983).

moment qu'il connaît le code. Le terme même de tramwaylogie ne permet d'ailleurs d'envisager cette morale que sous forme d'un *logos*, mais guère d'un *ethos*.

Envisager l'automatisation de la morale ne peut donc se faire de manière affirmative mais seulement de manière interrogative. Poser l'automatisation de la morale comme question nous permet alors d'emprunter des voies distinctes. Premièrement, elle nous conduit à nous interroger sur le type d'éthique sous-jacente à la morale automatisée et de montrer les limites de cette dernière. Deuxièmement, elle conduit également à se demander pourquoi on pourrait considérer comme moralement souhaitable de déléguer la morale à l'IA. Jusqu'à preuve du contraire, non seulement une telle nécessité n'a pas été démontrée; pis, elle viendrait à menacer inévitablement certains éléments clés de la vie morale: la liberté, le *care* et la responsabilité notamment.

BIBLIOGRAPHIE

- Descombes, Vincent (2004). *Le complément de sujet*. Paris, Gallimard.
- Foot, Philippa (1967). «The Problem of Abortion and the Doctrine of Double Effect», dans *Virtues and Vices and Other Essays in Moral Philosophy*, Oxford Review, n° 5, p. 204-217.
- Gilligan, Carol (2008). *Une si grande différence*. Paris, Flammarion.
- Greenspan, Patricia (1983). «Moral Dilemmas and Guilt», dans *Philosophical Studies*, vol. 43, n° 1, p. 117-125.
- Habermas, Jürgen (1986). *Morale et Communication*. Paris, Le Cerf.
- Habermas, Jürgen (1990). «Justice and solidarity: on the discussion concerning stage 6», *The Moral Domain: Essays in the ongoing discussion between philosophy and the social sciences*. Cambridge, The MIT Press.
- Hauser, Mark et Cushman, Fiery (2000). «The Moral Sense Test». Harvard University, Department of Psychology. Consulté en ligne.
- Kohlberg, Lawrence (1981). *Essays on Moral Development (volume I): The Philosophy of Moral Development*. San Francisco, Harper & Row Publishers.
- Kohlberg, Lawrence, Levine, Charles et Hewer, Alexandra (1983). *Moral Stages: A Current Formulation and a Response to Critics*. New York, Karger.
- Kohlberg, Lawrence (1984). *Essays on Moral Development (vol II): The Psychology of Moral Development: The Nature and Validity of Moral Stages*. San Francisco, Harper & Row, p. 245-246.
- Leichter-Flack, Frédérique (2015). *Qui vivra qui mourra: Quand on ne peut pas sauver tout le monde*. Paris, Éditions Albin Michel.
- Mead, George H. (1962 [1934]). *Mind, Self, and Society: From the Standpoint of a Social Behaviorist*. Chicago, The University of Chicago Press, p. 154.

- Mikhail, John (2011). *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge, Cambridge University Press.
- Nisbett, Richard et Wilson, Thimothy D. (1977). «Telling More Than We Can Know: Verbal Reports of Mental Processes», dans *Psychological Review*, vol. 84, n° 3, p. 231-257.
- Nurock, Vanessa (2008). *Rawls: Pour une démocratie juste*. Paris, Michalon.
- Piaget, Jean (1932). *Le jugement moral chez l'enfant*. Paris, Presses universitaires de France.
- Rawls, John (1997). *Théorie de la justice*. Paris, Seuil.
- Rawls, John (1999). «The Sense of Justice», *Collected Papers*. Cambridge, Harvard University Press.
- Rawls, John (2003). *La justice comme équité*. Paris, La Découverte.
- Remy, Mark (2017). «Variations of the Trolley Problem», New-York, The New-Yorker. Consulté en ligne.
- Thomson, Judith J. (1976). «Killing, Letting Die, and the Trolley Problem», *The Monist*, vol. 59, Oxford University Press, p. 204-217.
- Wren, Edward (1990). *The Moral Domain: essays in the ongoing discussion between philosophy and the social science*. Cambridge, MIT Press, p. 3-14.

CHAPITRE 3

À propos du statut moral des robots

Tomislav BRACANOVIĆ

INTRODUCTION

Dans les débats éthiques concernant les robots et la robotique, les thèmes prédominants des discussions concernent les implications et les conséquences de l'introduction des robots dans plusieurs secteurs de la vie humaine. Un autre sujet qui attire beaucoup d'attention touche la tentative de créer des robots capables de réguler leur fonctionnement de façon éthique, surtout en ce qui concerne leur interaction avec les êtres humains. La connexion entre ces deux enjeux est claire : si l'augmentation de la présence de robots crée de nouvelles et uniques menaces aux êtres humains, cela fait sens d'imposer non seulement des restrictions légales ou morales à ceux qui les conçoivent et produisent, mais de voir également si les robots eux-mêmes peuvent être programmés de façon éthiquement acceptable. Le statut moral des robots reste le sujet qui reçoit le moins d'attention. Ce texte examine les stratégies possibles pour discuter des pour et des contre concernant le statut moral de plusieurs types de robots. Il souligne que des robots à venir – des robots et biorobots dits cognitifs – devront être évalués comme candidats pour l'acquisition d'un statut moral. Leur développement futur est susceptible de défier certaines de nos croyances et valeurs morales enracinées.

1. LE STATUT MORAL ET SES THÉORIES

«Le statut moral» est un terme technique en philosophie qui est défini comme «l'aptitude d'un sujet à être considéré digne de souci moral direct» et comme «la nature ou le degré de la capacité d'un être comme un fondement de réclamation envers les agents moraux» (Johnson, 1999, p. 590). Avoir un statut moral, en autres termes, «signifie être une entité envers laquelle les agents moraux ont, ou peuvent avoir, des obligations morales. Si une entité a un statut moral, alors nous ne pouvons pas la traiter d'une façon quelconque; nous sommes moralement obligés de respecter ses besoins, ses intérêts et son bien-être, à l'intérieur de nos débats.» (Warren, 1997, p. 3) Une façon plus familière d'exprimer la même idée est de parler de quelqu'un ou de quelque chose qui a des droits moraux (qui ne doivent pas nécessairement coïncider avec des droits légaux), comme le droit de ne pas être lésé ou tué.

L'attribution du statut moral comporte des conséquences pratiques et réelles, principalement pour les êtres à qui un tel statut est nié ou accordé. La détermination du droit ou non-droit, au statut moral, pour quelqu'un ou quelque chose, signifie déterminer s'il est - sinon complètement, au moins partiellement - membre de notre communauté morale et vers qui nous avons des engagements moraux. Pour utiliser la métaphore bien connue de Peter Singer (1981) «le cercle d'extension», la question du statut moral est la question du sujet ou des êtres qui se trouvent à l'intérieur du cercle des êtres intimement précieux et qui sont protégés moralement, et de ceux qui restent en dehors de ce cercle. L'évolution de l'espèce humaine et l'histoire de la civilisation humaine - bien que loin d'être moralement idéales - peuvent être perçues comme un processus d'expansion, plus ou moins progressif, de ce cercle. «La sphère de l'altruisme s'est étendue de la famille et la tribu vers la nation, la race, et maintenant à tous les êtres humains», dit Singer (1981, p. 170) et il souligne que «le processus devrait s'étendre [...] pour inclure tous les êtres qui ont des intérêts, de toutes les espèces».

Il est probablement vrai que la «sphère de l'altruisme» ou la «sphère des êtres avec un statut moral» est aujourd'hui plus large que jamais, du moins en théorie que parce qu'elle englobe tous les êtres humains, y compris les groupes qui ont souvent été historiquement défavorisés, tels les femmes, les enfants ou les membres d'autres nations et races. Les problèmes liés au statut moral sont cependant loin d'être résolus

et causent toujours des débats moraux et politiques sérieux. Le statut moral des embryons et des fœtus suscite toujours des discussions houleuses, mais cette ferveur s'accompagne aussi des discussions au sujet d'autres candidats pour le statut moral, comme les animaux non humains, des espèces entières, les organismes clonés, les cellules souches embryonnaires, les générations futures ou les écosystèmes. Est-il possible que la robotique produise un autre candidat légitime à l'obtention d'un statut moral : les robots ? Est-il possible de penser aux robots comme de possibles porteurs de statut moral ? Avant d'essayer de répondre à ces questions, nous ferons un bref rappel des quelques théories fondamentales et des critères du statut moral, en nous basant sur Warren (1997, 2003) qui distingue :

- 1) *La théorie de l'humanité génétique.* Le statut moral, et l'ensemble des droits moraux correspondant, appartient seulement aux membres de l'espèce Homo sapiens. En autres termes, il est moralement incorrect d'infliger du mal à tout être génétiquement humain.
- 2) *La théorie de la responsabilité de l'agent moral.* Le statut moral appartient «aux agents moraux» (parfois confondus avec les «personnes») en tant qu'êtres libres et raisonnables, capables d'autonomie ou de gestion autonome de leur comportement, sur la base de principes moraux.
- 3) *La théorie de la vie organique.* Le fait d'être un vivant du point de vue organique est la caractéristique cruciale qui confère le statut moral. Tous les êtres vivants – simplement parce qu'ils sont vivants – ont le même statut moral et devraient avoir le même niveau de protection morale.
- 4) *Théorie de la sensibilité morale.* La sensibilité, ou la capacité d'éprouver le plaisir et la douleur, est nécessaire pour obtenir le statut moral. Ici, les principes essentiels sont que le plaisir est intrinsèquement bon tandis que la douleur est intrinsèquement mauvaise. Notre devoir moral est donc de ne pas infliger de douleur à ceux qui sont capables de l'éprouver, indépendamment des espèces ou des capacités mentales.
- 5) *Théories basées sur les relations.* Le statut moral dépend du fait d'être en relation avec quelqu'un ou quelque chose. Selon certaines théories, pour avoir un statut moral, il faut être inclus dans une relation de soins (*care*) – donner ou recevoir des soins. Selon d'autres théories, approuvées par plusieurs écoles d'éthique

environnementale, le statut moral d'une entité dépend de ses « rapports écosystémiques » avec d'autres entités vivantes et non vivantes.

Le point commun de toutes ces théories, selon Warren (1997 et 2003), c'est qu'elles ont « un seul critère » : on choisit une caractéristique qui est nécessaire et suffisante pour obtenir le statut moral. Cette approche parcimonieuse crée cependant des problèmes pour toutes ces théories. Fondamentalement, le problème est qu'une théorie peut être soit sous-inclusive, et elle assigne le statut moral à trop peu d'êtres, soit surinclusive, et elle assigne le statut moral à trop d'êtres. Si le fait d'être organiquement vivant, par exemple, est considéré comme nécessaire et suffisant pour obtenir le statut moral, toutes les formes de vie, des êtres humains aux organismes unicellulaires, auraient alors le même statut moral. Cet « égalitarisme biologique radical » (Warren, 1997, p. 34) est absurde et impraticable parce qu'on ne peut pas maintenir la vie humaine sans détruire au moins quelques êtres vivants. Le meilleur exemple de l'extrême sous-inclusivité est la théorie « de la responsabilité de l'agent moral ». Elle est trop restrictive – et paradoxale – parce qu'elle nie pratiquement le statut moral aux êtres qui ne sont pas (entièrement) raisonnables, mais qui sont généralement considérés comme ayant un statut moral (comme les nouveau-nés humains ou les personnes intellectuellement handicapées) ou qui sont des candidats sérieux pour en obtenir (comme certains animaux non humains très intelligents).

À cause des difficultés de l'approche unicritère du statut moral, l'opinion dominante parmi les philosophes semblerait être le « multicritère », c'est-à-dire que « les critères du statut moral doivent inclure certaines propriétés intrinsèques, y compris la vie, la sensibilité et l'identité individuelle ; ainsi que certaines propriétés relationnelles, qui incluent parfois de faire partie d'une communauté sociale ou biologique particulière » (Warren, 1997, p. 21). Cette tendance philosophique vers l'interprétation « multicritère » du statut moral semble être accompagnée par une tendance socioculturelle favorisant la pluralité des interprétations du statut moral d'une grande variété d'entités. La solidité théorique de certaines interprétations peut être fortement mise en doute (par exemple quand le statut moral est interprété de façon extrêmement large afin d'inclure des entités comme les montagnes ou les fleuves). Cependant, un motif compréhensible derrière toute « interprétation large » du statut moral est souvent l'idée que, lorsqu'il est question de sujets moraux, il est préférable d'errer du côté moralement

sécuritaire, c'est-à-dire qu'il vaut mieux assigner le statut moral aux êtres qui ne le méritent pas, plutôt que de ne pas assigner de statut moral aux êtres qui le méritent. Donc, notre « communauté morale » se développera probablement afin d'inclure graduellement quelques êtres et entités au-delà de nos propres espèces.

2. LES ROBOTS ?

En prenant en considération les théories, les critères et les tendances mentionnés ci-dessus, est-il sensé d'assigner un statut moral à certains types de robots ? Selon Keith Abney (2012, p. 50) : « Allons-nous agrandir la communauté morale pour inclure nos camarades (artificiels), ou refuserons-nous aux robots le droit de devenir une nouvelle sorte de nos enfants – ceux qui sont nés, non pas biologiquement, mais par des techniques de fabrication ? » Les opinions concernant cette question sont divergentes. Par exemple, Abney (2012, p. 40) lui-même croit que « nous pouvons affirmer sans risque que, pour le proche avenir, les robots n'auront aucun droit » ; selon Susan Anderson (2011), l'existence des caractéristiques nécessaires pour le statut moral sera difficile à établir chez les robots ; et selon Peter Asaro (2012, p. 180), « nous ne sommes pas susceptibles de faire face à une telle situation de sitôt ». Cependant, Colin Allen et Wendell Wallach (2012, p. 58) croient que, « du moins en théorie, les agents artificiels pourraient atteindre la capacité morale véritable dans le futur, avec des responsabilités et des droits qui seront comparables à ceux des humains » ; Steve Petersen (2012, p. 283) croit qu'« il est possible de créer des robots ayant un intérêt éthique – même de créer des personnes artificielles » ; et selon Kevin Warwick (2012, p. 329), « ce ne sera pas très long avant que de tels robots soient réalisés réellement ». Examinons maintenant quelques raisons favorables ou non à l'attribution du statut moral aux robots.

Une des raisons principales contre l'attribution du statut moral aux robots est le fait que ce ne sont que des machines ou des outils. Les robots ne devraient pas avoir de statut moral, puisque les machines et les outils n'en ont aucun. Un point important est que même les robots très complexes sont seulement programmés pour exécuter un nombre limité de tâches dans un nombre limité d'environnements stables et prévisibles. Ces robots, bien qu'habituellement désignés sous le nom d'« autonomes », ne sont pas autonomes dans le sens philosophique, parce que leur répertoire comportemental entier est préprogrammé

(déterminé) par leurs concepteurs humains. Leur autonomie – en tant que critère principal de la théorie «la responsabilité de l'agent moral» – est pratiquement inexistante. Ils ne sont pas différents des automates célèbres – comme le «Joueur de flute» ou le «Canard en digestion» – construits au XVIII^e siècle par Jacques de Vaucanson, et utilisés plus tard par Immanuel Kant (2002 [1788], p. 128) en tant qu'illustrations commodes du manque d'autonomie et du genre particulier de liberté («la liberté transcendantale») qu'il considérait comme indispensable pour le comportement moral authentique. L'impression que les entités, tels les automates ou les robots, soient en mesure de prendre des décisions autonomes de leur propre libre arbitre, peut être forte et suggestive. Cependant, ce n'est qu'une illusion créée par une tendance psychologique humaine d'anthropomorphiser et imaginer des intentions chez les objets inanimés. Ceci ne devrait pas nous induire à penser qu'ils possèdent quelque sorte de liberté parce que leur «liberté», pour utiliser les mots de Kant, n'est «guère meilleure que la liberté d'un tournebroche, qui effectue des mouvements sans aide lorsqu'il a été remonté» (Kant 2002 [1788], p. 123).

Une autre raison importante contre l'attribution du statut moral aux robots est le fait qu'ils soient incapables, en tant qu'êtres (non vivants) artificiels, d'éprouver le plaisir et la douleur ou toute autre émotion associée, par exemple la crainte, la colère ou la joie. L'intuition sensée derrière cette affirmation, qui est particulièrement accentuée par les philosophes utilitaristes, est la suivante: si un être est incapable de percevoir des expériences, qu'elles soient agréables ou douloureuses, alors il ne peut avoir aucun désir ou intérêt – y compris l'intérêt de ne pas être moralement lésé. Il est impossible d'associer des tâches morales directes à un tel être, de la même façon qu'il est impossible d'associer des tâches morales directes aux objets inanimés comme les pierres ou le sable. Soulignons que cette conception n'est pas exclusivement une idée utilitariste. On la retrouve également dans la vision mécaniste, souvent attribuée à René Descartes, qui affirme que la vivisection est moralement justifiée parce que les animaux sont seulement des automates qui ne peuvent pas sentir de douleur (l'attribution de cette pensée à Descartes est discutée dans Cottingham, 1978). Le message de base est celui-ci: puisque les robots sont des automates non vivants, ils ne peuvent être lésés en aucune manière; on ne peut pas leur faire de tort, et donc, leur assigner un statut moral n'aurait aucun sens, puisqu'ils ne peuvent pas être lésés.

Toutes ces raisons contre le statut moral des robots sont plausibles : il est absurde d'exiger un statut moral pour des machines qui effectuent simplement ce pour quoi elles ont été programmées par leurs concepteurs humains, de la même façon qu'il est absurde d'exiger un statut moral pour des entités qui ne sont pas vivantes et qui ne peuvent percevoir aucune expérience ou avoir un intérêt moralement pertinent. Cependant, la question du statut moral des robots n'aurait jamais été soulevée, si la robotique n'avait pas avancé au-delà des robots mentionnés jusqu'ici. Ce qui rend la question de leur statut moral intéressante philosophiquement est le fait que la robotique prépare déjà le terrain vers de nouveaux types de robots auxquels il sera peut-être plus difficile de nier le statut moral. Deux de ces genres de robots méritent une attention particulière : les robots et biorobots cognitifs.

3. LES ROBOTS COGNITIFS

Une façon pratique de décrire « les robots cognitifs » est de les comparer avec les robots classiques ou « déterministes » (on adopte ici la distinction terminologique « cognitif-déterministe » du rapport sur *l'éthique de la robotique* par la COMEST de l'UNESCO de 2017). Bien que les robots soient « autonomes » pratiquement par définition, ils le sont seulement dans la mesure où ils peuvent exécuter des tâches spécifiques sans être dirigés par un être humain. Cette « autonomie », comme nous l'avons vu, est non seulement techniquement limitée à un certain nombre de tâches dans des environnements fortement structurés, mais elle est réellement inexistante du point de vue philosophique, parce que toutes les actions de ces robots sont strictement préprogrammées par leurs concepteurs humains. Cependant, l'autonomie des robots cognitifs semblerait être différente. À la différence des robots déterministes qui ont besoin de programmation humaine pour effectuer de nouvelles tâches ou résoudre de nouveaux problèmes, les robots cognitifs peuvent apprendre par l'expérience (leur propre expérience, ainsi que l'expérience des autres) et adapter leur fonctionnement à de nouvelles circonstances inattendues dans leur environnement. Les robots cognitifs et leurs méthodes d'apprentissage uniques sont possibles grâce à la fusion de la robotique avec les sous-champs de l'intelligence artificielle, comme l'« apprentissage machine », l'« apprentissage profond » ou « les réseaux neurologiques ». Les détails techniques de ces prouesses sont au-delà de la portée de ce texte, mais trois aspects philosophiques pertinents de l'apprentissage machine tels

que récapitulés dans le rapport de la Royal Society de 2017, devraient être mentionnés :

- 1) L'apprentissage machine est la technologie qui permet aux systèmes d'apprendre directement depuis les exemples, les données et l'expérience. [...] Ces systèmes peuvent [...] apprendre à effectuer des processus complexes avec l'apprentissage des données, plutôt qu'en suivant des règles préprogrammées. [...] En raison de ces avancées, les systèmes qui luttent pour réaliser des résultats précis il y a seulement quelques années peuvent maintenant surpasser les humains dans des tâches spécifiques. (Royal Society, 2017, p. 16)
- 2) L'apprentissage machine permet d'effectuer des tâches d'une telle complexité que les effets désirés ne pourraient pas être spécifiés dans les programmes basés sur des processus étape par étape, créés par des humains. L'élément de l'apprentissage crée également des systèmes qui peuvent être adaptatifs, et qui continuent à améliorer l'exactitude de leurs résultats après qu'ils aient été déployés. (Royal Society, 2017, p. 19)
- 3) Une fois l'apprentissage terminé, beaucoup de systèmes d'apprentissage machine deviennent des « boîtes noires » dont les méthodes sont précises, mais difficiles à interpréter. Bien que de tels systèmes puissent produire des résultats statistiquement fiables, l'utilisateur ne pourra pas nécessairement expliquer comment ces résultats ont été produits ou quelles caractéristiques particulières d'un cas ont été importantes pour la prise de décision finale. (Royal Society, 2017, p. 93)

Les robots, qui sont programmés par apprentissage machine, ont apparemment des similitudes significatives avec les êtres humains, particulièrement leur capacité d'apprendre par l'expérience, l'exactitude de leur raisonnement dans certains contextes et un degré d'inexplicabilité des résultats finaux dans leur apprentissage et leur raisonnement. Cette similitude ne devrait pas nous surprendre, parce que les processus d'apprentissage de ces robots sont fréquemment modélisés sur l'apprentissage par l'expérience des nouveau-nés humains, qui apprennent par l'expérience pendant qu'ils se développent et grandissent (« robotique développementale ») ou selon la façon dont plusieurs organismes vivants apprennent ou s'adaptent à leur environnement respectif (« robotique évolutionnaire »). Les robots cognitifs pourraient aussi être interprétés comme les exemplaires de

l'idée historiquement influente – habituellement attribuée à John Locke – selon laquelle l'esprit est « une page blanche » (*tabula rasa*) qui n'a aucune connaissance innée ou préfixée (« préprogrammée »), mais qui apprend à partir de son expérience et son environnement. La vérité de l'idée de « page blanche », particulièrement si on l'interprète de manière à inclure les connaissances générales et les principes moraux, est souvent considérée comme quelque chose de moralement et politiquement souhaitable. Ceci impliquerait que tous nos maux moraux et sociaux ne sont pas produits par notre nature humaine génétiquement fixée, et qu'ils pourraient être supprimés par des réformes appropriées de notre environnement social, politique et culturel. Les robots cognitifs – avec leur capacité « d'apprentissage direct depuis les exemples, les données, et l'expérience » – pourraient être une étape unique vers cet idéal.

Est-ce qu'on peut considérer ces caractéristiques et capacités, par exemple l'apprentissage par l'expérience, une façon de surpasser les êtres humains dans certaines tâches complexes, de même que l'inexpliquabilité des résultats finaux de leur raisonnement, comme des raisons suffisantes pour justifier l'attribution du statut moral aux robots cognitifs ? On pourrait encore raisonnablement répondre par la négative. On pourrait préciser, par exemple, que l'apprentissage et le raisonnement chez les robots cognitifs ne sont toujours rien d'autre qu'une imitation de l'apprentissage et du raisonnement réel (c'est-à-dire humain) et qu'ils n'ont aucune caractéristique telle que la conscience, le libre arbitre ou la compréhension. Il semblerait normal d'insister sur l'idée que ces caractéristiques sont indispensables pour tracer la ligne entre les robots cognitifs et les êtres humains dans l'attribution du statut moral. Il y a cependant plusieurs réponses possibles. Voici les schémas de deux stratégies possibles pour reconduire la question du statut moral des robots cognitifs.

Une de ces stratégies est de préciser que les choses telles que la conscience, le libre arbitre et la compréhension sont des problèmes de philosophie éternels, sans solutions immédiates. Selon Wallach et Allen (2009, p. 66), elles ont certaines « connotations magiques » qui sont difficiles ou impossibles à expliquer, même si elles sont considérées comme des propriétés uniquement humaines. Les philosophes sont d'ailleurs en désaccord, non seulement au sujet de leur nature exacte, mais aussi à propos de leur existence même. Par exemple, une position influente appelée « le déterminisme dur » rejette complètement l'existence du libre arbitre. Son point de base est plutôt simple et convaincant : tout ce qui se produit dans le monde – y compris nos croyances, nos désirs

et actions – est causé par des événements précédents qui suivent les lois de la nature. Puisque la chaîne de ces causes s'étend dans le passé, avant notre naissance, et puisque nous ne pouvons pas changer les lois de la nature, toutes nos croyances, nos désirs et actions sont déterminés de façon accidentelle et donc de façon non libre. Une autre option que le déterminisme métaphysique est le « déterminisme faible ». C'est une théorie de « compatibilité » qui affirme que le libre arbitre et la causalité universelle existent. Cette compatibilité est réalisée en rétrécissant la définition de la liberté à la capacité de choisir des actions particulières par nous-mêmes, sans coercition ou contrainte externe, sur la base de nos propres croyances et désirs (qui ont eux aussi des causes dans le passé, avant notre naissance).

La discussion détaillée concernant les différentes positions sur le libre arbitre nous amènerait bien au-delà de la portée de ce texte, mais le point important est le suivant : il est problématique d'invoquer le libre arbitre afin de distinguer les êtres humains des robots, parce que leurs libertés respectives pourraient être soit inexistantes (« déterminisme dur ») ou avoir une même nature « étroite » (« déterminisme faible »). La liberté des robots cognitifs est limitée par leur conception technologique et les lois de la nature pertinentes, alors que la liberté humaine est limitée par sa conception évolutive et les lois de la nature pertinentes. Bien que le corps humain et la connaissance humaine soient beaucoup plus complexes que les corps et connaissances robotiques, il n'y a aucun ingrédient « magique » (« libre arbitre ») qui les rend essentiellement différents d'un autre être qui existe en ce monde déterminé de façon causale. Si, d'autre part, un tel ingrédient existe, il n'est probablement pas « magique » ou principalement inexplicable, parce que les robots ne pourraient pas l'avoir.

Kant lui-même – en dépit de sa revendication de la théorie « du libre-arbitre de l'agent moral » et de la « liberté transcendantale » en tant que l'« indépendance de tout ce qui est empirique et ainsi de la nature comme telle » (2002 [1788], p. 123) – se rendait compte que l'existence d'une telle liberté, tout comme l'existence de l'action motivée par des raisons purement morales, ne pouvait jamais être prouvée en théorie. Des stratégies sceptiques semblables sont aussi possibles avec les autres propriétés humaines, qu'on prétend uniques, telles que la conscience et la compréhension. L'essence de ces stratégies est captée avec vivacité dans un dialogue du film *I, Robot (Moi, un Robot)* (2004). Le détective de police interroge un robot : « Vous n'êtes qu'une machine. Une imitation de la vie. Un robot peut-il écrire une symphonie ? Un

robot peut-il transformer une toile en chef d'œuvre?» et le robot lui répond : «Et vous, en êtes-vous capable?». En autres termes, nous pourrions avoir de bonnes raisons pour donner le bénéfice du doute aux robots cognitifs, en ce qui concerne leur statut moral.

Une stratégie intéressante pour plaider en faveur du statut moral des robots cognitifs est de s'appuyer sur les «arguments des cas marginaux», souvent utilisés en éthique appliquée, surtout dans les discussions concernant les droits des animaux. On peut trouver deux versions. Une version se rapporte aux animaux non humains : certains animaux non humains, comme les chimpanzés et les dauphins, sont moins rationnels que certains robots cognitifs très évolués. Si ces animaux non humains sont des candidats pour le statut moral grâce à leur rationalité, alors les robots cognitifs très évolués devraient aussi devenir des candidats. L'autre version se rapporte aux êtres humains : certains êtres humains, comme les nouveau-nés et les personnes qui ont des incapacités intellectuelles graves, sont moins rationnels que certains robots cognitifs très évolués. Si ces êtres humains ont un statut moral, alors les robots cognitifs très évolués devraient aussi en avoir. La rationalité des robots cognitifs (définie par les capacités d'apprentissage, de raisonnement et de prise de décision) ne doit pas nécessairement excéder la rationalité type de l'espèce humaine : il est en effet suffisant qu'elle soit supérieure à la rationalité de certains groupes d'êtres humains. Bref, si quelques animaux très intelligents ont un statut moral (ou sont du moins des candidats sérieux pour l'obtenir), et si les êtres humains avec une rationalité diminuée (ou inexistante) ont un statut moral, pourquoi les robots cognitifs ne devraient-ils pas avoir de statut moral ou être des candidats sérieux pour l'obtenir ? Tant que l'on considère la rationalité comme une caractéristique essentielle pour conférer le statut moral à un être, la cohérence élémentaire exigera que les questions comme ces dernières obtiennent une réponse positive.

4. BIOROBOTS

Un moyen pour rejeter les suggestions précédentes, concernant le statut moral des robots, est de préciser que les robots – indépendamment du niveau de leurs capacités cognitives – ne peuvent pas avoir de statut moral à cause d'un simple fait : ils ne sont pas vivants. Cette objection est plausible. Supposons, comme déjà mentionné, qu'être vivant soit essentiel pour avoir un intérêt quelconque, y compris des

intérêts moraux, les entités non vivantes ne peuvent avoir d'intérêt et, par conséquent, aucun statut moral. Cependant, la robotique contemporaine a déjà fait des avancées qui pourraient rendre cette objection plus faible qu'elle ne semble. Si l'on considère que la création des robots cognitifs a été inspirée seulement par les systèmes biologiques, humains ou animaux, un nouveau genre de robots est en vue qui est réellement, à un degré significatif, un système biologique. On les désigne habituellement sous le nom d'«organismes cybernétiques» ou, plus de façon populaire, des «cyborgs». «Un organisme cybernétique» est typiquement défini comme un «être humain dont le corps est composé, en grande partie ou même pour la plupart, d'éléments robotiques, mais qui est toujours biologiquement vivant» (Gibilisco, 2003, p. 59). Cependant, les organismes cybernétiques ne doivent pas nécessairement être des humains parce qu'un être «hybride» quelconque, qui est composé de pièces biologiques et non biologiques, peut être considéré cybernétique.

Le rapport entre les composants biologiques et non biologiques des organismes cybernétiques peut varier. Roger Clarke (2011, p. 261) crée une distinction utile entre «les humains améliorés-machine» et «les machines biologiquement améliorées». Quand on considère «les humains améliorés-machine», le mot «cyborg» peut sembler incorrect ou exagéré, car il se rapporte aux êtres humains ordinaires, auxquels certaines capacités ont été simplement reconstituées ou améliorées, avec la technologie robotique. Les exemples typiques de ces «cyborgs» sont les personnes avec des stimulateurs cardiaques, des membres artificiels, des implants cochléaires, mais aussi celles à qui on a reconstitué ou amélioré la vue avec la technologie robotique (Tzafestas, 2016, p. 179-183). Dans le contexte de ce texte, «les machines améliorées biologiquement» sont plus intrigantes. C'est le type de cyborg qui ne sera jamais pleinement un être humain ou un autre animal; c'est principalement une machine, mais qui contient quelques composants biologiques indispensables à son fonctionnement. Les organismes cybernétiques de ce genre, «les machines améliorées biologiquement» ou les «biorobots», sont toujours en phase de développement, mais quelques pas ont déjà été faits dans le sous-champ de la robotique normalement dénommé «biorobotique» ou «robotique douce».

Un des exemples mieux connus d'une «machine améliorée biologiquement» est le robot à roues doté d'un cerveau biologique, qui a été développé en 2010 par Kevin Warwick et ses collègues à l'Université de Reading. Ce robot a donc un cerveau biologique composé de

neurones du cortex neural d'un rat, qui est utilisé pour contrôler ses déplacements et éviter les obstacles (Warwick, 2012). Un autre exemple intéressant est un robot miniature qui marche, qui a été créé en 2014 par une équipe de recherche de l'Université de l'Illinois, dirigée par Rashid Bashir. Ce robot n'a aucun «cerveau» biologique, mais plutôt un système de déplacement biologique, construit avec des pièces mécaniques et des cellules de cœur et de muscles de squelette de rat (Cvetkovic *et al.*, 2014). Un autre exemple intéressant est une raie robotique, créée en 2016 par Kit Parker et ses collègues de l'Université de Harvard. Elle a été créée avec des pièces mécaniques, mais elle contenait aussi des cellules de cœur de rat, modifiées génétiquement, qui lui ont permis de nager et de suivre des sources lumineuses qui l'entouraient (Park *et al.* 2016).

Toutes ces «machines améliorées biologiquement» sont très primitives et ne sont en aucune façon comparables à un être humain ou à un «humain amélioré-machine», surtout si on considère leurs capacités d'apprentissage, de raisonnement et de prise de décisions. Ils ne sont certainement même pas près de devenir des agents moraux dans le sens exigé par la théorie «du libre-arbitre de l'agent moral». Cependant, on pourrait dire que cela n'est pas une raison de leur refuser le statut moral et de les traiter comme s'ils étaient seulement des choses ou des machines. Bien que les êtres capables de l'agent moral aient typiquement un statut moral (comme la plupart des êtres humains), les capacités d'agent moral ne sont pas considérées comme essentielles pour le statut moral (une discussion utile concernant cette distinction se trouve dans le contexte de la robotique dans Floridi, 2011). Comme nous l'avons observé, beaucoup de gens considéreraient ce critère comme trop restrictif parce qu'il nie le statut moral non seulement à tous les animaux non humains, mais à certains êtres humains aussi. Du même coup, le statut moral des biorobots ne devrait pas nécessairement dépendre de leurs capacités d'apprentissage, de raisonnement et de prise de décisions. Si le fait d'être vivant ou sensible – comme le suggèrent certaines autres théories de statut moral – est suffisant pour obtenir un statut moral, les biorobots, particulièrement ceux qui sont capables de répondre aux stimulations de l'environnement et de s'y adapter, pourraient réellement devenir des candidats sérieux.

Y a-t-il une façon de repousser cette suggestion ? Une possibilité est le fait que les biorobots ne peuvent pas avoir de statut moral parce qu'ils sont créés artificiellement. Ici un problème se crée, parce que beaucoup d'êtres créés artificiellement ont déjà un statut moral complet ou du

moins un certain genre de statut moral. Les animaux domestiques bénéficient d'un certain degré de protection morale (par exemple contre la cruauté ou contre certaines méthodes d'élevage industriel), et ceci est indépendant du fait qu'ils soient créés artificiellement (par exemple par sélection artificielle ou modification génétique). Un autre exemple mentionné par Petersen (2012, p. 287) pourrait également être pertinent : celui des êtres humains créés par des méthodes comme la fertilisation *in vitro*, l'implantation artificielle, la maternité de substitution, etc. Il serait extrêmement étrange – pour ne pas dire moralement répugnant – d'affirmer que ces êtres humains n'aient aucun droit au statut moral ou que leur statut soit inférieur au statut des êtres humains « créés naturellement ». « L'objection d'artificialité » n'est donc pas un obstacle en soi au statut moral des biorobots.

Une autre objection possible à l'assignation du statut moral aux biorobots est qu'un peu de tissu vivant, intégré dans une machine, n'est pas assez pour conférer un statut moral. Cette objection est plus plausible parce qu'elle nous empêche de glisser dans « l'égalitarisme biologique radical » paradoxal et impraticable. Il y a cependant plusieurs réponses possibles. On ne devrait pas oublier que pour beaucoup de gens « un peu de tissu vivant » est souvent assez pour faire une grande différence morale. Il suffit de penser au statut moral de l'embryon ou du fœtus, partagé non seulement par les conservateurs occidentaux, mais par beaucoup de groupes religieux et culturels dans le monde. Beaucoup de gens, libéraux et conservateurs, ont tendance à avoir une aversion morale forte pour certaines pratiques qui impliquent « un peu de tissu vivant », comme le clonage, la recherche de cellules souches embryonnaires ou la vente des organes. « Un peu de tissu vivant », en autres termes, pourrait facilement devenir important pour la perception publique des biorobots et de leur statut moral. D'ailleurs, si cette quantité devient « un peu plus importante », ou si le tissu est humain, la réticence d'assigner le statut moral aux biorobots diminuerait probablement. Warwick offre l'illustration hypothétique suivante :

La situation pourrait changer avec un robot (conscient) dont le cerveau est créé avec des neurones humains, surtout s'il y en a des milliards. Ce robot aurait plus de cellules de cerveau qu'un chat, un chien ou un chimpanzé, et probablement plus que beaucoup d'êtres humains. Il y a des règlements, des normes et des lois, pour ce genre d'animaux dans la plupart des pays. L'animal doit être respecté et traité au moins raisonnablement. On doit aussi s'occuper de ses besoins. On doit le promener, lui offrir des

grands espaces, ou bien ils peuvent exister naturellement, dans la nature sauvage, sous aucun contrôle humain. Un robot avec un cerveau de neurones humains aura sûrement au moins ces droits, et même peut être plus? On ne peut sûrement pas simplement les traiter comme des choses dans un laboratoire? De plus, si la porte d'un incubateur est laissée ouverte et un robot meurt par mort cérébrale, alors quelqu'un doit être tenu responsable et faire face aux conséquences. (Warwick, 2012, p. 329)

Il y a deux façons de créer des organismes cybernétiques: la première est d'ajouter des éléments biologiques à un tout (mécanique) non biologique; la seconde est de remplacer les éléments biologiques (les organes ou le tissu) d'un organisme vivant avec des parties artificielles, mais fonctionnellement équivalentes. Si ce dernier procédé implique des êtres humains («humain amélioré-machine» comme résultat final), la question du statut moral ne se posera probablement même pas. Le statut moral de ces êtres serait incontestable parce que, selon Wallach et Allen (2009, p. 33), «le composant humain d'un cyborg fournirait l'élément sensible des considérations morales, bien qu'un souci moral reste que la technologie implantée ne puisse pas interférer avec l'autonomie et les facultés morales de l'être humain». Patrick Lin (2012, p. 8) suggère «qu'un cerveau humain conscient (et son corps) a probablement des droits humains, et si l'on remplace certaines parties de ce cerveau avec autre chose, sans déranger ses fonctions, il devrait conserver ces droits». Cependant, il ne semblerait y avoir aucune raison, à première vue, pour ne pas appliquer la même logique dans les cas inverses: si des organismes cybernétiques sont créés dans la direction opposée, c'est-à-dire en ajoutant des éléments biologiques à un tout (mécanique) non biologique. Si un jour «les machines améliorées biologiquement» acquièrent les mêmes (ou presque les mêmes) caractéristiques que les êtres vivants ou sensibles, alors elles devraient avoir le même statut moral que ces êtres ou, au moins, le même droit de réclamer un tel statut.

CONCLUSION

Les robots et biorobots cognitifs sont susceptibles de défier un certain nombre de nos croyances, logées en profondeur, concernant le statut moral d'une grande gamme d'êtres, y compris les êtres humains. Puisque les sciences et technologies contemporaines ont tendance à converger, la convergence entre la robotique, l'intelligence artificielle

et la bioscience est aussi probable, ce qui pourrait mener à la création de robots qui contiennent des composants biologiques animaux ou humains et qui sont capables d'apprentissage sophistiqué et de raisonnement dans des domaines moralement pertinents de la vie. De tels robots brouilleront non seulement la « séparation homme-machine » (Warwick 2015) en général, mais aussi la séparation entre les êtres avec et sans statut moral. En autres termes, le développement de la robotique, selon Gianmarco Veruggio et Keith Abney (2012, p. 350), « exigera probablement l'examen minutieux, sinon la révision complète, de certains de nos concepts opposés, tels que la personne (être moral) contre la machine, la liberté contre le déterminisme, ou l'autonomie intelligente contre algorithme ».

Face à l'amélioration technique future des caractéristiques et des capacités des robots, un certain nombre d'arguments philosophiques, tels ceux discutés dans ce texte, joueront sûrement un rôle important quand le temps viendra où l'on devra attribuer un statut moral aux robots. Il serait peu réaliste, cependant, de s'attendre à ce que ce changement de statut moral se produise en raison des seuls arguments philosophiques. Les effets de deux prédispositions psychologiques des êtres humains seront probablement également décisifs : la première est d'anthropomorphiser et de développer de l'affection pour certains types de robots, et la seconde, de percevoir les robots comme une menace au bien-être ou même à l'existence de quelqu'un. Ces effets s'intensifieront sûrement parce que les robots, contrairement aux animaux non humains qui ont pratiquement disparu de nos vies quotidiennes, deviennent presque omniprésents dans des secteurs aussi différents que le lieu de travail, la santé, l'éducation, le ménage, le divertissement et les rapports sexuels. Dans de telles circonstances, le fardeau de la preuve pourrait facilement passer des partisans de l'attribution du statut moral aux robots à leurs adversaires.

Traduction : Carroll Guérin-Lafèche

Révision : Marie-Hélène Parizeau

BIBLIOGRAPHIE

- Abney, Keith (2012). « Robotics, ethical theory, and metaethics : a guide for the perplexed », dans Patrick Lin, George Bekey et Keith Abney (dir.), *Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 35-52.
- Allen, Colin et Wallach, Wendell (2012). « Moral machines: contradiction in terms or abdication of human responsibility? », dans P. Lin, G. Bekey et K. Abney (dir.),

- Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 55-68.
- Anderson, Susan L. (2011). «The unacceptability of Asimov's Three Laws of Robotics as a basis for machine ethics», dans P. Lin, G. Bekey et K. Abney (dir.), *Machine Ethics*. Cambridge University Press, p. 285-296.
- Asaro, Peter M. (2012). «A body to kick, but still no soul to damn: legal perspectives on robotics», dans P. Lin, G. Bekey et K. Abney (dir.), *Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 169-186.
- Clarke, Roger (2011). «Asimov's Laws of Robotics: Implications for information technology», dans P. Lin, G. Bekey et K. Abney (dir.), *Machine Ethics*. Cambridge University Press, p. 254-284.
- Cottingham, John (1978). «'A brute to the brutes?' Descartes treatment of animals», dans *Philosophy*, vol. 53, n° 206, p. 551-559.
- Cvetkovic, Caroline, Raman, Ritu, Chan, Vincent *et al.* (2014). «Three-dimensionally printed biological machines powered by skeletal muscle», dans *Proceedings of the National Academy of Sciences*, vol. 111, n° 28, 30 juin 2014, p. 10125-10130.
- Floridi, Luciano (2011). «On the morality of artificial agents», dans P. Lin, G. Bekey et K. Abney (dir.), *Machine Ethics*. Cambridge University Press, p. 184-212.
- Gibilisco, Stan (2003). *Concise Encyclopedia of Robotics*. New-York, McGraw-Hill.
- Johnson, E (1999). «Moral status», dans Robert Audi (dir.), *The Cambridge Dictionary of Philosophy*. Cambridge University Press, p. 590.
- Kant, Immanuel (2002 [1788]). *Critique of Practical Reason*. Traduction par Werner S. Pluhar. Cambridge/Indianapolis, Éditions Hackett.
- Lin, Patrick (2012). «Introduction to robot ethics», dans P. Lin, G. Bekey et K. Abney (dir.), *Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 3-15.
- Park, Sung-Jin., Gazzola, Mattia., Park, Kyung Soo *et al.* (2016). «Phototactic guidance of a tissue-engineered soft-robotic ray», dans *Science*, vol. 353, n° 6295, p. 158-162.
- Petersen, Steve (2012). «Designing people to serve», dans P. Lin, G. Bekey et K. Abney (dir.), *Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 283-298.
- Royal Society (2017). *Machine Learning: The Power and Promise of Computers that Learn by Example*. Londres, The Royal Society.
- Singer, Peter (1981). *The Expanding Circle: Ethics, Evolution and Moral Progress*. New Jersey, Princeton University Press.
- Tzafestas, Spyros G. (2016). *Roboethics: A Navigating Overview*. Éditions Springer International Publishing Switzerland.
- UNESCO (2017). «Report of COMEST on robot ethics». Paris, UNESCO. Consulté en ligne.
- Veruggio, Gianmarco et Abney, Keith (2012). «Roboethics: The applied ethics for a new science», dans P. Lin, G. Bekey et K. Abney (dir.), *Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 347-363.
- Wallach, Wendell et Allen, Colin (2009). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press.

- Warren, Mary E. (1997). *Moral Status: Obligations to Persons and Other Living Things*. Oxford, Clarendon Press.
- Warren, Mary E. (2003). «Moral status», dans Raymond G. Frey et Christopher H. Wellman (dir.), *A Companion to Applied Ethics*. Oxford, Blackell, p. 439-450.
- Warwick, Kevin (2012). «Robots with biological brains», dans P. Lin, G. Bekey et K. Abney (dir.), *Robot Ethics: The Ethical and Social Implications of Robotics*. Londres, MIT Press, p. 317-332.
- Warwick, Kevin (2015). «The disappearing human-machine divide», dans Jan Romportl, Eva Zackova et Jozef Kelemen (dir.), *Beyond Artificial Intelligence: The Disappearing Human Machine Divide*. Éditions Springer International Publishing Switzerland, p. 1-10.

CHAPITRE 4

Quel(s) droit(s) pour la robotique ?

Pistes de réflexion sur la gouvernance par le droit du développement de la robotique interactive

Charles-Étienne DANIEL

INTRODUCTION

Qualifiés de « caméléons juridiques » dans un récent rapport déposé conjointement devant l'Assemblée nationale et le Sénat de France, les robots interactifs¹ soulèvent différents défis au regard de leur encadrement par le droit (Le Déaut et Sido, 2016, p. 59). En effet, à l'image d'autres technologies de rupture comme les biotechnologies et les nanotechnologies, la robotique est susceptible d'être employée pour

1. Le terme « interactif » est préféré dans ce texte à d'autres appellations couramment retrouvées dans la littérature telles que « robots de service » ou encore « robots sociaux » (Campa, 2016). Nous entendons désigner sous ce terme les différents dispositifs robotiques destinés à interagir avec les humains en vue d'accomplir une diversité d'applications civiles.

une pluralité d'usages, et ce, dans plusieurs secteurs d'activités : soins et services de santé, agroalimentaire, industriel, transports, services à la clientèle, etc. Différentes machines aux capacités techniques d'autonomie décisionnelle potentiellement évolutive se retrouveront à côtoyer l'homme dans la réalisation des tâches professionnelles et domestiques. Les différents impacts tant positifs que négatifs qui résulteront de l'introduction de robots dans nos sociétés posent dès lors la question de leur encadrement normatif. Comment peut-on baliser le développement de la robotique, afin que l'homme puisse profiter des avantages engendrés par l'introduction de machines de plus en plus sophistiquées, tout en évitant la réalisation de conséquences adverses et indésirables ?

Face au développement de robots interactifs dotés de facultés d'autonomie et d'apprentissage variables, l'adéquation du recours à l'ensemble des règles juridiques actuelles pour les encadrer apparaît toutefois incertaine (Palmerini, 2013 ; Benssoussan et Benssoussan, 2015). Placé devant cette interrogation, le droit mobilisé comme dispositif de gouvernance, c'est-à-dire comme système de règles destiné à orienter et encadrer les différentes conduites en société (Legault et Patenaude, 2007), se retrouve dès lors remis en cause. De quelle manière les règles de droit positif peuvent-elles s'adresser à des machines possédant diverses formes d'intelligence artificielle ? Est-ce que les règles juridiques applicables à l'homme peuvent également être transposées aux machines ? Sinon, quel nouveau droit devrait-on concevoir pour régir la conduite des robots ?

En réaction aux transformations sociales résultant du développement technologique, plusieurs membres de la communauté juridique ont maintes fois souligné l'apparente incapacité du droit à s'adapter au progrès technoscientifique (Marchant et Wallach, 2013 ; Mandel, 2013). Plus précisément, le droit positif est souvent considéré comme lent ou inadapté dans la réponse normative qu'il peut apporter dans l'encadrement des nombreux impacts soulevés par le développement technologique (Marchant, 2011). C'est pour pallier ces limites d'un droit trop rigide et à la remorque des avancées technologiques que de nouvelles formes de gouvernance plus flexibles et mieux adaptées aux multiples changements engendrés par la technoscience sont invoquées par certains juristes (Marchant et Wallach, 2013).

Dans le contexte de l'essor de la robotique interactive, nous retrouvons les mêmes critiques à l'égard de l'encadrement normatif fourni par le droit pour baliser le développement de robots dotés de capacités

techniques d'apprentissage et d'autonomie décisionnelle (Palmerini *et al.*, 2014). Plutôt que de s'en remettre aux règles juridiques traditionnellement retrouvées dans les lois et les règlements, certains préconisent le recours à d'autres instruments normatifs, comme les normes techniques ou les codes et chartes éthiques (Benssoussan et Benssoussan, 2015). De tels instruments normatifs associés à la catégorie du droit souple – ou *soft law* – offrent ainsi une solution de remplacement aux régimes juridiques du droit positif, en proposant des réponses taillées sur mesure et plus rapides pour faire face aux défis entraînés par l'essor de la robotique interactive. À titre d'illustration, certaines lignes directrices ou chartes contenant des principes éthiques (Riek et Howard, 2014) sont proposées pour encadrer le design, la conception et la fabrication de robots destinés à interagir avec une clientèle vulnérable.

La possibilité de recourir à des instruments de *soft law* pour encadrer le développement de la robotique remet toutefois en question, d'une certaine manière, la prépondérance du droit positif comme mode de gouvernance. Quelle place et quel statut peut-on attribuer au *soft law* dans l'encadrement normatif de la robotique interactive ? De quelle manière celui-ci pourrait-il offrir des réponses aux critiques soulevées à l'encontre du droit pour baliser l'essor de robots de plus en plus sophistiqués ?

Pour mieux analyser les implications du recours au *soft law* dans la gouvernance de la robotique interactive, nous effectuerons d'abord au cours d'une première section un survol des principales critiques adressées au droit (I). C'est en précisant les différentes limites du droit sous-jacentes aux propos et critiques retrouvés dans le discours des juristes qu'il nous sera possible d'identifier également quelques solutions intéressantes récemment proposées pour répondre aux insuffisances ainsi soulevées. Dans une deuxième section, nous examinerons plus précisément la nature des autres formes de droit retrouvées dans le concept général de *soft law* et mobilisées par ces réponses normatives différentes (II). Ceci nous permettra d'effectuer par la suite une brève analyse de la place qui pourrait être accordée aux différents instruments normatifs proposés pour encadrer le développement de la robotique.

1. LES LIMITES DU DROIT DANS L'ENCADREMENT DU DÉVELOPPEMENT DE ROBOTS INTERACTIFS ET LES SOLUTIONS DE REMPLACEMENT PROPOSÉES

Grâce aux récents progrès réalisés dans le domaine de l'intelligence artificielle, la robotique bénéficie d'un essor renouvelé depuis ces dernières années et davantage d'attention est accordée à son développement. Principalement destinés à l'exécution de tâches simples, répétitives ou dangereuses (*dull, dirty or dangerous*), les robots ont longtemps été confinés dans des environnements clos et restreints, au sein desquels des tâches précises sont effectuées par un personnel spécialisé (Fletcher et Webb, 2017). La diffusion, l'amélioration des performances et la diminution du coût des senseurs, actionneurs et microprocesseurs ont favorisé la recherche et le développement de robots destinés à des usages de plus en plus diversifiés (CERNA, 2014). Ceux-ci étant susceptibles d'être opérés par des néophytes et de se retrouver dans des environnements publics, les robots interactifs posent toutefois un nombre croissant de défis techniques au regard de leur design et de leur conception. Parmi ceux-ci, Christensen *et al.* soulignent notamment la difficulté qu'ont les robots à s'adapter à un environnement changeant, imprévisible et incertain (2016, p. 6).

C'est dans ce contexte que les capacités techniques d'autonomie décisionnelle et d'apprentissage du robot acquièrent une importance de plus en plus décisive, puisqu'il est particulièrement difficile de prévoir comment celui-ci doit réagir dans une myriade de situations différentes (Christensen *et al.*, 2016, p. 14). L'accroissement de ces capacités octroyées aux robots soulève toutefois d'importantes questions sur les plans éthiques et juridiques, lesquelles sont directement liées aux caractéristiques d'un robot interactif. En effet, l'ajout de la « dimension immatérielle liée à l'autonomie décisionnelle et à la capacité d'interaction » du robot pose la question de l'adéquation du droit face à la complexité des interactions qu'il présente avec l'homme et aux conséquences liées à son apprentissage (Benssoussan et Benssoussan, 2015, p. 5). Bien que plusieurs auteurs réfutent l'idée voulant que la robotique se développe dans un vide juridique des plus complets, l'application des cadres juridiques au développement de robots interactifs soulève néanmoins certaines difficultés (Le Déaut et Sido, 2016 ; Nevejans, 2017 ; Nevejans, 2016 ; Leenes *et al.*, 2017). Après avoir identifié les difficultés susceptibles d'être rencontrées par le droit face au développement de la robotique (1), nous examinerons les différentes solutions de remplacement proposées pour répondre aux limites ainsi soulevées (2).

1.1 Les limites du droit face au développement de la robotique

Comme le résume Calo (2015, p. 534), la robotique présente des défis inédits au regard de son encadrement, en ce qu'elle incorpore une dimension immatérielle – liée à l'autonomie et l'apprentissage – dans un artefact matériel. Par conséquent, l'émergence de comportements totalement inattendus et imprévisibles (« *unpredictable by design* ») pose différents problèmes pour plusieurs règles de droit fondées sur la prévisibilité du comportement et du raisonnement humains (Calo, 2015, p. 542). Ces problèmes sont d'autant plus critiques, puisqu'ils entraînent un risque de préjudices et de dommages physiques. Par ailleurs, le design octroyé au robot est susceptible d'induire chez l'homme une pluralité d'effets psychologiques et émotionnels liés à la perception de la machine (Scheutz, 2012, p. 210-211). Cette perception brouille dès lors les conceptions traditionnelles retrouvées en droit relativement à la distinction entre les personnes physiques et les biens inanimés (Teubner, 2006). Ces caractéristiques inhérentes à la robotique interactive suscitent ainsi un vif débat parmi la communauté juridique relativement à la capacité du droit à pouvoir encadrer de manière effective son développement. Parmi les différentes remises en causes retrouvées dans les débats animant les réflexions juridiques, nous pouvons regrouper en deux grandes catégories les divers questionnements soulevés face à la suffisance du recours au droit positif pour encadrer l'essor de robots interactifs.

La première catégorie de critiques et d'interrogations concerne l'adéquation des cadres juridiques actuels au développement de robots de plus en plus autonomes. Comme le soulignent Matsuzaki et Lindemann (2016, p. 505), la conduite des robots autonomes doit être régie par un ensemble de règles garantissant la santé, la sécurité et le bien-être des citoyens auxquels ils se destinent. Pour établir leur encadrement, il importe de qualifier d'abord les robots sur le plan juridique, ce qui permet par la suite de déterminer la nature des règles juridiques leur étant plus particulièrement applicables (Le Déaut et Sido, 2016). Cependant, la définition même d'un robot pose problème, étant donné la diversité des morphologies, applications et usages imprévus possibles (CERNA, 2014). D'une part, la signification des termes employés diffère entre les univers techniques et juridiques et les définitions retrouvées pour certaines catégories réglementaires – comme les véhicules aériens, automobiles ou les appareils médicaux – apparaissent peu ou prou adéquates à encadrer les robots qui semblent s'y associer (Le Déaut et Sido, 2016; Nevejans, 2017 ;

Palmerini *et al.*, 2014). À titre d'exemple, le consortium européen de recherche RoboLaw (Palmerini *et al.*, 2014, p. 178-179, 191) souligne la difficulté de définir un robot personnel de soins ou une prothèse selon les réglementations actuellement en vigueur. Si un même robot accomplit plusieurs tâches, quel régime juridique doit-on lui octroyer et peut-il cumuler plusieurs définitions?

D'autre part, le robot est qualifié sur le plan juridique comme un bien meuble et ne possède aucun droit de la personnalité ni aucun patrimoine (Matsuzaki et Lindemann, 2016; Richards et Smart, 2016). Néanmoins, à quel point cette distinction demeure-t-elle pertinente face à l'essor de robots de plus en plus autonomes, polyvalents et dotés de capacités d'apprentissage? Cette question revêt un intérêt d'autant plus pertinent à la lumière de la distinction récemment apportée par le *Code civil du Québec* relativement au statut juridique accordé aux animaux, lesquels sont dorénavant reconnus comme des « êtres doués de sensibilité ayant des impératifs biologiques » (art. 898.1 C.c.Q.)². Outre les considérations ontologiques relatives à la conception et à la représentation de l'homme dans son environnement, cette réflexion se retrouve au cœur des débats soulevés dans la communauté juridique à l'égard de l'encadrement juridique du robot. En effet, la responsabilité du fait des gestes et décisions commis par le robot demeure actuellement attribuée aux fabricants ou aux propriétaires et utilisateurs, bien qu'elle pourrait être de plus en plus difficile à établir (Beck, 2016). La pluralité des intervenants qui prennent part à la conception et à la fabrication du robot (« *many hands problem* ») de même que le recours à la communauté *Open source* (Calo, 2011) pourraient compliquer la preuve de l'attribution de la responsabilité parmi l'ensemble des acteurs impliqués (Matsuzaki et Lindemann, 2016; Dahiyat, 2010; Beck, 2016). Peut-on dès lors tenir ces parties responsables des agissements imprévisibles d'une machine de plus en plus autonome? D'un côté, en attribuant aux fabricants la responsabilité des gestes commis par le robot, ceux-ci se retrouvent à supporter le fardeau financier étroitement lié à l'innovation, dans les cas où des indemnités sous forme de dommages-intérêts devront être versées aux victimes ayant subi des préjudices (Hubbard, 2016). D'un autre côté, la diffusion et

2. Avant les amendements effectués dans le *Code civil du Québec*, les animaux tant domestiques que sauvages étaient considérés comme des biens meubles, au même titre que le sont par exemple une table ou un crayon. En sus des modifications législatives apportées, la *Loi sur le bien-être et la sécurité de l'animal* (LRQ, c. B-3.1.) apporte des précisions supplémentaires au regard des protections juridiques leur étant attribuées.

la commercialisation de robots pourraient être compromises, dans l'éventualité où les consommateurs néophytes refuseraient d'être tenus responsables des agissements de robots imprévisibles. C'est dans ce contexte de nature économique que plusieurs juristes s'interrogent sur l'adéquation des règles de responsabilité civile pour assurer la sécurité du développement de la robotique, puisqu'elles ne semblent pas offrir une solution satisfaisante face à l'autonomie grandissante des robots (Matsuzaki et Lindemann, 2016 ; Kearnow, 2016 ; Benssoussan et Benssoussan, 2015 ; van den Berg, 2011).

La seconde catégorie de critiques et d'interrogations concerne plutôt le respect des règles juridiques actuelles par la machine et l'ensemble des intervenants impliqués dans son processus de fabrication. Autrement dit, plusieurs auteurs et rapports s'interrogent sur l'application et le respect de certaines règles juridiques jugées essentielles. En effet, outre le risque de dommages physiques que pourrait causer un robot, celui-ci est également susceptible de porter atteinte à certains droits fondamentaux, tels que le respect de l'autonomie et de la dignité d'une personne, sa vie privée ou encore sa liberté de choix (Palmerini *et al.*, 2014). Comme le souligne Calo (2012, p. 41), la robotique interactive présente un risque d'atteinte à la vie privée inédit, en ce que les robots domestiques pourraient exercer une surveillance étroite et continue des personnes et obtenir des informations personnelles en ayant accès à des lieux privés jusqu'alors restreints ou encore par le fait de confidences indues. Les robots personnels de soins destinés à interagir avec une clientèle particulièrement vulnérable – telle que les personnes âgées ou présentant divers handicaps physiques ou mentaux – présentent également un défi face au respect des droits fondamentaux (Tamburrini, 2016 ; Sharkey et Sharkey, 2012). De surcroît, le recours grandissant à des robots personnels pour assister les personnes âgées dans leurs tâches quotidiennes augmente le risque de leur « objectification », ceux-ci étant susceptibles d'être davantage délaissés par le personnel soignant ou leur famille (Sparrow, 2016). Dans ce contexte où la robotique cohabitera d'une manière de plus en plus intime avec l'homme, la sauvegarde et le respect des droits fondamentaux par l'ensemble des intervenants impliqués dans le processus de conception et du design des robots deviennent dès lors essentiels (CERNA, 2014 ; Palmerini *et al.*, 2014). Dans cette optique, le défi consiste ainsi à assurer l'adhésion des concepteurs, designers et fabricants aux règles de droit et valeurs démocratiques établies par les différentes cultures et sociétés. Comme l'ont précisément remarqué

à cet égard Matsuzaki et Lindemann (2016, p. 512), le contexte social et culturel diffère grandement d'une société à une autre – entre, par exemple, l'Europe et le Japon –, ce qui est susceptible de compliquer davantage la conception de robots destinés à un marché mondial.

Que peut-on alors retenir des différentes critiques animant les discussions dans le domaine juridique à l'égard de la robotique interactive? Parmi l'ensemble des propos mentionnés, nous retrouvons principalement trois conséquences affectant le recours au droit pour réguler son développement. En premier lieu, les questions concernant l'application des régimes et catégories juridiques actuels illustrent un phénomène de « déconnexion normative » (Brownsword, 2008), en vertu duquel le contexte très particulier de la technologie est en dissonance avec le contenu normatif des règles de droit (Bennett Moses, 2011 ; Bennett Moses, 2013). En d'autres termes, l'application des règles de droit – de nature plus générale – s'effectue difficilement étant donné la spécificité de la technologie qu'elles tentent ainsi d'encadrer. C'est en ce sens que certains soulignent l'incohérence de l'application des règles de responsabilité civile aux robots autonomes (désignée sous le terme « *responsibility gap* »), ces dernières n'ayant pas été initialement énoncées pour tenir compte d'artefacts possédant une forme d'autonomie décisionnelle (Benssoussan et Benssoussan, 2015). En second lieu, l'élaboration de règles juridiques plus spécifiques à la robotique exige le respect des procédures formelles d'édiction du droit par les autorités publiques, ce qui entraîne différents délais et ce que Marchant (2011, p. 24) nomme un phénomène d'*ossification du droit*. Par exemple, la lenteur des procédures requises pour l'élaboration des normes relatives à la conception et l'utilisation sécuritaires de robots opérant dans certains milieux, comme celui des voitures autonomes, peuvent notamment freiner l'essor de leur développement (Palmerini *et al.*, 2014 ; Leenes *et al.*, 2017). Positionné en aval du développement de la robotique, le droit se retrouve ainsi à fournir de la sorte une gouvernance essentiellement réactive (House of Commons, 2016). En dernier lieu, le constat effectué au regard des risques d'atteintes à la vie privée, au regard de l'autonomie, de la dignité et de la liberté de choix illustre en quelque sorte une adhésion variable des concepteurs, designers et autres intervenants aux règles juridiques protégeant ces principes fondamentaux. L'attention devant être accordée au respect de ces droits durant toutes les phases de la conception des robots s'avère essentielle, car leur acceptabilité sociale pourrait en être autrement compromise (Leenes *et al.*, 2017, p. 30-31 ; Palmerini *et al.*, 2014, p. 202).

1.2 Les solutions de remplacement proposées pour encadrer la robotique interactive

Étant donné l'ampleur des enjeux éthiques, légaux et sociaux soulevés par le développement de la robotique interactive, différentes propositions normatives ont récemment été suggérées pour accompagner les cadres juridiques actuels. Parmi celles-ci, deux exemples s'illustrent plus particulièrement : une normalisation technique propre à la robotique mobile et de services, de même que la *Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique* (2015/2103(INL)).

Comme l'observe Nevejans (2017, p. 343), si la normalisation s'est révélée très active dans le secteur industriel depuis de nombreuses années, elle reste encore marginale pour les autres secteurs de la robotique. Certains organismes de normalisation tels que l'Organisation internationale de normalisation (ISO), l'Institute of Electrical and Electronics Engineers (IEEE), l'American National Standards Institute (ANSI) ou le British Standard Institute (BSI) ont toutefois amorcé l'élaboration de différents standards techniques propres aux robots et aux interactions et communications humains-machines. Parmi les standards élaborés, plusieurs offrent des définitions sous forme d'une nomenclature caractérisant divers robots sur le plan technique. Par exemple, la norme ISO 19649:2017 énonce un vocabulaire technique propre aux robots mobiles, tandis que la norme ISO 8373:2012 énonce un vocabulaire destiné de façon plus générale aux robots et composantes robotiques. Certaines normes présentent également différentes exigences de sécurité, telles que la norme ISO 13482:2014 (Robots et composantes robotiques – Exigences de sécurité pour les robots de soins personnels) ou encore la norme ISO 13850:2015 (Sécurité des machines – Fonction d'arrêt d'urgence – Principes de conception).

Par ailleurs, le British Standard Institute a publié en 2016 une première norme technique axée sur des considérations éthiques, intitulée BS 8611 : 2016 (*Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems*). Cette norme vise ainsi à prodiguer des lignes directrices pour l'identification des préjudices potentiels que les différents usages de divers robots, notamment ceux pour les soins personnels, médicaux et industriels, pourraient occasionner sur le plan éthique (« *identification of potential ethical harm* »). Elle comporte également des lignes directrices additionnelles visant

à éliminer, sinon réduire le risque associé à la survenance de ces préjudices éthiques à des seuils acceptables (« *eliminate or reduce the risks associated with these ethical hazards to an acceptable level* »). Comme le précise le résumé du document technique, la considération des enjeux éthiques liés aux différents usages de robots se retrouve ainsi explicitement mise en exergue, en ce que les lignes directrices s'appliquent dès la phase du design du robot (BSI, 2017). Dans la même veine, l'IEEE Standards Association a récemment constitué un groupe de travail pour l'élaboration de huit standards³ également destinés à intégrer la considération des enjeux éthiques que peuvent engendrer les robots dès la phase de leur conception (IEEE, 2017 ; Schlenoff, 2012).

D'autre part, la Résolution 2015/2103 (INL) adoptée le 16 février 2017 par le Parlement européen souligne également « l'importance fondamentale » de considérer les différents enjeux éthiques et juridiques soulevés par la nouvelle révolution industrielle qu'entraînera la robotique, sans pour autant étouffer l'innovation (para. B). Dans la résolution adressée à la Commission européenne pour la mise en œuvre de règles de droit civil encadrant la robotique, le Parlement « considère qu'il est essentiel que l'Union actualise et complète son cadre juridique, le cas échéant, en se fondant sur des principes éthiques qui puissent refléter la complexité du sujet que constituent la robotique et ses nombreuses implications sociales, médicales et bioéthiques » (para. 11). Ce dernier préconise ainsi la création d'un « cadre éthique clair, rigoureux et efficace » pour l'ensemble du processus de développement, l'utilisation et la modification des robots qui compléterait les recommandations retrouvées dans la Résolution adoptée (para. 11). À cet égard, le Parlement européen demande à la Commission de présenter une « proposition d'instrument législatif sur les aspects juridiques du développement et de l'utilisation de la robotique et de l'intelligence artificielle à un horizon de 10 ou 15 ans, combinée à des instruments non législatifs, tels que des lignes directrices et des codes de conduite » (para. 51). C'est dans cette optique que le Parlement européen propose en annexe de sa résolution l'élaboration d'un

3. Précisons que les différentes normes spécifiquement élaborées par l'IEEE pour encadrer le développement technique de l'intelligence artificielle et de la robotique sont les suivantes : P7000 – *Model Process for Addressing Ethical Concerns During System Design*; P7001 – *Transparency of Autonomous Systems*; P7002 – *Data Privacy Process*; P7003 – *Algorithmic Bias Considerations*; P7004 – *Standard for Child and Student Data Governance*; P7005 – *Standard for Transparent Employer Data Governance*; P7006 – *Standard for Personal Data Artificial Intelligence (AI) Agent* et P7007 – *Ontological Standard for Ethically Driven Robotics and Automation Systems* (IEEE, 2017).

instrument législatif évolutif qui anticiperait les évolutions prévisibles à moyen terme de la robotique en matière de responsabilité civile et la création d'une charte éthique adressée aux chercheurs, praticiens, utilisateurs et concepteurs pour garantir le respect d'exigences éthiques (2015/2103(INL), Annexe).

Destinées à accompagner de différentes manières les cadres juridiques, ces solutions remettent en question la place bien établie qu'occupe le droit dans l'encadrement qu'il offre face à la robotique (Stradella, 2013). Quelle considération peut-on alors donner à ces normes alternatives sur le plan du droit ? De quelle manière ces dernières peuvent-elles *réellement* accompagner les normes juridiques ? Face aux limites soulevées à l'encontre d'un seul recours au droit positif pour baliser le développement de la robotique, il convient d'examiner plus attentivement le rôle pouvant être accordé à ces normes alternatives.

2. LE RECOURS AUX NORMES ALTERNATIVES POUR GOUVERNER LE DÉVELOPPEMENT DE LA ROBOTIQUE INTERACTIVE

Qualifiées de *soft law*, ces propositions comportent une nature juridique particulière et distincte des normes de droit positif retrouvées dans les lois ou les règlements nécessitant certaines définitions et clarifications sur le plan normatif (1). Toutefois, ce terme vague et général peut se définir plus précisément selon les formes de droit mou et souple identifiées par l'étude annuelle du Conseil d'État français, en 2013. À la lumière de la précision que nous fournirons à l'égard de la nature et des caractéristiques de ces différentes formes de droit, il sera par la suite possible d'examiner plus précisément la place pouvant leur être accordée dans la gouvernance du développement de la robotique (2).

2.1 La nature et les caractéristiques des différentes formes de droit retrouvées dans la gouvernance par le droit

Ne faisant pas l'objet d'un consensus clair et unanime parmi les juristes, le *soft law* représente une catégorie hétéroclite de normes retrouvées en marge des règles du droit positif traditionnel, souvent qualifié de droit « dur » (Hachez, 2012). Une première caractéristique fondamentale distingue les normes hétérogènes du *soft law* des règles de droit positif : bien qu'elle semble s'apparenter à une règle de droit,

une norme de *soft law* ne comporte aucune obligation juridique et, par conséquent, n'entraîne aucune sanction si un comportement observé n'est pas conforme à son énoncé. Autrement dit, le respect des normes de *soft law* demeure une entreprise principalement laissée à la volonté de ses destinataires d'adhérer à son contenu. Plutôt que de commander ou d'imposer une intention ou une action comme l'énoncent les règles de droit dur, celles du *soft law* vont plutôt le recommander, le suggérer ou le déclarer (Thibierge, 2009a). Puisqu'aucune autorité ne peut ordonner et forcer leur application, la condition essentielle de la mise en œuvre des normes de *soft law* repose nécessairement sur l'adhésion des destinataires auxquels elles s'adressent.

Constatant la place grandissante qui lui est accordée depuis quelques années, le Conseil d'État français a effectué dans son étude annuelle de 2013 une analyse approfondie du *soft law*, en créant une doctrine guidant son recours par les autorités publiques. Cette étude annuelle est particulièrement intéressante, puisque cet organe constitutionnel vient officialiser pour une première fois le rôle que peuvent jouer les normes alternatives dans la régulation sociale. Outre l'absence de création de droits ou d'obligations juridiques, deux critères supplémentaires sont soulevés par le Conseil d'État pour caractériser des normes de *soft law* : leur élaboration doit comporter un certain formalisme ou encore un degré de structuration, tandis que leur contenu doit permettre d'orienter, baliser ou modifier le comportement de ses destinataires (2013, p. 63). En d'autres termes, un instrument de *soft law* se reconnaît par la teneur et la présentation formelle et structurée de son contenu.

Par ailleurs, le rapport du Conseil d'État est autrement digne d'intérêt, puisqu'il offre la perspective aux pouvoirs publics d'inclure le *soft law* – qu'il qualifie de droit souple – dans la famille du droit. Pour ce faire, le Conseil d'État trace une frontière entre un droit souple non reconnu par le droit dur – qualifié de *droit mou* – et un « vrai » droit souple, entièrement reconnu par le droit dur (2013, p. 21, 69-71 ; Legault, 2016, p. 279). Pleinement susceptible de se transformer en droit souple et d'acquérir ainsi une juridicité formelle, le droit mou doit pour ce faire bénéficier d'une forme de reconnaissance officielle attribuée par le droit dur. Plus précisément, le passage du droit mou au droit souple s'effectue par une reconnaissance *directe* du droit dur, notamment par une délégation législative ou réglementaire à son égard, ou encore *indirecte*, lorsqu'un juge l'invoque comme moyen d'interprétation auxiliaire de la règle de droit (Conseil d'État, 2013, p. 72-74). Examinons

brèvement de quelle manière peut s'illustrer cette reconnaissance par le droit dur.

Le cas des normes — ou standards — techniques représente l'un des meilleurs exemples d'une reconnaissance directe par le droit dur que sont susceptibles de recevoir de tels instruments normatifs. Comme le précise Lanord Farinelli (2005, p. 738), une norme technique est établie « avec la coopération et le consensus ou l'approbation générale de toutes les parties intéressées et [elle est] fondée sur les résultats conjugués de la science, de la technologie et de l'expérience », alors qu'elle est approuvée par un organisme qualifié sur le plan national, régional ou international. Le recours aux normes techniques permet ainsi d'harmoniser sous la houlette d'un même standard la diversité d'exigences légales et réglementaires susceptibles de se retrouver pour un domaine ou un secteur d'activité particulier au niveau national ou international (Ouedraogo, 2013, p. 163). Par exemple, les normes initialement élaborées par l'Association canadienne de standardisation (CSA), un organisme privé sans but lucratif, sont d'application volontaire. Cependant, le contenu de plusieurs dispositions réglementaires québécoises du *Règlement sur la santé et la sécurité du travail* (RLRQ, c. S-2.1., r. 13) renvoient explicitement à des exigences de sécurité pour divers équipements prévues par les normes CSA, ce qui leur confère alors un caractère obligatoire⁴. L'acquisition d'une obligatorité par voie de délégation réglementaire permet alors de « diminuer le volume des textes réglementaires, d'accroître leur stabilité et de faciliter l'innovation » (Conseil d'État, 2013, p. 13).

Les normes de conduite et éthiques, appartenant *a priori* au droit mou et n'étant donc pas obligatoires, peuvent également bénéficier de différentes formes de reconnaissance par le droit dur. Lorsqu'ils sont officiellement adoptés par les ordres professionnels, les codes de déontologie destinés à orienter la conduite de leurs membres acquièrent alors une portée obligatoire, puisque de tels instruments sont édictés et adoptés selon la même procédure formelle qu'un règlement⁵. Par ailleurs, certaines normes de conduite regroupées dans divers codes et

4. À titre d'illustration, le second alinéa de l'article 45 du *Règlement sur la santé et la sécurité du travail* prévoit que l'équipement individuel de protection respiratoire d'un employé « doit être choisi ajusté, utilisé et entretenu conformément à la norme Choix, entretien et utilisation des respirateurs, CSA Z94.4-93. Un programme de protection respiratoire doit être élaboré et mis en application conformément à cette norme. »

5. C'est en ce sens que le *Code de déontologie des avocats*, R.R.Q., ou le *Code de déontologie des médecins* R.R.Q., énoncent des normes de conduite pour leurs membres.

chartes éthiques peuvent également bénéficier d'une forme de reconnaissance indirecte et sont ainsi susceptibles de se transformer en droit souple. Comme le précise le Conseil d'État, la jurisprudence française a démontré que dans certaines circonstances, des instruments de droit souple peuvent être invoqués devant un juge et faire l'objet d'un contrôle juridictionnel par le tribunal (2013, p. 80).

Les normes de conduite peuvent également aider le juge à interpréter le comportement attendu d'une personne, en vue d'apprécier par exemple une faute invoquée dans le cadre d'un procès en responsabilité civile. Nous retrouvons ainsi quelques exemples à cet égard dans la jurisprudence québécoise. Dans l'affaire *Imbert de Friberg c. Bouchard* (2010 QCCS 1703, para. 64), le juge a précisé que les normes de déontologie « s'avèrent d'une utilité indiscutable dans l'analyse de la faute », bien que leur non-respect n'entraîne pas systématiquement la responsabilité civile du professionnel. La décision *Audet c. Landry* (2009 QCCS 3312) a poussé davantage la réflexion sur le rôle que jouent les normes de conduite en responsabilité civile. Dans cette affaire, la juge s'est précisément basée sur de telles normes pour déterminer que la conduite adoptée par l'intimé dans la préparation d'un rapport d'expertise avait été fautive puisqu'elle portait atteinte à la réputation du demandeur. Alors qu'il était interrogé pendant l'audience sur les qualités essentielles que doit posséder un expert, l'intimé Landry avait souligné qu'il n'était pas membre de la Société des experts en évaluation médico-légale du Québec et qu'il n'était donc pas tenu de se plier à leur code de déontologie. En réfutant cette conclusion qualifiée de simpliste, la Cour a précisé dans sa décision que « ce n'est pas parce qu'un professionnel choisit de ne pas y adhérer [à la Société des experts] qu'il peut décider d'ignorer les règles de l'art en matière d'expertise qu'elle énonce » (para. 90) Bien que la juge a reconnu que le contenu du code de déontologie ne pouvait lier le Tribunal, les normes de conduite qui y sont énoncées peuvent néanmoins l'aider à interpréter ce qui constitue l'état des règles de l'art (para. 144). C'est en ce sens que le Tribunal a conclu que la conduite de l'intimé n'avait pas respecté les règles de l'art et que leur manquement entraînait l'établissement d'une faute engageant sa responsabilité civile.

Il se dégage ainsi de l'analyse du Conseil d'État une redéfinition de la conception traditionnelle de la normativité du droit héritée de Kelsen, selon laquelle les seules règles juridiques obligatoires et sanctionnées constituent le « vrai » droit. À l'inverse, le rapport propose d'ouvrir la

conception du droit à ce qu'il nomme une échelle graduée de la normativité, en vertu de laquelle le droit souple s'imbrique progressivement au droit dur, tout en laissant au droit mou l'occasion d'y trouver une place par le biais d'une reconnaissance directe ou indirecte (Conseil d'État, 2013, p. 68). En d'autres termes, les formes de droit – mou, souple et dur – deviennent alors autant de nuances distinctes dans un univers élargi de la normativité du droit, au sein duquel celui-ci s'exprime sous différentes textures, qu'elles soient sous forme déclaratoire, recommandatoire ou obligatoire (Thibierge, 2009a, p. 153).

En décloisonnant la normativité du droit au-delà de la seule considération des règles juridiques obligatoires et contraignantes, le droit devient – pour employer l'expression de Thibierge – multiple: « multiplicité des ordres, des sources et des normes », où se côtoient et s'influencent mutuellement les formes dure, souple et molle du droit (2016, p. 124). Cette reconnaissance de plus en plus manifestée par les ordres juridiques à l'égard du droit mou et souple nous offre ainsi l'occasion d'examiner plus précisément à quel titre les normes alternatives proposées pour encadrer le développement de la robotique peuvent accompagner les cadres législatifs et réglementaires actuels.

2.2 La considération des propositions dans l'encadrement normatif de la robotique

Comme le précise le Conseil d'État (2013, p. 91), les droits mou et souple permettent d'appréhender certains phénomènes émergents qui se multiplient dans le monde contemporain, qu'il s'agisse d'évolutions technologiques ou de mutations sociales. Aisé à faire évoluer, voire se transformer progressivement en droit dur, le droit mou permet d'établir des « orientations préalables sur des phénomènes qui ne sont pas encore tout à fait bien cernés » (Conseil d'État, 2013, p. 91). Face aux nombreux enjeux sociaux, éthiques et légaux soulevés par la robotique qui témoignent d'un tel contexte de nouveauté, comment peut-on apprécier le rôle du droit mou invoqué à titre d'alternative au droit dur ?

Pour mieux situer le rôle des différentes formes de droit mobilisées face au développement de la robotique, deux critères évaluatifs peuvent être employés afin d'examiner plus précisément les forces et les limites du recours aux droits dur, mou et souple: leur efficacité et leur légitimité sur le plan démocratique (Lenoble, 2002). Intimement

liés aux conditions de mise en œuvre du droit comme mode de gouvernance, ces deux critères offrent de mieux comprendre les fonctions «normatrices» (Conseil d'État, 2013, p. 56) des formes dure, souple ou molle du droit invoquées pour réguler la société. D'une part, le critère d'efficacité du droit peut se comprendre comme étant l'analyse des différents facteurs déterminant sa capacité à pouvoir s'attaquer au problème ou à l'enjeu social qu'il entend régir. Ces facteurs peuvent être notamment reliés au contexte particulier d'application d'une norme ou encore liés à l'adhésion suscitée chez les destinataires auxquels elle s'adresse (Lenoble, 2002; Thibierge, 2009b; Vézina, 2013). Par exemple, à quel point la norme est-elle *réellement* prise en compte par les citoyens? Comment répond-elle au problème ciblé, de quelle manière est-elle appliquée et en quoi représente-t-elle – ou non – une réponse satisfaisante et adéquate? D'autre part, le critère de légitimité peut se définir comme étant la vocation du droit à fournir un sens ou à devenir une référence en vue d'orienter ou de diriger les conduites de ses destinataires (Thibierge, 2009b; Thibierge, 2012). Dans cette optique, la légitimité d'une norme peut s'évaluer à la lumière de différents facteurs tels que «la qualité et l'autorité de son auteur, sa place dans la hiérarchie des normes, la nature de l'instrument qui la porte [...], l'intention qui a présidé à son élaboration, la régularité de son élaboration, la légitimité de son contenu, etc.» (Thibierge, 2009b, p. 822). Tous ces éléments participent ainsi à la construction du sens attribué à la norme par les destinataires auxquels elle s'adresse.

Appliqués aux solutions proposées que sont les normes techniques et la Résolution 2015/2103 (INL) du Parlement européen relative à la robotique, ces deux critères d'analyse nous offrent ainsi de mieux apprécier la place occupée par le droit mou dans la gouvernance de la robotique. Nous l'avons vu, les deux types de solutions de remplacement visent principalement à répondre aux limites et difficultés susceptibles d'être rencontrées par le droit pour encadrer le développement de la robotique interactive. Face aux possibles conséquences d'une déconnexion des règles du droit dur avec son contexte d'application, de la lenteur de leur élaboration ou de l'adhésion variable des destinataires à leur contenu normatif, nous examinerons à quel point le droit mou peut représenter une solution efficace et légitime à ces problèmes reprochés au droit dur.

Sur le plan de l'efficacité, les normes techniques constituent une solution pouvant remédier respectivement aux problèmes de lenteur et d'adhésion des destinataires aux règles de droit. Ainsi, le contenu des

normes techniques est régulièrement révisé et mis à jour, l'ensemble des normes ISO et CSA faisant par exemple l'objet d'une révision par les différents comités d'experts tous les cinq ans afin de garantir leur actualité et leur pertinence pour le marché (ISO, 2017 ; CSA Group, 2017). Par ailleurs, « la normalisation est un facteur de progrès puisque, au regard des avantages, les entreprises sont poussées à s'y conformer » (Nevejans, 2017, p. 342). En effet, la conformité d'un produit ou d'un service à une norme technique représente un avantage concurrentiel pour les entreprises. Une forte dynamique d'adhésion des entreprises au contenu des normes est ainsi suscitée lorsqu'un standard est suffisamment respecté pour devenir une référence sur le marché (Conseil d'État, 2013, p. 105). La norme technique constitue aussi un facteur de compétitivité pour les entreprises, car l'atteinte des objectifs proposés leur offre de rationaliser l'ensemble du cycle de production d'un produit, tout en assurant un gage de qualité et de sécurité, ce qui renforce du même coup la confiance des consommateurs dans le produit (Nevejans, 2017 ; Ouedraogo, 2013 ; Lanord Farinelli, 2005). En l'absence de disposition réglementaire concernant explicitement la robotique interactive, la normalisation présente de ce fait un fort potentiel au regard de son efficacité pour l'élaboration de balises destinées à encadrer plus précisément les enjeux de santé et de sécurité propres à la robotique. Adaptées aux différentes formes et utilisations des robots, les normes techniques pourraient d'ores et déjà poser les jalons d'un encadrement adapté à la réalité des acteurs concernés. Selon l'état du développement, cet encadrement volontaire serait alors susceptible de se muer ultérieurement en réglementation contraignante par le biais d'une délégation législative.

Pour leur part, les diverses mesures normatives proposées dans la Résolution 2015/2103(INL) du Parlement européen sur les règles de droit civil présentent une solution intéressante pour répondre efficacement au problème de la déconnexion du droit face au contexte de la robotique. En effet, le Parlement européen suggère d'adapter et de préciser les règles classiques de responsabilité civile par la mise en place d'un texte législatif de droit dur destiné à évoluer sur un horizon de 10 à 15 ans, lequel serait accompagné de lignes directrices et d'un code de conduite (para. 51). Cet instrument évolutif semble s'inspirer des règles juridiques de type « *sunset* » (Palmerini, 2013, p. 17), lesquelles comportent une durée de vie « normative » limitée requérant une évaluation périodique de la pertinence de leur contenu. Comme le remarque Rodotà (2013, p. 30), la logique traditionnelle prévalant

lors de l'élaboration d'un texte législatif est alors complètement transformée, dans la mesure où les règles énoncées ne s'appuient plus sur un objectif prédéfini, mais entrent plutôt dans une phase d'expérimentation législative. En ce sens, l'exigence normative d'évolution du contenu retrouvée dans le texte présenterait un certain degré d'efficacité en offrant aux règles de responsabilité civile une adaptabilité au contexte mouvant de la robotique interactive. La détermination de la faute et du lien de causalité pourrait dès lors s'ajuster selon l'état du développement des capacités techniques d'autonomie des robots. Cette clarification du régime juridique de responsabilité civile stimulerait du même coup l'innovation et la mise en marché des produits issus du développement de la robotique interactive (Calo, 2011 ; Nevejans, 2017 ; Hubbard, 2016).

De surcroît, un tel instrument législatif comporterait une utilité indéniable dans le cadre d'un procès en responsabilité civile, lequel offrirait au juge une compréhension plus fine du contexte hautement technique et propre à la robotique dans l'adaptation des règles juridiques applicables (Nevejans, 2017, p. 704). Dans la même veine, les lignes directrices et la charte éthique de la robotique accompagnant le texte législatif offriraient de mieux baliser les conduites et les décisions prises par les programmeurs, chercheurs, ingénieurs et utilisateurs dans l'établissement de leur responsabilité en cas de dommages commis par le robot. Ces instruments ne seraient toutefois opposables qu'à l'égard de leurs adhérents et non à des tiers externes, tels que les acheteurs (Nevejans, 2016, p. 30). Comme la Résolution insiste sur l'importance d'une imputabilité de la responsabilité civile à un humain et non à une machine vu l'état actuel de leur développement (para 56), de tels documents normatifs viendraient préciser davantage le critère juridique de comportement d'une personne raisonnablement diligente dans la conception ou l'utilisation d'un robot. L'efficacité des lignes directrices serait également renforcée par la possibilité qu'un juge prenne en compte dans sa décision les lignes directrices pour mieux déterminer l'imputation d'une faute dans l'éventualité d'un dommage commis par le robot.

Néanmoins, toute l'efficacité des normes techniques et de la charte éthique sur la robotique repose sur la nécessaire reconnaissance qu'en font les destinataires concernés par ces normes et le sens qu'ils attribuent à de tels instruments normatifs (Legault, 2016, p. 291 ; Bernatchez, 2016, p. 234). Dans la mesure où ces instruments de droit mou sont d'application volontaire, c'est leur légitimité qui participe à

la construction de ce sens. Plus précisément, la légitimité des normes techniques et de la charte éthique dépend étroitement de l'implication des parties prenantes et de la transparence du processus de leur élaboration, en divulguant par exemple les divers intérêts en jeu (Conseil d'État, 2013, p. 121). L'accessibilité des normes techniques demeure également un facteur important, leur acquisition entraînant souvent d'importants coûts pour leurs destinataires (Conseil d'État, 2013, p. 116).

Au-delà de ce bref examen réalisé à l'égard des propositions soulevées par le Parlement européen, il convient également de s'interroger plus amplement sur le rôle que pourra véritablement jouer le droit mou dans la gouvernance de la robotique. Comme le souligne le Parlement européen, la charte éthique sur la robotique n'est destinée qu'à accompagner subsidiairement le droit dur, en proposant des mesures guidant les usages de robots interactifs dès la phase de leur conception (2015/2103(INL), Annexe). Dans la mesure où la gouvernance juridique de la robotique s'établit de façon prépondérante sur le recours à l'instrument législatif évolutif et les normes techniques pour déterminer des balises sécuritaires applicables aux robots comme produits commercialisés, quelle importance sera réellement attribuée à cette charte éthique ?

Cette interrogation est d'autant plus critique puisque la gouvernance juridique du risque technologique par le droit réglementaire s'inscrit traditionnellement dans une distinction entre les enjeux EHS (« *environmental, health and security issues*») et les enjeux ELS (« *ethical, legal and social issues*») (Patenaude *et al.*, 2015). Au cœur d'une telle logique de gouvernance, une attention prépondérante est accordée aux risques sécuritaires et sanitaires, lesquels se retrouvent alors évalués à la lumière des bénéfices économiques de la technologie (Zei, 2013 ; Patenaude et Legault, 2014). Les « autres » enjeux ELS se retrouvent alors analysés après coup, selon les usages d'un produit technologique déjà jugé sécuritaire et mis en marché (Patenaude *et al.*, 2015). Les revendications sociales exprimées face au développement des biotechnologies et la commercialisation des OGM ont démontré que cette gouvernance juridique ne peut plus se légitimer démocratiquement sur la distinction entre les enjeux EHS et ELS (Patenaude et Legault, 2014 ; Jasanoff, 2003). Plus récemment encore, le débat public national français sur les nanotechnologies de 2009-2010 a illustré la remise en cause de l'acceptabilité sociale et éthique du mode traditionnel de gouvernance juridique, lequel ne peut désormais plus faire l'économie

d'une considération ouverte à l'ensemble des enjeux soulevés (Daniel, Legault et Bernier, 2015 ; Legault, 2015 ; Daniel, 2015).

Il sera donc intéressant de suivre les travaux d'élaboration du texte législatif évolutif et de la charte éthique sur la robotique pour déterminer à quel point ces instruments normatifs se cantonnent dans la distinction EHS/ELS ou s'ils s'ouvrent au contraire à une considération plus inclusive de l'ensemble des enjeux posés par la robotique, tant à l'égard du produit représenté par le robot qu'à l'égard de ses usages.

CONCLUSION

Par l'analyse fournie dans cette contribution, nous avons voulu tracer les premiers jalons d'une double réflexion axée tant sur la nature que sur le rôle des instruments normatifs alternatifs invoqués dans la régulation du développement de la robotique interactive. En premier lieu, nous avons condensé l'ensemble des critiques principalement reprochées au droit positif traditionnel – qualifié de droit dur – en trois catégories. Face aux problèmes reprochés au droit dur relativement à sa lenteur, à la déconnexion de ses règles ou au manque d'adhésion de ses destinataires, nous avons présenté différentes alternatives récemment proposées et mises de l'avant pour y répondre. Plutôt que de figer ces propositions dans la catégorie vague du *soft law* leur étant généralement attribuée, nous avons établi qu'une distinction beaucoup plus fine pouvait leur être accordée selon les catégories de droit mou et de droit souple identifiées par le Conseil d'État français. Ainsi, la normativité conférée à de tels instruments alternatifs s'établit de manière dynamique, en suivant un continuum graduant différentes formes de droit selon lequel l'intensité des obligations contenues dans ces instruments variera en fonction de la reconnaissance directe ou indirecte leur étant attribuée par le droit dur. Autrement dit, nous avons démontré que les deux solitudes traditionnelles (droit et non-droit) sont beaucoup plus nuancées qu'il ne le paraît au premier regard. En vertu de cette normativité plus fluide et mouvante, la nature du droit mou est ainsi appelée à se transformer selon le rôle qui lui est attribué dans la gouvernance par le droit.

En second lieu, nous nous sommes davantage attardés à la place pouvant être accordée à ces différentes propositions normatives alternatives et au rôle du droit mou dans la régulation du développement de la robotique. Pour mieux comprendre et apprécier la pertinence

de leur recours comme moyens alternatifs de régulation, nous avons employé deux critères d'évaluation : l'efficacité qu'ils sont susceptibles de présenter, de même que leur légitimité. D'une part, les normes techniques tout comme l'instrument législatif évolutif et la charte éthique sur la robotique présentent une efficacité qui leur offrirait d'aborder les trois problèmes du droit identifiés dans la première partie. Dans la mesure où le droit mou est mobilisé comme auxiliaire pour préparer le droit dur au développement de la robotique, ces instruments normatifs comportent une utilité et une efficacité qui leur permettrait alors d'établir des balises préalables concernant la sécurité des produits robotiques développés et commercialisés. D'autre part, l'efficacité de ces propositions demeure soumise *a priori* aux conditions régissant leur élaboration. Néanmoins, la légitimité de ces normes de droit mou peut aussi s'apprécier sur le plan démocratique, en vertu duquel elles apparaissent uniquement mobilisées dans le cadre de la gouvernance traditionnelle du droit axée sur la distinction entre les enjeux EHS et ELS.

De cette façon, le droit mou offre une certaine efficacité sous forme de souplesse et de flexibilité pour préparer la gouvernance à l'essor grandissant de robots interactifs. Étant donné la reconnaissance potentielle que peut lui attribuer le droit dur, les propositions alternatives normatives présentées dans cette contribution sont actuellement mobilisées dans le même paradigme de gouvernance que ceux ayant été identifiés dans la régulation des biotechnologies et des nanotechnologies. En effet, les normes techniques sont majoritairement invoquées pour mieux circonscrire les enjeux de sécurité, alors que la charte sur la robotique se destine à encadrer les autres enjeux éthiques et sociaux propres aux usages du robot, une fois ce dernier commercialisé. La légitimité démocratique de ce mode de gouvernance ayant été fortement critiquée, nous entendons poursuivre et approfondir nos travaux à l'égard du rôle que pourrait jouer le droit mou dans une meilleure prise en compte de l'ensemble des enjeux soulevés par le développement de la robotique interactive.

BIBLIOGRAPHIE

- Beck, Susanne (2015). «The problem of ascribing legal responsibility in the case of robotics», *AI and Society*, vol. 31, n° 4, p. 473-481.
- Bennett Moses, Lyria (2013). «How to Think about Law, Regulation, and Technology: Problems with 'Technology' as a Regulatory Target», *Law, Innovation and Technology*, vol. 5, n° 1, p. 1-20.

- Bennett Moses, Lyria (2011). « Agents of Change: How the Law 'Copes' with Technological Change », *Griffith Law Review*, vol. 20, n° 4, p. 763-794.
- Benssoussan, Alain et Benssoussan, Jérémy (2015). *Droit des robots*. Paris, Larcier.
- Bernatchez, Stéphane (2016). « La pratique de l'opération normative : les conditions de possibilité de la gouvernance par le droit », dans Louise Lalonde et Stéphane Bernatchez, avec la collaboration de Georges Azzaria (dir.), *La norme juridique « reformatée », Perspectives québécoises des notions de force normative et de sources revisitées*. Sherbrooke, Éditions Revue de Droit de l'Université de Sherbrooke, p. 203-245.
- British Standard Institute (BSI) (2016). « BS 8611:2016. Robots and robotic devices. Guide to the ethical design and application of robots and robotics systems », 30 avril 2016. Éditions BSI. Consulté en ligne.
- Brownsword, Roger (2008). *Rights, Regulation and the Technological Revolution*. Oxford University Press.
- Calo, Ryan (2015). « Robotics and the Lessons of Cyberlaw », *California Law Review*, vol. 103, n° 3, p. 513-564.
- Calo, Ryan (2012). « Robots and Privacy », dans Patrick Lin, Keith Abney et George Bekey (dir.), *Robot Ethics*. Cambridge, MIT Press, p. 187-202.
- Calo, Ryan (2011). « Open Robotics », *Maryland Law Review*, vol. 70, n° 3, p. 571-613.
- Campa, Riccardo (2016). « The Rise of Social Robots: A Review of the Recent Literature », *Journal of Evolution & Technology*, vol. 26, n° 1, Février 2016, p. 106-113.
- Christensen, Henrik *et al.* (2016). *Next Generation Robotics: Version 1*. Éditions Computing Community Consortium, p. 1-22. Consulté en ligne.
- Commission de réflexion sur l'Éthique de la Recherche en sciences et technologies du Numérique d'Allistene (2014). *Éthique de la recherche en robotique: Rapport n° 1 de la CERNA*. Novembre 2014, p. 1-58. Consulté en ligne.
- Conseil d'État (2013). *Le droit souple*. Paris, La documentation Française.
- CSA Group (2017). « Standards Development ». Consulté en ligne.
- Dahiyat, Emad Abdel Rahim (2010). « Intelligent agents and liability: is it a doctrinal problem or merely a problem of explanation? », *Artificial Intelligence and Law*, vol. 18, n° 1, p. 103-121.
- Daniel, Charles-Étienne (2015). « Gouvernance du développement de la robotique interactive : enjeux éthiques et juridiques », *Ethica*, vol. 20, n° 1, p. 97-135.
- Daniel, Charles-Étienne, Legault, Georges A. et Bernier, Louise (2015). « La régulation des nanotechnologies, le débat national français et le dialogue social : nanomonde, grandes attentes normatives? », *Lex Electronica*, vol. 20, n° 1, Hiver 2015, p. 93-125.
- Fletcher, Sarah et Webb, Paul (2017). « Industrial Robot Ethics: The Challenges of Closer Human Collaboration in Future Manufacturing Systems », dans Aldinhas Ferreira Maria Isabel *et al.* (dir.), *A World with Robots. International Conference on Robot Ethics: ICRE 2015*, coll. « Intelligent Systems, Control and Automation: Science and Engineering », vol. 84. Éditions Springer, p. 159-169.
- Hachez, Isabelle (2012). « Le *soft law*: qui trop embrasse mal étroit? », dans Isabelle Hachez *et al.* (dir.), *Les sources du droit revisitées. Théorie des sources du droit*, vol. 4. Publications des Facultés universitaires Saint-Louis, p. 51-100.

- House of Commons (2016). «Robotics and artificial intelligence», dans *Science and Technology Committee: Fifth Report of Session 2016-17*. p. 1-42. Consulté en ligne.
- Hubbard, Patrick F. (2016). «Allocating the risk of physical injury from “sophisticated robots”: Efficiency, fairness, and innovation», dans Ryan Calo, Michael A. Froomkin et Ian Kerr (dir.), *Robot Law*. Northampton, Edward Elgar, p. 25-48.
- Institute of Electrical and Electronics Engineers (IEEE) (2017). «The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems». Consulté en ligne.
- International Organization for Standardization (ISO) (2017). «How we develop standards». Consulté en ligne.
- Jasanoff, Sheila (2003). «Technologies of humility: citizen participation in governing science», *Minerva*, vol. 41, p. 223-244.
- Kearnow Curtis E. A. (2016). «The application of traditional tort theory to embodied machine intelligence», dans Ryan Calo, Michael A. Froomkin et Ian Kerr (dir.), *Robot Law*. Northampton, Edward Elgar, p. 51-77.
- Lanord, Farinelli Magali (2005). «La norme technique: une source de droit légitime?», *Revue Française de Droit Administratif*, n° 4, Juillet 2005, p. 738-751.
- Law and Ethics*, November 2014, p. 1-215. Consulté en ligne
- Le Déaut, Jean-Yves et Sido, Bruno (2016). «Les robots et la loi», dans *Rapport de l'Office parlementaire d'évaluation des choix scientifiques et technologiques: compte rendu de l'audition publique du 10 décembre 2015 et de la présentation des conclusions du 3 mars 2016*, n° 3551, Assemblée nationale et n° 570, Sénat, p. 1-79. Consulté en ligne.
- Leenes, Ronald, *et al.* (2017). «Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues», *Law, Innovation and Technology*, vol. 9, n° 1, p. 1-44.
- Legault, Georges A. (2016). «Théoriser l'effectuation du droit: force normative et performativité juridique», dans Louise Lalonde et Stéphane Bernatchez avec la collaboration de Georges Azzaria (dir.), *La norme juridique «reformatée». Perspectives québécoises des notions de force normative et de sources revisitées*. Sherbrooke, Éditions Revue de Droit de l'Université de Sherbrooke, p. 267-293.
- Legault, Georges A. (2015). «Enjeux éthiques du développement technologique», *Ethica*, vol. 20, n° 1, p. 49-72.
- Legault, Georges A. (2011). «L'éthique appliquée, la médiation et l'insuffisance du droit: enjeux de gouvernance», dans Louise Lalonde et Stéphane Bernatchez (dir.), *La place du droit dans la nouvelle gouvernance étatique*. Sherbrooke, Éditions Revue de Droit de l'Université de Sherbrooke, p. 19-75.
- Legault, Georges A. et Patenaude, Johane (2007). «Au-delà des critiques adressées aux comités d'éthique de la recherche: un choix de gouvernance», *Journal international de bioéthique*, vol. 18, n° 4, Décembre 2007, p. 17-45.
- Lenoble, Jacques (2002). «Droit et gouvernance. Pour une procéduralisation contextuelle du droit», *Revue Canadienne Droit et Société*, vol. 17, n° 1, p. 1-38.
- Mandel, Gregory N. (2013). «Emerging technology governance», dans Gary E. Marchant, Kenneth W. Abbott et Braden Allenby (dir.), *Innovative Governance Models for Emerging Technologies*, Northampton, Edward Elgar, p. 44-62.

- Marchant, Gary E. et Wallach, Wendell (2013). «Governing the governance of emerging technologies», dans Gary E. Marchant, Kenneth W. Abbott et Braden Allenby (dir.), *Innovative Governance Models for Emerging Technologies*, Northampton, Edward Elgar, p. 136-152.
- Marchant, Gary E. (2011). «The Growing Gap Between Emerging Technologies and the Law», dans Gary E. Marchant, Kenneth W. Abbott et Braden Allenby (dir.), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight*, coll. «*The International Library of Ethics, Law and Technology*», vol. 7. New York, Éditions Springer, p. 19-33.
- Matsuzaki, Hironori et Lindemann, Gesa (2016). «The autonomy-safety-paradox of service robotics in Europe and Japan: a comparative analysis», *AI & Society*, vol. 31, n° 4, November 2016, p. 501-517.
- Nevejans, Nathalie (2017). *Traité de droit et d'éthique de la robotique civile*. Bordeaux, LEH Édition.
- Nevejans, Nathalie (2016). *Règles européennes de droit civil en robotique*, Direction générale des politiques internes, Département thématique C: Droits des citoyens et affaires constitutionnelles, Commission des affaires juridiques du Parlement européen, PE 571.379, Octobre 2016, p. 1-33. Consulté en ligne.
- Ouedraogo, Awalou (2013). «Standard et standardisation: la normativité variable en droit international», dans *Revue québécoise de droit international*, vol. 26, n° 1, p. 155-186.
- Palmerini, Erica, et al. (2014). *Regulating Emerging Robotic Technologies in Europe: Robotics facing*
- Palmerini, Erica (2013). «The interplay between law and technology, or the RoboLaw project in context», dans Erica Palmerini et Elettra Stradella (dir.), *Law and Technology. The Challenge of Regulating Technological Development*. Pise, Pisa University Press, p. 7-24.
- Patenaude, Johane et al. (2015). «Framework for the Analysis of Nanotechnologies' Impacts and Ethical Acceptability: Basis of an Interdisciplinary Approach to Assessing Novel Technologies», dans *Science and Engineering Ethics*, vol. 21, n° 2, Avril 2015, p. 293-315.
- Patenaude, Johane et Legault, Georges A. (2014). «A proposal for an E3LS – ethical, environmental, economic, and social – approach to the regulation of nanomedicine», dans Dora Porto et al. (dir.), *Bioética: saúde, pesquisa, educação*, vol. 2. Brésil, Conselho Federal de Medicina, Sociedade Brasileira de Bioética, p. 77-99.
- Parlement européen (2017). «Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL))». Consulté en ligne.
- Richards, Neil M. et Smart, William D. (2016). «How should the law think about robots?», dans Ryan Calo, Michael A. Fromkin et Ian Kerr (dir.), *Robot Law*, Northampton, Edward Elgar, p. 3-22.
- Riek, Laurel D. et Howard, Don (2014). «A Code of Ethics for the Human-Robot Interaction Profession», dans *We Robot 2014 Conference*, p. 1-10. Consulté en ligne
- Rodotà, Stefano (2013). «Technology and regulation: a two-way discourse», dans Erica Palmerini et Elettra Stradella (dir.), *Law and Technology. The Challenge of Regulating Technological Development*. Pise, Pisa University Press, p. 27-36.

- Scheutz, Matthias (2012). «The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots», dans Patrick Lin, Keith Abney et George Bekey (dir.), *Robot Ethics*. Cambridge, MIT Press, p. 205-221.
- Schlenoff, Craig *et al.* (2012). «An IEEE Standard Ontology for Robotics and Automation», dans *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol.7, n° 12, Octobre 2012. Portugal, Vilamoura, p. 1337-1342.
- Sharkey, Noel et Sharkey, Amanda (2012). «The Rights and Wrongs of Robot Care», dans Patrick Lin, Keith Abney et George Bekey (dir.), *Robot Ethics*. Cambridge, MIT Press, p. 267-282.
- Sparrow, Robert (2016). «Robots in aged care: a dystopian future?», *AI & Society*, vol. 31, n° 4, Novembre 2016, p. 445-454.
- Stradella, Elettra (2013). «Approaches for regulating robotic technologies: lessons learned and concluding remarks», dans Erica Palmerini et Elettra Stradella (dir.), *Law and Technology. The Challenge of Regulating Technological Development*. Pise, Pisa University Press, p. 335-357.
- Tamburrini, Guglielmo (2016). «On the ethical framing of research programs in robotics», *AI & Society*, vol. 31, n° 4, Novembre 2016, p. 463-471.
- Teubner, Gunther (2006). «Rights of Non-Humans? Electronic Agents and Animals as New Actors in Politics and Law», *Journal of Law and Society*, vol. 33, n° 4, p. 497-521.
- Thibierge, Catherine (2016). «La force normative, un des apports conceptuels pour une “théorie ouverte du droit”», dans Louise Lalonde et Stéphane Bernatchez (dir.), *La norme juridique «reformatée». Perspectives québécoises des notions de force normative et de sources revisitées*. Sherbrooke, Éditions Revue de Droit de l'Université de Sherbrooke, p. 99-127.
- Thibierge, Catherine (2012). «“Force normative” et “Validité plurielle”: Des alliées pour une “théorie ouverte du droit”?», dans Isabelle Hachez *et al.* (dir.), *Les sources du droit revisitées. Théorie des sources du droit*, vol. 4. Bruxelles, Publications des Facultés universitaires Saint-Louis, p. 457-499.
- Thibierge, Catherine (2009a). «Rapport de synthèse», dans *Association Henri Capitant des amis de la culture juridique française, Le droit souple: Journées nationales Tome XIII / Boulogne-sur-Mer*. Paris, Dalloz, p.141-161.
- Thibierge, Catherine (2009b). «Conclusion», dans Claude Thibierge (dir.), *La force normative. Naissance d'un concept*. Paris, L.G.D.J., p. 813-838.
- Van den Berg Bibi (2011). «Techno-elicitation: Regulating behaviour through the design of robots», dans Bibi van den Berg et Laura Klaming (dir.), *Technologies on the stand: Legal and ethical questions in neuroscience and robotics*. Nijmegen, Wolf Legal Publishers, p. 417-436.
- Vézina, Christine (2012). «Dans l'angle mort de l'effectivité du droit: une exploration de l'effectivité internormative», dans Georges Azzaria (dir.), *Les cadres théoriques et le droit. Actes de la 2^e Journée d'étude sur la méthodologie et l'épistémologie juridiques*. Cowansville, Éditions Yvon Blais, p. 115-136.
- Zeï, Astrid (2013). «Shifting the boundaries or breaking the branches? On some problems arising with the regulation of technology», dans Erica Palmerini et Elettra Stradella (dir.), *Law and Technology. The Challenge of Regulating Technological Development*. Pise, Pisa University Press, p. 167-204.

CHAPITRE 5

Du robot interprète de la volonté de l'homme au robot réducteur de l'humain

Hoda NEHMÉ

1. INNOVATIONS ET MUTATIONS

En parcourant l'histoire de l'humanité aujourd'hui, nous pourrions avancer certaines affirmations difficiles désormais à contourner ou à laisser dans l'oubli, tant elles ont marqué le cheminement de notre vie en tout temps et en tout lieu. Les différentes innovations issues de la créativité humaine n'ont jamais cessé de se multiplier et de prouver leur fécondité en idées-machines, comme des rouleaux compresseurs, en réponse à une modernité hissée ici et à une modernité émergente là. L'homme façonné autrefois par la religion prend apparemment fin, cédant ainsi la place à l'homme rationnel, libre, autonome, responsable, qui raisonne, expérimente, vérifie et qui théorise, non sans questionnement, une vérité scientifique voulue prouvée et exacte. Dieu est ainsi renvoyé de la scène publique, la religion fait maigre figure ici dans les temps modernes, bien qu'elle remonte ailleurs, de manière fracassante à la surface.

Phénomène expansionniste, la modernité traverse les frontières étanches, s'établit en maître à penser et secoue, par ses progrès, la société, sans halte ni interruption. N'étant pas faite uniquement d'idées,

la modernité exponentielle s'accompagne de révolutions technologiques et industrielles et d'une percée scientifique inédite qui se mue en pouvoir économique, politique, médiatique et en une forme de Foi en le progrès et en la personne capitaliste. S'imposant en raison universelle, la modernité n'a pas manqué à sa mission. La preuve en est, et longtemps nous l'avons constaté, sans la négliger ni la nier, que la modernité a réussi à partager la société humaine en deux :

- a) une société numérique se targuant de suprématie grâce à sa jouissance en TIC, en exploitations scientifiques, en robotique, etc., et mettant ainsi au rancart valeurs, normes, principes, et éthique en faveur d'une suprématie technologique, fabrique de nouveaux dieux générateurs de la nouvelle société oisive ;
- b) une société qui n'est pas fatiguée de se sentir étrangère à tout ce que la modernité véhicule comme philosophie de vie et essor scientifique.

S'enchaîne inévitablement une conception du monde différente, enclenchée depuis le XVIII^e siècle, qui situe l'avenir de l'humanité en fonction du progrès technologique, robotique et économique. Est-ce que la modernité souhaiterait fonder la vie sur une forme de totalitarisme aliénant ? Les modernes se rejettent-ils en eux-mêmes ? Se dirigent-ils vers un post humanisme par exemple ?

2. ÉTAT DES LIEUX : CONFRONTATIONS ENTRE PARTISANS ET OPPOSANTS

Avouons-le et sans crainte, les choses de la modernité remuent toutes les sociétés, développées soient-elles ou en voie de développement. Notre entrée dans le troisième millénaire est tout autre. L'homme postmoderne a introduit le monde dans le chaos au point de dire que l'homme scientifique s'approprie les ingrédients nécessaires à une révolution contre l'homme, révolution qui prépare sa disparition lente mais probablement sûre, livrant ainsi la terre et la vie au gré d'un enjeu technologique susceptible de jeter dans l'oubli le jardin d'Eden et de rendre caduc le contenu des bibliothèques.

S'agit-il d'un intérêt qui s'étaye sur le fait que les humains sont fatigués d'être eux-mêmes ? La technologie, la robotique fourniraient-elles le dernier idéal du moi ? Après le désastre des deux premières guerres mondiales, et par suite des tragiques situations historiques qui

discriminent l'homme dans ses actes prémédités, comme Hiroshima, Auschwitz, la tragédie de Septembre 2001 à New York, le « printemps arabe », les guerres en Afghanistan et en Irak, les guerres inachevées du monde arabo-israélien, etc., l'humain semble œuvrer en faveur d'une rupture qui ferait advenir l'inédit, le transhumanisme. La parole à ce stade se dysfonctionne, elle n'est plus une thérapie, elle perd son énergie cinétique, l'outil se substitue à la parole, la culture numérique ne se gêne pas de porter atteinte au langage, d'accabler l'humain, de plus en plus, par l'automatisme, et de devenir le symbole des sociétés qui offrent à leur clientèle des machines, mettant ainsi l'intelligence humaine à l'épreuve d'interrompre la mécanisation.

Serait-ce une nouvelle forme d'idéologisation numérique qui discrimine l'homme par rapport à l'intelligence artificielle? Ou serait-il question d'une civilisation corrodée par un matérialisme chronique et perforée par un conformisme qui assomme son génie? Ou sommes-nous au seuil d'une vision du futur surprenante, inattendue, imprédictible, gérée par le groupe américain devenu le maître d'Internet, Google, dont le business d'information n'est en fait qu'un élément de contrôle et de maîtrise supplémentaire des personnes humaines?

La question de fond serait de percevoir si, au cœur d'une numérisation obsédante et d'une robotisation bientôt généralisée, il s'agit toujours d'une science-fiction, alors que

- Samuel Bultler avait prédit « qu'il n'y a aucune raison d'être inquiet au sujet du bonheur futur de l'homme tant qu'il sera de quelque utilité aux machines » (Bultler, 1872);
- de nombreux économistes, actuellement, nous préviennent contre la possibilité incontournable de la domination de l'humain par la machine;
- de nombreux industriels n'ont pas la moindre crainte d'atteindre la société robotisée même si le travail; « ils ne croient que le bien économique et le repos de l'être humain face à des travaux durs et agressifs » (Frédéric Joignot, 2016);
- Martin Rees, l'astronome royal du Royaume-Uni, se réjouit d'être à l'aube de l'ère postmoderne et affirme que « le destin de l'espèce humaine est tranché: notre rôle au sein de l'évolution aura été de faire advenir les machines » (Paul Jorin, 2015);
- les acteurs sociaux s'interrogent si l'homme est prêt à se supprimer pour devenir l'éventuelle victime de la mécanisation;

- l'exploration de multiples régulations et transformations au niveau de la vie sociale à l'heure où les multiples visages du robot, voire ceux de n'importe quelle machine, sont classés dans le robot humanoïde;
- les incroyables prophéties de Google qui investit massivement dans le futur de l'humanité, de l'intelligence artificielle, des nano biologies, dans ce qu'on appelle couramment, l'homme augmenté qui repousse les limites des forces de l'âge.

Des confrontations sérieuses entre partisans d'une robotisation, qui s'ancrent, bon gré mal gré, dans notre quotidien et qui invitent à contenir la transformation de notre façon de vivre au travail, de nos interactions, de notre mobilité en raison d'un problème, non seulement économique, mais d'un modèle social en mutation à plus d'un niveau, (Chazelle, 2015); et opposants à l'évolution robotique et à son expansion démesurée, en quête de la personne humaine, de sa place menacée au cœur des mutations et, à qui répond-on que, grâce à la robotisation, le monde serait de plus en sécurité, voire en meilleure condition de vie, et ce, en raison, à titre d'exemple, d'un globule rouge robotisé, minuscule, mille fois plus puissant que sa version biologique, qui permettra au corps humain de se réparer tout seul et à l'esprit d'entrer en contact avec des ordinateurs, etc. Et, en guise de justification, on vous livre, et toujours à titre d'exemple, le modèle avancé par le Darpa (Département de la Défense américaine), assurant la défense et la sécurité de la société par des robots soldats, des super-soldats, tant que la robotisation est maîtrisée et ne nuit pas au développement de l'être humain.

Si la robotisation n'est pas menaçante, pourquoi donc ce déferlement de forums, de séminaires, de symposiums, de colloques, de concours d'idées, pour interroger la science et pour aiguïser une prise de conscience? Pour quelle raison Raja Chatila aurait dit: «si les gens ne profitent pas des grains de productivité générés par les robots, nous n'aurons pas seulement à faire face à un problème économique mais à un problème sociétal»? (Chatila, 2015)

L'enjeu n'est pas ordinaire, l'idéologie de la mécanisation gouverne la société sans scrupule, la déstabilise, change sa relation à l'environnement, à elle-même, à la nature, au travail, au droit, à la consommation, à la santé. Et c'est sans compter toutes les nouvelles avenues à explorer, par exemple les nanorobots, les exosquelettes, les implants cérébraux, etc. Le bouleversement n'est pas simple, la prophétie de Samuel Bultler

se lit dans l'accroissement massif d'humains sans lendemain, et, paradoxalement, dans ce paysage non lointain de bonheur insoutenable qui verrait une réduction pesante et rapide de la population mondiale – à titre d'exemple, un enfant par femme au maximum, plus ou moins un milliard d'êtres humains sur terre. Nous sommes au seuil d'une prise de conscience conflictuelle. Et, dans une telle perspective, nous n'exagérons pas dans notre vision futuriste évoquant la disparition de l'espèce humaine en raison d'une inéluctable propension de la révolution numérique. Il nous faut trouver la recette d'une philosophie de vie qui préserverait la dignité de l'homme avant que l'automatisation ne le réduise à son incompetence, ou à son inaptitude à lutter ou à aller à contrecourant en rival désarmé sinon de ses illusions. Nombreuses sont les interrogations et pointilleuses sont les réponses. Sommes-nous en mesure d'hypothéquer l'avenir? Est-ce que c'est la première fois que l'essor scientifique transforme notre vie? Ou dirions-nous que cette fois cet essor nous ébranle?

L'incontournable développement robotique

De l'automatisation industrielle à la transformation numérique à l'œuvre depuis les années 2000, de profondes transformations ont marqué le développement de la société ici et là.

Les États et leurs gouvernements encouragent leurs industriels, leurs laboratoires scientifiques, leur génie technologique à la numérisation et à la robotisation. Des fortunes colossales sont versées pour parvenir à une robotisation chargée de produire des machines intelligentes couramment connues sous l'appellation de *travailleurs de demain*, de *compagnons de service* et, par extension, de *futurs collègues*.

C'est sur les champs de bataille majeurs que sont l'Irak et l'Afghanistan que les États-Unis se sont remis en question quant à l'échec des interventions extérieures dans lesquelles les opérations militaires ont trouvé en elles-mêmes leur propre fin. De même, nous ne saurions pas occulter d'autres champs de bataille, qui témoignent dans les opérations du « *printemps arabe* » du nombre de drones militaires, à l'œuvre dans tel ou tel autre camp en confrontation, en tant qu'outils vecteurs de victoire sur *l'ennemi* présumé ici et là. Et, c'est en réponse à une telle évidence que l'État américain, d'une part, et tout État, par ailleurs, friand de puissance ou de stabilité politico-économique, finance les recherches pour la production des drones militaires, des robots de surveillance et de défense de l'État et de la société. Le paysage mondial

de nos jours, en matière de robotique, n'appartient plus uniquement à l'Occident. C'est une activité épidémique : là où elle se développe, « elle remet en cause la typologie et le classement actuels des puissances militaires et industrielles, modifie l'équilibre des puissances, recompose le panorama stratégique national et international, et exhorte à une redistribution de la puissance économique entre les pays traditionnellement dominants et les pays émergents dont nombreux maîtrisent les technologies de l'automatisation, de la miniaturisation et de l'informatique » (Les écoles de Saint-Cyr Coëtquidan, 2017).

Est-ce que les promoteurs d'une telle initiative n'y voient pas une éventuelle déresponsabilisation ou un possible désengagement de l'État ? Interrogation hâtive, peut-être, mais elle appréhende la destinée manifeste de la technologie en faveur d'un exercice légitime de la violence par un quelconque État, lequel serait capable de restructurer n'importe quel champ de conflit à son profit sous prétexte de prévention et de neutralisation des conflits. Cette parenthèse a présenté un ensemble de systèmes de robots et s'est interrogée sur la validité et la crédibilité de tout système de robots.

3. L'INCONTOURNABLE NÉCESSITÉ D'UNE IDENTITÉ JURIDIQUE DE « LA PERSONNE ÉLECTRONIQUE »

Aussi paradoxal que cela puisse paraître, en Europe comme aux États-Unis, comme en Asie, et comme partout où la robotique s'amplifie, politiciens, juristes, fiscalistes, éthiciens se penchent sur la question d'une législation chargée de réguler « *la question robot* » et ses retombées socioéconomiques. On propose déjà la conception d'une sorte de code cybernétique qui regroupe les droits et les devoirs des esclaves en métal. C'est aussi une manière d'imposer un code de conduite éthique à l'industrie qui développe les robots. Il est demandé à ce code de pouvoir réguler droit, responsabilité d'erreur, accident de parcours, de santé, impact social, perte de travail, etc. ; au point de soumettre le robot à l'impôt. Il s'agit d'aller vers la création d'un statut de « *la personne électronique* » et d'engager sérieusement le chapitre d'une taxe robots (Verset, 2017). Xavier Oberson, avocat, fiscaliste et professeur à la faculté de Droit, présente ses réflexions au sujet « d'une taxe robots », lors de la leçon d'ouverture, le 21 février 2017, à Uni Dufour.

Si les robots sont conçus dans la perspective de remplacer les êtres humains dans nombre d'activités, y compris celles

considérées jusqu'ici à l'abri de ce genre de développement, l'attention sera portée sur les éventuelles pertes de recettes fiscales pour l'Etat, sur le problème de financement des infrastructures et des assurances sociales. De même, le problème de chômage conséquent et croissant, au fur et à mesure que le chiffre en robots grossit, pose la question de compensation. La taxe semblerait un moyen efficace pour limiter les effets négatifs de la robotique. (Oberson, 2017)

Toute une littérature juridico-économique préventive se conçoit avant l'inflation en robots pour éviter la catastrophe à venir, une sorte de guerre psychique, sociale, économique qui placerait tout État concerné dans une impasse. On imagine alors une telle situation où la robotique confirme sa suprématie sur la place publique, en l'absence d'une taxe sur le revenu, ou de propositions de cotisations sociales que verseraient les entreprises sur la base d'un revenu théorique qu'elles auraient calculé, ou d'une TVA sur les services qu'effectueraient éventuellement les robots, etc.

Nous constatons, à la lecture des pistes de réflexion juridiques et économiques, qu'il n'est plus permis à l'État qui favorise la multiplication des robots dans tous les domaines de faire fi d'une coordination internationale pour légitimer une *taxe robots*.

Bien qu'une taxe robots constitue une bouée de sauvetage à confirmer dans l'avenir proche, il n'en demeure pas moins que la société n'éviterait pas le handicap d'être, d'exister, de produire, de donner sens à la vie.

4. GENÈSE DE LA ROBOTIQUE

Dans l'hier récent comme dans l'aujourd'hui, le monde employé, ouvrier, agricole, fonctionnaire, militaire, médiatique, etc., est sérieusement touché par la mécanisation. Des caisses automatiques au supermarché aux robots agricoles dans les champs, aux ordinateurs qui peuvent à terme « effectuer toutes les tâches auxquelles un raisonnement logique ou un modèle statistique peut apporter une réponse » (Julia, 2017), aux électroménagers, aux iPhone, iPad, iPod, tablettes numériques, à Internet, etc. : autant de formes de robots, qui occupent une place, de plus en plus substantielle, dans notre quotidien. Notons, toutefois, que la secousse sociale provoquée, autrefois, par l'industrialisation avancée s'était inscrite dans le développement

des courants philosophiques dont la tendance avait été d'étayer des systèmes politico-économiques progressistes, socialistes, marxistes, communistes, capitalistes, démocratiques, etc., donnant ainsi lieu à la réflexion humaine d'interroger le progrès scientifique et industriel et ses retombées sur le bien-être de l'humanité.

Aujourd'hui, bien que la machine au service de l'homme n'ait pas disparu et bien que ce ne soit point la machine à l'intérieur d'une usine ou d'une maison ou d'une entreprise qui pose problème ou qui inquiète, toujours est-il que lorsque la machine dépasse l'interprétation de la volonté de l'homme et se transforme en robot humanoïde qui, décrit-on, peut transpirer comme un être humain, il est impensable qu'il laisse l'homme sans inquiétude, le livrant, sans atermoiements, à l'incertitude. L'enjeu majeur est là.

Sur le plan militaire: Que dirions-nous lorsque le drone militaire est la machine qui prend seule une partie des décisions concernant son vol? (Bugeon, 2014) Ne décide-t-il pas son action? Et les effets pervers de l'action militaire provoquée par le drone? Avons-nous compté ou imaginé le nombre de victimes face au pilote ou aux soldats sauvés, puisque remplacés par le drone? Et la guerre à distance qu'effectue une machine programmée minutieusement pour reconnaître le *target* n'est, tout compte fait, qu'une opération qui situe l'humain hors de cette équation. Or, les problèmes qui nous interpellent et nous secouent se concentrent dans les interrogations suivantes: Quels sont les instruments mis à contribution dans de telles conditions de conflit? La Charte des Droits de l'Homme? Les Conventions de Genève? Les protocoles de 1977? La souveraineté des Nations? Comment mettre à l'œuvre le principe de distinction pour un drone sur le champ de bataille, le principe de proportionnalité par rapport aux avantages militaires et le principe de responsabilité? Il me vient à l'esprit, à titre d'exemple, la difficulté au Yémen de prévoir les traumatismes, ou en Syrie, ou à Gaza ou dans n'importe quel lieu du monde où la guerre se fait à distance. Une telle approche ne peut que nous révolter contre cette forme d'attaque qui, en écartant le pilote, dissipe le dernier sentiment humain sur un champ de bataille au nom de la victoire de la machine qui extermine l'être humain sans regrets ni remords.

Toute une crise de culture s'installe en créant des espaces de vie où se muent des antagonistes manipulés ici par le pouvoir scientifique et là par la culture du désespoir et du terrorisme comme seuls moyens en l'absence d'humanisme, du respect de la vie et d'éthique.

Sur le plan médical: Lorsque la machine devient un robot chirurgical qui fait grossir le chiffre d'affaires de l'entreprise concernée, à l'exemple du Robot Rosa (TM) aux États-Unis, en Asie, en Europe, etc., il n'y a plus moyen de contrôler le nombre des patients, des médecins-chirurgiens et des systèmes de santé, qui augmente incessamment et qui confère au Robot Rosa (ou à n'importe quel robot similaire) une notoriété internationale.

Lorsque le logiciel Watson d'IBM se veut fiable à 90 % dans ses diagnostics pour les cancers du poumon, et lorsque le distributeur des médicaments (tel le Hubot Tug, implanté dans plus de 140 hôpitaux américains) menace la fonction infirmière et celle d'aide-soignant, à quoi faut-il s'attendre sur le plan de la protection de la santé?

Lorsque les recherches de Google montrent, d'une part, que, dans une vingtaine d'années, le cerveau humain pourrait être lié à un cerveau synthétique virtuel, et la pensée humaine sera, par voie de conséquence, un mélange de pensée biologique et non biologique, et, d'autre part, que la longévité de la vie humaine serait possible grâce aux exploits de la science nanobiologique qui permettrait de faire reculer notre mort, qui serait plus fiable, le médecin ou le robot?

Les réponses foisonnent dans tous les sens et donnent lieu à un flux d'idées qui remettent en cause l'essor scientifique contre nature, le clonage, la manipulation et le transfert des membres humains, la transformation de sexe, l'exploitation génétique, et, de surcroît, la disparition consciente, licite et illicite de l'humain, au profit d'une société économique formée du binôme science-industrie, générateur de la théologie de l'écrasement optimiste.

Bien que certaines études rejettent l'idée de la disparition du médecin ou de l'infirmier dans un avenir proche, il n'en demeure pas moins que la robotique médicale est en plein essor dans le secteur chirurgical comme dans celui de la réhabilitation, et la résistance à cette progression est inutile du moment que les systèmes sont indispensables à la prise de décision en médecine (Dortier, 2017). Et, à l'heure où la possibilité est donnée à tout un chacun d'être acteur de sa propre santé, il suffit de consulter les téléphones intelligents qui regorgent d'applications pour contrôler les différents types de maladies (ihealth, isommeil, idiabète, etc.).

Sur le plan domestique: Et que dit-on du robot ami et domestique qui peuple les salons et qui s'affirme en compagnon et en support occultant subséquent parents, amis, frères et sœurs, et tout un milieu de

vie, en faveur d'un contrat homme-machine qui, à moyen et à long terme, devient la règle du vivre-ensemble et provoque, à notre sens, l'enlèvement dans un rapport dénué de sentiments et la suppression de l'autre en tant qu'être avec lequel on parle, on s'avoue, on se dit et on se défait du fardeau pesant de la vie. Les robots se proposent comme une extension de la famille et se perçoivent comme des pièces de rechange greffées à notre société, ou comme un nouveau type d'acteur social, œuvrant à la transformation relationnelle, et mettant fin aux émotions et aux pulsions de vie. Pouvons-nous évaluer et estimer quelle révolution s'effectue dans la mémoire d'une personne âgée lorsqu'elle se retrouve à la merci d'un robot? Quelle philosophie sous-tend cette forme de soutien sinon cette insistance à supprimer la mémoire humaine ou à la déraciner pour l'implanter dans de nouvelles conditions de vie qui ne ressemblent pas à celles de l'itinéraire historique de la personne.

Sur le plan de l'emploi: Que dire du champ de travail? Nous ne saurons pas négliger une évidence traduite dans une nouvelle restructuration du marché de l'emploi et dans une refonte également de l'ensemble des relations professionnelles et personnelles. Approchons la question à partir des statistiques tirées des différentes études alarmantes portant sur le changement structurel de l'emploi :

- l'IFR (International Federation of Robotics) projette plus de 1,3 million de robots industriels actifs dans le monde à partir de 2018.
- Il est prévu d'ici 2025 la disparition de plus de 3 millions d'emplois en France, par exemple. Un taux de chômage de 10 % agite déjà l'équilibre social: que ferait un État si 30 % ou plus des travailleurs du pays étaient remplacés par des systèmes robotisés? Nos sociétés sont-elles préparées à supporter l'ingression de tels taux de chômage, alors que la recherche juridique relative à l'amplification robotique sur le marché de l'emploi est, à cette heure, encore timide?
- «Le dirigeant d'une multinationale gagne aujourd'hui 450 fois plus que le salaire moyen dans son entreprise, grâce au remplacement de l'homme par la machine, alors qu'autrefois Ford gagnait 40 fois plus que ses ouvriers [...]» (Nasi, 2015)
- Actuellement, il est mentionné que 147 sociétés contrôlent 40 % de la richesse totale du réseau influant sur l'économie mondiale (Nasi, 2015). Il est impensable, sur ce plan, d'occulter le

phénomène de la concentration de la richesse vers le patrimoine et celui de la réduction de sa distribution par le travail. Aucun secteur n'échappe à la robotique, y compris la presse « dont la mise en scène éditoriale est déjà sous le contrôle des robots ».

La conséquence immédiate de l'omniprésence robotique est le chômage en amplification démesurée ou les solutions économiques génératrices de questionnements à l'infini. Mais, ce qui demeure indiscutable, c'est qu'il est nécessaire de déployer des efforts pour repenser la société sans travail comme norme. Le travail, perçu par la société comme un facteur d'identité et d'intégration sociale, est en retrait au profit d'une société oisive en devenir.

Tous les fondements de la vie en société seront examinés. Dorénavant, ce n'est plus *comment revenir à l'économie de plein emploi?* qui interpelle politiciens, juristes, économistes, ou gouvernements, mais *comment se préparer à l'économie de l'emploi rare?*

Si la société en pleine robotisation ne développe pas une idéologie autour du travail rare et de l'oisiveté à venir, la société postmoderne, hautement développée et qualifiée techniquement, sombrera dans la grande misère humaine et, probablement à un tel moment critique et éprouvant, il n'y aura ni Victor Hugo, ni Karl Marx, ni Hannah Arendt, ni Simone Weil, ni un héros hybride homme-dieu pour soutenir l'humanité.

Nous avons traité quelques propos relevant du mythe technologique et nous sommes enclin à reconnaître le progrès ascendant qui a jalonné l'histoire de l'humanité, assumé des améliorations et engagé des adaptations et des transformations désormais incontournables.

Si nous nous mettons à énumérer les progrès qui ont épargné à l'être humain des métiers désagréables ou avilissants, la liste serait interminable, à titre d'exemple, le porteur d'eau, le cocher, le pétrisseur à pain, la domestique, le magasinier, etc. Les machines de substitution ont permis une mutation sociale non sans profit tant sur le plan de la productivité que sur celui de l'emploi. La robotique qui annonce une avancée dans le fonctionnement industriel, agricole, médical, chirurgical, éducatif, médiatique, etc., rend certainement de grands services aux sociétés.

Conscientisation: réfléchissons avec éthique sur le robot sexuel en devenir, revisitons le clonage, examinons les nanorobots, repensons le

nucléaire, parcourons les différentes études relatives à la robotisation et mettons en exergue une prise de conscience préventive, susceptible de préserver l'Homme, de lui assurer un développement intégral et de préserver la nature, le système écologique, autrement dit, préserver la VIE dans toutes ses dimensions et tous ses aspects. À ce sujet me revient à l'esprit cet ancien adage: si la vie ne vaut rien, rien ne vaut la vie, et la vie demeure notre objectif.

Ce n'est pas la profession ni le robot collègue ou travailleur de demain qui posent seuls problèmes. C'est en fait le bouleversement social, nous dirions même civilisationnel, qui compte. Car si dans l'avenir proche on pourra guider un fauteuil pour handicapés, commander une prothèse, conduire une voiture, rédiger un éditorial, prodiguer un médicament, offrir un diagnostic parfait, rien qu'en sollicitant les ondes émises par notre cerveau, qui empêcherait alors qu'on transmette nos influx érotiques ou nos sentiments ou notre savoir de la même manière?

5. VERS QUEL HOMME S'ACHEMINE-T-ON ?

À cette question certains répondent: vers *l'homme désincarné*, libéré des vicissitudes de l'enveloppe physique, sorte d'intelligence artificielle, qu'un sursaut éducatif prévisionnel, qui aurait permis éventuellement à l'employé menacé par le robot d'exploiter les nouvelles technologies pour se trouver un travail, se vide de son sens.

D'autres avancent que nous nous acheminons vers une vie éternelle de l'esprit et de la connaissance par le biais de la mémoire dans un Big data planétaire (Fathi, 2016).

Sommes-nous au seuil d'une civilisation qui voit en Dieu une ancienne forme d'intelligence artificielle? Le posthumanisme reprend le vieil adage de Descartes: renoncer au savoir acquis, rejeter le savoir traditionnel, construire avec sa raison naturelle son savoir, quitter les matières d'éveil et s'orienter vers la science.

Une fracture civilisationnelle s'impose et une nouvelle histoire humaine commence. Nous proposons encore une fois à l'humanité en progrès d'aller à la rencontre de l'homme ancêtre, de se référer à la puissance de la mémoire, aux travaux écrits par tant d'humains qui ont donné à la civilisation humaine la possibilité de dire et de se dire.

Nous partageons avec tous ceux qui étudient la robotisation ou qui s'interrogent à son sujet cet espoir de repenser le savoir transmis qui

constitue la Mémoire de l'Humanité dans sa diversité culturelle et de donner à ce savoir les possibilités de revivification pour éviter que la mécanisation nous ôte la grâce de Vivre et d'Être Corps et Esprit, Mémoire et Innovation.

CONCLUSION

Se convaincre que le recours aux technologies et à la robotisation est incontournable, parce que celles-ci sont coextensives, est une illusion qui occulte l'extrême pauvreté de l'automatisme.

Croire fermement que les animaux robots destinés à distraire et à amuser les enfants contribuent à leur assurer un environnement chaleureux, où ils se sentent aimés, signifie cautionner cet attachement narcissique amplifié et symptôme irréversible d'isolement.

Insister sur la nécessité de l'usage des robots émotionnels, largement répandus dans la société contemporaine comme forme de remédiation à la carence affective, est une preuve de perte de capacités relationnelles avec l'autre et une forme d'isolation sociale audacieuse à l'ère des réseaux de la communication et de l'information;

Prétendre que le *robot de service* est un partenaire appelé à partager une culture avec l'humain, pose un problème d'aliénation. Rejeter l'idée que le robot s'impose dans notre paysage culturel et social, serait un aveuglement ou une inconscience de la réalité des choses. Recevoir la promesse du super soldat, de l'aide-soignant, comme une progression vers un monde meilleur, occulte sciemment la mauvaise interprétation du super soldat et toutes les neurosciences qui prouvent que le robot aide-soignant pour les personnes âgées ou délaissées n'est certes pas la meilleure solution qui leur est offerte en matière de remèdes et de soins. Revisiter l'évaluation de la robotique prostatique à titre d'exemple, pour découvrir que, suite aux comparaisons effectuées entre opération avec médecin et opération avec robot, aucun progrès n'a été signalé. Ignorer les robots à l'heure actuelle semble un phénomène utopique.

Alors? Si nous considérons que : a) les robots ont permis la sortie des tâches subalternes; b) la machine n'est plus perçue comme un outil, mais comme un modèle désirable; c) la performance ne se confond pas nécessairement avec l'Intelligence; d) dénoncer la robotisation n'est pas notre principal objectif; e) repenser l'ordinaire est bien plus

intéressant que la mystification des choses externes; f) la revalorisation du normal, et la réévaluation de nos valeurs sont bien plus importantes dans notre vie que la focalisation sur le super héros; g) le cerveau n'est pas le Big Data, mais c'est un réseau instable constamment renouvelé et une force constante complexe et imprédictible; h) l'avancement se fait grâce à un cerveau sociétal et non à un implant ou à une augmentation chirurgicale.

Toutes ces analyses et ces observations complexes mettent face à face les impensés idéologiques de la robotisation et la difficile définition de l'humain à venir.

De l'Homme moderne à l'Homme postmoderne, à celui posthumaniste, à l'humain de demain, lequel aura la chance de perdurer?

À notre avis, ce sera l'Homme qui aura compris qu'il ne peut s'approprier la vérité et que l'espace-temps et l'espace-lieu sont deux données pour réaliser une part de notre humanité en besoin constant de savoir pour la découvrir.

BIBLIOGRAPHIE

- Association relative à la télévision européenne (ARTE) (2016). *Nos collègues, les robots*. Documentaire en ligne: <https://www.arte.tv/fr/>
- Baynes, Sian et Ross, Jen (2007). The «Digital Native» and «Digital Immigrant»: a dangerous Opposition. Conférence annuelle de la Society of Research into Higher Education (SRHE), University of Edinburg.
- Butler, Samuel (2005 [1872]). *Erewhon*. Consulté en ligne.
- Bugeon François, «Les robots vont-ils nous mettre au chômage?» in *Libération*, le 27 octobre 2014.
- Chatila Raja, (2015). <https://www.meta-media.fr/2015/07/07/les-robots-vont-non-seulement-impacter-leconomie-mais-aussi-modifier-notre-societe.html>
- Chazelle, Barbara (2015). «Les robots vont non seulement impacter l'économie, mais aussi transformer notre société», dans Méta-Média. Consulté en ligne
- Davidenkoff, Emmanuela (2014). *Le tsunami numérique*. Collection, Essais-Documents, Éditions StockGanascia.
- Dessibourg, Olivier (2016). «Des scientifiques rappellent l'urgence de légiférer pour encadrer l'évolution rapide de l'intelligence artificielle», *Le Devoir*. Consulté en ligne.
- Dortier, Jean-François (2012). «L'intelligence artificielle, espoirs et réalisations», dans Jean-François Dortier (dir.), *Le cerveau et la pensée*. Éditions Sciences humaines.
- Dortier, Jean-François (2015). «Les robots vont-ils tuer les emplois?», dans *Revue Sciences Humaines*, n° 274, Septembre-October 2015. Consulté en ligne.
- Fathi (2016). «Résistance Inutile», en commentaire sous l'article «Les robots vont-ils tuer les emplois?» Écrit par Jean-François Dortier. Consulté en ligne.

- Ganascia, Jean-Gabriel (2017). *Le mythe de la singularité, Faut-il craindre l'intelligence artificielle?* Collection : science ouverte, Paris, Éditions du Seuil.
- Gualtieri, Julia (2017). « Quand la numérisation s'attaque aux classes moyennes », dans *Digital Society Forum*. Consulté en ligne
- Guichard, E. (2011). « Le mythe de la fracture numérique », dans *Regards croisés sur l'internet*. Villeurbanne, Presses de l'ENSIB, p. 69-100.
- Joignot, Frédéric (2016). « Robotisation générale », *Le Monde*, 16 novembre 2016.
- Jost, Céline (2009). « Robotique et intelligence artificielle », dans *Valoria*. Master en recherche informatique encadré par Dominique Duhaut. Consulté en ligne.
- Jorin, Paul « Emploi et robotisation : progrès ou prolétarisation de la société », *Le Monde.fr*, 11 décembre 2015.
- Karsenti, Thierry, et Collin, Simon (2013). « Usages des technologies en éducation : analyse des enjeux socioculturels », *Éducation et francophonie*, vol. XLI, Printemps 2013.
- Langier, Edouard (2014). « Après l'informatisation et la numérisation, la robotisation », *Le nouvel économiste*. Consulté en ligne.
- Les écoles de Saint-Cyr Coëtquidan (2017). « Robotisation de coercition ». Consulté en ligne.
- L'Organisation de coopération et de développement économiques (OCDE) (2011). *L'enseignement supérieur à l'horizon 2030 - Volume 2: Mondialisation, La recherche et l'innovation dans l'enseignement*. Paris, Éditions OCDE.
- Nasi, Margherita (2015). « Emplois et robotisation : progrès ou prolétarisation de la société », *Le monde*, 14 décembre 2015. Consulté en ligne.
- Oberson Xavier, Une taxe robot pour éviter le naufrage de l'État social, in *Service de Communication, le journal de l'UNIGE*, n° 127, 3 février 2017.
- Oudeyer, Pierre-Yves (2009). « Les grands défis de la robotique au 21^{ème} siècle », *Science, Technologie et société*. Consulté en ligne.
- Poellhuber, Bruno et Boulanger, Raymond (2001). *Un modèle constructiviste d'intégration des TIC : rapport de recherche*. Consulté en ligne.
- Reverd, Christophe (2016). « Quelle est la place de la robotique pédagogique au sein de l'éducation? », *Vitrine technologique éducation*. Consulté en ligne.
- Robin, Jean-Pierre (2016). « Robots et nouvelles technologies feront perdre 5 millions d'emplois d'ici à 2020 », *Le Figaro*, 18 janvier 2016. Consulté en ligne.
- Sciences humaines (2014). « Changer le travail : 20 pistes pour améliorer la qualité de vie au travail », *Revue Sciences Humaines*. Grands Dossiers n° 36, Septembre-October-Novembre 2014. Consulté en ligne.
- Sylvain (2015). « Les robots partout », en commentaire sous l'article *Les robots vont-ils tuer les emplois?* Écrit par Jean-François Dortier. Consulté en ligne.
- Verset, Jean-Claude (2017). « Le robot disposera d'une identité juridique : « la personne électronique » », dans la rubrique *Info* de la Radio-Télévision belge de la Communauté française. Consulté en ligne.

CHAPITRE 6

Ingénierie neuromorphique et intelligence artificielle : construire un oiseau qui ne sait possiblement pas voler

Frédéric DUBOIS

« Apprendre le secret du vol d'un oiseau a été beaucoup comme apprendre le secret de la magie d'un magicien : une fois qu'on connaît le truc et qu'on sait où regarder, on voit des choses qu'on n'avait pas remarquées lorsqu'on ne savait pas exactement où regarder. »

(trad. libre, Orville Wright,
How we invented the airplane, p.16)

Les théories computationnelles de l'esprit (Putnam, 1992 ; 2002 ; Fodor, 1974) – incluant, dans un sens plus large, le domaine des neurosciences computationnelles (NSC) en général (Hodgkin et Huxley, 1952 ; Marr, 1982 ; Schwartz, 1990) – se sont imposées depuis la deuxième moitié du XX^e siècle avec l'idée qu'il serait envisageable dans le futur de reproduire les facultés cognitives humaines en algorithme calculable par ordinateur. Suivant leur leitmotiv, l'ordinateur serait au cerveau ce que l'esprit serait à un programme informatique. Évoquant dans le même élan l'ordinateur et le cerveau, la formule a bien de quoi en

faire rêver plus d'un. En effet, sans y prendre garde, on peut facilement interpréter cette ritournelle comme une relation d'identité entre le cerveau et l'ordinateur (cerveau = ordinateur), à comprendre tout simplement que le cerveau ne serait rien de plus qu'une machine sophistiquée de traitement symbolique de « 1 » et de « 0 ». Après tout, la cybernétique (Wiener, 1948) a déjà depuis longtemps défini le cerveau et les organes sensoriels biologiques comme des outils autorégulés de traitements d'informations. Le fonctionnalisme computationnel n'est finalement qu'un pas de plus vers l'affirmation de cette idée que le corps humain est une « machine », bien que complexe. En somme, si le cerveau *est* bel et bien un ordinateur, et que, subséquemment l'esprit *est* un programme informatique, qu'est-ce qui nous empêcherait de reproduire les mécanismes de la Raison dans le cuivre et le silicium ? Pourrait-on alors matérialiser le rêve de la science-fiction et discuter dans un futur proche avec une machine consciente ?

Sauf si, justement, l'optique était de défendre une position « forte¹ » de l'identité cerveau-ordinateur (ce qui ne sera pas notre cas), plusieurs théoriciens de l'esprit se voient pourtant assez réfractaires à l'idée de réduire le cerveau à des processus physico-chimiques algorithmisés, ce qui ferait par le fait même abstraction de toute expérience qualitative (*qualia*) de la conscience humaine. En fait, le projet initial inspiré du leitmotiv cerveau-ordinateur peut dès le départ nous sembler quelque peu trompeur lorsque l'on considère que les pionniers des théories computationnelles que sont Hilary Putnam (2002) et Jerry Fodor (1974) étaient de fervents défenseurs d'un physicalisme non réductionniste. De toute évidence, la formule originelle du cerveau-ordinateur ne visait pas directement à la réalisation du robot conscient futuriste. Beaucoup plus terre à terre, leur objectif visait plutôt à invalider les théories comportementales et les thèses de corrélation forte entre le cerveau et le mental² dominant les sciences de l'après-guerre. Ainsi, l'analogie

-
1. L'opposition entre une théorie « forte » (*strong AI*) et une théorie « faible » (*weak AI*) de l'intelligence artificielle, qu'utilise notamment John Searle (1980), conçoit, d'un côté, les théories avançant la possibilité de concevoir une machine pensante et consciente au même titre que l'humain et, de l'autre, les théories voyant plutôt les comportements des machines comme des imitations des actions humaines, sans pour autant que la simulation ne signifie la reproduction réelle du phénomène.
 2. La thèse de la corrélation entre le mental et le physique a été notamment défendue par John J. C. Smart dans les années 1950. Selon cette thèse, chaque événement mental (disons, ressentir une douleur) peut être réduit à un phénomène physico-chimique (disons, pour la douleur, l'activation des fibres C) ayant lieu au même moment dans l'organisme. Bref, nos sensations et nos émotions ne seraient *rien de plus* que des phénomènes neurochimiques ayant lieu dans notre cerveau. Nous pourrions alors conclure, à la manière des

cerveau-ordinateur devait servir d'argument en faveur d'une théorie nouvelle, à savoir le « fonctionnalisme³ » dit « computationnel ». Selon cette théorie, le monde serait un vaste réseau de relations où chaque élément imbriqué joue un rôle (fonction) dans la régulation des systèmes. Pour le fonctionnalisme, le cerveau n'est pas seulement un réceptacle d'entrées (*input*) et de sorties (*output*), comme le voulaient les théories béhavioristes du début du XX^e siècle, même si ces deux dernières notions demeurent centrales au fonctionnement du cerveau fonctionnaliste. Le fonctionnalisme ajoute à l'équation l'ensemble des composantes internes du cerveau (mémoire, système nerveux, organes sensoriels, etc.) pour les placer dans un vaste réseau d'interactions causales, où chacune des composantes joue son rôle dans la maintenance du système neuronal. La cohérence du système assure alors une fonction unifiée, celle qui, finalement, décrit tout autant le cerveau humain que l'ordinateur : la simple tâche de traiter l'information reçue. Nous faudrait-il alors nuancer la relation d'identité cerveau-ordinateur qui nous a servi à introduire ce chapitre ? En fin de compte, s'il existe pour le fonctionnalisme un rapprochement entre le cerveau et l'ordinateur, c'est bien sur le plan de la tâche analogue que tous deux réalisent, qui se résumant encore une fois au traitement d'informations. Simplement dit, pour le fonctionnalisme computationnel, le cerveau *est* comme un ordinateur, puisque tous deux font la *même* chose, soit servir de support matériel à l'esprit (ou à un programme informatique) qui traite de manière algorithmique les tâches du quotidien.

Cette reformulation fonctionnaliste du cerveau et de l'esprit ne réduit certainement pas les espoirs de réaliser un être artificiel conscient. Au contraire, le fonctionnalisme computationnel contribue de manière

théoriciens éliminativistes (Churchland, 1981 ; Churchland, 1986), que toute affirmation sur les états mentaux ne découlant pas des neurosciences relève inévitablement de pseudosciences.

3. Le terme « fonctionnalisme » renvoie à plusieurs écoles, disciplines et méthodes d'analyses différentes. Il va de soi qu'un sociologue ou un anthropologue – on penserait ici à l'influence de Bronislaw Malinowski (1884-1942) ou de Radcliff-Brown (1881-1955) – pourrait comprendre différemment cette référence au « fonctionnalisme », qu'un philosophe de l'esprit associerait peut-être plutôt à des auteurs comme Hilary Putnam et Jerry Fodor. Notre interprétation du fonctionnalisme s'approche d'ailleurs beaucoup de celle d'Hilary Putnam, et nous concevons au courant de cet article le terme « fonctionnaliste » comme synonyme de théorie computationnelle de l'esprit. S'il va encore de soi que, d'un auteur à un autre, certaines nuances permettent des rapprochements théoriques qui seraient plutôt interdits selon d'autres auteurs, cette note métrologique ne devrait pas pour autant nuire à notre propos. En effet, nous cherchons à rendre compte d'un esprit général de l'époque contemporaine en prenant comme témoin l'idée du cerveau-ordinateur.

décisive à rendre cette idée plus *vivante* que jamais. En effet, les travaux d'Hilary Putnam et Jerry Fodor dans les années 1960 et 1970 mettaient déjà l'accent sur la possibilité de libérer l'esprit du domaine exclusif des processus physico-chimiques biologiques. L'idée à retenir de leur fonctionnalisme computationnel – et qui deviendra l'une des idées les plus universellement acceptées, comme le rapportait l'influent philosophe analytique Jaegwon Kim (1992) – sera celle de la « réalisabilité multiple » des propriétés mentales. En résumé, on viendrait à comprendre qu'il n'est jamais nécessaire d'imiter la nature, molécule par molécule afin de reproduire artificiellement les phénomènes naturels, tout comme l'ordinateur n'a pas besoin d'être fait de matières organiques pour calculer. C'est alors en vertu « isomorphisme fonctionnel », pour reprendre les termes de Putnam (1992), à savoir que deux entités différentes remplissent des fonctions analogues, que l'on se permet finalement de faire le rapprochement entre le cerveau et l'ordinateur.

Comme nous le verrons dans la première partie de ce texte, le recours à l'analogie fonctionnaliste précède de loin l'invention de l'ordinateur. Nous nous contenterons de reculer qu'au début du XX^e siècle, plus précisément en 1903 alors que les frères Orville et Wilbur Wright feront décoller le premier engin volant motorisé. Ce sera ici l'analogie avion-oiseau qui retiendra notre attention, une analogie que Jerry Kaplan résume dans les mêmes termes du fonctionnalisme introduit plus haut : « oui, les avions ont des ailes, qui ont sûrement été inspirées des ailes des oiseaux, mais les avions ne battent ni ne plient leurs ailes et leur système de propulsion est complètement différent des oiseaux, tant par leur portée, par l'altitude que, en fait, par à peu près tout à leur propos » (trad. libre, Kaplan, 2016, p. 16). Ce détour par l'histoire de l'aviation nous permettra ainsi d'ouvrir quelques pistes de réflexions sur ce qui motive l'usage de l'analogie dans les développements technologiques.

Notre questionnement deviendra alors : pourquoi l'ordinateur serait-il un cerveau artificiel, alors que nul n'oserait avancer qu'un avion est un oiseau artificiel ; encore moins dire que l'oiseau est un avion biologique ? Qu'est-ce qui nous permet ainsi ce glissement sémantique lorsqu'on en vient à définir l'esprit humain ? Dans la deuxième partie de ce chapitre, nous reprendrons justement le jalon de la réalisabilité multiple et du fonctionnalisme, en nous plongeant dans l'univers de l'ingénierie neuromorphique. En bref, cette discipline porte la prétention de pouvoir reproduire artificiellement les fonctions du cerveau

biologique humain dans la machine-ordinateur. Cette thèse découle naturellement des prémisses de l'isomorphisme fonctionnel dans la mesure où, si le fait qu'une tâche (fonction) ne soit pas limitée à être réalisée par une structure physique fixe (disons, l'acte de penser qui serait limité aux cerveaux biologiques humains), qu'est-ce qui empêcherait qu'une autre structure physique (de cuivre et de silicium?) puisse être apte à soutenir les fonctions de la conscience ?

Nous concluons notre recherche en troisième partie en offrant des pistes de réflexions supplémentaires afin de mieux saisir notre objet d'étude. À l'instar de la philosophie de Gilbert Simondon (2012), qui nous servira d'ailleurs momentanément de guide, l'instant de cette troisième partie, nous laisserons entendre que la question de l'objet technique gagnerait à être adéquatement revisitée pour ces justes considérations ontologiques. Il nous restera alors à démêler de la question du cerveau-ordinateur le réel du fantasme, de la possibilité théorique à la possibilité réelle d'espérer voir dans un futur rapprocher l'ordinateur robot partager avec nous, humain, ce mystère qu'est la conscience.

1. LE BATTEMENT DES AILES : UN ISOMORPHISME FONCTIONNEL

L'avion est aujourd'hui une réalité presque banale. Pourtant, la machine volante a longtemps relevé du domaine de l'imaginaire, objets de contes et de légendes fantastiques. Des tapis magiques du Moyen-Orient aux manteaux de plumes de la déesse nordique Freyja, passant (plus localement) par les canots enchantés de la chasse-galerie canadienne jusqu'aux (plus classiquement) ailes de cire d'Icare, les cultures du monde entier semblent toutes avoir dans leur répertoire au moins une référence mythique au vol. Munis de capes et d'arsenaux emplumés, les premiers « fous volants » de l'Histoire ont aussi tenté de réaliser ce rêve vieux comme Hérode en s'élançant des collines et des hautes tours du Monde connu⁴. L'idée derrière ces tentatives hasardeuses – que l'on sait aujourd'hui être vaines – était

4. Nous retenons l'histoire d'Eilmer de Malmesbury, surnommé le « moine volant », qui, vers l'an 1000, eut l'idée de s'élaner du haut de la tour de l'abbé Malmesbury, recouvert de plumes de la tête aux pieds. Avant lui, en Andalousie, vers 875, Abbas Ebas Firnas Abou Kassem se serait élancé – avec succès, selon l'historien médiéviste Lyhn White (1978) – d'une colline avec la même intention de voler à l'aide d'un planeur.

qu'en battant suffisamment fort des bras emplumés, l'on imiterait le mouvement de l'oiseau et, subséquemment, l'on réussirait à s'envoler. En fin de compte, cette manière de faire aura conduit à bien plus d'os brisés qu'à des heures de vol effectives. Pour les plus optimistes de ces « fous volants », l'échec n'était pas tant compris comme le signe d'un rêve impossible à réaliser, mais plutôt comme la résultante d'une quelconque faute technique (p. ex. un mauvais choix de plumes⁵), voire parfois de la simple mauvaise fortune (p. ex. un coup de vent malchanceux) (Marck, 2012). En fin de compte, les échecs s'accumuleront pendant des siècles de sorte que, « dans cette histoire, écrit Daniel Parrochia, les réussites n'étaient que des légendes, et tous les départs des "redéparts"⁶ » (Parrochia, 2003, p. 28). Le vol restera ainsi pendant des siècles une fantasmagorie, que l'homme de science raisonnable se saurait prendre au sérieux...

1.1 Vers une théorie fonctionnaliste du vol

Fred C. Kelly, ami et biographe des frères Wright, introduit l'ouvrage *How we invented the airplane* en rappelant une idée selon lui fautive, mais déjà répandue à peine quelques décennies après l'invention de l'avion. « Une erreur commune, écrit-il, est de croire que la réussite du vol avait été retardée principalement par le manque de moteurs légers opérationnels. Fait étant que les frères Wright avaient résolu la plus grande partie du problème du vol – équilibre et contrôle, et la conception adéquate des ailes – dès leur planeur de 1902, sans du tout de moteur. » (trad. libre, Wright, 1953, p. 3) Kelly répondait alors à une idée courante selon laquelle les grandes inventions techniques n'ont été rendues possibles qu'en raison (exclusive) du niveau d'avancées techniques d'une époque donnée. Certes, nous ne pouvons pas nier l'importance d'un savoir technique dans l'invention d'une nouvelle machine. Or, le point intéressant, et qui est certainement beaucoup plus déterminant dans l'invention de l'avion, aura été, selon Kelly, une question non pas pratique, mais plutôt théorique. Comme l'explique

5. Comme le rapporte l'historien Bernard Marck, cela a notamment été le cas de John Damian, un Écossais qui, en 1507, après s'être brisé le fémur à la suite d'un essai de vol, reconnaît avoir commis l'erreur « d'avoir préféré des plumes de poulet à la parure de l'aigle! » (Marck, 2012, p. 14)

6. Ce passage fait directement référence à l'analyse qu'a faite Gaston Bachelard de l'épistémologie du vol en préface de l'ouvrage de Duhem, Jules (1959). *Histoire des origines du vol à réaction*, Paris, Nouvelles éditions latines.

Kelly, « les expérimentateurs précédents n'ont échoué qu'à seulement la connaissance des principes de l'aérodynamisme, ce que de leur côté les frères Wright ont laborieusement réussi » (trad. libre, Wright, 1953, p. 3). Comme on peut le lire dans le témoignage qu'offre Orville Wright (1953), le succès des frères Wright aura en effet beaucoup plus été dû à leur passion pour un savoir contemplatif, qui les poussa à remettre en question les *a priori* qui pendant des siècles ont agi comme obstacle à la réalisation du vol.

De toute évidence, si l'humain a pu maîtriser les techniques de vol, ce n'est certes pas parce qu'il a appris à battre plus rapidement des ailes. En effet, l'erreur des précurseurs du vol – et que n'auraient pas commise les frères Wright – aura été de penser l'acte de voler comme une question de mouvement mécanique rapide d'un corps emplumé sans considérer à leur juste valeur les principes scientifiques de l'aérodynamisme. À la question « pourquoi, durant tout ce temps, des hommes vont-ils tomber et se tuer en essayant vainement de voler ? » Parrochia répond dans un chapitre consacré à ce problème :

Parce qu'à la question « comment les oiseaux volent ? », ils répondent, comme Aristote, de manière biaisée, victimes qu'ils sont des apparences : pour voler, il faut battre des ailes préalablement emplumées, parce que, bien que ce mouvement ne soit pas naturel, le privilège de la plume, comme celui de l'air ou du feu, est la légèreté, et que les corps légers, comme on le sait, sont supposés aller plutôt vers le haut. (Parrochia, 2003, p. 28)

Le biais aristotélien en question aurait été celui de concevoir le vol non pas à partir des principes de l'aérodynamisme – certes, encore sommaires à son époque –, mais comme celui d'un mouvement (*kinesis*) des ailes grâce auxquelles l'oiseau, de force, pouvait s'orienter dans les airs (Parrochia, 2003, p. 22). Une deuxième erreur aristotélienne aurait par la suite été de comprendre le mouvement dans l'air selon les mêmes lois que le mouvement sous l'eau. En effet, Aristote aurait conçu les nageoires des poissons (« *ptérugia* », en grec) comme l'équivalent aquatique des ailes (« *ptéruques* ») des oiseaux (Parrochia, 2003, p. 25). Suivant la poussée d'Archimède, on savait déjà à l'époque d'Aristote qu'un corps physique plongé dans l'eau subira une force verticale directement proportionnelle à la masse du même corps physique. S'il avait été de même pour un corps se déplaçant dans l'air, on se serait alors attendu à ce que puisse voler tout corps possédant une masse suffisamment basse pour vaincre la force verticale de l'air. En

bref, en vertu des repères théoriques de l'époque, la légèreté du corps semblait, en toute conséquence logique, être une condition nécessaire pour s'envoler.

Finalement, sans le savoir, les pionniers du vol faisaient face à un véritable « obstacle épistémologique », au même sens que l'entendait Gaston Bachelard dans *La formation de l'esprit scientifique* (1947). Certes, l'oiseau vole grâce à ses ailes et, par ailleurs, l'hélium de la montgolfière est plus léger que l'air. Or, nous le savons aujourd'hui, voler ne requiert pas pour autant qu'en tout lieu il soit nécessaire ni de battre des ailes ni d'être plus léger que l'air. Pourtant s'il peut nous sembler presque ridicule aujourd'hui de dire que battre des ailes rapidement n'est pas suffisant pour voler, c'est bien parce qu'au cours de l'histoire technique, il y a eu un changement important dans la manière même de concevoir le monde, ce qui a en retour rendu possible l'invention de l'avion plus de deux millénaires après le biais d'Aristote. Nous qualifions ici cette conception moderne de « pensée fonctionnaliste », dont les succès théoriques et pratiques seront dus à sa méthodologie d'opérationnalisation de fonctions. Pour demeurer dans le domaine de l'aviation, prenons à titre d'exemple la définition moderne du vol que donne le dictionnaire Larousse. Le vol y est défini comme l'« action de voler ; [une] suite de mouvements d'ailes coordonnés au moyen desquels les oiseaux, les insectes se soutiennent dans les airs et s'y déplacent ». Le « vol » demeure un mouvement (*kinesis*), mais dans l'optique que ce mouvement (est) un « moyen » en vue de réaliser une fin particulière, celle de se soutenir et de se déplacer dans l'air. Le « vol » n'est donc pas seulement le « mouvement », qui n'est qu'un moyen pour réaliser le « maintien dans les airs » (fonction réelle du vol) *en fonction* des lois de l'aérodynamisme. L'idée de « légèreté » n'est conservée que dans un sens littéraire et poétique, à savoir « Vol : (Littéraire) Déplacement d'un objet léger dans l'air ». Désenchanté, l'esprit scientifique se concentre plutôt sur la définition de l'action comme celle de la réalisation de principes physiques réels (et non pas métaphoriques, issus d'une mauvaise lecture de la poussée d'Archimède). Il s'agira encore une fois, dans le cas du vol, des principes aérodynamiques, que les frères Wright ont dû étudier et maîtriser pour faire décoller et maintenir dans les airs le célèbre *Flyer*.

1.2 Isomorphisme et multiréalisabilité

Comme nous l'avons souligné, le fonctionnalisme ouvre grand la porte à la perspective de reproduire certaines fonctions caractéristiques d'une structure physique particulière (disons, un oiseau) dans un artefact technique (disons alors un avion). Dans cette optique, l'on dira que «l'oiseau est comme un avion» puisque tous deux réalisent la même fonction de voler. Dès lors, une mise en garde s'impose. Dire qu'un oiseau *est comme* un avion ne pose pas une relation d'identité stricte de type «oiseau = avion», du même type qu'«eau = H₂O». En effet, nul ne saurait prétendre qu'un avion *est* un oiseau, ou qu'en retour l'oiseau *est* un avion. «Est comme» ne réfère donc pas à une relation de synonymie, mais plutôt à une analogie, c'est-à-dire au rapprochement de deux choses distinctes en vertu de similitudes. Ces similitudes en question ne se trouvent évidemment pas dans les propriétés physiques de l'oiseau et de l'avion – deux entités évidemment très différentes – ni tant dans leurs formes qui diffèrent aussi sur plusieurs aspects. Quitte à nous répéter, précisons qu'il s'agit bien entendu d'une similitude de fonctions. Nous dirons alors que ces deux entités se recourent sous la grande catégorie des «êtres-volants».

Les propriétés qui unissent sous une même catégorie l'oiseau et l'avion ont alors moins à voir avec leurs particularités physiques, même si elles jouent un certain rôle évident dans l'exécution du vol. Les propriétés mises en rapport sont plutôt des propriétés dites fonctionnelles. Le fonctionnalisme définira de manière formelle ces propriétés de sorte que: Pour une chose x , avoir F (ou être un F) = $_{\text{def}}$ pour x , avoir une propriété P telle que $C(P)$, où $C(P)$ est une spécification de la tâche causale que P est censée remplir dans x (Kim, 2008, p. 136).

Nous reprendrons plus loin cette même définition formelle pour montrer comment le cerveau-ordinateur y est introduit. Pour le moment, il nous faut retenir du fonctionnalisme une première idée, que Putnam nomme l'isomorphie fonctionnelle: «deux systèmes sont fonctionnellement isomorphes s'il y a une correspondance entre les états de l'un et les états de l'autre qui préserve les relations fonctionnelles» (Hilary Putnam, 1992, p. 91). L'isomorphie fonctionnelle met ainsi une expression sur cette idée que le fonctionnalisme cherche à exprimer en rapprochant analogiquement tantôt l'oiseau et l'avion, plus tard le cerveau et l'ordinateur. L'idée se résume ainsi: deux structures physiques différentes, l'une biologique (un cerveau) et l'autre technologique

(ordinateur), sont *identiques* si, en dépit des différences sur le plan de leurs structures physiques, les propriétés fonctionnelles demeurent réciproques.

On nommera à son tour «réalisabilité multiple» (ou «multiréalisabilité») cette deuxième idée selon laquelle deux types différents de structures physiques peuvent réaliser un même état interne. Comme l'écrivait déjà Alan Turing en 1950, «aucun ingénieur ou chimiste ne prétend être capable de produire un matériau que rien ne distingue de la peau humaine» (Turing, 1999, p. 136-137). Le projet en question demeure bien entendu de créer l'équivalence fonctionnelle en utilisant d'autres matériaux, probablement, dans notre cas, le cuivre et le silicium. Turing continue :

Dans le système nerveux, les phénomènes chimiques sont au moins aussi importants que les phénomènes électriques. Dans certains ordinateurs le système de mémorisation est principalement acoustique. On voit donc que l'utilisation de l'électricité n'est qu'une similarité très superficielle. Si nous souhaitons découvrir de telles similarités, nous devrions plutôt chercher des analogies mathématiques de fonction. (Turing, 1999, p.144)

En somme, si la réalisabilité multiple est avérée, le fait qu'un ordinateur moderne fonctionne sous tension électrique alors que son ancêtre du XIX^e siècle, la machine de Babbage, utilisait plutôt des cartes perforées ne change en effet rien au fait que – sauf peut-être sur le plan de l'efficacité, mais nous y reviendrons – ces deux structures physiques distinctes réaliseront les mêmes directives (disons, par exemple, de calculer la somme de «4» et «2»). La multiréalisabilité des fonctions mentales est certes cruciale à cette idée qu'un même état interne puisse être réalisé dans deux structures tout aussi différentes qu'un cerveau biologique puisse l'être d'un ordinateur. En résumé, à partir du fait observable que l'oiseau et l'avion réalisent de manière multiple l'acte de voler, pouvons-nous aussi conclure qu'un ordinateur en cuivre et en silicium puisse venir jouer le même rôle qu'un organe biologique, qui plus est – et sûrement de manière plus controversée que tout autre organe – jouer le rôle d'un cerveau apte à penser consciemment? Force est logiquement d'admettre que si la réalisabilité multiple est avérée, la faculté de penser, un état interne (ou mental), ne peut être le domaine exclusif des structures biologiques, encore moins seul de l'être humain. Comme l'écrit Putnam, sans même devoir se référer à l'idée d'une machine artificielle pensante, «si nous pouvons trouver un

seul prédicat psychologique qui peut être clairement appliqué à la fois à un mammifère et à une pieuvre (par exemple «avoir faim»), mais dont le «corrélat» physico-chimique est différent dans les deux cas, la théorie de l'état cérébral s'effondre⁷», (Putnam, 2002, p. 281). L'argument de la réalisabilité multiple entend en fin de compte s'attaquer à une sorte de «chauvinisme neural» (Block, 1992) voulant que l'humain soit le seul animal doté de neurones à posséder des états mentaux. Pourtant, le saut de l'animal à la machine ne va pas pour autant de soi, dans la mesure où il soulève la question de la machine –contrairement à l'animal non humain –, il ne nous apparaît pas comme étant une question de niveau de conscience. Bien que cet enjeu ne soit pas l'objet de notre présente recherche, nous laissons quand même cette piste ouverte à la réflexion.

2. LE FONCTIONNALISME ET L'INGÉNIERIE NEUROMORPHIQUE

2.1 Le cerveau-ordinateur

Dans la section précédente, nous avons posé le cadre théorique dans lequel a été pensée, dans les termes du fonctionnalisme computationnel, notamment l'invention technique de l'avion. Il nous reste maintenant à présenter comment cette même logique s'est transposée aux développements en intelligence artificielle. L'ingénierie neuromorphique consacre sans aucun doute les efforts les plus aboutis quant à la quête de la réalisation d'une architecture artificielle neuro-inspirée. Bâtissant sur les thèses de l'isomorphisme fonctionnel et de la multiréalisabilité des propriétés mentales, il ne reste plus à l'ingénierie neuromorphique qu'à trouver la fonction qui pourrait sans équivoque être appliquée autant au cerveau qu'à une architecture artificiellement conçue. Cette quête s'est amorcée dès le départ par la schématisation des cellules de base du système nerveux : les neurones. On en connaît suffisamment sur les neurones pour expliquer non seulement leur composition, mais aussi leur fonctionnement précis quant au rôle qu'ils semblent jouer dans l'organisation du cerveau. Les neurones – comme le voudrait un cybernéticien – transmettent des influx (information)

7. Par «théorie de l'état cérébral», Putnam entend ici la théorie de l'identité des types de Smart (1959).

sous forme à la fois électrique et chimique. Puisque nous avons déjà déterminé le rôle des neurones, l'on pourrait s'attendre, en vertu de la multiréalisabilité des fonctions, à ce que ce même rôle de messenger puisse être tenu analogiquement par une structure physique distincte. Pour reprendre la définition formelle du fonctionnalisme, penser (ou ressentir (un) quelconque état mental) à l'aide de nos neurones s'exprimerait de sorte que :

Pour une chose x (comme un cerveau), avoir la propriété fonctionnelle F de penser (à comprendre la définition cybernétique de « traitement d'information » = def pour x , avoir une propriété physique P (comme un neurone) telle que $C(P)$, où $C(P)$ est une spécification de la tâche causale (c'est-à-dire le rôle de transmission d'informations) que P est censée remplir dans x .

Sans plus de doute, le traitement de l'information – selon la définition précédente – devrait être réalisable dans une machine reproduisant les mêmes principes de traitement d'information. Comme le conseillait Turing (1999 p. 144)⁸, il demeurerait simplement à trouver les fonctions mathématiques sous-jacentes au fonctionnement des neurones. Les lois et les principes de l'électricité (ex. : la loi d'Ohm) devraient alors suffire à cette tâche, sachant que le traitement de signaux effectué par les neurones implique une activité électrique, mais quitte à délaissier la complexité des interactions chimiques dans le cerveau. D'ailleurs, nos systèmes électroniques fonctionnent déjà suivant ces mêmes principes de manipulations de signaux électriques. L'ordinateur, en tant que système électronique des plus évolués, s'insère du même coup en toute convenance dans la définition fonctionnaliste de la pensée préalablement consacrée :

Pour une chose x (comme un ordinateur), avoir la propriété de penser (F) (qui se définit comme « traitement d'information ») = def pour (l'ordinateur), avoir une propriété P (transistors, condensateurs et résistances) telle que $C(P)$, où $C(P)$ est une spécification de la tâche causale (traiter l'information) que P (les transistors, les condensateurs et les résistances) est censée remplir dans x (l'ordinateur).

8. Voir citation plus haut dans Turing, Alan M. (1950). «Computing Machinery And Intelligence», dans *Mind*, 49, p. 433-460. Traduction française dans : Girard, Jean-Yves (1999). *La Machine de Turing*. Éditions du Seuil, Paris.

Ainsi, l'ordinateur se présenterait donc comme le candidat par excellence afin de remplir les fonctions d'être artificiel pensant. Il ne suffirait plus qu'à brancher et programmer *de la bonne manière* ce dispositif électronique pour espérer simuler la pensée. Beaucoup plus simple à dire, bien entendu, qu'à faire... N'empêche que certains modèles poseront les bases en vue de schématiser les plans de cette future architecture pensante, le modèle Hodgkin et Huxley figurant comme l'un de ces modèles de référence.

2.2 Les modèles du cerveau-ordinateur.

Alan L. Hodgkin et Andrew F. Huxley (1952) ont proposé dans les années 1950 leur schématisation électronique des fibres neuronales (figure 1), modèle qui deviendra non seulement une référence dans le domaine, mais qui orientera, depuis les travaux fondateurs de Carver Mead (1989), les recherches en ingénierie neuromorphique. D'autres modèles se sont également imposés au fil du temps, à commencer par le modèle du perceptron par Rosenblatt (1958), le modèle *Integrate and Fire* (IF) et *Leaky Integrate and Fire* (LIF) (Renaud *et al.*, 2010), le modèle de FitzHugh–Nagumo (FitzHugh, 1961 ; Nagumo *et al.*, 1962), le modèle Hindmarsh–Rose (1984), le modèle Galves–Löcherbach (2013), ou le modèle d'Izhikevich (2003) et encore d'autres variantes en constant développement. Dans l'ensemble, tous les différents modèles reprennent les prémisses du modèle Hodgkin et Huxley, soit de chercher dans la fonction mathématique un moyen de reproduire artificiellement un système neuronal. Or, chacun de ces modèles ajoute sa pierre à l'édifice de l'ingénierie neuromorphique, soit en venant répondre aux lacunes d'un autre modèle, soit en s'imposant tout simplement comme étant le plus propice, pour des raisons techniques, à certains usages particuliers. En effet, contrairement aux méthodes de simulation par ordinateur, comme le rappelle Renaud et ses collègues, « les réseaux neuronaux basés sur un support matériel sont consacrés à un modèle donné » (trad. libre, Renaud *et al.*, 2010, p. 905) et ne peuvent ainsi plus être aisément remodelés à la guise d'un programmeur une fois sa configuration concrète réalisée. Ainsi, certains modèles seront tantôt privilégiés pour reproduire telles fonctions neuronales précises alors que, selon le contexte, telles autres fonctions neuronales se prêteront mieux à la schématisation par un

autre type de modèle⁹. La question du choix du modèle devient alors plutôt technique et stratégique, sans pour autant déroger à la logique sous-jacente à leur utilisation première : la possibilité de schématiser, pour ensuite reconstruire artificiellement, les fonctionnalités cognitives du cerveau humain.

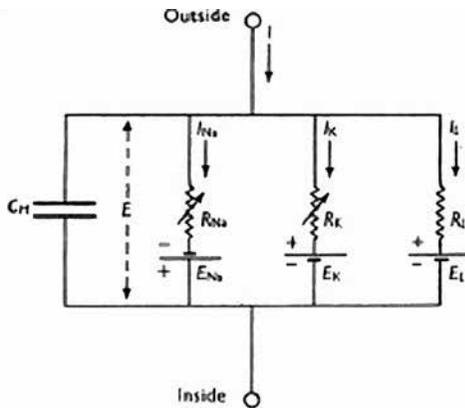


FIGURE 1 Un schéma de Hodgkin et Huxley tel qu'on le retrouve dans l'article de 1952. Chaque composante d'un neurone a été remplacée par une pièce électronique équivalente, d'un condensateur (C) à des résistances (R) et des sources de tensions (E). Observons également l'introduction d'une entrée et d'une sortie. Le neurone artificiel devrait donc être apte à recevoir un influx électrique (entrée), à transformer cette donnée pour finalement traduire cette nouvelle information en sortie, comme le ferait le cerveau fonctionnaliste.

Comme présenté ci-dessus, le schéma de Hodgkin et Huxley nous permet d'appliquer les principes de base de l'électronique afin de contrôler la tension aux bornes de la résistance (U), la valeur de la résistance (R) et l'intensité du courant (I) distribué sur l'ensemble du circuit neuro-inspiré ($U = RI$). Dans sa forme ainsi simplifiée, le processus n'est pas plus sorcier que ce qui se trouve déjà dans nos appareils électroniques (ordinateur compris). Or, faut-il encore en premier lieu prendre en compte les variantes externes au circuit (par

9. Comme le répertorie Antoine Joubert (2013), les différents modèles neuromorphiques sont d'ailleurs simultanément utilisés dans les laboratoires de recherches mondiaux. Pour une liste de ces projets, et une courte explication de leur application, voir Joubert, Antoine (2013). *Neurone analogique robuste et technologies émergentes pour les architectures neuromorphiques*. Thèse de doctorat soutenue à l'Université Grenoble.

exemple, le fait que la valeur d'une résistance change en fonction de la température) et l'inévitable dégradation des matériaux. Il s'agit ici d'un problème d'acclimatation au milieu externe, élément que nous laissons de côté dans le cadre de cette présente analyse, mais qui n'est pas moins crucial au mode d'existence des objets techniques. En effet, les objets techniques évoluent dans des milieux dans lesquels la fixité – propre à l'objet technique automatisé – devient un handicap majeur à leur pérennité. La grande capacité d'adaptation des êtres biologiques fait d'eux des êtres déjà plus parfaits (ou concrets, pour utiliser le vocabulaire qui sera tantôt celui de Simondon).

En deuxième lieu, le modèle semble ici assez simple dans la mesure où il ne représente qu'un seul neurone individuel et isolé. Le cerveau en contient des milliards en constant rapport les uns avec les autres et agencés d'une manière dont l'on ne comprend encore que la surface. L'idée de reproduire une intelligence artificielle consciente en accumulant ces neurones individuels doit aussi tenir compte de l'agencement et de la mise en rapport des neurones. Nous ajoutons par le fait même une nouvelle condition à la réalisation de la conscience artificielle : la structuration de *la bonne manière* des composantes de base (ici, des milliards de neurones) nécessaire à la réalisation de la tâche causale à accomplir. Cette idée va d'ailleurs de pair avec la philosophie physicaliste de laquelle est issue la théorie fonctionnaliste computationnelle. Après tout, comme l'écrit Hawkes, « le physicalisme, comme l'est le fonctionnalisme, est l'affirmation que l'esprit est entièrement composé de parties physiques non conscientes et non rationnelles » (Hawkes, 2017). Or, ce même constat demeure par ailleurs un défi explicatif de taille pour les théories physicalistes. Comment un monde de particules physiques, sans conscience et sans sens, peut-il donner naissance à cette chose que l'on nomme la conscience ? En somme, il semblerait difficile de concevoir que la conscience ne soit que la résultante d'une simple accumulation au hasard de molécules – tout aussi ordonnées soient-elles finalement –, qui, au-delà d'un certain seuil, émergeraient tout d'un coup en une structure pensante. Nous pourrions tout autant regarder d'un œil suspect toute idée de recréer une intelligence artificielle consciente de par la simple accumulation de pouvoir de calcul. En fait, si la conscience n'était que puissance de calcul, force serait alors de constater que les superordinateurs des grandes universités mondiales sont déjà *plus* conscients que les humains.

3. QU'EST-CE QUE L'OBJET TECHNIQUE ?

3.1 Multiréalisabilité et causalité instrumentale

Les conséquences de la multiréalisabilité sont nombreuses, mais nous nous concentrerons pour notre propos sur l'une de ces conséquences précises : celle de la réduction du principe de causalité à une causalité instrumentale. Afin de rendre cette réduction évidente, reconsidérons en premier lieu l'argument de la réalisabilité de la même manière que le font Thomas W. Polger et Lawrence A. Shapiro dans leur ouvrage *The Multiple Realization Book* (2016). Ces derniers illustrent la multiréalisabilité en prenant l'exemple d'une série d'ouvre-bouteilles (Polger et Shapiro, 2016, p. 24). Comme le définit – d'ailleurs, de manière fonctionnelle – le dictionnaire Larousse, un ouvre-bouteille se résume tout simplement à « un outil pour ouvrir des bouteilles ». Imaginons que nous avons devant nous une série de quatre ouvre-bouteilles dont la seule différence soit leur couleur respective. Nous aurions ainsi un ouvre-bouteille jaune, un bleu, un rouge, puis un vert. Nous conviendrions déjà intuitivement que cette différence physique n'a pas d'impact sur la réalisation de la tâche causale à remplir. La couleur demeure une propriété accidentelle qui ne joue aucun rôle dans la réalisation de la tâche désirée, soit d'ouvrir une bouteille. Nous ne serions alors pas tentés de dire que nous sommes en présence d'une réalisabilité multiple. On peut dès lors déjà conclure que certaines propriétés d'un objet (ici, la couleur des ouvre-bouteilles) ne sont pas pertinentes dans la mesure où la couleur ne fait aucune différence à la réalisation de la fonction de l'objet. L'accident (en termes aristotéliens) doit alors être compris comme ce qui ne fait pas de différence dans la réalisation causale d'une tâche.

Imaginons encore une série d'ouvre-bouteilles, mais cette fois composée d'ouvre-bouteilles ayant tous certaines caractéristiques physiques différentes. Nous pourrions alors imaginer que certains sont de forme cylindrique alors que d'autres sont plutôt cubiques ; que certains sont faits en bois alors que d'autres sont en aluminium. En dépit de toutes ces différences physiques, tous ces ouvre-bouteilles peuvent sans aucun doute remplir la tâche pour laquelle ils ont été conçus, c'est-à-dire, encore une fois, « ouvrir des bouteilles ». Dès lors, on remarque que ni la forme ni la matière n'ont un grand rôle à jouer dans la réalisation d'une tâche causale. Certes, les choix de la matière et de la forme d'un objet influenceront l'efficacité d'exécution (pensez simplement à

comparer une machine de Turing à un ordinateur contemporain!). Cependant, comme la définition aristotélicienne de l'accident le laisse entendre, l'accident est «ce qui peut appartenir à une seule et même chose, et aussi ne pas lui appartenir» (Aristote, *Topiques*, 102 b 6-7). Il conviendrait alors de conclure que la forme et la matière sont des propriétés accidentelles de l'objet fonctionnalisé. Bref, toutes causes formelles et matérielles – toujours pour parler en termes aristotéliciens – n'auraient plus de raison d'être que si elles sont relayées à une cause instrumentale. En d'autres mots, ne seraient pertinentes la forme et la matière que si elles pouvaient, de manière instrumentale, remplir la description de l'emploi attribué à l'objet-outil. En résumé, la réalisabilité multiple, en réduisant l'essence d'une chose à sa description de l'emploi, c'est-à-dire le rôle joué comme instrument, réduit du même coup son spectre causal à la seule causalité instrumentale.

3.2 De l'abstraction à la réalisation concrète de l'objet technique

Certes, montrer que le fonctionnalisme ne s'accorde pas avec le cadre d'analyse aristotélicien n'invalide pas pour autant son propos. Après tout, de Hume qui rejetait la notion de substance (Hume, 1962) jusqu'à Bachelard qui invitait tout simplement à dépasser cette même notion aristotélicienne (Bachelard, 1968), bien des philosophes modernes se sont efforcés de se distancer de la pensée du Stagirite. De plus, on pourrait aisément rétorquer que nos présents objets d'étude sont, après tout, des objets-outils et que, par conséquent, ils se prêtent naturellement à être dirigés vers la causalité instrumentale. Même l'organe biologique, si l'on se réfère à l'étymologie grecque du mot, soit *organa* comme «outil», peut porter vers cette interprétation. Finalement, force est de constater que, même chez Aristote, l'instrumentalité fait partie intégrante de l'essence d'une chose. Mais encore... elle n'en fait que «partie».

Comme le montrait bien Martin Heidegger dans son célèbre essai *La Question de la Technique* (1958), l'on fait souvent vite de réduire la question ontologique de l'outil à sa dimension instrumentale, en oubliant que la causalité signifie bien plus. «De même, l'essence de la technique n'est absolument rien de technique.» (Heidegger, 1958, p. 9). Il précède de cette explication : «Quand nous recherchons l'essence de l'arbre, nous devons comprendre que ce qui régit tout arbre en tant qu'arbre n'est pas lui-même un arbre qu'on puisse rencontrer parmi les autres arbres.» (Heidegger, 1958, p. 9). Notre problème s'apparente à

cette méditation heideggérienne. Réduire la définition d'un objet outil à sa cause instrumentale nous apparaît au mieux circulaire, ou pis n'offre qu'une définition fractionnaire qui écarte la notion essentielle de l'objet technique en soi. Qu'est-ce qui demeure universellement de l'objet technique et qui puisse (comme l'arbre d'Heidegger) régir tout objet technique en tant qu'objet technique ?

À l'instar de Simondon dans *Du mode d'existence des objets techniques* (2012), nous pourrions reposer la question en repartant du point de départ qui était celui du fonctionnalisme, soit de penser les traits communs des objets techniques *en fonction* de l'hétérogénéité de leur mode de fonctionnement. Or, Simondon y décelait rapidement le problème qu'il présentait ainsi :

un moteur à vapeur, un moteur à essence, une turbine, un moteur à ressort ou à poids sont tous également des moteurs ; pourtant, il y a plus d'analogies réelles entre un moteur à ressort et un arc ou une arbalète qu'entre ce même moteur et un moteur à vapeur ; une horloge à poids possède un moteur analogue à un treuil, alors qu'une horloge à entretien électrique est analogue à une sonnette ou à un vibreur. L'usage réunit des structures et des fonctionnements hétérogènes sous des genres et des espèces qui tirent leur signification du rapport entre ce fonctionnement et un autre fonctionnement, celui de l'être humain dans l'action. Donc, ce à quoi on donne un nom unique, comme, par exemple, celui de moteur, peut être multiple dans l'instant et peut varier dans le temps en changeant d'individualité. (Simondon, 2012, p. 21-22)

Ce passage pose les limites à la définition des objets techniques centrés sur un argument de multiréalisabilité comme l'entendent les théoriciens fonctionnalistes. Pour cause, les usages pratiques par lesquels le fonctionnalisme définit ses objets d'étude ne présentent finalement qu'« une spécificité illusoire, pour reprendre les mots de Simondon, car aucune structure fixe ne correspond à un usage défini » (Simondon, 2012, p. 21). Il s'agit ici de faire ressortir la difficulté de définir l'objet technique en rapport à son individualité ou sa différence au sein d'un système de catégorisation basé sur l'hétérogénéité du fonctionnement technique des objets. Pour Simondon, l'objet technique est un objet qui porte en lui sa propre genèse, c'est-à-dire qu'il se situe dans un *monde historique* où se côtoie une multitude de facteurs déterminants de son devenir. Or, cette historicité n'est pas pour autant une genèse linéaire. « L'être technique, poursuit alors Simondon, évolue par convergence

et par adaptation à soi ; il s'unifie intérieurement selon un principe de résonance interne.» (Simondon, 2012, p. 23) Pour mieux comprendre l'évolution et l'adaptation technique au sein de sa genèse, considérons ce passage :

Le moteur d'automobile d'aujourd'hui n'est pas le descendant du moteur de 1910 seulement parce que le moteur de 1910 était celui que construisaient nos ancêtres. Il n'est pas non plus son descendant parce qu'il est plus perfectionné relativement à l'usage ; en fait, pour tel ou tel usage, un moteur de 1910 reste supérieur à un moteur de 1956. Par exemple, il peut supporter un échauffement important sans grippage ou coulage, étant construit avec des jeux plus importants et sans alliages fragiles comme le régule ; il est plus autonome, possédant un allumage par magnéto. D'anciens moteurs fonctionnent sans défaillance sur des bateaux de pêche, après avoir été pris sur une automobile hors d'usage. C'est par un examen intérieur des régimes de causalité et des formes en tant qu'elles sont adaptées à ces régimes de causalité que le moteur d'automobile actuel est défini comme postérieur au moteur de 1910. Dans un moteur actuel, chaque pièce importante est tellement rattachée aux autres par des échanges réciproques d'énergie qu'elle ne peut pas être autre qu'elle n'est. (Simondon, 2012, p. 23)

Les nécessités du milieu interne à l'objet technique agissent sans aucun doute sur le devenir de l'objet technique. L'objet technique devient tel ou tel en raison de contraintes internes qui forcent l'ingénieur à penser la solution, non pas en vertu de l'idée abstraite du départ, mais en vertu des contraintes concrètes et réelles du système en évolution. L'objet devient alors de plus en plus concret dans la mesure où il vient à intégrer une complexité de causalités internes réciproques. L'objet abstrait, quant à lui, demeure partes extra partes où les relations causales de réciprocité y sont encore faibles. Sur la même lancée, Simondon ajoute que « dans le moteur ancien, chaque élément intervient à un certain moment dans le cycle, puis est censé ne plus agir sur les autres éléments ; les pièces du moteur sont comme des personnes qui travailleraient chacun à leur tour, mais ne se connaîtraient pas les unes les autres » (Simondon, 2012, p. 24). La concrétisation dont parle Simondon n'est finalement que la convergence des fonctions dans l'unité structurale, c'est-à-dire « la réduction précise de cette marge entre les fonctions des structures plurivalentes ». La convergence des fonctions en un système de causalité réciproque concrétise

l'objet technique puisqu'elle le spécifie: l'objet technique ne peut plus être autrement que ce qu'il est, car il est parfaitement uni dans ses fonctions. Quel serait donc alors le problème ontologique du fonctionnalisme – en considérant ici la cybernétique d'où découle la définition fonctionnaliste du cerveau-ordinateur? Simondon donne encore une réponse à ce sujet:

Ce qui risque de rendre le travail de la Cybernétique partiellement inefficace comme étude interscientifique [...], c'est le postulat initial de l'identité des êtres vivants et des objets techniques auto-régulés. Or, on peut dire seulement que les objets techniques tendent vers la concrétisation, tandis que les objets naturels tels que les êtres vivants sont concrets dès le début. Il ne faut pas confondre la tendance à la concrétisation avec le statut d'existence entièrement concrète. Tout objet technique possède en quelque mesure des aspects d'abstraction résiduelle; on ne doit pas opérer le passage à la limite et parler des objets techniques comme s'ils étaient des objets naturels. Les objets techniques doivent être étudiés dans leur évolution pour qu'on puisse en dégager le processus de concrétisation en tant que tendance; mais il ne faut pas isoler le dernier produit de l'évolution technique pour le déclarer entièrement concret; il est plus concret que les précédents, mais il est encore artificiel [...] C'est selon cette voie seulement que le rapprochement entre être vivant et objet technique a une signification véritable, hors de toute mythologie. (Simondon, 2012, p. 59-60)

Cette réflexion mène à un dernier point, que nous ne pourrions qu'effleurer faute de temps et d'espace. Il s'agit de la codétermination réciproque des milieux humains (et culturels) sur la genèse de l'objet technique. L'objet technique a été jusqu'ici présenté selon ses déterminants externes (environnement) et internes (contraintes physiques internes au système). Pourtant, l'objet technique demeure aussi une conception humaine, qui concrétise alors un peu de soi dans la machine. Selon la formule centrale de l'ouvrage de Simondon, la machine a une part « d'humain méconnu ». Il nous faudrait peut-être s'empresser de chercher à connaître cet inconnu, en tant que ce qu'il est réellement. Pendant ce temps, consciente ou non, l'intelligence artificielle devient de plus en plus médiatrice (avec ou sans notre consentement) de notre quotidien, et par le fait même partie intégrante de nos milieux humains.

CONCLUSION

Notre questionnement initial visait de manière générale l'engouement entourant l'émergence de l'intelligence artificielle. Cette question se résumait à savoir pourquoi plusieurs acceptent si facilement de faire le rapprochement entre le cerveau et l'ordinateur (au sens même d'une relation d'identité cerveau-ordinateur), alors que nul n'oserait dire que l'avion est un oiseau artificiel ? Pourquoi refusons-nous de dire que l'oiseau n'est pas un avion, sans pourtant hésiter à rapprocher le cerveau à l'ordinateur ? C'est finalement toute une philosophie fonctionnaliste de l'intelligence artificielle qui aura été mise en lumière tout au long de ce chapitre. Bien que cette philosophie puisse servir de pierre d'assise aux développements d'intelligences artificielles toujours plus efficaces, ses limites théoriques – et sûrement aussi pratiques – peuvent continuer à donner de l'eau au moulin aux détracteurs de l'idée d'une machine-pensante.

En revanche, devrions-nous arrêter de conceptualiser la machine-pensante qu'en raison des difficultés théoriques qu'elle pose ? Les historiens du vol nous le rappellent à propos de l'« invention du vol » : si l'être humain restera si longtemps enchaîné au sol, c'est bien parce qu'il se refusa de penser la possibilité même d'une machine volante – plus communément appelée aujourd'hui un « avion » – en dehors du mythe et du fantôme (Niccoli, 2003 ; Marck, 2012 ; Polacco, 2013). Encore à la fin du XIX^e siècle, alors que les frères Wright entamaient la conception des premiers plans de l'avion, non seulement le public, mais aussi une grande partie de la communauté scientifique s'entêtaient toujours à croire que la machine volante était « aussi impossible que le mouvement perpétuel » (Wright, 1953, p. 20).

En même temps, il semble que cette méprise historique permet plus souvent de justifier en toute impunité l'engouement pour la recherche d'une intelligence artificielle consciente, une mission nous apparaissant – comme aux scientifiques de l'époque des frères Wright – aussi impossible que fabuleuse. Dans cette narration, les « fous » d'hier seront les pionniers des histoires de demain. Si l'on suivait à la lettre la rhétorique ci-dessus, nous devrions alors concéder que toute avancée technologique ne serait qu'une question de temps, un temps qu'on ne cesse d'annoncer comme toujours plus rapproché. Face à cette position optimiste toujours en suspens, il resterait plus qu'à attendre et regarder le progrès se dérouler naturellement devant nos yeux.

Pourtant, au-delà du fantasme, la concevabilité¹⁰ théorique sur lequel repose ces espoirs de reproduire artificiellement un cerveau n'assure en rien sa réalisabilité pratique. Il y a en effet une différence entre la possibilité qu'une chose puisse en théorie exister, et qu'une chose puisse en pratique être réalisée. L'argument de la réalisabilité multiple prenait justement comme point de départ la première position, à savoir qu'il est possible (en théorie) de trouver dans l'univers un corrélat mental différent, mais qui pourtant réalise nos mêmes états mentaux humains. Cependant, la possibilité théorique qu'une chose *puisse* être devrait-elle être suffisante pour dire que la réalisation de cette chose soit vraie ou qu'elle puisse se réaliser? Nous aurions tendance à répondre par la négative à cette question. Autrement, comme le voudrait le fameux paradoxe probabiliste du singe savant, nous serions encore à attendre qu'un chimpanzé réussisse par hasard à écrire Hamlet de Shakespeare. Nous pourrions tout autant postuler qu'il est tout à fait concevable, en théorie, de construire un cerveau complet en briques Lego, même si cela demandait un nombre incommensurable de briques, et peut-être même nécessitait la taille complète de l'univers pour être réalisé¹¹. Pourtant, il est peu probable que ce cerveau-lego se réalise, soit-il tout simplement impossible à créer dans la pratique.

En conclusion, le cerveau-ordinateur apparaît à bien des égards comme ce cerveau-lego : une possibilité peut être théorique sans pour autant être réalisable dans la pratique. Y a-t-il donc des enjeux de l'intelligence artificielle plus pressants, et surtout réels, auxquels nous devons faire face? Poser la question, c'est bien sûr y répondre. Pendant que nous attendons d'un côté l'avènement de la machine-consciente, de l'autre côté notre quotidien se voit pris en charge toujours un peu plus jour après jour par des algorithmes –aussi efficaces soient-ils – qui n'ont finalement que très peu à voir avec l'intelligence humaine. Comme l'avion qui ne *sait* pas qu'il vole, peut-être devrions-nous être plutôt préoccupé du fait que des machines remplissant des tâches analogues aux milliers de décision qu'un humain prend à chaque jour, soient appelées à agir et à décider pour nous, sans *savoir* ce qu'elles font...

10. Nous entendons ici l'argument de la concevabilité de la même manière que l'utilisait déjà Descartes dans les Méditations pour démontrer la possibilité d'une existence distincte de son corps.

11. L'exemple du cerveau-lego est repris de l'article de Hawkes (2017) qui faisait remarquer que le cerveau-lego est une conséquence logique du physicalisme qui veut que l'esprit soit la résultante de la composition de parties physiques non conscientes et non rationnelles correctement agencées.

BIBLIOGRAPHIE

- Aristote (1956). *Parties des animaux*. Traduction française Pierre Louis. Paris, Belles Lettres.
- Aristote (1967). *Topiques*. Traduction française par Jacques Brunschwig. Paris, Belles Lettres, 2 v.
- Bachelard, Gaston (1947). *La formation de l'esprit scientifique: contribution à une psychanalyse de la connaissance objective*. Paris, Vrin, 256 p.
- Bachelard, Gaston (1968). *Le nouvel esprit scientifique*. Paris, Presses universitaires de France, 179 p.
- Block, Ned (1992). «Le fonctionnalisme face au problème des qualia», dans *Les Études philosophiques, n° 3, La théorie computationnelle de l'esprit*. Paris, Presses universitaires de France, p. 337-369.
- Churchland, Patricia (1986). *Neurophilosophy: toward a unified science of the mind-brain*. Cambridge, MIT Press, 546 p.
- Churchland, Paul M. (1981). «Eliminative Materialism and the Propositional Attitudes», *Journal of Philosophy*, vol. 78, n° 2, p. 67-90.
- Descartes, René (1974). *Discours de la méthode: suivi des Méditations*. Verviers, Marabout Université, 182 p.
- Duhem, Jules (1959). *Histoire des origines du vol à réaction*. Paris, Nouvelles éditions latines.
- Fitzhugh, Richard (1961). «Impulses and Physiological States in Theoretical Models of Nerve Membrane», *Biophysical journal*, vol. 1, p. 445-466.
- Fodor, Jerry (1974). «Special Sciences, or the Disunity of Science as a Working Hypothesis», *Synthese*, vol. 28, p. 97-115.
- Galves, Antonio et Löcherbach, Eva (2013). «Infinite Systems of Interacting Chains with Memory of Variable Length – A Stochastic Model for Biological Neural Nets», *Journal of Statistical Physics*, vol. 151, n° 5, p. 896-921.
- Hawkes, Gordon (2017). «How Many Lego Bricks to Build a Mind?», *A Philosopher's Take*. En ligne : <https://aphilosopherstake.com/> (page consultée le 12 avril 2017).
- Heidegger, (1958). *Essais et conférences*. Paris, Gallimard, 349 p.
- Hindmarsh J. L., et Rose R. M. (1984). «A model of neuronal bursting using three coupled first order differential equations», *Proc. R. Soc. London, Ser. B.*, vol. 221, p. 87-102.
- Hodgkin, Alan L. et Huxley, Andrew F. (1952). «A quantitative description of membrane current and its application to conduction and excitation in nerve», *The Journal of physiology*, vol. 117, p. 500-544.
- Hume, David (1946). *Traité de la nature humaine: essai pour introduire la méthode expérimentale dans les sujets moraux*. Paris, Aubier, 2 v., 766 p.
- Izhikevich, Eugene M. (2003). «Simple model of spiking neurons», *IEEE Transactions on neural networks*, vol. 14, p. 1569-1572.
- Joubert, Antoine (2013). *Neurone analogique robuste et technologies émergentes pour les architectures neuromorphiques*. Thèse de doctorat soutenue à l'Université de Grenoble.
- Kaplan, Jerry (2016). *Artificial Intelligence: What Everyone Needs to Know*. Oxford University Press, 165 p.

- Kim, Jaegwon (1992). « Multiple Realization and the Metaphysics of Reduction », *Philosophy and Phenomenological Research*, vol. 52, n° 1, p. 1-26.
- Kim, Jaegwon (2008). *Philosophie de l'esprit*. Traduction par David Michel-Pajus, Mathieu Mulcey et Charles Thérét. Paris, Éditions Ithaque.
- Marck, Bernard (2012). *Histoire de l'aviation*. Paris, Flammarion.
- Marr, David (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco, W.H. Freeman.
- Mead, Carver (1989). *Analog VLSI and neural systems*. VLSI systems series, Addison-Wesley.
- Nagumo J., Arimoto S., and Yoshizawa S. (1962). « An active pulse transmission line simulating nerve axon », *Proceedings of the IRE*, vol. 50, n° 10, p. 2061-2070.
- Niccoli, Riccardo (2003). *Le rêve d'Icare: histoire de l'aviation*. Paris, Gründ, 320 p.
- Parrochia, Daniel (2003). *L'homme volant. Philosophie de l'aéronautique et des techniques de navigation*. Seyssel, Éditions Champ Vallon, 317 p.
- Polacco, Michel (2013). *Une brève histoire de l'aviation*. Paris, JC Béhar, 157 p.
- Polger, Thomas W. et Shapiro, Lawrence A. (2016). *The Multiple Realization Book*. Oxford, Oxford University Press, 258 p.
- Putnam, Hilary (2002). « La nature des états mentaux », traduction par Jean-Michel Roy, dans Fiset, D. et Poirier, P. (dir.), *Philosophie de l'esprit, psychologie du sens commun et sciences de l'esprit*. Paris, Vrin, p. 269-287.
- Putnam, Hilary (1992). « Philosophy and Our Mental Life », dans *The Philosophy of Mind: Classical Problems/ Contemporary Issues*. Cambridge, MIT Press.
- Renaud, Sylvie., Tomas, Jean et al. (2010). « PAX: A mixed hardware/software simulation platform for spiking neural networks », *Neural networks: the official journal of the International Neural Network Society*, vol. 23, p. 905-916.
- Rosenblatt, Frank (1958). « The perceptron: A probabilistic model for information storage and organization in the brain », *Psychological Review*, vol. 65, p. 386-408.
- Schwartz, Eric (1990). *Computational neuroscience*. Cambridge, MIT Press.
- Simondon, Gilbert (2012). *Du mode d'existence des objets techniques*. Paris, Flammarion.
- Smart, John J. C. (1959). « Sensations and Brain Processes », *Philosophical Review*, vol. 68, p. 141-156.
- Turing, Alan M. (1950). « Computing Machinery And Intelligence », dans Girard, Jean-Yves, *La Machine de Turing*. Paris, Éditions du Seuil.
- White, Lynn (1978). « Eilmer of Malmesbury, an eleventh-century aviator: a case study of technological innovation, its context and tradition », dans Lynn White (dir.), *Medieval religion and technology*, p. 59-73.
- Wiener, Robert (1961). *Cybernetics or Control and Communication in the Animal and the Machine*. Cambridge, MIT Press, 212 p.
- Wright, Orville (1953). *How we invented the airplane*. New York, McKay, 78 p.

CHAPITRE 7

Robotique développementale et intelligence artificielle « incorporée » : questions éthiques du point de vue du robot

Alexandre PITTI

INTRODUCTION

Le développement des robots humanoïdes, dans les sociétés, pose plusieurs problèmes, en ce qui concerne leur acceptation et leur utilité. Bien qu'ils soient conçus de façon réaliste, leurs interactions physiques ou sociales avec les êtres humains demeurent faibles, et ils ne représentent que des machines rigides, contrôlables et prévisibles au sein d'un environnement restreint. Dans les environnements non restreints, ils peuvent se transformer en automates fascinants et imprévisibles, et leur manipulation est risquée à cause de leurs capacités limitées, sur les plans de la coordination et de l'adaptation.

L'utilité des robots humanoïdes est encore incertaine à court terme, mais les progrès de l'intelligence artificielle, qui développe progressivement un comportement plus autonome et plus adaptable, font qu'elle semble inéluctable à long terme. Le problème de l'éthique des robots est donc légitime, car ceux-ci sont en contact avec des êtres

humains. Comme pour toute nouvelle technologie, l'introduction des robots ne laisse personne indifférent et suscite beaucoup de réactions (et fantaisies), parfois trop pessimistes et parfois trop optimistes, en ce qui concerne leurs habiletés réelles et les risques potentiels qu'ils représentent.

Le genre d'interactions possibles varie, selon que le comportement du robot est imprévisible, désagréable et frustrant, ou, à l'autre bout du spectre, non répétitif, agréable et s'ajustant aux circonstances. Mais comment savoir si nos interactions avec ces jouets de conception avancée sont une source d'émancipation ou plutôt une source d'aliénation? On ne peut pas prévoir les résultats, car cela dépend de notre humeur, de l'utilisation du robot et de son objectif, ainsi que de sa conception. Le développement des technologies numériques (ordinateurs, cellulaires et Internet) peut servir de représentation pour ce qui est de l'évolution future du monde des robots, ainsi que de ses normes et règlements. Internet a déjà changé la façon dont on communique et travaille.

Internet a aussi changé la façon dont nous organisons notre vie sociale, ainsi que les limites de notre sphère privée. Comme cela s'est déjà produit avec les technologies précédentes, il est tout à fait probable qu'une réglementation soit définie par la société au sujet des robots dans les maisons. Quand les robots franchiront le seuil de nos foyers, par exemple, ils pourront récolter plus d'informations sur nos habitudes qu'aucune autre technologie, et pourtant, le problème de la confidentialité n'a pas encore été abordé. Par ailleurs, les robots peuvent maintenant presque agir de façon autonome, dans leur environnement physique et numérique.

Puisque leur corps aura une ressemblance humaine et leurs actions sembleront intentionnelles, ils pourraient causer des problèmes ou des blessures chez ceux qui n'y sont pas encore habitués, tels que les enfants et les personnes âgées. De plus, comme pour les jeux vidéo, la littérature de science-fiction et le cinéma ont toujours suggéré que nous nous confondrons avec les robots et que nous ne pourrions plus les distinguer parmi nous dans le futur. Il n'y a là rien de nouveau et cela ne se limite pas aux robots.

René Girard a développé une théorie de la mimésis, ou du désir mimétique, selon laquelle les êtres humains projettent leurs émotions, désirs et intentions sur les autres (Girard, 1976-1977). Girard avait compris que notre autonomie est «illusoire» et que nos désirs sont ceux des

autres. Puisque les êtres humains sont toujours renfermés dans la mimésis, ils se sentent en même temps attirés et rejetés par ce qui leur ressemble et les imite. Dans cet état de mimétisme réciproque, il est difficile de comprendre qui est l'original et qui l'imitateur, car la distance entre eux disparaît. Les robots et l'intelligence artificielle sont nos miroirs modernes, qui nous renvoient nos fantaisies, selon la théorie de la mimésis.

Nous aimerions toutefois percevoir une tendance plus positive chez les robots. L'intelligence artificielle (IA) est modelée sur notre intelligence ; ces algorithmes peuvent être utilisés pour mieux nous comprendre et mieux évaluer nos limites, tout en améliorant notre potentiel physique et cognitif. On pourrait aussi développer nos propres habiletés d'apprentissage en étant accompagnés au sein de notre vie quotidienne. Les robots sont des machines qui nous offrent un reflet et qui nous demandent comment nous définissons l'intelligence ou comment nous voudrions que les autres se comportent envers nous ; il serait peut-être bon d'interroger notre propre éthique. Puisque les robots peuvent apprendre et transformer leur comportement grâce à l'interaction, ils pourraient finir par nous ressembler encore plus, en imitant notre comportement, nos habitudes et règles sociales, nos préjugés et nos tabous.

Les robots pourraient améliorer notre apprentissage plus que toute autre machine, car nos mécanismes d'apprentissage sont implicites et puissants et qu'ils sont influencés par nos interactions sociales (Marshal et Meltzoff, 2015). Nous allons analyser ici la question de l'éthique des robots sur le plan de leur interaction avec les êtres humains, ainsi que la quantification et la capacité de codification ou d'apprentissage de comportements éthiques, à l'intérieur d'un système autonome. L'éthique robotique est une recherche interdisciplinaire qui se trouve au croisement de l'éthique appliquée et de la robotique. Nous voulons l'analyser à travers notre connaissance de l'apprentissage du cerveau humain et du développement infantile.

1. DE LA CORPORÉITÉ

Les ingénieurs conçoivent l'intelligence comme une conséquence directe de leur conception des comportements « intelligents » chez les robots. Elle influence les actions et interactions des robots et, par conséquent, l'éthique de leur comportement. Nous sommes d'avis que l'intelligence ne se base pas sur l'accumulation de données, comme

dans un entrepôt ou dans un ordinateur, mais plutôt sur la manière dont on perçoit et comment on choisit de définir ces données. Les données importantes ne peuvent pas être dissociées du support physique de cette information: selon la théorie de l'information de Shannon, l'information a une réalité physique. Puisque toute connaissance est acquise à travers la perception que nous en avons au sein de notre environnement, nous suggérons que le « corps » est un vecteur de soutien important pour le développement de l'information et de l'intelligence.

D'autre part, l'information est évidemment limitée, incomplète et subjective. Nous pensons donc que l'intelligence de la corporéité est un principe de conception prérequis pour tout agent qui souhaite réaliser un comportement intelligent, tel le comportement éthique (Pfeifer et Scheier, 1999; Pfeifer et Bongard, 2006; Pfeifer et Pitti, 2012).

Cette proposition peut paraître évidente ou simpliste, mais il n'y a qu'à travers le corps qu'un agent puisse ressentir sa propre existence, ainsi que celles des autres.

La perception active représente la capacité de pouvoir prédire et anticiper les conséquences sensorielles futures de nos propres actions et, parfois, de celles d'autrui (Clark, 2015).

C'est du moins ce que l'on reconnaît en ce qui concerne le développement cognitif des enfants. Des recherches effectuées par Andrew Meltzoff et Jacqueline Nadel ont montré que l'imitation est profondément enracinée dans nos gènes, puisque les enfants apprennent en imitant les autres, dès leur naissance (Nadel et Butterworth, 1999; Meltzoff et Decety, 2003). Andrew Meltzoff a poursuivi cette idée et développé un paradigme connu sous le nom de l'hypothèse « comme moi », illustrée par les bébés qui interprètent et s'approprient les actions des autres en termes de « comme eux », en les intégrant dans leur répertoire personnel d'actions incarnées (Meltzoff, 2007). Dès un très jeune âge, ils peuvent interpréter les intentions d'autrui mentalement, en les observant et en comparant leurs actions avec les leurs. Cette proposition constitue une hypothèse de travail intéressante à développer dans le cas des robots: on pourrait les doter de capacités cognitives et leur permettre de comprendre les intentions des êtres humains en analysant leurs corps à travers le principe « comme eux ». En acquérant plus d'expérience à propos d'eux-mêmes et d'autrui, ils peuvent acquérir la notion de ce que représentent le soi et les autres, pour appréhender la question de l'éthique.

Nous analyserons plus loin les liens entre l'intelligence incarnée et la façon dont l'information est acquise et représentée dans les parties du cerveau connectées à l'agence, aux représentations de soi par rapport à autrui, à l'empathie, à la théorie des cerveaux selon le système des neurones miroirs, pour déterminer comment cela peut influencer la façon dont les roboticiens conçoivent l'architecture cognitive chez les robots quant aux comportements éthiques.

Cette parenthèse sur la personnification est importante pour expliquer notre approche et souligner la différence entre ce qu'on appelle l'IA faible, qui perçoit les comportements intelligents produits dans un robot comme des simulations fabriquées par des ingénieurs, et l'IA forte, soit l'idée que certains aspects des comportements intelligents peuvent être appris et testés à travers les expériences de perceptions personnelles des robots.

Nous ne croyons pas que les ingénieurs doivent codifier des règles abstraites directement (p. ex. des règles éthiques) dans la mémoire du robot, comme pour un programme d'ordinateur et pour des algorithmes d'IA anciens (incluant les algorithmes « à propagation avant » comme les réseaux d'apprentissage profond), car toute situation est perçue à travers le corps et devrait être apprise à travers ses connexions sensorimotrices particulières.

Comme les dynamiques du cerveau et du corps sont tellement complexes, le cerveau personnifié est illimité, car il est toujours en interaction avec son environnement, tandis que l'ordinateur est un système fermé, qui travaille dans un monde symbolique abstrait, en dehors de la réalité physique (Lungarella et Sporns, 2005). Par conséquent, les deux stratégies pour interpréter l'information sont très différentes.

L'information traitée par le cerveau est distribuée, non synchronisée, analogique et robuste quant au bruit; les ordinateurs, au contraire, traitent l'information de façon centralisée, synchronisée, numérique, symbolique et sont particulièrement sensibles aux erreurs. Étant donné que le corps sert d'interface pour toute déduction perceptuelle, les robots peuvent ancrer leurs mémoires de perception, transcrites dans le « langage » parlé à travers leurs circuits sensorimoteurs. Nous sommes d'avis que toute formulation de comportement éthique chez les robots devrait être transcrite sur le plan sensorimoteur, à travers la personnification, et non sur le plan symbolique.

2. DÉVELOPPEMENT ET APPRENTISSAGE

Cette différence entre les systèmes ouverts et fermés est importante lorsque l'on conçoit des robots actifs, aussi longtemps que les humains font partie de l'équation. Comme dans le cas de la cybernétique (Wiener, 1948), Edgar Morin explique que ce qui caractérise les systèmes complexes est leur habileté à organiser leur propre comportement à travers leurs actions, qui transforment simultanément leur organisation interne (Morin et Le Moigne, 1999).

Cette réflexivité des procédés ouverts basés sur la rétroaction est ce qui rend tous les systèmes adaptifs et robustes par rapport aux changements. Si toute action est aléatoire et peut produire du bruit chez le robot, elle produit aussi de nouveaux phénomènes propres aux systèmes dynamiques, tels l'émergence et l'organisation personnelle, qui servent à améliorer l'adaptabilité et la créativité.

Si l'interaction est ouverte plutôt que fermée, toute situation est potentiellement nouvelle. C'est un processus de coconstruction qui transforme le soi et autrui à travers le temps. Ceci exige donc un système d'apprentissage adaptif qui puisse apprendre à anticiper et à comprendre les intentions d'autrui. Morin affirme ensuite que toute connaissance devrait comprendre une certaine subjectivité reliée à l'observation des autres, pour interpréter les problèmes éthiques :

Toute connaissance (et conscience) qui ne peut concevoir l'individualité, la subjectivité, qui ne peut inclure l'observateur dans son observation, est infirme pour penser tous problèmes, surtout les problèmes éthiques. Elle peut être efficace pour la domination des objets matériels, le contraire des énergies et les manipulations sur le vivant. Mais elle est devenue myope pour appréhender les réalités humaines et elle devient une menace pour l'avenir humain. (Morin, 2014, p. 65).

Nous pensons que ce genre de réflexivité basée sur la rétroaction est l'objectif de tout système autonome, qui transcrit l'information en connaissance. Puisque le bruit et l'erreur intrinsèque sont inhérents à toute interaction avec l'environnement, même chez les robots, l'adaptation produite par l'erreur – ou ce qu'on appelle l'apprentissage renforcé critique – peut mener à des comportements plus robustes et même à la détection de situations nouvelles et imprévisibles.

En pensant à la façon dont les agents évoluent dans un environnement dynamique, on peut aider à concevoir des systèmes autonomes qui

seraient capables de gérer les incertitudes et les imprévus. Les agents ouverts devraient donc afficher des propriétés d'apprentissage à travers leurs essais et leurs erreurs, de façon à examiner les situations à travers leurs actions. Selon la proposition d'Alan Turing, tout robot devrait apprendre progressivement, comme un enfant le fait à travers ses essais et erreurs en interagissant avec des êtres humains, ainsi que par l'imitation (Turing, 1950).

De la même façon, nous proposons de ne pas codifier de règles chez les robots, mais de les laisser apprendre les règles éthiques à travers leur expérience et leur interaction avec les personnes, comme un processus de développement. Autrement dit, pour nous, l'éthique devrait représenter le résultat du développement du robot et de ses interactions avec les autres, plutôt qu'une notion codifiée de façon abstraite, programmée séparément et en dehors du système décisionnel sensorimoteur du robot, comme pour les modules branchés.

L'apprentissage développemental devrait donc conduire au comportement adaptif, dont l'éthique ne représente qu'un aspect de la personification, qui est intégré sur le plan de ses propres actions et de sa compréhension de la scène.

Plutôt que d'essayer de produire un programme pour simuler le cerveau adulte, ne serait-il pas mieux d'essayer de simuler le cerveau d'un enfant ? Si celui-ci est ensuite soumis à un programme éducatif, on pourrait obtenir un cerveau adulte [...]. Notre Espoir est qu'il y ait peu de mécanismes dans un cerveau d'enfant, pour que ce soit facile à programmer. La quantité de travail éducatif que nous pouvons envisager, comme première approximation, devrait être semblable à celui d'un enfant humain. (Turing, 1950, p. 456)

Cette idée est à la base de la robotique développementale, qui a comme but principal de modéliser le développement de processus cognitifs progressivement complexes, au sein de systèmes naturels et artificiels, et de comprendre comment émergent ces processus à travers l'interaction physique et sociale (Asada *et al.*, 2001 ; Metta *et al.*, 2003). Cette approche considère le corps entier comme machine computationnelle uniformisée, qui peut apprendre toutes les expériences perceptives, y compris l'interaction avec les autres, à partir des actions. Cet objectif se base sur la découverte du système de neurones-miroir du cerveau (Rizzolatti *et al.*, 2001), qui donne un rôle important à l'intégration

multimodale et sensorimotrice de l'agence et au développement d'une théorie de la pensée chez les enfants (Gallese, 2005).

Les chercheurs en robotique développementale voudraient pouvoir reproduire ces architectures cognitives chez les robots pour profiter de ces possibilités (Kuniyoshi, 2014; Cangelosi et Schesinger, 2015). Le roboticien Minoru Asada a proposé un scénario de développement ancré dans le partage et l'apprentissage des valeurs éthiques chez les robots, telle l'empathie, à travers le développement de la conscience de soi en relation avec les autres (Asada, 2015). On trouve la capacité de prévoir les conséquences de nos propres actions à la base de ces architectures cognitives, de façon spéculative (Asada *et al.*, 2011).

3. INFÉRENCE ACTIVE

Être personnifié implique qu'on dépend, sur le moment, des événements et de la coordination sensorimotrice pour constater que nos actions ont un impact réel. Puisque toutes les situations sont différentes et se basent sur un contexte, il est important d'équiper tout système autonome d'une capacité de comprendre l'état dans lequel il se trouve, et d'adapter un plan en conséquence, selon le contexte actuel ou l'objectif. Comme chez les enfants, cette capacité d'« autocalibration », comme l'a nommée le chercheur en développement cognitif Philippe Rochat, est essentielle aux robots pour qu'ils puissent réaliser des comportements intelligents et développer leur sentiment de maîtrise (Rochat, 2003, 2011). Leur maîtrise sur les choses, ou sur l'agence, offre plus de prévisibilité aux personnes qui interagissent avec eux, ainsi qu'un sentiment d'aise, car la machine se comporte comme prévu. L'agence ou l'autocritique se basent sur les signaux d'anticipation d'erreurs qui analysent l'activité sensorielle courante, selon les prévisions du cerveau. Selon lui, le cerveau essaie de minimiser les différences et les anticipations d'erreurs de façon continue, pour pouvoir s'adapter à la situation courante, et ceci correspond au sentiment de l'agence ou de la prise de conscience.

Cette habileté du cerveau à prévoir le codage (codage prévisionnel) est très importante pour la compréhension de la scène, la lecture de pensées, ainsi que la représentation de soi et l'interaction avec les autres. Beaucoup de chercheurs ont défendu cette idée récemment, dont Andy Clark (Clark, 2015), Karl Friston et autres (Friston et Kiebel, 2009; Apps et Tsakiris, 2014). Jusqu'à présent, aucun robot n'a pu reconnaître une

scène ou se reconnaître lui-même dans le miroir, aucun n'a pu réussir à saisir un objet de façon satisfaisante, comprendre où se trouve sa propre main, ou même apprendre à anticiper les intentions des autres.

4. LES RÉGIONS SOI-AUTRUI DU CERVEAU ET LES NEURONES-MIROIR

Selon Anil Seth, le réseau du soi à l'intérieur du cerveau humain peut apprendre les signaux intéroceptifs en supposant des causes cachées (inférence active) ainsi que des erreurs (apprentissage par erreur) quand il est possible (ou non) de les prévoir (Seth, 2013). Seth a identifié le cortex insulaire antérieur (CIA) comme circuit qui compare et participe à l'inférence intéroceptive, en étant utile pour l'apprentissage par erreur dans le réseau du soi. Par extension, on a proposé que le circuit comparatif du CIA affiche le développement d'un réseau pour la théorie de l'esprit (habileté de lecture de la pensée) en utilisant le codage prévisionnel et l'apprentissage par erreur, pour l'inférence extéroceptive, en particulier (Frith et Frith, 2003).

Nous pensons que des règles abstraites de haut niveau (p. ex. les règles éthiques) peuvent être apprises à travers l'interaction personnifiée et créées à un niveau très primitif, par l'imitation de robots équipés de cette architecture cognitive basée sur le codage prévisionnel.

Les chercheurs neuroscientifiques ont examiné d'autres régions du cerveau, pour situer l'habileté de lire la pensée, tel que le système de neurones miroirs, perfectionné par les roboticiens développementaux. Dans le cerveau humain et chez le singe, le système des neurones miroirs est placé dans les circuits partagés du pariétal et dans les zones temporales et motrices et il est formé par les connexions réciproques et anatomiques qui travaillent en parallèle pour transformer l'information sensorimotrice (Ferrari *et al.*, 2009; Murata et Ishida, 2007). Quoique les fonctions primaires des robots soient dirigées vers leur interaction avec le monde physique, la neuroscience moderne peut attribuer un rôle beaucoup plus grand à leur engagement dans le monde social en transformant l'information multimodale sensorielle des objets en commandes motrices. La preuve principale vient des trouvailles de Rizzolatti *et al.* au sujet d'une classe particulière de neurones dans la zone 75 du singe, qui agit soit quand le singe effectue une action, soit quand il observe quelqu'un d'autre l'effectuer (Rizzolatti, *et al.*, 1996; Gallese *et al.*, 1996). La métaphore d'un mécanisme de miroir,

entre les appareils sensoriels et moteurs, pour produire une action et comprendre celles des autres, a été retenue pour nommer cette classe spéciale de neurones, qui se trouvent surtout dans les circuits moteurs ainsi que dans le cortex pariétal (Keysers, 2004; Gallese *et al.*, 2004). Ces neurones miroirs sont des neurones multimodaux qui relient des signaux à des systèmes visuels, proprioceptifs, auditaires et somato-topiques (Sakata *et al.*, 1997; Caggiano *et al.*, 2009).

Certains d'entre eux (dans la zone ventrale interpariétale) sont impliqués dans la représentation de l'espace à portée de main, l'espace périphérique personnel, qui codifie une image physique à la surface de la peau, ayant pour but de localiser les positions relatives des parties du corps et des objets aux alentours, dans des coordonnées centrées sur la personne (Rizzolatti *et al.*, 1997).

Par extension, ces neurones similitiroirs décrivent non seulement la manière dont le corps interagit physiquement dans son environnement, mais aussi la façon dont le soi incarné crée des liens sociaux avec les autres. Ces caractéristiques des neurones miroirs ont des implications profondes. En effet les neurones miroirs pourraient nous fournir des informations sur la création, depuis le niveau neural de la représentation du corps lui-même, des aptitudes cognitives plus élevées, comme l'empathie (Decety et Sommerville, 2007; Decety et Jackson, 2004; Bufalari *et al.*, 2007). Ils pourraient également fournir un éclairage sur la théorie de l'esprit (Fogassi *et al.*, 2005; Fujii *et al.*, 2008), sur les racines du langage (Rizzolatti et Arbib, 1998) et même sur la conscience corporelle (Keysers et Gazzola, 2006; Rizzolatti et Fabbri-Destro, 2008).

Il y a donc beaucoup d'intérêt quant au modelage des caractéristiques de structure du cerveau chez les robots, qui pourraient être utiles pour soutenir les comportements éthiques.

5. ALGORITHMES D'APPRENTISSAGE EXPLICATIFS

La codification prévisionnelle est un paradigme intéressant pour les robots éthiques, car ses racines sont ancrées dans l'inférence bayésienne. La théorie bayésienne est utilisée pour mettre à jour la probabilité d'une hypothèse, quand des preuves ou informations deviennent disponibles. Dans ce cadre, un robot peut mettre à jour ses hypothèses actuelles concernant les états actuels, ou selon le contexte pour les signaux entrants, ou basé sur ses actions dans son

environnement. Toute erreur ou tout conflit peut mettre à jour l'auto-jugement du robot en ce qui concerne une hypothèse, qui peut aussi être mieux expliquée à un utilisateur. Après tout, on ne comprend vraiment quelque chose que quand on peut l'expliquer. Ceci nous amène à la conception de robots qui apprennent en pratiquant, à partir de leurs erreurs et de leurs réussites, et découvrent les règles causales, comme s'ils étaient de petits scientifiques, ce que l'on se propose de faire avec les enfants (Gopnik *et al.*, 2000, Meltzoff, 2007, Tenenbaum, 2011). Richard Feynman affirmait : « *I cannot understand what I cannot create.* » (« Je ne comprends pas ce que je ne puis créer. ») Suivant ce constat, le principe de l'action engagée et de l'organisation de la connaissance à travers l'action est un choix nécessaire pour obtenir des robots qui apprennent par la pratique.

Nous pouvons ainsi affirmer que les algorithmes bayésiens sont un exemple des algorithmes d'apprentissage explicatifs. Puisque les hypothèses actuelles peuvent être évaluées et exprimées à l'utilisateur, ce genre d'apprentissage est différent de l'apprentissage avec les algorithmes boîte-noire, où l'état interne n'est pas disponible aux utilisateurs. Les processus bayésiens permettent de qualifier les incertitudes, en évaluant quantitativement ce qui est imprévu ou imprévisible. Certaines architectures d'IA sont conçues correctement pour déduire les causes cachées, comme les machines Boltzmann, les autoencodeurs ou les Réseaux bayésiens. De nouvelles implémentations ont aussi été proposées pour l'ingénierie inverse, comme les réseaux antagonistes génératifs (Goodfellow *et al.*, 2014), de même que des implémentations de codification prévisionnelle (Spratling, 2016).

Parce qu'ils imaginent les causes cachées de l'état actuel du robot, la codification prévisionnelle et le cadre bayésien peuvent être utilisés pour concevoir le retracement des décisions du robot. Lorsqu'il y a un problème, il faut laisser le robot agir, et peut-être échouer dans certaines situations, en essayant plusieurs hypothèses. C'est seulement ainsi que les algorithmes de codification prévisionnelle peuvent encourager l'accès ouvert aux architectures cognitives des robots, pour une meilleure compréhension de l'état interne du robot et un meilleur contrôle pour l'utilisateur.

6. UNE SCIENCE D'APPRENTISSAGE

Dans un article original, certains scientifiques développementaux préconisent la « fondation d'une nouvelle science de l'apprentissage » qui serait à la frontière de l'éducation, de la médecine, des sciences cognitives et de l'apprentissage par machine (Meltzoff *et al.*, 2009).

Tel un miroir de nous-mêmes, les robots pourraient apprendre des hypothèses du monde progressivement, avec les incertitudes et le bruit, les essais et les erreurs, pour nous aider à comprendre nos propres erreurs et échecs, ainsi que les préjugés que nous introduisons et propageons inconsciemment.

Cette approche plus horizontale concernant les capacités des robots, et plus ouverte concernant leurs interactions avec leur environnement, serait peut-être moins efficace que les algorithmes actuels qui se spécialisent autour d'une seule tâche. Mais elle pourrait aussi permettre plus facilement de généraliser et serait plus intéressante à utiliser en raison des nouveautés qu'elle génère et du travail d'équipe qu'elle favorise. Nous pouvons assumer que, à la longue, les algorithmes de la curiosité, favorisant dans les machines autonomes un apprentissage à vie, pourraient également avoir des effets bénéfiques sur les personnes qui interagissent avec ces machines, afin d'améliorer leurs capacités d'apprentissage pour les tâches cognitives et pour les sports, et ce, non pas seulement chez les enfants ou les gens âgés. Nous pensons que c'est l'étape à franchir pour maîtriser la complexité et la capacité d'adaptation des robots autonomes.

BIBLIOGRAPHIE

- Apps, Matthew et Tsakiris, Manos (2014). « The free-energy self: a predictive coding account of self-recognition », *Neurosci Biobehav Rev*, n° 41, p. 85-97.
- Asada, Minoru, Hosoda, Koh *et al.* (2001). « Cognitive developmental robotics as a new paradigm for the design of humanoid robots », *Robotics and Autonomous System*, vol. 37, p. 185-193.
- Asada, Minoru, Hosoda, Koh., Kuniyoshi, Yashuo *et al.* (2011). « Cognitive developmental robotics: a survey », *IEEE Transactions on Autonomous Mental Development*, vol. 1, n° 1, p. 12-34.
- Asada, Minoru (2015). « Development of artificial empathy », *Neuroscience Research*, vol. 90, p. 41-50.
- Bufalari, Ilaria, Aprile, Taryn *et al.* (2007). « Empathy for pain and touch in the human somatosensory cortex », *Cerebral Cortex*, vol. 17, p. 2553-2561.

- Caggiano, Vittorio, Fogassi, Leonardo *et al.* (2009). « Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys », *Science*, vol. 324, p. 403-406.
- Cangelosi, Angelo et Schlesinger, Matthew (2015). *Developmental Robotics*. MIT Press.
- Clark, Andy (2015). *Surfing Uncertainty Prediction, Action, and the Embodied Mind*. Oxford, Oxford University Press.
- Decety, Jean et Jackson, Philip L. (2004). « The functional architecture of human empathy », *Behavioral and Cognitive Neuroscience Reviews*, vol. 3, p. 71-100.
- Decety, Jean et Sommerville, Jessica A. (2007). « Shared representations between self and other: a social cognitive neuroscience view », *Cognitive Sciences*, vol. 7, n° 12, p. 527-533.
- Ferrari, Pier Francesco, Paukner, Annika., Ruggiero, Angela *et al.* (2009). « Interindividual differences in neonatal imitation and the development of action chains in rhesus macaques », *Child Development*, vol. 80, n° 4, p. 1057-1068.
- Fogassi, Leonardo, Ferrari, Pier Francesco *et al.* (2005). « Parietal lobe: from action organization to intention understanding », *Science*, vol. 308, p. 662-667.
- Friston, Karl et Kiebel, Stefan (2009). « Predictive coding under the free-energy principle », *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 364, p. 1211-1221.
- Frith, Uta et Frith, Christopher (2003). « Development and neurophysiology of mentalizing », *Phil. Trans. R. Soc. Lond. B*, n° 358, p. 459-473.
- Fujii, Naotaka, Hihara, Sayaka et Iriki, Atsushi (2008). « Social cognition in premotor and parietal cortex », *Soc. Neurosci*, vol. 3, p. 250-260.
- Gallese, Vittorio, Fadiga, Luciano *et al.* (1996). « Action recognition in the premotor cortex », *Brain*, p. 593-609.
- Gallese, Vittorio, Keysers, Christian et Rizzolatti, Giacomo (2004). « A unifying view of the basis of social cognition », *Cognitive Sciences*, vol. 8, n° 9, p. 396-403.
- Gallese, Vittorio (2005). « Embodied simulation: from neurons to phenomenal experience », dans *Phenomenology and the Cognitive Sciences*, vol. 4, p. 23-48.
- Girard, René (1976). *Deceit, Desire, and the Novel: Self and Other in Literary Structure*. Baltimore, Johns Hopkins University Press.
- Girard, René (1977). *Violence and the Sacred*. Paris, Grasset.
- Goodfellow, Ian J., Pouget-Abadie, Jean *et al.* (2014). *Generative Adversarial Networks*. arXiv:1406.2661.
- Gopnik, Alison., Meltzoff, Andrew et Kuhl, Patricia K. (2000). *The Scientist in the Crib What Early Learning Tells Us About the Mind*. New York, Publisher William Morrow Paperbacks.
- Keysers, Christian (2004). « Demystifying social cognition, a hebbian perspective », dans *Cognitive Sciences*, vol. 8, p. 501-507.
- Keysers, Christian et Gazzola, Valeria (2006). « Towards a unifying neural theory of social cognition », *Brain Research*, vol. 156, p. 379-401.
- Kuniyoshi, Yashuo (2014). « Humanoid robotics and neuroscience: Science, engineering and society », *Learning from Examples: Imitation Learning and Emerging Cognition*. Boca Raton, CRC Press/Taylor & Francis, p. 233-249.
- Lungarella, Max et Metta, Giorgio *et al.* (2003). « Developmental robotics: a survey », *Connection Science*, vol. 15, n° 4, p. 151-190.

- Lungarella, Max et Sporns, Olaf (2005). « Information self-structuring: Key principle for learning and development », *Proceedings of the 4th Int IEEE Conference on Development and Learning*, vol. 11, p. 25-30.
- Marshal, Peter J. et Meltzoff, Andrew (2015). « Body maps in the infant brain », *Trends in Cognitive Sciences*, vol. 19, n° 9, p. 499-505.
- Meltzoff, Andrew et Decety, Jean (2003). « What imitation tells us about social cognition a rapprochement between developmental psychology and cognitive neuroscience », *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, vol. 358, p. 491-500.
- Meltzoff, Andrew (2007). « Infants' causal learning Intervention, observation, imitation », dans Alison Gopnik et Laura Schulz (dir.), *Causal learning Psychology, philosophy, and computation*. Oxford, Oxford University Press, p. 37-47.
- Meltzoff, Andrew (2007). « "Like me": a foundation for social cognition », *Developmental Science*, vol. 10, n° 1, p. 126-134.
- Meltzoff, Andrew., Kuhl, Patricia K. *et al.* (2009). « Foundations for a New Science of Learning », *Science*, vol. 325, n° 5938, p. 284-288.
- Morin, Edgar et Le Moigne, Jean-Louis (1999). *L'Intelligence de la complexité*. Paris, Éditions L'Harmattan.
- Morin, Edgar (2014). *La méthode 6: l'éthique*. Paris, Le Seuil. 288 p.
- Murata, Akira et Ishida, Hiroaki (2007). « Representation of bodily self in the multimodal parieto-premotor network », dans Shintaro Funahashi (dir.), *Representation and Brain*. Tokyo, Springer.
- Nadel, Jacqueline et Butterworth, George (1999). *Imitation in Infancy*. Cambridge, Cambridge University Press.
- Pfeifer, Rolf et Scheier, Christian (1999). *Understanding Intelligence*. Cambridge, MIT Press.
- Pfeifer, Rolf et Bongard, Josh (2003). *How the Body Shapes the Way We Think, A New View of Intelligence*. Coll. « A Bradford Books », Cambridge, MIT Press.
- Pfeifer, Rolf et Pitti, Alexandre. (2012). « La révolution de l'intelligence du corps ». Paris, Manuella Editions.
- Rizzolatti, Giacomo, Fadiga, Luciano *et al.* (1996). « Premotor cortex and the recognition of motor actions », *Cognitive Brain Research*, vol. 3, p. 131-141.
- Rizzolatti, Giacomo., Fadiga, Luciano *et al.* (1997). « The space around us », *Science*, vol. 277, p. 190-191.
- Rizzolatti, Giacomo et Arbib, Michael A. (1998). « Language within our grasp », *Neuroscience*, vol. 21, p. 188-194.
- Rizzolatti, Giacomo, Fogassi, Leonardo et Gallese, Vittorio (2001). « Neurophysiological mechanisms underlying the understanding and imitation of action », *Nature Reviews Neuroscience*, vol. 2, n° 9, p. 661-670.
- Rizzolatti, Giacomo et Fabbri-Destro, Maddalena (2008). « The mirror system and its role in social cognition », *Curr. Opin. Neurobiol.*, vol. 18, p. 17984.
- Rochat, Philippe (2003). « Five levels of self-awareness as they unfold early in life », *Consciousness and Cognition*, vol. 12, p. 717-731.
- Rochat, Philippe (2011). « The self as phenotype », *Consciousness and Cognition*, vol. 20, n° 1, p. 109-119.
- Sakata, Hideo et Taira, Masato (1997). « The parietal association cortex in depth perception and visual control on hand action », *Neurosciences*, vol. 20, p. 350-357.

- Seth, Anil K. (2013). « Interoceptive inference, emotion, and the embodied self », *Trends in Cognitive Science*, n° 17, p. 565-573.
- Spratling, Michael W. (2016). « Predictive Coding as a Model of Cognition », *Cognitive Processing*, vol. 17, n° 3, p. 279-305.
- Tenenbaum, Joshua B., Kemp, Charles *et al.* (2011). « How to grow a mind statistics, structure, and abstraction », *Science*, vol. 331, n° 6022, p. 1279-1285.
- Turing, Alan (1950). « Computing machinery and intelligence », *Mind*, vol. LIX, n° 236, p. 433-460.
- Wiener, Norbert (1948). *Cybernetics or Control and Communication in the Animal and the Machine*. Cambridge, MIT Press.

CHAPITRE 8

Le vivant et la machine en biologie de synthèse : entre ingénierie et bricolage

Guillaume PELLETIER

Il est courant, dans le cadre d'une réflexion sur les robots, de chercher à évaluer dans quelle mesure un être artificiel pourrait éventuellement se rapprocher par certains de ses attributs des caractéristiques que nous prêtons au vivant. On aborde ainsi généralement cette réflexion à partir du point de vue de la machine, en considérant la possibilité que celle-ci, en se rapprochant du vivant, en vienne à effacer ou à rendre caduques les frontières qui nous permettraient de distinguer ces deux entités. C'est pourquoi il nous semble pertinent de considérer ici le chemin inverse, en examinant le point de vue de certains ingénieurs qui, dans le domaine de la biologie de synthèse, font au contraire des efforts considérables pour rapprocher le vivant des machines qui nous sont les plus familières. Ainsi se demande-t-on par exemple, depuis le tout premier congrès de la biologie de synthèse en 2004, comment l'on pourrait réduire suffisamment la complexité du vivant pour éviter ou surmonter des problèmes liés à l'émergence, à l'autoréplication ou aux aléas du contexte cellulaire, dans le but souvent explicite de concevoir des vivants analogues à nos machines électroniques. Il nous paraît assez ironique, à ce titre, que les caractéristiques du vivant que l'on cherche souvent à atteindre ou à imiter en robotique soient considérées

par certains biologistes de synthèse comme autant d'obstacles à un traitement technique de la vie.

La mise en perspective de ces deux attitudes nous semble être un moyen efficace de susciter une réflexion philosophique sur les relations entre l'organisme et la machine. Bien que toutes deux reconnaissent à la vie une certaine imprévisibilité, l'une y voit un modèle de conception technique alors que l'autre y voit un obstacle à la construction de machines. En nous appuyant sur une réflexion amorcée par le philosophe Georges Canguilhem dans le texte « Machine et organisme » (1946), nous proposerons dans ce texte de voir entre ces deux conceptions de l'objet technique non pas une opposition, mais une continuité biologique.

Nous débuterons par décrire le point de vue orthodoxe des ingénieurs de la biologie de synthèse, selon lequel une bonne machine est une machine qui répond à certains critères de rigidité et de prévisibilité. Nous expliciterons cette conception à travers la présentation de différents principes d'ingénierie, en nous concentrant sur l'exemple de la standardisation des « pièces » biologiques ; pratique qui à ce jour occupe une place importante dans les travaux effectués en biologie de synthèse. À partir de ce portrait général, nous nous questionnerons sur la vision du vivant, de la machine et de la technique que traduisent ces efforts, en appuyant notre réflexion sur certains travaux de Georges Canguilhem. Nous montrerons enfin qu'il est possible, au-delà des ambitions premières de la biologie de synthèse, de retrouver dans plusieurs de ses pratiques les plus récentes une attitude « bricoleuse » qui, en instrumentalisant les propriétés de la vie sans nécessairement chercher à les réduire entièrement à l'analyse, actualise la thèse d'une continuité entre la machine et l'organisme.

1. QU'EST-CE QU'UNE BONNE MACHINE ?

La spécificité de la biologie de synthèse en tant que discipline a souvent été caractérisée par l'application de principes d'ingénierie dans la sphère des sciences de la vie (voir par exemple : Brent, 2004 ; Heinemann et Panke, 2006 ; Andrianantoandro, 2006 ; Voigt, 2007). Bien que l'appellation « biologie de synthèse » recouvre aujourd'hui une multiplicité de méthodes et de visées, il demeure que son approche la plus répandue a pour ambition explicite de réduire ou de surmonter les complications techniques que soulèvent les systèmes vivants, afin

de rendre l'ingénierie des réseaux génétiques analogue, par exemple, à la conception de circuits électroniques.

Lorsque l'ingénieur dessine le schéma d'un circuit électronique, il connaît les paramètres généraux de chaque composante ; il connaît leurs effets, leur charge maximale, leur compatibilité, etc. Il maîtrise ainsi un ensemble de règles qui lui permettent de prévoir assez fidèlement, derrière son ordinateur, le fonctionnement du dispositif électronique qu'il construira. Même si la mise en œuvre du dispositif en question peut exiger certains ajustements, les principes de l'électronique sont de manière générale assez bien établis pour que l'on puisse prédire les résultats avec une certaine fiabilité. Or, c'est l'ambition d'en arriver à une telle maîtrise des systèmes génétiques qui motive aujourd'hui une bonne part des travaux effectués en biologie de synthèse. Ultimement, on espère arriver à établir des règles de conceptions standardisées qui permettront de concevoir des réseaux de gènes avec une prédictibilité plus proche de celle que connaît l'ingénieur en électronique. Ces dispositifs génétiques, une fois intégrés à des organismes, accompliront alors diverses fonctions utiles, telles que la synthèse de composés pharmaceutiques ou la production de biocarburants.

Cette version de la biologie de synthèse, par sa visée même, sous-entend que la non-ressemblance entre les systèmes vivants et les systèmes électroniques constitue le premier obstacle sérieux au traitement ingénieur d'un être vivant. En fait, cette difficulté initiale est généralement admise d'emblée par les biologistes de synthèse. Non seulement la quantité d'information que contient une cellule est-elle très importante, mais cette information ne connaît pas le même genre de barrières que constitue par exemple un fil de cuivre. Les macromolécules en suspension dans le cytoplasme, en plus d'avoir parfois des fonctions multiples, comportent une part d'imprévisibilité due au caractère stochastique de leurs mouvements (Heams, 2009). À ces difficultés viennent s'ajouter certains phénomènes proprement biologiques tels que l'émergence, l'auto-réplication ou encore la possibilité d'une mutation génétique. Cette complexité, il va sans dire, rend très difficile tout projet de modélisation fidèle à la réalité. C'est pour ces raisons, notamment, que les biologistes de synthèse se sont jusqu'à maintenant limités à concevoir des circuits génétiques relativement simples, introduits dans des organismes modèles bien connus, comme *Saccharomyces cerevisiae* (la levure de boulanger) ou la bactérie *Escherichia coli* (Bensaude-Vincent et Benoit-Browaeys, 2011).

Les efforts de la biologie de synthèse, jusqu'à maintenant, peuvent être interprétés comme des tentatives de contourner ces difficultés. Si tous les biologistes de synthèse ne s'accordent pas sur la possibilité éventuelle de caractériser fidèlement tous ces échanges d'informations, la plupart sont cependant confiants quant à la possibilité de « surmonter » cette complexité. Les chercheurs Heinemann et Panke, dans un article justement intitulé « Synthetic biology – putting engineering into biology », résument bien cette idée centrale :

[...] the ability to engineer biology in a directed and successful manner is still rather limited today and as a consequence, the complexity of things we can efficiently make is still quite small. Synthetic biology with its engineering vision aims to overcome the existing fundamental inabilities in system design and system fabrication, by developing foundational principles and technologies to ultimately enable a systematic forward-engineering of (parts of) biological systems for improved and novel applications. (Heinemann et Panke, 2006)

Ce que l'on voudrait souligner ici, c'est que la complexité du vivant, dans ce genre de vision, est considérée comme un *obstacle* à la conception de réseaux génétiques fonctionnels. Selon cette optique, l'objet technique devrait répondre à des règles strictes de contrôle et de prévisibilité, comme dans le cas de la conception de circuits électroniques. Considérons par exemple les propos de Drew Endy, pionnier de la biologie de synthèse au MIT, qui affirmait dans une entrevue : « I hate emergent properties. I like simplicity. I don't want the plane I take tomorrow to have some emergent property while it's flying¹ ». Au-delà de l'ironie de ces propos, il y a une vision de ce que doit être une machine. Une fonction échappant à un contrôle étroit et rigide n'a pas sa place dans une machine. Le vivant, dans l'esprit de la biologie de synthèse, est trop imprévisible. Il doit être modifié, purgé de ce genre d'aléas, afin d'être une bonne machine.

Ce n'est donc pas un hasard s'il est fréquent de trouver dans les écrits de scientifiques comme Drew Endy une opposition tranchée entre la rationalité de l'ingénieur et l'irrationalité des résultats de l'évolution. Les vivants actuels y sont en effet souvent dépeints comme étant l'aboutissement imparfait d'un enchaînement aveugle et irréflecti, auquel pourraient s'opposer les systèmes optimaux imaginés par

1. Entrevue de 2008 accordée à *the Edge Foundation* (Brockman, 2008, 19 février)

l'être humain. Cette opposition, étudiée par le philosophe et historien de sciences Michel Morange (2013), n'est pas sans s'appuyer sur la vision de l'évolution comme «bricolage», notion déjà évoquée par Darwin, mais surtout popularisée par François Jacob dans les années 1970. Au sens de ces auteurs, la métaphore du bricolage n'a pas une connotation péjorative, mais sert uniquement à mettre en lumière le fait que l'évolution ne peut, comme l'ingénieur, revenir au tableau noir et repenser le «*design*» d'un organisme. Comme le dit François Jacob, l'évolution «[...] travaille sur ce qui existe déjà, soit qu'elle transforme un système ancien pour lui donner une fonction nouvelle, soit qu'elle combine plusieurs systèmes pour en échafauder un plus complexe» (Jacob, 1981, p. 70). Or, en mettant l'accent sur la métaphore du bricolage et sur les limitations du pouvoir de l'évolution, certains ingénieurs en viennent, en déformant un peu les propos de ces auteurs, à souligner l'imperfection, l'incomplétude et le caractère non optimal des résultats de la sélection naturelle. Il résulterait de cette longue histoire d'échafaudages et de réutilisations une complexité incompréhensible et inadaptée au travail de l'ingénieur. Appliquer l'ingénierie au vivant reviendrait alors, en suivant la métaphore jusqu'au bout, à insérer la rationalité dans un monde bricoleur ou, pour reprendre les mots de Drew Endy, à «se libérer de la tyrannie de l'évolution²».

Plus concrètement, pour en arriver à surmonter l'imprévisibilité et la complexité caractérisant le vivant, les principaux artisans de la biologie de synthèse prônent l'adoption de principes ou de «règles de composition» qui guideront l'assemblage des dispositifs génétiques. On retrouvera, au premier rang de ces principes, le «raffinage» (*refinement*) et la standardisation des composantes biologiques (Endy, 2005; Canton, Labno, et Endy, 2008). Le raffinement fait référence au travail préalable de caractérisation et de manipulations qui doit être effectué sur différentes molécules (par exemple, une protéine ou une section d'ADN), afin qu'elles puissent être utilisées par les biologistes de façon plus ou moins prévisible. La standardisation, quant à elle, joue le même rôle qu'en électronique : au même titre que l'ingénieur électronique n'a pas à fabriquer ses propres condensateurs, le biologiste de synthèse n'aura pas idéalement à inventer ses propres pièces, mais pourra les commander en ligne à partir de banques de données. Ce processus de standardisation est effectué par des organisations telles que le Registry of Standard Biological Parts ou la BioBricks Foundation, qui

2. Rapporté dans une entrevue au *New Yorker*, le 28 septembre 2009 (Specter, 2009).

enregistrent et collectionnent des séquences d'ADN pour l'usage de toute personne ou organisation ayant les qualifications requises pour s'en servir. À ce jour, le Registry of Standard Biological Parts possède par exemple plus de 20 000 composantes biologiques à la disposition des biologistes de synthèse³.

Le but avoué de l'application de ces règles de composition est l'atteinte d'une certaine « abstraction », c'est-à-dire d'une séparation entre l'étape de conception et l'étape de fabrication. On espère ainsi que l'ingénieur puisse se concentrer sur le *design* de nouveaux dispositifs biologiques en se fiant uniquement aux fiches descriptives de composantes préalablement caractérisées et fabriquées (Mackenzie, 2010). Un exemple éloquent est celui du logiciel *Gene designer*, un outil à l'usage de biologistes amateurs ou professionnels disponible gratuitement sur Internet⁴. Ce logiciel dont l'interface est accessible et analogue à celles des logiciels utilisés en électronique permet de copier-coller, sélectionner ou modifier de manière variée des séquences de gènes afin de concevoir rapidement et efficacement de nouveaux circuits fonctionnels. Une fois le circuit modélisé, il ne reste plus qu'à commander sur le web, par le biais du logiciel, les biobriques requises. Ici encore, cette méthode est tout à fait analogue à celle utilisée en électronique : la plupart des disciplines du génie ont en effet leurs propres logiciels visant à modéliser les circuits et à aider à la recherche de composantes sur le net, le logiciel *Digi-Key* étant l'un des exemples les plus connus.

Ainsi, il est possible d'identifier dans plusieurs discours et pratiques des biologistes de synthèse une certaine vision de la technique et de la machine opposant la rationalité de l'être humain et de ses machines aux aléas du monde vivant. Par la dichotomie évolution-ingénieur se trouvent formulées, d'une part, la puissance de la raison, qui est explicitement valorisée et mise en contraste avec le reste du monde non humain et, d'autre part, la nécessité technique de considérer le vivant comme une machine, c'est-à-dire de le soumettre à cette raison en lui conférant prévisibilité et stabilité.

3. http://parts.igem.org/Main_Page

4. <https://www.dna20.com/resources/genedesigner>

2. DE L'ORGANISME À LA MACHINE

Il est utile de s'arrêter ici pour engager une réflexion proprement philosophique sur ces différentes visions du vivant et de la machine. Pour ce faire, il nous semble pertinent de recourir à la thèse avancée par le philosophe français Georges Canguilhem dans « Machine et organisme ». Dans cet article issu d'une conférence présentée en 1946, Canguilhem examine les relations entre la machine et l'organisme, non pas en se demandant dans quelle mesure la machine peut servir d'explication à l'organisme, mais en s'interrogeant sur le sens même de l'activité technique qui donne lieu à des machines. Ainsi, si la machine peut dans une certaine mesure servir d'*explication* à un vivant ou à ses fonctions, la *compréhension* de l'origine des machines ne se fait pas sans référer à une dimension vitale, c'est-à-dire à une action technique orientée en fonction de besoins authentiquement organiques. Cette thèse d'une continuité biologique entre l'organisme et la machine, comme nous le verrons, est susceptible d'éclairer certains des problèmes épistémologiques soulevés plus haut.

Avant de présenter cette continuité, il est cependant opportun d'en présenter les pôles, car un certain nombre de différences entre la machine et l'organisme identifiées par Canguilhem résonnent avec les caractéristiques que leur prêtent également les biologistes de synthèse. Canguilhem note par exemple la *rigidité* du fonctionnement des machines par rapport à celui des organismes. Une machine, au sens classique du terme, répond invariablement aux normes de fonctionnement imaginées par son constructeur. C'est pourquoi l'on peut diviser une machine, la réduire à une série de pièces qui, une fois soumises à l'analyse, permettront d'en comprendre le fonctionnement global : « Dans la machine, il y a vérification stricte des règles d'une comptabilité rationnelle. Le tout est rigoureusement la somme des parties. L'effet est dépendant de l'ordre des causes. » (Canguilhem, 2003, p. 149). Inversement, là où la machine est rigide, l'organisme fait plutôt preuve de plasticité. Pour illustrer cette plasticité, on peut prendre l'exemple du développement embryonnaire, exemple souvent mobilisé par Canguilhem. Rien ne révèle plus l'incroyable souplesse du vivant que la croissance et le développement d'un embryon, et c'est ainsi souvent pour mettre en lumière cette spécificité du vivant, irréductible à l'analyse, que Canguilhem a recours à cette discipline dans plusieurs de ses textes. Le finalisme de la machine est fixé à l'avance lors de sa construction, alors que la genèse d'un être vivant est une histoire qui ne peut être rigoureusement prédite à partir des

premières étapes du développement embryonnaire. Là où la machine se conforme à un plan rigide établi d'avance, le vivant possède une plus grande latitude et explore une foule de potentialités.

Ces différences majeures que note Canguilhem entre la machine et l'organisme ne doivent cependant pas nous faire perdre de vue la continuité qui les relie et qui constitue le point central de sa thèse. En effet, la différence effective de ces deux pôles n'implique pas que nous ayons affaire à deux ordres de réalité distincts. Il ne s'agit pas alors pour Canguilhem de réduire cette différence au profit du mécanisme, comme lorsque Descartes affirmait que les rouages d'une montre sont des mouvements aussi naturels que la fructification des arbres. Cette continuité ne s'effectue pas dans le sens cartésien d'une réduction de l'organisme à la machine, mais bien dans la direction inverse. Pour le dire autrement, il ne s'agit pas d'affirmer que les organismes sont des machines éminemment complexes, mais d'affirmer que les machines sont des outils biologiques extériorisés. Pour saisir la portée de cette thèse, considérons le passage suivant de « Machine et organisme » :

Mais, de notre point de vue, nous pouvons et nous devons inverser le rapport de la montre et de l'arbre et dire que les roues dont une montre est faite afin de montrer les heures, et, d'une façon générale, toutes les pièces des mécanismes montés pour la production d'un effet d'abord seulement rêvé ou désiré, sont des produits immédiats ou dérivés d'une activité technique aussi authentiquement organique que celle de la fructification des arbres et, primitivement, aussi peu consciente de ses règles et des lois qui en garantissent l'efficacité, que peut l'être la vie végétale. (Canguilhem, 2003, p. 155)

Canguilhem inverse ainsi la proposition cartésienne en donnant pré-séance au vital plutôt qu'au mécanique. Ce qui fait que les montres ne sont pas si différentes des arbres et du vivant en général, c'est qu'elles sont justement issues du vivant. Si on peut bien expliquer des machines avec les seules lois de la physique, on ne peut toutefois pas en comprendre le sens et la finalité sans inscrire les mécanismes fabriqués par l'être humain dans une histoire de la vie.

C'est ainsi que la proposition de Canguilhem s'appuie sur une réflexion sur la technique et sur sa relation avec la science. La technique, en tant qu'elle est l'effort d'un vivant pour répondre aux besoins et aux exigences de son milieu, ne peut être réduite à la science sans commettre d'anachronisme. Le savoir-faire acquis dans une connaissance

pratique a en ce sens quelque chose d'aussi organique que la fructification des arbres; il ne peut être analysé et divisé sans que l'on perde de vue sa finalité. Cette préséance historique de l'effort vital par rapport à la connaissance scientifique explique d'ailleurs la « témérité » caractérisant la technique⁵. Dans bien des cas, c'est cette dernière qui agit en premier sur le monde et qui oriente ensuite la science vers les problèmes à résoudre. L'exemple de la médecine est à ce titre paradigmatique: c'est l'état du patient, la dévalorisation existentielle de sa maladie, qui suscite la pratique médicale et qui, après coup, rend possible une médecine théorique.

Inscrire la technique dans un processus biologique n'implique cependant pas l'adhésion à un vitalisme fort, c'est-à-dire un vitalisme percevant le vivant comme une exception aux lois de la matière. Canguilhem, il faut le souligner, ne posait pas de différence ontologique entre la machine et l'organisme, au sens où il ne posait pas deux ordres d'entités répondant à des lois différentes. Il ne cherchait cependant pas pour autant, en faisant cela, à affirmer le pouvoir explicatif absolu de la physico-chimie. Il s'agissait plutôt pour lui de montrer que la technique, par sa dimension biologique, n'est pas réductible à la connaissance. Il posait ainsi simplement, presque méthodologiquement, une priorité absolue au fait que nous sommes avant tout des *vivants* constructeurs de machines.

On pourrait ainsi reformuler cette thèse de la façon suivante: à partir d'un certain niveau de complexité biologique, décomposer et analyser ne suffit plus. Il devient nécessaire, pour saisir la vie, de se mettre à son niveau. Or, à ce niveau, il suffit plus de faire appel à des lois physiques; il faut aussi faire appel à des normes biologiques. La normativité vitale, thème central de l'œuvre de Canguilhem, représente ce niveau d'explication. Là où une loi physique décrit un état de fait vrai en tout lieu et en toute circonstance, une norme décrit, pour reprendre les termes de Canguilhem, une « exigence » à l'égard de l'expérience. La vie ne fait pas que subir, comme les objets inertes, les conditions de son milieu; elle exclut et préfère certains états de fait. Le phénomène technique et, par extension, la construction des machines ne peuvent être saisis qu'en faisant appel à ce niveau d'explication.

5. « Si la témérité d'une technique, ignorante des obstacles qu'elle rencontrera, n'anticipait constamment sur la prudence de la connaissance codifiée, les problèmes scientifiques, qui sont des étonnements après avoir été des échecs, seraient bien peu nombreux à résoudre. » (Canguilhem, 2013, p. 81)

Il serait sans doute possible d'argumenter que certaines de nos machines, notamment dans certains courants d'intelligence artificielle et en robotique, sont plus «plastiques» et moins «rigides» que les machines décrites par Canguilhem. Il faut cependant rappeler que ces distinctions entre rigidité et plasticité ne changent au fond rien à la thèse de Canguilhem, au sens où même si nos machines étaient en tout point semblables à nous, cela ne rendrait pas moins irréductible leur *origine* vitale. Il importe, pour notre propos, d'insister sur ce point. Oui, nous dit-il, les machines actuelles sont distinguables des vivants par leur rigidité; mais il n'y a aucune raison de croire que cela demeurera ainsi. Rigidité et plasticité ne sont au fond que les pôles d'un continuum et ne représentent pas deux entités ontologiquement distinctes. Dans une note de bas de page à la toute fin de son texte, Canguilhem semble clairement envisager la possibilité que cette distinction puisse tendre à s'effacer. En faisant référence à des disciplines encore jeunes, la «*bionic*» et le «*bio-engineering*», il souligne qu'une «nouvelle espèce d'ingénieurs» a vu le jour au Massachusetts Institute of Technology, et qu'elle trouve dans le vivant un modèle technologique fécond (Canguilhem, 2003, p. 163). Cette dernière réflexion, comme nous le verrons maintenant, est plus que jamais d'actualité avec la biologie de synthèse.

3. LE BRICOLAGE ET L'EXPLORATION TECHNIQUE EN BIOLOGIE DE SYNTHÈSE

Les nouvelles entités qui sont fabriquées dans le cadre de la biologie de synthèse jouent sur les tensions entre les deux pôles du continuum décrit par Canguilhem. Étant à la fois machines et êtres vivants, ils nous font douter par leur existence même de la rigidité des frontières entre ces catégories. Pourrait-on par exemple appeler «machine» au sens classique du terme une éventuelle cellule synthétique qui porterait un tant soit peu de l'autonomie que nous attribuons généralement aux êtres vivants? Inversement, que signifierait «créer» la vie, si l'objet obtenu – par exemple un circuit génétique minimal et fonctionnel – répondait aux critères stricts de prévisibilité que l'on est en droit d'attendre d'une machine inerte? Comme on peut le voir avec ces exemples, il semble bien que la machine, pour se rapprocher de ce que nous entendons généralement comme «vivant», se voie contrainte d'abandonner un peu de sa rigidité.

Il n'est cependant pas évident que la possibilité de rapprochement entre nos machines et le vivant envisagée par Canguilhem corresponde à la vision que les biologistes de synthèse ont de leur propre travail. Comme nous l'avons vu, plusieurs discours semblent plutôt traduire la volonté de se distancier du vivant pour se rapprocher des machines plus classiques. Pourtant, il semble bien y avoir, dans le travail concret des biologistes de synthèse, une dimension technique échappant à la stricte analyse. Le vivant dans son intégralité sert en effet aujourd'hui à la fois de modèle, d'outil et de milieu technique pour bien des ingénieurs. La complexité du vivant, plutôt que d'être entièrement maîtrisée ou surmontée, est couramment instrumentalisée d'une façon qui correspond très peu au paradigme de l'ingénieur véhiculé dans les discours. Cette dimension exploratoire, qui nous semble être davantage admise et assumée dans certains travaux plus récents, ouvre la porte à une vision de la technique et du vivant correspondant à celle de Canguilhem, au sens où la vie y est prise à la fois comme point de départ et comme condition de l'activité technique.

C'est dans un tel esprit que Maureen O'Malley propose par exemple de voir dans la biologie de synthèse une part de ce que les informaticiens appellent du « *kludging* », qui pourrait être traduit par « bricolage », « bidouillage » ou, en bon québécois, « patentage ». Pour ce qui est du domaine informatique, cette expression désigne une façon d'opter pour une attitude très pragmatique en focalisant sur les résultats (par exemple, le bon fonctionnement d'un logiciel) même si cela est fait aux dépens de l'élégance ou même de la rationalité des moyens employés. De fait, la plupart des programmes informatiques nécessitent après leur conception un certain nombre de correctifs (couramment appelés « *patch* ») visant à surmonter les erreurs et les irrégularités. Pour ce faire, le « *kludging* » est un moyen efficace et couramment utilisé, qui s'avère bien souvent corriger les défauts d'un logiciel sans même que l'on comprenne nécessairement pourquoi cela a fonctionné. En laissant une liberté d'action aux programmeurs ou aux bricoleurs de tout genre sur les moyens utilisés pour parvenir à leurs fins, non seulement peut-on arriver à une solution satisfaisante face à un problème complexe, mais on peut aussi dans certains cas y arriver plus rapidement.

De la même façon, les dispositifs conçus par les biologistes de synthèse ont souvent besoin de « *kludging* ». Par exemple, l'un des exploits techniques de la biologie de synthèse – souvent considéré comme son plus grand succès – est la modification par l'équipe de Jay Keasling en 2006 d'une souche de levure, afin que celle-ci synthétise un composé

précurseur de l'artémisinine (Ro *et al.*, 2006). L'intérêt de cette réussite est que l'artémisinine est une molécule efficace dans le traitement du paludisme, et c'est pourquoi d'ailleurs l'équipe de Keasling bénéficia d'une bourse de 42 millions de dollars de la fondation Bill & Melinda Gates pour réaliser cette synthèse. Or, pour en arriver à modifier le métabolisme cellulaire de la levure afin qu'elle effectue les réactions désirées, il ne suffisait pas d'y insérer les bons gènes et d'attendre de voir le résultat: les chercheurs ont dû effectuer un bon nombre de modifications *ad hoc* visant à réguler ces réactions, le plus souvent d'une façon exploratoire et en étant bien loin de saisir tous les rouages menant à la synthèse désirée (O'Malley, 2009). Reste que les travaux de Keasling n'en demeurent pas moins une réussite: les composés précurseurs de l'artémisinine ont bel et bien été synthétisés par les levures, ouvrant la porte à la possibilité d'une production de masse de ce traitement contre le paludisme. Les techniques utilisées, bien que relevant du «*Kludging*», n'en demeurent pas moins d'une efficacité admirable. Plus encore, elles sont à envisager comme étant à la source de solutions innovantes, et Maureen O'Malley n'hésite pas d'ailleurs à en parler comme d'un «art»: «*Kludging should not be interpreted as a failure of synthetic biology, but as a highly creative and effective process.*» (O'Malley, 2009)

Cette façon d'envisager le travail technique des biologistes de synthèse donne une tout autre signification à la métaphore du «bricolage» en la débarrassant de son caractère péjoratif. Au fond, même une technique aussi sophistiquée que la biologie de synthèse, présentée comme l'irruption de la rationalité dans le monde vivant, doit faire face à des problèmes concrets, explorer à tâtons une foule de possibilités, et acquérir une forme d'expérience pratique. La distinction claire entre la rationalité de l'ingénieur et l'exploration bricoleuse du monde vivant est donc plus brouillée que l'on voudrait le croire. Cela ne signifie pas que la connaissance rationnelle, scientifique, n'informe pas la technique d'une façon peut-être plus manifeste en ingénierie ou en biologie de synthèse. Mais il est certain que les discours sur la biologie de synthèse ont trop souvent tendance à dénigrer sa dimension «savoir-faire», conçue comme un aspect superficiel du travail de l'ingénieur, alors qu'elle a quelque chose de constitutif de ce travail.

Certains éléments indiquent cependant que la biologie de synthèse pourrait s'orienter vers une reconnaissance de cette dimension technique non réductible à la science. Il est de plus en plus admis que même les composantes standardisées fournies aux biologistes de

synthèse ont généralement besoin de ce genre d'intervention bricoleuse lorsqu'elles doivent être assemblées dans un système complexe. Roberta Kwok, dans un article publié dans *Nature* en 2010, énonçait en ce sens cinq « dures vérités » auxquelles devait maintenant faire face la biologie de synthèse. Malgré les nombreuses promesses énoncées depuis le début des années 2000, il s'avère par exemple que les composantes censées être standardisées restent en grande partie très mal caractérisées. De plus, même dans les cas où les composantes sont très bien connues, leur intégration dans un système plus large cause souvent des conséquences imprévisibles (Kwok, 2010). Dans un autre article plus récent, Manuel Porcar, chercheur à l'Université de Valencia, soutient que la fameuse standardisation promise par la biologie de synthèse se révèle aujourd'hui insuffisante, mais que cette insuffisance n'est pas forcément à voir comme un échec (Porcar, 2016). Après plus de 15 ans de standardisation, il s'avère en effet toujours impossible de concevoir sur papier un réseau métabolique qui serait véritablement prédictible. Pourtant, cela n'empêche pas les ingénieurs d'en construire avec succès. Aussi Porcar soutient-il une attitude pragmatique qui renoncerait, du moins en partie, à l'attitude rationaliste défendue par les fondateurs de la biologie de synthèse.

Il semble ainsi que l'on puisse parfois s'accommoder ou même profiter de l'imprévisibilité du vivant. Les bactéries modifiées issues de la biologie de synthèse possèdent une relative indépendance par rapport à leurs « constructeurs », ne serait-ce qu'en raison de leur éventuelle capacité à se reproduire et à se transformer. Elles sont à la fois des objets techniques et des êtres vivants qui échappent en partie à notre contrôle. C'est dans cet ordre d'idées que Catherine et Raphael Larrère affirment, dans *Penser et agir contre la nature* (2015), que les biologistes de synthèse « pilotent » la nature tout en croyant la contrôler. Plutôt que d'envisager la technique comme un moyen de domination de la nature, le pilotage implique de considérer celle-ci comme un partenaire avec lequel il faut composer. Comme dans les cas paradigmatiques de l'agriculture, de la domestication des animaux ou de la fermentation contrôlée, la notion de partenaire implique la reconnaissance que certains paramètres sont indépendants de nous et échappent à un contrôle rigide. Il ne s'agit alors pas d'imposer notre volonté à une nature passive, mais plutôt d'orienter des phénomènes déjà présents, de « faire-faire » et de « faire-avec » (Larrère et Larrère, 2015, p. 15).

Ainsi, le fait que nos machines tendent aujourd'hui à être difficilement discernables de la vie ne montre pas que nous aurions accompli une

réduction du vivant grâce au pouvoir de notre connaissance. Bien au contraire, ce rapprochement mène à la démonstration inverse; il consiste en un certain aveu des limites de l'analyse à saisir à elle seule le vivant et les techniques qui l'instrumentalisent. Cela ne signifie pas que nous devrions abandonner les principes d'ingénierie qui sont légitimement promus dans les discours des biologistes de synthèse. Rendre compte de la normativité du vivant n'était d'ailleurs pas, pour Canguilhem, une façon d'invalider la connaissance objective: « Ces affirmations n'entraînent aucune interdiction », disait-il après avoir soutenu que les « formes vivantes » ne pouvaient être saisies dans une division (Canguilhem, 2003, p. 14). De même, la biologie de synthèse a sans doute tout à gagner à formaliser, à analyser et à standardiser. Mais cela n'élimine pas le fait que ses applications sont désormais vivantes, et qu'en cela il est tout à fait légitime d'assumer l'antériorité de notre activité technique sur la connaissance des mécanismes du vivant.

Il demeure donc selon nous pertinent de rendre compte de cette dimension autonome et indépendante que sont appelées à occuper nos machines. Une foule de techniques tendent aujourd'hui à être à la fois vivantes et artificielles. Certes, cette dualité n'est pas nouvelle. Les humains n'ont par exemple pas attendu la connaissance de la génétique pour modifier les vivants qu'ils employaient pour leur activité technique. Mais cette dualité entre vie et artifice est particulièrement accusée avec les biotechnologies. C'est le vivant dans son intégralité qui, par le biais de ces nouvelles technologies, est instrumentalisé d'une façon radicalement nouvelle. L'organisme devient lui-même, dans sa totalité, un milieu technique. Ce retournement vers le vivant, sans être un renoncement à la connaissance analytique, montre néanmoins les limites d'une conception cartésienne de la technique attachée à connaître et à agir de façon prévisible sur une matière passive. De même, les « machines » de la biologie de synthèse, étant eux-mêmes vivantes, ne peuvent guère être aujourd'hui perçues comme la simple application d'un savoir théorique.

CONCLUSION

Nous avons vu que, selon Canguilhem, le sens biologique est perdu lorsqu'on tente de réduire un fait vital à l'analyse de ses constituants. Les vivants correspondent à des « formes » et à des « totalités » qui « [...] peuvent être saisies dans une vision, jamais dans une division » (Canguilhem, 2003, p. 14). De la même façon, les biologistes de synthèse

renoncent en partie à diviser lorsque vient le temps d'instrumentaliser la vie à des fins pratiques. Ils orientent des tendances plus qu'ils n'inventent rationnellement des mécanismes. Des fonctions déjà présentes dans le vivant sont ainsi modifiées pour s'aligner aux finalités techniques de l'être humain. Loin d'être une vision de la technique comme instanciation d'une connaissance, il s'agit d'une technique véritablement « à l'école de la vie ». Renonçant en partie à l'exigence rationaliste voulant que l'on divise avant de fabriquer, cette approche n'attend pas toujours la science pour tirer avantage du vivant, agissant ainsi avec la témérité qui selon Canguilhem témoignait de l'origine vitale de la technique.

À cet égard, le continuum que met en lumière Canguilhem entre l'organisme et la machine nous semble plus pertinent qu'un clivage strict entre « naturel » et « artificiel ». De plus en plus, certains objets seront appelés à brouiller ces frontières traditionnelles, précisément parce que les constructions humaines emprunteront davantage à la normativité typiquement associée à la vie. La normativité de la vie suppose une certaine autonomie, une finalité propre à elle-même. Inversement, la rigidité des machines suppose un contrôle plus étroit des paramètres, permis notamment par la standardisation des pièces. Entre les deux, il n'y a pas de coupure, mais une extériorisation progressive de la finalité en dehors de l'être humain, allant dans le sens de l'organique à la technique. En plaçant les différentes réalisations de la biologie de synthèse le long de ce continuum, on comprend mieux le sens de l'activité technique, et on évite en partie les difficultés liées à l'attribution d'un statut d'artificialité. Une machine n'est pas plus « artificielle » qu'un animal ayant subi un grand nombre de sélections artificielles, ce qui ne veut pas dire qu'il n'y a pas pour autant un large écart entre ces deux entités. Une machine n'a pas de finalité inhérente; elle ne peut pas tomber malade, instituer des normes, apprécier, ni déprécier son expérience. Il vaut peut-être mieux s'efforcer de comprendre où, entre ces deux extrêmes, se situe l'entité technique que l'on étudiera, qu'il s'agisse d'une cellule synthétique, d'une bactérie modifiée ou d'un robot autonome.

BIBLIOGRAPHIE

- Andrianantoandro, Ernesto., Basu, Subah, *et al.* (2006). « Synthetic biology: new engineering rules for an emerging discipline », *Molecular Systems Biology*, vol. 2, Mai 2016, p. 1-14.

- Bensaude-Vincent, Bernadette et Benoit-Browaey, Dorothée (2011). *Fabriquer la vie: Où va la biologie de synthèse?* Paris, Éditions du Seuil.
- Brent, Rebecca (2004). «A partnership between biology and engineering», *Nature Biotechnology*, vol. 22, n° 10, Octobre 2004, p. 1211-1214.
- Brockman, John (2008). «Engineering biology: A talk with Drew Endy», *Edge*, Février, p.19.
- Canton, Barry, Labno Anna et Endy Drew (2008). «Refinement and standardization of synthetic biological parts and devices», *Nature Biotechnology*, vol. 26, n° 7, p. 787-793.
- Canguilhem, Georges (1937). «Descartes et la technique», dans *Travaux du IX^e congrès international de philosophie* (II^e partie). Paris, Hermann.
- Canguilhem, Georges (2003). *La connaissance de la vie*. Paris, Vrin.
- Canguilhem, Georges (2013). *Le normal et le pathologique*. Paris, Presses universitaires de France.
- Endy, Drew (2005). «Foundations for engineering biology», *Nature*, vol. 438, n° 7067, Novembre 2005, p. 449-453.
- Heams, Thomas (2009). «De quoi la biologie synthétique est-elle le nom?», dans Thomas Heams, Phillippe Huneman, Guillaume Lecointre et Marc Silberstein (dir.), *Les mondes darwiniens: l'évolution de l'évolution*. Paris, Syllepse, p. 637-686.
- Heinemann, Matthias et Panke Sven (2006). «Synthetic biology – putting engineering into biology», *Bioinformatics*, vol. 22, n° 22, Novembre 2006, p. 2790-2799.
- Jacob, François (1981). *Le jeu des possibles: Essai sur la diversité du vivant*, Paris, Fayard.
- Kwok, Roberta (2010). «Five hard truths for synthetic biology», *Nature*, vol. 463, Janvier 2010, p. 288-290.
- Larrère, Catherine et Larrère, Raphaël (2015). *Penser et agir avec la nature*. Paris, Éditions La Découverte.
- Mackenzie, Adrian (2010). «Design in synthetic biology», *Biosocieties*, vol. 5, n° 20, p. 180-198.
- Morange, Michel (2013). «Comparison between the work of synthetic biologists and the action of evolution: engineering versus tinkering», *Biological Theory*, vol. 8, n° 4, Mai 2013, p. 318-323.
- O'Malley, Maureen A. (2009). «Making knowledge in synthetic biology: design meets kludge», *Biological Theory*, vol. 4, n° 4, Novembre 2009, 378-389.
- Porcar, Manuel (2016). «Synthetic biology: From having fun to jumping the gun», *Nanoethics*, vol. 10, Janvier 2016, p. 105-109.
- Rheinberger, Hans-Jörg (1995). «Beyond nature and culture: a note on medicine in the age of molecular biology», *Science in context*, vol. 8, n° 1, p. 249-263.
- Ro, Dae-Kyun., Paradise, Eric M. et al. (2006). «Production of the antimalarial drug precursor artemisinic acid in engineered yeast», *Nature*, vol. 440, n° 13, Avril 2006, p. 940-943.
- Specter, Michael (2009). «A life of its own: Where will synthetic biology lead us?», *The New Yorker*, Septembre 2009.
- Voigt, Christopher A. (2006). «Genetic parts to program bacteria», *Current Opinion in Biotechnology*, vol. 17, n° 5, Octobre 2006, p. 548-557.

CHAPITRE 9

Robotisation et aliénation : de l'objectivation de l'homme à la valorisation de la machine

Chantale PILON

La fulgurance des avancées technoscientifiques qui ont cours depuis les derniers siècles, la présence toujours plus marquée et diversifiée des robots dans notre quotidien, l'omniprésence des représentations de corps améliorés, etc., laissent entendre que ce qui n'était jadis que pure fantaisie est en voie de prendre corps. Le fantasme d'un monde dominé par les robots et les cyborgs, marquant le **passage vers une humanité transformée technologiquement où l'homme aura quitté son enveloppe charnelle pour devenir un être hybride, augmenté de puces et de prothèses, à mi-chemin entre la machine et le réseau**, ne nous a jamais paru si vraisemblable.

Bien qu'il soit dénoncé depuis longtemps, ce processus de « mécanisation » semble s'accroître de manière inédite. L'humain ne serait plus seulement comme une machine, il pourrait maintenant « devenir un robot » (Gunkel, 2012 ; Carr, 2014). De sorte que, indépendamment de l'enthousiasme ou des craintes que cette perspective suscite, le moment où la fiction rejoint la réalité paraît frapper à nos portes.

Pour mieux saisir l'essence des enjeux liés à cette problématique, nous souhaiterions traiter de cette question très actuelle et de ses incidences

sur nos existences, en soulignant d'abord le renversement de valeurs que trahit cet engouement pour la robotisation et le phénomène plus large de la technologisation de la société, dans lequel elle s'inscrit.

En effet, cet intérêt suppose, d'un côté, que l'identification à la machine, traditionnellement entendue de façon péjorative et réductrice, est aujourd'hui perçue positivement, jusqu'à devenir un idéal à atteindre. De l'autre, il implique que l'humanité ait développé un degré suffisamment élevé d'insatisfaction face à sa nature pour souhaiter sa propre transformation et développer des moyens technologiques (l'« *human engineering* ») afin de s'émanciper des limites de sa condition.

Ce qui nous amène à nous demander si la robotisation est bien la source de cette insatisfaction ou le simple moyen destiné à y répondre. En suivant cette idée, notre attention se portera moins sur les conséquences de la robotisation de l'humain et de la société que sur ce qui nous y a conduits. Il nous semble, en effet, plus aisé de comprendre vers où nous nous dirigeons lorsque nous comprenons comment et pourquoi nous en sommes arrivés là. Ce déplacement de la question ouvre la voie à une relecture du problème au regard de la situation historique qui l'a fait naître et de l'idéologie sociale qui préside à son développement.

Certains éléments donnent à penser que ce désir de robotisation n'est pas inédit mais qu'il s'enracine plutôt dans le processus d'objectivation de l'humanité, conditionnel à l'essor du système capitaliste. Ainsi, les inquiétudes suscitées par ce désir de transformation s'expliqueraient moins par la nouveauté de sa présence historique que par le sentiment de vulnérabilité et d'« immaîtrise » éprouvé face à la marche du progrès, qui annonce sa réalisation prochaine. Autrement dit, la démesure du pouvoir technique ne serait pas la cause de cette volonté; elle s'inscrirait plutôt dans l'extension de la logique capitaliste, comme son corolaire.

Pour mieux saisir l'origine de cette hostilité grandissante face à la nature humaine et de cette inclination à développer des technologies toujours plus sophistiquées afin d'y remédier, il importe d'examiner les présupposés et les mécanismes de la conception idéologique responsable de ces tendances conjointes. Cela engage une réflexion sur leurs affinités électives, c'est-à-dire la façon dont, d'une part, la logique capitaliste et, d'autre part, la pensée technicienne entrent, à partir d'une certaine affinité de sens, dans un rapport d'attraction, d'influence et de renforcement réciproques.

Les avenues explorées devraient nous permettre de comprendre : Comment le désir de domination se lie à l'objectivation de l'humain ? Pourquoi ce processus de réduction et cet idéal, dont les standards sont constamment redéfinis et repoussés par les technologies, ont-ils engendré un sentiment d'insuffisance face au corps et ont-ils alimenté le désir d'émancipation technoscientifique des limites de la nature humaine ? Et, pourquoi la quête de pouvoir produit de l'« immaîtrisé » ? En conclusion, cet examen des directions que nous sommes en train de prendre quant à notre futur doit ouvrir la voie à un champ de résolution permettant de nous réapproprier ces tendances et de les réorienter vers une destination plus saine et prometteuse.

1. L'IDÉOLOGIE DE LA DOMINATION ET DÉSHUMANISATION

D'emblée, soulignons que la technologie vise toujours à répondre à un besoin pratique précis auquel elle doit apporter une solution technique efficace. En tant que moyen, elle se présente comme un projet qui ne contient pas sa justification en lui-même. À cet égard, elle est débitrice de l'idéologie qu'elle vise à réaliser, qui préside à ses développements et les rend possibles par le lexique analytique et les cadres conceptuels dans lesquels ils sont présentés et justifiés. Il convient donc de s'y intéresser.

Tout d'abord, qu'est-ce qu'une idéologie ? Les idéologies sont des systèmes de conviction opératoires qui résultent d'une praxis sociale spécifique et se rencontrent dans des systèmes d'idées, des pratiques et des formes d'habitus¹. Ces idées se tiennent et se forment systématiquement dans des rapports déterminés. Elles concrétisent notre relation au monde et, partant, l'horizon d'interprétation dans lequel nous nous comprenons, dans lequel nous comprenons les conditions et les structures sociales, ainsi que la manière dont nous évoluons. Les formes idéologiques dominantes, qui exercent une influence décisive sur toutes les manifestations de la vie, ne visent pourtant ni l'équité, ni la connaissance, ni la vérité. Le concept d'idéologie apparaît lorsqu'il est question de représentations imaginaires et non de savoir. En effet, une représentation sera dite « idéologique » lorsqu'elle sera fautive,

1. L'habitus peut être défini comme des dispositions psychiques qui peuvent être influencées par l'éducation, mais qui ne sont ni inconscientes ni soustraites à l'action de la volonté. Elles sont des cadres ou des guides, dont le sujet peut se détacher avec plus ou moins d'aisance (Boudon, 2003, p. 113).

c'est-à-dire lorsque le système social, parce qu'il nuit au développement de la véritable nature de l'homme, compromet sa dignité². Il ne s'agit cependant pas là d'une simple erreur. Le statut singulier de l'idéologie relève du fait qu'elle ne se limite pas à désigner le caractère erroné des représentations inadéquates de la réalité. Les enjeux de l'idéologie ne sont pas d'ordre « théorique » (c.-à-d. orientés vers la connaissance et l'élimination des erreurs), ils sont politiques (c.-à-d. rattachés à la domination du champ de bataille théorique). Le discours idéologique vise à harmoniser la vie sociale en créant des entraves sémantiques qui affectent la conscience et contribuent au maintien des rapports sociaux de domination. En cela, il se révèle être un acte destiné à les légitimer.

Or, pour dominer, il faut identifier des subordonnés et s'en distinguer. La notion de pouvoir ne va pas sans celle d'ascendant, de maîtrise, de domination et, corrélativement, de dépendance et de subordination. C'est ce qui la distingue de celle d'autorité³. Par principe, la domination ne peut être que la domination d'une minorité, c'est-à-dire exercée par une minorité, dans l'intérêt d'une minorité.

La négation ou la neutralisation de l'humanité, nécessaire à l'instauration du rapport de domination objectif, s'opère notamment par le biais d'une :

- a) dissociation radicale, qui considère autrui comme fondamentalement étranger et inconnaissable ;
- b) assimilation à soi, qui consiste à rendre semblable ou identique à soi, en transformant ou en incorporant à sa propre substance une substance étrangère ou des éléments vivants d'un type déterminé qui, par là même, sont réduits à néant.

En effet, qualifier autrui d'« inférieur » ou de subordonné, c'est refuser de le considérer comme un semblable pour en faire quelqu'un

2. Par exemple, une représentation idéologique de la vie d'un pommier pourrait être celle d'un pommier qui ne produit pas de pommes. Cette représentation de la vie d'un pommier, même si elle traduit correctement une situation réelle particulière, est considérée comme fautive car l'image qu'elle nous en présente n'est pas représentative de la vérité du pommier et ne permet pas de saisir sa véritable nature.

3. Cette dernière renvoie au sage, par exemple, celui duquel l'on observe volontiers les préceptes sans devoir y être contraint, une personne qui jouit d'une grande considération, qui inspire le respect, l'admiration, s'impose d'elle-même et qui, par sa valeur reconnue, peut servir de référence. La vérité a une valeur d'autorité, qui ne requiert pas de plaire à ses adhérents. En raison de son incontestable exactitude, elle s'impose d'elle-même et ne laisse d'autre choix que de se rendre à l'évidence.

d'étranger, de lointain, de radicalement différent, au point qu'on puisse l'employer comme un instrument ou le concevoir comme un animal, en effaçant la distance qui l'en sépare pour en faire une proie. Selon Hannah Arendt, la déshumanisation consiste à opérer une réduction qui se traduit par une négation de « l'élément proprement humain (la spontanéité et l'expression de la différence d'un individu à l'autre) » (Ménissier, 2008, p. 4) et culmine dans l'engendrement d'une sorte d'indistinction, « une uniformisation de la pluralité humaine » (Ménissier, 2008, p. 3) qui « transforme les individus singuliers en masse informe » (Ménissier, 2008, p. 3), des « généralisations nihilistes ». Bref, il s'agit de dénier à « l'homme sa pluralité naturelle de condition » (Ménissier, 2008, p. 7) et d'amenuiser « jusqu'à l'extrême l'espace dans lequel celle-ci s'exprime normalement » (Ménissier, 2008, p. 7). Cette déposition de la liberté ramène l'humain à la fonction d'un corps agissant, le réduit à une condition bioanthropologique, où toute possibilité d'agir⁴ lui est refusée (sa vie se bornant alors à de simples comportements), et lui ôte toute prise sur son histoire⁵.

La domination renvoie au « droit du plus fort » (notion autocontradictoire, car la force est anti-juridique), à un régime de l'arbitraire absolu. À cette occasion, la relation de contrainte exercée par un être sur un autre se réduit au pouvoir despotique, qui devient une propriété individuelle et où la volonté d'un seul vaut comme si elle était la loi de tous. Elle consiste à utiliser la force comme moyen de pression sur la volonté de l'autre pour qu'il accepte que nos objectifs deviennent ses objectifs et pour le contraindre à agir, au risque de porter atteinte à son intégrité. Menaçant ou détruisant l'unité de l'être, la domination implique de nier que l'autre est un être de raison. Elle s'exerce alors moins sur un sujet que sur un objet, dans la mesure où le corps dominé est déssubjectivé, assimilé et colonisé.

Elle exige donc une asymétrie, soit une « appréciation différentielle de la vie humaine⁶ », et, en raison de cette asymétrie, une répartition inéquitable des droits, des ressources, des pouvoirs, du fardeau socio-économique, etc. Il importe, en effet, de comprendre que la misère relève

4. Au sens de l'agir collectif et concerté.

5. Car cela revient, en définitive, à fermer la perspective historique elle-même.

6. Cette idée implique qu'on attribue une valeur différente à la vie humaine, que certaines vies ou certains êtres humains soient considérés comme plus ou moins importants que d'autres.

moins d'un manque de ressources que d'un déséquilibre produit par la répartition inéquitable de la richesse.

En tant que facteur de renforcement de l'inégalité parmi les hommes, l'exercice du pouvoir est assimilable à une panoplie de techniques de domination, qui apparaissent sous la forme de contraintes sociales et culturelles plus ou moins consciemment intégrées au travers des comportements individuels. Différents aspects de la vie sociale se présentent alors comme une forme de violence détournée, comme l'effet d'organisations ou d'institutions d'ordre socioculturel et politique particulièrement restrictives et répressives envers certaines catégories de la population.

2. RATIONALISATION, DÉVALORISATION ET AUTO-INSTRUMENTALISATION

Nous tenterons maintenant de montrer que la logique capitaliste et le processus de rationalisation de la société, en tentant de rationaliser l'irrationnel, ont conduit à un renversement radical de valeurs, qui apprécie la vie humaine en vertu de critères issus de la physique et de l'ingénierie. Ce nouvel étalon conduit à une dévalorisation de l'humanité et encourage l'identification à la machine, la promotion de tout ce qui nous en rapproche et une auto-instrumentalisation.

Le problème essentiel auquel se heurte la société bourgeoise tient au fait que le discours promotionnel, celui qui a accompagné l'ascension au pouvoir de la bourgeoisie, s'est profilé à l'horizon de l'éloge de la raison, de la démocratisation et des droits de l'homme. Cette société ne connaît plus de hiérarchies naturelles. Son pouvoir autocratique s'inscrit dans une histoire sans finalité et se trouve sans légitimité extérieure, car aucune morale ne le transcende et aucune providence ne l'ordonne. Le Dieu qui ratifiait l'aristocratie est mort. Conséquemment, dans un système de type capitaliste où l'État a englobé la société civile, il ne peut être question de véritable autonomie du politique. Ses actions sont limitées par la nécessité de les entériner démocratiquement. Il est donc impératif, pour le maintien du pouvoir, que les dominés consentent à l'être et se conforment à ces prescriptions « volontairement ». Le défi consiste alors à justifier rationnellement le désir pulsionnel (donc irrationnel) de dominer, qui commande l'assujettissement d'une frange, arbitrairement déterminée, de l'humanité. Toute l'ambiguïté est là. D'une part, le départage entre les dominants

et les dominés ne peut être que partial, puisque, entre membres d'une même espèce, il n'existe pas de distinction d'essence. Rien ne justifie naturellement une taxinomie et une systématisation de cet ordre. D'autre part, l'obéissance de l'homme à l'homme n'est jamais entièrement fondée sur le droit, sur un principe absolument rationnel ni susceptible d'être parfaitement admis par les deux parties. Dès lors, comment établir un rapport de domination rationnel, « volontairement consentie », entre des êtres présumés égaux ?

Pour parvenir à légitimer l'asymétrie, c'est le prétexte du mérite, fondé sur l'égalité des chances, qui fut invoqué. Ce dernier implique :

- a) d'ignorer les spécificités des individus jusqu'à l'indistinction afin de les concevoir initialement comme des entités neutres et équivalentes, ramenant ainsi l'égalité à la mêmeté ;
- b) que le mérite est une donnée quantifiable et mesurable ;
- c) que la dignité humaine se situe dans l'efficacité et la performance (des critères empruntés à la physique et l'ingénierie).

La méritocratie prône ainsi l'idée « que le meilleur gagne » en regroupant indistinctement tous les êtres dans la même catégorie, les opposant les uns aux autres, dans le cadre d'une concurrence matérielle permanente. Ici, la compétition atomise, concentre l'attention sur le particulier et fait écran à la logique d'ensemble. Or, en associant la réussite au mérite, la méritocratie associe la misère à la médiocrité et parvient non seulement à cautionner l'exploitation des uns par les autres, mais de surcroît à la rendre méritoire, en imputant aux exploités la responsabilité de leur asservissement.

La rationalisation, fondée sur l'altération fallacieuse des notions de liberté et d'égalité, conduit donc à de nouvelles formes de déshumanisation, qui caractérisent la société marchande (soit celle où l'on ne produit plus pour consommer sa propre production, mais exclusivement pour vendre des marchandises). La transformation formelle correspond à l'élimination de l'ensemble des entraves politico-sociales et des vestiges de type patriarcal, politique et religieux, par lesquelles se modifie aussi le rapport de domination existant. Cette relation d'exploitation se substitue aux régimes de subordination traditionnels, pour faire place à un nouveau système, qui se manifeste aussi sous une forme politique. La contrainte ne repose plus sur un rapport fixe de domination d'ordre sociopolitique ou personnel. Il s'agit d'une relation

de subordination « objective », fondée sur la dépendance économique et uniquement déterminée par les différentes « fonctions économiques ».

Sous l'angle de l'accord contractuel (précédant le procès de production), les échangistes sont, théoriquement, relativement sur le même palier et se font face comme possesseurs de marchandises, n'ayant entre eux qu'un rapport purement monétaire. Une chose est offerte par l'un, demandée par l'autre, et l'échange comble les deux parties, de sorte que chacun y trouve son compte. L'entente relationnelle est d'ordre pragmatique, strictement utilitaire (c.-à-d. attaché à sa fonction et borné à sa réalisation) et impersonnel. Elle ne concerne ni la personne qui offre ni celle qui demande. Elle porte exclusivement sur ce qui est offert et ce qui est demandé, sans tenir compte des personnes impliquées.

Toutefois, il convient de s'intéresser au caractère paradoxal de l'entente capitaliste : Comment une entente relationnelle volontaire, destinée à satisfaire l'intérêt commun, peut-elle conduire à un rapport de domination ? Pour répondre à cette question, il convient de souligner trois points importants :

1. Le rapport d'échange financier concerne l'entente sur un montant, à partir du montant offert et du montant demandé ;
2. La domination capitaliste est basée sur le profit de l'exploitation. Il s'agit d'un système où « devenue une fin en soi, la richesse matérielle se développe en opposition à l'homme et à ses dépens » (Marx, 1867). De fait, il est bien connu que, dans un échange équitable (p. ex., quatre 25 cents contre un dollar), personne ne s'enrichit. Pour s'enrichir, il faut donner moins que la valeur réelle de l'objet. Conformément à cette idée, le succès de l'entreprise découle de sa capacité à amener autrui à accepter l'échange le plus inéquitable possible afin d'en retirer un maximum de profit. Or, cela n'est possible que si l'on s'arroge le droit de ne pas s'embarasser de ceux aux dépens desquels cet enrichissement se construit, d'où l'intérêt de l'entente impersonnelle ;
3. La fabrication du capital, son mode de production et d'appropriation, se présente comme un acte de violence.

Du point de vue idéologique et juridique, l'entente capitaliste provient essentiellement du processus d'expropriation des masses laborieuses de leurs propriétés, qui aboutit à un grand transfert de titres, d'une catégorie de personnes à une autre, en créant un fonds colossal pour

payer les emplois et le péculat. C'est ce qu'on appelle la concentration ou la centralisation du capital. Sur le plan économique, sa production repose sur l'instauration d'un nouveau mode de fabrication, qui génère un régime de domination, exigeant que les producteurs, jusque-là indépendants et travaillant pour eux-mêmes, perdent cette autonomie traditionnelle. Cette lutte contre l'autonomie procède directement du « problème de la liberté » ou du consentement volontaire. De fait, il va de soi qu'une personne indépendante, autosuffisante, n'a aucune raison d'accepter la soumission ou l'exploitation. Pour se résigner à de telles choses, il faut que le besoin l'y contraigne. Le capitalisme doit donc continuellement créer de la rareté et des besoins pour provoquer des demandes. Il entretient une situation de précarité et de dépendance, afin de maintenir l'être dans une situation de besoin. L'entente objective entre partenaires libres et égaux signifie en substance que : tout le monde est libre d'offrir ce qui lui chante, chacun serait toujours libre d'accepter ou de refuser, même lorsque sa survie en dépend ! La neutralisation donne lieu à une oppression perverse, car impure, mélangée, générée et en partie dissimulée derrière une apparence de légitimité, un déguisement inconsistant de l'ordre légal, qui cache maladroitement la réalité de relations sociales où seule prévaut la force brute.

C'est ainsi que le capitalisme transforme entièrement les combinaisons sociales et les procédés techniques du travail. Ce type d'organisation se distingue notamment par la création d'un « marché du travail » où la marchandise vendue est le travail. Cette création procède du fait que, dans la société marchande, d'un côté, les détenteurs du capital s'accaparent et monopolisent les ressources, dont l'accès est dorénavant prohibé par la privatisation. De l'autre, le seul moyen d'obtenir ce qui est devenu autrement inaccessible est la transaction monétaire. L'argent devient ainsi l'élément clé, indispensable, qu'il est obligatoire de posséder pour se procurer les ressources essentielles.

Alors qu'auparavant, l'artisan créait des objets qu'il vendait au marché, il est maintenant entièrement démuné, spolié de tout. Conséquemment, le prolétaire, qui doit, néanmoins, absolument trouver quelque chose à « vendre » pour acquérir de l'argent, n'a plus rien d'autre à offrir que « lui-même ». Il ne peut vendre que des périodes de servitude, soit le droit, momentané, d'être « utilisé », au bénéfice d'un tiers. C'est ainsi que le principe d'auto-instrumentalisation généralement associé à la prostitution (c.-à-d. le fait d'accepter de se vendre, de vendre ses

services et l'usage temporaire de son corps pour la jouissance d'autrui) se retrouve au cœur du système capitaliste.

3. L'HOMME-MACHINE ET L'ORGANISATION SCIENTIFIQUE DU TRAVAIL

Nous montrerons maintenant qu'avec l'organisation scientifique du travail (OST), l'objectivation de la force de travail de l'ouvrier – par rapport à l'ensemble de sa personnalité –, qui s'accomplit lors de sa vente comme marchandise, se transforme en réalité quotidienne durable. Nous verrons qu'en réduisant le travailleur, devenu simple exécutant, à sa force mécanique, l'OST dissout la frontière entre l'humain et la machine, et conduit le corps de l'être déqualifié à l'obsolescence.

Sur le plan pratique, la concentration du capital et la création d'un « marché du travail » donnent lieu à une concentration systématique des forces de travail. Avec elle se met en place une forme de gestion et d'organisation scientifique du travail, destinée à optimiser la productivité et les profits de l'entreprise. L'OST doit permettre d'identifier la façon la plus efficace de découper et de structurer le travail en vue de réduire « la flânerie », d'éliminer les « temps morts », et de déterminer la « meilleure et unique façon » de procéder. L'OST exige la clarification de deux dimensions complémentaires, l'une horizontale, l'autre verticale. La première concerne la décomposition du processus de production en une suite de tâches précises (qui seront respectivement confiées à un ouvrier « spécialisé ») et la délimitation précise des paramètres d'intervention de chaque opérateur. La seconde vise à établir une distinction stricte entre les tâches intellectuelles et celles dites d'exécutions. Conformément à cette idée, tout travail intellectuel doit être réservé à la direction, qui hérite du mandat, d'une part, de concevoir et d'organiser le travail, d'autre part, de former et de superviser les travailleurs. Le processus de rationalisation de l'entreprise (destiné à l'optimiser) va de pair avec sa mécanisation continue et le recours à des technologies toujours plus sophistiquées. Il commande, d'abord, de rompre avec les procédés classiques de production pour pouvoir calculer le processus du travail, puis de soumettre ces procédés à un examen minutieux en observant et en analysant les temps opératoires, grâce à une décomposition fine des gestes et à leur chronométrage, afin d'accroître leur efficacité. De fait, on ne peut parvenir à la rationalisation que par la décomposition la plus précise de chaque ensemble

complexe en ses éléments, par l'étude des lois partielles spécifiques de la production. En l'occurrence, pour chaque aspect de la tâche d'un travailleur, la direction doit : remplacer les méthodes artisanales par d'autres, développées scientifiquement ; standardiser les pratiques et faire en sorte que les hommes ainsi que l'ensemble des moyens requis s'y adaptent ; s'assurer qu'il soit exécuté suivant les principes et les techniques prédéterminées ; établir un processus rigoureux de sélection et de formation des travailleurs.

On assiste donc, avec l'OST, à une pluralité de ruptures contre-intuitives. Sont ainsi introduites dans le monde du travail : a) un morcellement extrême des opérations ; b) une séparation radicale entre ceux qui conçoivent et ceux qui produisent. Contrairement aux méthodes artisanales, l'OST réduit les praticiens à de simples exécutants dans d'immenses entreprises mécanisées. L'OST disloque la relation traditionnelle des producteurs au produit fini comme une réalisation entière, et privilégie le développement de compétences techniques n'exigeant pas une connaissance approfondie du processus d'ensemble de la part des exécutants. Elle transforme les tâches en une série d'opérations mécaniques précises, parcellarisées et abstraites, que le travailleur exécute machinalement, à l'instar de réflexes conditionnés, sans trop comprendre leurs fonctions. La productivité et le profit de l'entreprise augmentent au prix de tâches toujours plus redondantes, insensées et abrutissantes, qui n'offrent guère de défis et de fierté. Plus le travail est simplifié, répétitif et impersonnel, moins l'employé est coûteux et plus il devient agile, productif, anonyme et remplaçable.

La fragmentation du processus de production et la médiatisation continuelle des opérations par la technique accentuent constamment l'écart entre le producteur et le produit, rivent la conscience aux faits et entraînent un déséquilibre entre les sens, qui engendre une perte des repères et une désorientation. De fait, lorsque les effets des pratiques dépassent un certain degré de médiation, ils commencent à se brouiller aux yeux du praticien et à se soustraire tant à sa représentation qu'à sa perception. Plus ce degré augmente, plus l'appareil se complique, s'embrouille et moins l'être humain le comprend, de sorte que « chacun perd l'idée de l'objet à faire, par la division des travaux » (Alain, 1936, p. 33).

Parallèlement, plus le sentiment d'avoir prise sur ce processus et celui d'y participer diminuent, et plus l'homme se sent étranger face à un

système qui, sous l'effet de l'abstraction logique et de la technologisation, se déshumanise et s'autonomise. L'exécutant n'apparaît ni objectivement ni dans son comportement comme le véritable porteur de ce processus. Il est plutôt incorporé comme une partie mécanisée dans un grand système de la productivité mécanique, qu'il trouve devant lui, achevé, fonctionnant dans une totale indifférence par rapport à lui et aux lois auxquelles il doit se soumettre. L'absence de réciprocité produit un sentiment d'impuissance et une perte de sens. L'action de l'exécutant perd alors son caractère d'activité, et s'ensuit une attitude de résignation. Paradoxalement, l'être agissant est plongé dans une position de passivité devant un système qu'il ne contrôle pas et qui n'a rien à faire de lui, mais qui pourtant le contrôle et duquel il dépend. Conséquemment, avec la rationalisation, ce dont les hommes sont privés, c'est du « sentiment d'être un « agent » et de la possibilité de véritablement comprendre la portée réelle de ce qu'ils font. C'est ainsi que, chemin faisant, la conscience s'évade et que se dissout le principe de responsabilité.

Le recours à la technique, qui expose l'individu à des flux sensoriels et informationnels ininterrompus, accentue cette tendance. Les sens sont paradoxalement intensifiés et désarmés par la sollicitation permanente et le surplus d'information, ce qui produit des effets sur l'ensemble de l'être. La surexposition à un flux continu de sensations atténue les capacités psychiques (d'activité, de volonté, de choix, de réflexion, etc.) et entraîne une baisse de la sensibilité, du vouloir, de la perception et de l'impression des objets extérieurs. La convocation constante pousse à **se laisser porter par le mouvement machinal**, engendrant progressivement une sorte de mobilité, qui culmine dans la passivité et la soumission, voire l'effacement de soi. Plutôt que d'émaner du travailleur, cette activité incessante s'impose à lui de l'extérieur, l'active et l'épuise. En se décuplant, elle s'anesthésie. Les individus seraient alors plus enclins à accepter, subir, voire collaborer à des formes d'assujettissement plus sévères. Le mélange d'apathie et d'inattention tend à rendre les exécutants aveugles aux conséquences de leurs actes et désengagés (voir Haroche, 2009).

Avec l'OST, l'activité ouvrière est strictement considérée sous l'angle objectif de la mesure et de la mécanique du mouvement. L'analyse logique neutralise le facteur humain (la dimension qualitative des êtres et des choses), qui disparaît de l'équation, pour ne retenir que l'aspect quantifiable. La rationalisation du processus de fabrication et la distinction ferme entre les tâches intellectuelles et manuelles

ramènent l'humain sur le même palier que la machine. Les exécutants objectivés, homogénéisés, déshumanisés (c.-à-d. considérés sous l'angle de leur puissance productive) sont conçus comme du « matériel humain », « soumis à un pouvoir hautain et lointain qui ne voit en eux que moyen et matière » (Haroche, 2009, p. 14). En effet, dans le cadre de l'OST, l'ouvrier n'est pas là pour penser, mais pour exécuter à la lettre les gestes calculés et prévus pour lui. La subordination accentue le sentiment d'extériorité à l'égard d'une « mécanique, ajustée pour se passer de tout » (C. Haroche, 2009, p. 32-33), qui « se suffit à elle-même » (Haroche, 2009, p. 32-33) en évacuant l'espace de liberté, où s'exerce le pouvoir discrétionnaire. Ainsi, parce que la moindre part de liberté, qui conduit à réfléchir, ralentit le rythme de la production, la moindre liberté est pourchassée. En l'occurrence, le praticien est engagé dans une activité qui ne l'implique plus, qui est conçue de façon à neutraliser toute initiative ou possibilité de s'investir en caractérisant son action. L'accroissement de l'efficacité n'est possible que si l'on standardise ou dépersonnalise l'activité. La « puérité » devient l'attribut de l'exécutant, au sens où la division des tâches (intellectuelles et manuelles) et la complète prise en charge de l'ouvrier (qui paralyse son autonomie et le place dans un état de dépendance face à la direction) l'apparentent à une personne mineure sur le plan de l'état civil.

Cette approche tend naturellement à séparer l'âme du corps et l'intention de l'acte. De fait, au sein d'une relation dirigeant-exécutant, où l'action se réalise par procuration, ce dualisme prend corps puisqu'il renvoie à des êtres distincts. L'âme ou l'intention de celui qui ordonne sont alors transposées, greffées, sur un autre corps (celui de l'exécutant) qui en devient l'extension. C'est parce qu'il colonise des corps que le pouvoir a le bras long. Cette version du dualisme diffus neutralise le corps agissant pour lui imposer une volonté extérieure qui le dirige et à laquelle il doit se conformer. À cette occasion, l'homme n'est ni sollicité pour lui-même ni maître de lui-même. Il est la présence absente d'un corps désobjectivé. Il n'est plus l'incarnation irréductible de l'être agissant, il n'est plus « personne ». Il est tenu à l'écart d'un individu au statut toujours plus abstrait, devenu spectateur de sa vie. Il est l'expression du vouloir d'un tiers, protégé par la distance, qui n'est concerné par les « misères » de ce corps que d'un point de vue « comptable ». En ce sens, il s'agit moins d'un corps que d'une machine corporelle évidée. Le dualisme n'est plus d'ordre métaphysique, mais il décide du concret de l'existence. Cette situation d'investissement du corps alimente le sentiment d'étrangeté, de confusion et de dépersonnalisation, soit une

forme de dissociation désengagée. Déresponsabilisation et déposition de l'esprit discrétionnaire sont des phénomènes qui vont de pair.

4. L'OBSOLESCENCE DE L'HOMME

L'homme déqualifié est envisagé comme un mécanisme programmé et adapté aux opérations, au même titre que le reste de l'équipement, et ne s'en distingue que par la nature de son substrat. Dès lors, la frontière entre le monde des objets et celui des êtres humains s'efface, tout devient commutable en puissance puisque tout est envisagé sous le même angle. Dans la mesure où le travailleur est, « par nécessité, ramené au niveau de la force mécanique » (Alain, 1936), il s'agit moins de se contenter de sa forme, que de chercher à modifier ses assises pour accroître sa performance. On tente donc inlassablement d'augmenter le corps de l'exécutant pour renforcer son efficacité, de compléter, parfaire, accroître et renchérir technologiquement ses capacités physiologiques, comme une proposition toujours à affiner et à reprendre. Cette volonté exige de continuellement repenser et polir les moyens greffés au corps agissant, afin que ce dernier adhère aux projets nourris pour lui, et renvoie progressivement à une tâche impossible.

Cette approche engendre donc nécessairement un sentiment d'insuffisance face à un corps dont la seule dignité tient à son efficacité, laquelle s'impose à lui comme une norme dont les standards sont constamment redéfinis et repoussés par les technologies. De fil en aiguille, c'est l'humanité de l'être qui est discréditée en raison de l'intolérance grandissante face à ses limites naturelles et à ses défaillances potentielles. L'imperfection, les restrictions et la contingence, qui sont le propre de la vie humaine et les conditions de la liberté, font pâle figure devant le fantasme de l'illimitation technoscientifique.

L'objectivation de l'humain entraîne, du même coup et inversement, l'humanisation de la machine, et c'est ainsi qu'on rencontre, par exemple, des « téléphones intelligents ». Cette inversion s'accompagne d'un renversement radical de valeur, voulant que tout ce qui éloigne l'homme de la machine soit discrédité et que tout ce qui l'en rapproche soit valorisé. En situant la dignité humaine dans la performance et l'efficacité, on incite l'homme à imiter la machine et à expulser de lui-même ce qui l'en distingue et fait sa richesse. Or, l'humain est une créature qui n'est pas physiquement faite pour les impératifs croissants de rendement, d'efficacité, de vitesse, de compétition, etc., qui régissent

les sociétés occidentales. Que les impératifs capitalistes neutralisent, ou bestialisent, n'est guère étonnant. En effet, « l'utilité » domine le monde de la vie animal, car toute existence est conditionnelle à la satisfaction de ses besoins fondamentaux. Or, ce qui distingue l'humain de l'animal, c'est sa capacité à déborder le champ de la survie pour se tourner vers les activités biologiquement « inutiles », dans lesquelles s'exprime son humanité, tels les arts ou la culture. Il est ainsi permis de dire que « l'inutilité » est essentielle à l'homme, car c'est par elle que s'opère l'humanisation. L'ampleur culturelle des modifications technoscientifiques trahit ce diagnostic de l'insuffisance d'êtres appauvris dont on a neutralisé l'humanité pour ne retenir que la force mécanique, mais qui ne sont plus à la hauteur des machines qu'ils inventent. Les technosciences annoncent l'obsolescence du corps qui entrave l'expression de leur puissance.

En tentant de justifier et de légitimer rationnellement l'irrationnel, l'être humain se heurte à une série d'obstacles (eux-mêmes résultant de la logique de la domination), qui conduisent à la fabrication du non-contrôle. D'une part, contrairement à la légitimité transcendante, qui ne peut être mise en question, les justifications et les limites rationnelles ne constituent pas un frein efficace. Elles sont sans cesse repoussées et redéfinies, puisque le savoir progresse par la remise en cause et le dépassement des hypothèses précédentes. Dès lors, si elles parviennent à ralentir l'évolution d'un processus, elles n'arrivent cependant ni à l'enrayer ni à établir clairement des frontières à ne pas franchir, sous peine de modifier irrémédiablement l'essence humaine. D'autre part, si les idées sont impuissantes à changer le monde c'est parce que nous ne pouvons maîtriser que ce que nous connaissons et que l'insatiable volonté de puissance prend le pas sur la quête de la vérité. En effet, dans une démarche de connaissance, la domination est subordonnée à la vérité. L'important n'est pas d'avoir raison, mais de sortir de l'ignorance, au risque de voir nos convictions démenties, afin de se rapprocher toujours davantage de la vérité. Soulignons que l'opposition entre la démesure (*hubris*) et la mesure, ou la justice (*dikè*), était au centre des préoccupations des philosophes de la Grèce antique. Contrairement aux penseurs modernes, ces derniers considéraient l'illimité de manière négative et la limite de manière positive. Selon eux, c'est en luttant contre la démesure originaire (le chaos) que le monde se structure. En grec, l'illimité (*apeiron*) désigne aussi l'ignorance, tandis que la limite (*peiras*) définit les choses et fait surgir des figures du néant en traçant leurs contours. La raison est une mise en

forme et une maîtrise. Ce n'est que lorsque la pensée est gouvernée par des principes irrationnels qu'il est question de «démésure» et que la raison, qui cesse d'être raison, devient déraisonnable et ne parvient plus à donner la mesure. La médiation prend alors la forme du mythe (*muthos*). Les Grecs situent l'origine de la démesure dans la raison elle-même. Il s'agit d'une représentation imaginaire qui serait le fruit de l'intelligence. *L'hubris* s'enracine dans l'infamie, la superstition et l'intolérance. Pour eux, le mal, c'est la démesure à laquelle on n'aurait rien d'autre à opposer que la démesure elle-même. Insatisfait de son sort et avide de conquêtes, l'homme tragique, qui démonise le réel et insulte la nature, est victime de la démesure qui le détruit. Or, c'est seulement en quittant le domaine du mythe pour celui du logos, et en répondant à la démesure par la raison, que l'homme parvient à la pourfendre et à triompher des malheurs causés par la transgression des limites.

Le paradoxe de la volonté de puissance, qui n'a que faire de la vérité, est le suivant : c'est en cherchant désespérément à assouvir sa soif que la volonté de puissance échappe le pouvoir, et qu'elle fabrique du non-contrôle et de l'immaîtrisable. Le strict désir de domination fait écran à la justice, même calculatrice, qui fait elle-même écran à la sagesse. En faisant de la science, de la technique, de l'économie de marché et de la démocratie des complices du pouvoir, les idéologues et les acteurs politiques ont engendré la démesure, celle du vide de l'existence, de la massification sociale, de l'atomisation des individus, des formes procédurales, qui ont évacué tout contenu substantiel.

5. EN CONCLUSION

Notre réflexion a montré l'aspect hautement problématique de l'idéologie qui préside au développement des technologies et l'importance d'une prise de conscience sociale sur la direction que nous sommes en train de prendre quant à notre futur. La robotisation est un moyen qui, en soi, n'est menaçant que dans la mesure où il gagne en autonomie et sert une idéologie malsaine. C'est pourquoi les problèmes épineux qu'elle soulève doivent s'envisager en insistant sur leur caractère authentiquement politique, c'est-à-dire indissolublement historique et critique du réel. Afin de répondre aux défis nés du «tout technologique» et de nous assurer que la «robolution» ne s'opère pas contre l'humanité, mais avec elle et pour elle, nous considérons que les possibilités qu'elle offre doivent s'inscrire dans un projet de société.

Le point nodal de cette réflexion concerne donc la structure de la société contemporaine et la conception de l'humanité qui oriente son développement. Il semble, en effet, que l'unique moyen de comprendre la réalité sociale et d'agir sur elle est, d'abord, d'envisager l'être humain comme un être grégaire (collectif), actif, essentiellement lié à ses environnements, dont l'existence est de façon constitutive à la fois sociale et interdépendante. Contrairement à ce que suggère le capitalisme, cela suppose de reconnaître que l'espèce n'est pas biologiquement « programmée » pour évoluer en vase clos, que le nourrisson est entièrement vulnérable et dépendant et que le déploiement de l'individualité est lui-même le fruit d'un projet transindividuel (car l'être se développe à partir des interactions avec ses environnements et du bagage culturel qui est mis à sa disposition et avec lequel il va articuler ses capacités génétiques et neuropsychiques). Or, plus l'héritage culturel (qui constitue le terreau du développement personnel et façonnera son cerveau) est dense et fertile (héritage dont la fécondité dépend directement de la quantité et de la qualité des apports des membres de la collectivité), plus ses possibilités de développement personnel s'accroissent. En favorisant l'épanouissement de l'ensemble des individus, l'être favorise donc ses propres possibilités d'accomplissement. Ici ce n'est donc pas l'être qui existe, c'est l'humanité qui existe à travers lui.

Puis, il s'agit de repenser les avancées technoscientifiques sous l'angle de la nécessité (celle du monde socio-historique produit par et pour l'humanité), à partir de la praxis de l'être de besoin (qui s'enracine dans la lutte darwinienne pour l'existence et est conditionnée par la vie et la mort de l'organisme pour répondre aux conditions complexes de la préservation et de la reproduction de l'espèce).

Le but de la vie étant de vivre et de s'épanouir (pré-requis essentiel, conditionnel à toute chose et motif originel de la réunion des êtres en société), son assise fondamentale devient la biophilie. L'argument qui devrait guider cette réflexion devient alors plus écologique et que moral (c'est-à-dire qu'il n'est plus basé sur les normes du bien et du mal, qui peuvent être relatives et variables). Cette redéfinition axée sur l'existence humaine, devrait soumettre le mandat des technosciences à des considérations d'ordre vital selon les critères de l'autopréservation qui prennent la vie pour centre de réflexion, pour critère discriminatoire, sa protection puis, comme valeur suprême, son respect (celui de sa nature) et son épanouissement (conformément à la structure ontologique de son développement). En tant que consensus unificateur élémentaire,

la biophilie constitue une valeur valide difficilement récusable qui permet de présenter les enjeux de la robotisation dans les termes d'une théorie de la rationalité et de ses exigences en évitant les discussions byzantines sur le pluralisme des valeurs.

Sur la base de ces présupposés, il paraît inconséquent que la forme de l'organisation sociale (c'est-à-dire la manière dont la nature humaine trouve une forme d'expression concrète dans le monde social), destiné à répondre aux besoins de la vie et à favoriser son épanouissement, ne permette pas d'y parvenir, qu'elle prône l'individualisme et la compétition entre les besoins individuels et sociaux. Le capitalisme s'avère donc irrationnel, car, tel que nous l'avons vu, il s'agit d'un système :

- a) construit par l'humain mais pas pour l'humain, où la production n'est plus orientée vers la satisfaction des besoins vitaux, mais vers l'échange et la quête de profits (et qui pour cette raison parvient à ignorer l'humain) et où l'attention est concentrée sur les besoins et la croissance du marché (plutôt que sur les besoins de l'humain) ;
- b) qui fait abstraction des contenus qualitatifs spécifiques des êtres et des choses pour les remplacer par des déterminations quantitatives (comme le prix).

Bien qu'on puisse sembler à cent lieues de questions pratiques, les considérations qui précèdent visent pourtant un but qui l'est incontestablement. Il s'agit bien de restaurer une raison politique et une politique de la raison qui ouvre des questionnements sur l'universalité de l'humanité et sur son avenir. De fait, il ne tient qu'à nous de nous appartenir (c'est-à-dire de ne pas être les objets mais les sujets de la robotisation) et de mettre la robotisation au profit de l'humanité. Toutefois, nous n'y parviendrons qu'en nous appuyant sur une conception du genre humain qui se construit historiquement, dont les caractéristiques sont de plus en plus dignes d'elle-même, qui prend de plus en plus conscience d'elle-même, passant d'un rapport au monde « En soi » à une relation avec un monde « Pour soi », bref, une approche orientée vers la spécificité humaine et son épanouissement.

BIBLIOGRAPHIE

Alain (1936). *Mars ou la guerre jugée*. Collection idées, n° 178, Paris, Éditions Gallimard, 309 p.

- Boudon, Raymond, Cherkaoui, Mohamed *et al.* (2003). *Dictionnaire de sociologie*. Paris, Larousse.
- Carr Nicholas, 2014, *The Glass Cage. How our computers are changing us*, New York, Norton Compagny.
- Gunkel, David, 2012, *The machine question. Critical perspectives on AI, Robots, and Ethics*, Cambridge, MIT Press.
- Haroche, Claudine (2009). «L'absence d'interdit dans l'illimitation des sociétés contemporaines», *Droit et culture*, vol 57, n° 1.
- Marx, Karl (1867). «Résultats du procès de production immédiat», dans un chapitre inédit du *Capital*, section II : *La production capitaliste comme production de plus-value*. Consulté en ligne.
- Ménissier, Thierry (2008). «L'animalité comme limite et comme horizon pour la condition humaine selon Hannah Arendt», dans Jean-Luc Guichet (dir.), *Usages politiques de l'animalité*. Paris, L'Harmattan.

DEUXIÈME PARTIE

Les robots dans la société : enjeux éthiques et politiques

CHAPITRE 10

Les implications éthiques de la perte progressive du contrôle humain

John L. FINNEY

INTRODUCTION

Je crois que nous pouvons tous accepter la prémisse que nous sommes individuellement responsables d'un point de vue légal et éthique de nos actions, et que nous sommes imputables de nos actions commises envers quiconque ou envers quoi que ce soit. La responsabilité et l'imputabilité sont assez faciles à comprendre, lorsque la chose ou la personne en question se trouve à proximité. Mais que se passe-t-il lorsque la distance géographique qui nous sépare de l'objet de nos actions est plus importante? Qu'advient-il lorsque la relation éthique est interprétée par un dispositif technologique qui pourrait ne pas être complètement contrôlé par un humain?

Ce genre de situation soulève beaucoup de questions. En bref, comment pourrait-on déterminer l'effet de cet «intermédiaire» sur nos responsabilités éthiques et légales? Et, plus précisément, comment ces responsabilités pourraient-elles être transformées à travers la *complexité* de cette technologie intermédiaire, voire être remplacées par une *pleine autonomie* de ce système? Si nos responsabilités diminuent proportionnellement à la présence de cet intermédiaire, qui donc doit

assumer nos responsabilités ? Et comment ces situations transformées pourraient-elles influencer d'autres relations externes ?

Prenons comme exemple celui des conflits armés, où les systèmes robotiques sont de plus en plus présents. Ces derniers peuvent représenter de simples machines contrôlées à distance utilisées pour désamorcer des bombes ou des mines n'ayant pas explosé ; elles sembleraient tout à fait acceptables puisqu'elles sont défensives et ne mettent personne en danger. Mais il pourrait tout aussi bien s'agir de véhicules de terre, d'air ou de mer, pilotés à distance – éventuellement pleinement autonomes – et accompagnés d'armes « sublétales » et même létales.

1. CONVENTIONS INTERNATIONALES

Nous avons déjà des conventions en place qui gouvernent la façon dont on utilise les armes et qui sont acceptées au niveau international. Le droit international humanitaire (DIH) s'applique aux conflits armés, et le droit des droits de l'homme (DDH) s'applique aux autres situations. En ce qui concerne les nouvelles armes, l'article 36 du Protocole additionnel I aux Conventions de Genève exige des États de déterminer si l'emploi en serait interdit par ce protocole ou par toute autre règle du droit international. À celui-ci s'ajoutent la Déclaration universelle des droits de l'homme ainsi que la Charte des droits fondamentaux de l'Union européenne. Ces conventions sont-elles suffisantes pour réguler l'utilisation des armes robotisées, dont le principal « argument de vente » est la portée des actions, pouvant atteindre jusqu'à plusieurs milliers de kilomètres, au sein d'une zone de conflit ?

Si l'on considère la façon dont ces avancées technologiques ont remis en question la pertinence de ces conventions, il importe de souligner les points suivants. Premièrement, les Conventions de Genève ont été signées en 1949, bien avant l'ère des ordinateurs. Ceux-ci étaient encore à un stade très primitif. L'état des technologies de l'époque peut être représenté par le tourne-disque 45 tours, l'invention de l'année en 1949 ! Ensuite, si on passe directement aux Protocoles additionnels de 1977, à ce moment-là, les ordinateurs personnels n'en étaient qu'au stade initial. La capacité des technologies de l'information et des communications (TIC) était minime comparativement à aujourd'hui : la loi de Moore a été assez bien suivie depuis 1977, et les ordinateurs actuels sont probablement un million de fois plus puissants qu'à l'époque, si on se base sur la densité des transistors. Lorsqu'on analyse d'autres

facteurs ayant augmenté la capacité des machines, par exemple l'apprentissage automatique (*machine learning*), les TIC représentent aujourd'hui un tout autre monde. Quels sont donc les problèmes liés à ces nouvelles technologies, au regard du DIH et du DDH ?

Les principes de *distinction*, *proportionnalité* et *imputabilité* sont au cœur du DIH. Comment sont-ils évalués quand les actions des armes sont contrôlées à *distance* (1) ? Et comment le sont-ils quand les actions sont accomplies de façon *autonome* (2), c'est-à-dire quand l'élément humain est « hors circuit » et donc qu'il n'y a plus de véritable contrôle humain sur la machine ? Bien que ces systèmes ne semblent pas encore exister, il serait plus prudent de considérer les conséquences de leur développement futur afin d'anticiper les défis auxquels le DIH pourrait faire face. C'est ce que nous allons tenter de faire.

2. ARMES ROBOTISÉES CONTRÔLÉES À DISTANCE

Il n'est pas facile de *distinguer* un combattant d'un civil, dans un conflit, si on se trouve à distance. Avant de prendre une décision, le pilote d'un drone armé et un commandant sur le terrain disposent de données très différentes. Le passage des informations physiques aux bases de données implique une tout autre appréciation de la situation. L'augmentation de la complexité des bases de données avec lesquelles une décision ciblée doit être prise accroît la possibilité d'erreur d'identification. La qualité technique de l'image disponible et la qualité de l'intelligence numérique disponible pour le pilote sont deux éléments essentiels quant à la décision de faire feu. Faire voler un drone armé a été comparé à « faire voler un avion en regardant à travers une paille », ce qui a des conséquences évidentes quant aux limites de l'information géospatiale disponible lorsqu'on prend une décision de tir.

La distribution de la prise de décision au sein d'une équipe d'avion télépilote encourager « la pensée de groupe », c'est-à-dire que les équipes vont chercher activement des données justifiant le tir. Les systèmes de discrimination se basent sur l'objet cherché et l'identification à distance d'une cible est faite à partir de caractéristiques simplifiées. Le stéréotype approximatif de la cible représente principalement des jeunes militaires mâles. Un civil qui participerait directement aux combats peut être une cible légitime sous le DIH, mais un pilote à distance serait-il en mesure de prendre une décision fiable dans les circonstances et ne pas se tromper ?

Les délais inhérents des signaux de transmission entre l'opérateur et le véhicule, allant de 1.5 à 4 secondes, peuvent représenter un grand changement sur le terrain, entre la décision de tirer et le moment d'amorcer le déclic. La situation peut aussi changer très soudainement, et l'on pourrait tirer sur des non-combattants et les tuer. Certes, ce délai diminue avec l'évolution technologique, mais les limites physiques empêchent de réduire ce délai à zéro.

Sur le terrain, un combattant peut choisir de se rendre. Comment peut-il se rendre lorsqu'il fait face à un véhicule piloté à distance ?

Évaluer la *proportionnalité* dans le cas d'un véhicule piloté à distance devient encore plus problématique, car celui-ci cause des effets qui vont au-delà des victimes civiles. Il faudrait pouvoir évaluer les dommages collatéraux, mais ceci est difficile en raison de la distance. Bien que la précision des attaques de drones permette de minimiser les dommages à la propriété, il est difficile d'évaluer à distance les dommages psychologiques et les pénuries liées à la survie.

La recherche effectuée sur l'effet des attaques de drones au Yemen¹ illustre certains de ces problèmes en termes de proportionnalité. Soixante-et-onze pour cent des personnes interviewées souffraient de stress posttraumatique et de conditions psychologiques anormales sévères, telles la dépression, les réactions de panique, les réactions hystériques-somatiques, les réactions de peur et des réactions pathologiques de deuil. L'impact sur les enfants était particulièrement préoccupant : ils souffraient de problèmes d'attachement et présentaient de graves symptômes liés au bruit, des manques de concentration et de la perte d'intérêt pour des activités agréables. La fréquentation scolaire avait aussi chuté.

En ce qui concerne les impacts socioéconomiques, près de 50 % de la population était privée de sécurité alimentaire, de nombreuses familles n'avaient plus ni chef familial, ni accès aux services de santé de base, ni accès à l'éducation (les écoles étant fermées) ou aux funérailles à cause de la peur de représailles additionnelles. Quarante mille résidents, au sein d'une seule province, auraient abandonné leur maison en raison de leur crainte des attaques de drones.

Sur le plan légal, ces « dommages collatéraux » pourraient être jugés comme des « punitions collectives », lesquelles vont à l'encontre de

1. <http://www.channel4.com/news/drone-attacks-traumatising-a-generation-of-children>

l'interdiction de représailles. L'impact psychologique des attaques de drones sur les populations civiles est-il suffisamment reconnu par le DIH, ou est-il à peine pris en compte ? Que penser des impacts socio-économiques et des perturbations sociales qui n'ont apparemment aucune limite ?

Le contrecoup lié à l'utilisation d'armes contrôlées à distance est souvent compris par la radicalisation d'acteurs non étatiques. Un chercheur du Royal United Services Institute for Defense and Security Studies² a récemment dit :

L'utilisation de systèmes robotisés pourrait plaire aux politiciens, car il y aurait moins de retours de sacs mortuaires... Pour moi, la guerre est une activité humaine à la base, mais si l'on fait la guerre télécommandée à un autre groupe ou pays, ils finiront par trouver un moyen de nous atteindre chez nous, car c'est le seul moyen de répliquer.

L'utilisation d'armes commandées à distance pourrait entraîner une intensification du terrorisme hors des zones de combat ou des cyberattaques non imputables à l'État. Cette utilisation doit donc être examinée soigneusement.

En termes d'*imputabilité*, lorsque les forces armées régulières utilisent des avions téléguidés pour effectuer le type d'opérations qu'elles ont toujours fait, la position juridique du DIH semble peu problématique, car les troupes ont reçu l'entraînement au DIH et elles font partie d'une chaîne de commandement militaire.

Quoi qu'il en soit, l'utilisation de ce genre de machines téléguidées par les organisations de services secrets, ou les compagnies privées engagées par les gouvernements, demeure contestable sur le plan juridique. La CIA, aux États-Unis, dirige régulièrement des attaques de drones en dehors des zones de combat conventionnel, comme au Pakistan, en Somalie et au Yémen. Il n'est pas clair si ces attaques doivent être soumises au droit international humanitaire, comme proposé par les États-Unis, ou plutôt au droit des droits de l'homme. La légalité des attaques pourrait varier selon les cas. Dans un rapport aux Nations Unies rédigé par le Rapporteur spécial sur les exécutions extrajudiciaires sommaires ou arbitraires, au mois de mai 2010³, il a été suggéré

2. Elizabeth Quintana, in : <http://www.bbc.co.uk/news/business-41035201>

3. <http://www2.ohchr.org/english/bodies/hrcouncil/docs/14session/A.HRC.14.24.Add6.pdf>

que les exécutions effectuées par les drones de la CIA violaient les DHH et le DIH parce qu'elles n'étaient pas conformes aux exigences de transparence des procédures et de protection qui assurent la légalité et la justification des exécutions : « le manque de transparence donne aux États une licence d'exécution virtuelle inadmissible ». Certains des arguments du rapporteur concernent la notion du « droit de légitime défense » et de la légalité des attaques de drones sous l'article 51.

Notons également qu'un opérateur non militaire ne serait pas protégé par son statut de combattant dans son propre pays ; il représente donc une cible possible sur le plan juridique.

La *transparence* devient nécessaire pour établir les responsabilités. Dans le contexte de la guerre à distance, les DIH et DDH offrent déjà leur protection aux civils, mais le manque de transparence implique que les caractéristiques des contextes d'attaques de drones armés peuvent être contestées. De plus, l'évaluation de l'impact des attaques de drones et de leur légalité demeure difficile à déterminer *a posteriori*.

L'absence de transparence est caractéristique du déploiement de véhicules aériens armés. Un grand nombre d'activités au sein des États reste secret et l'information n'est souvent ni disponible au public ni même aux gouvernements démocratiques. La sécurité nationale est invoquée comme raison pour défendre la rétention d'informations. La rationalité et les raisons qui prouveraient la légalité de leur utilisation ne sont pas dévoilées par les gouvernements. Ceci soulève, inévitablement, des questions concernant l'identité des responsables des prises de décisions, ainsi que les raisons justifiant leur manque de transparence. Or, le manque de transparence fait en sorte que l'imputabilité reste vague et il entrave les enquêtes d'abus en regard du DDH.

La transparence est également absente de l'évaluation des armes robotisées en regard de l'article 36 (évoqué dans la section 2 ci-dessus). Bien que ces analyses aient été effectuées, leurs résultats ne sont guère disponibles. Alors que l'Article 36 exige que soit démontré que les armes *sont capables de répondre aux exigences* du DIH, il n'y a aucune transparence en ce qui concerne leur *utilisation*.

La question de la légalité du *consentement* au déploiement des véhicules armés robotisés ne semblerait pas différente de celle du déploiement d'autres moyens militaires. Certains ont argumenté en faveur de la légalité de leur utilisation dans des États qui ne peuvent se défendre face à un danger imminent ; cependant, cette menace doit être directe pour l'État qui prend l'initiative de l'attaque. Des enjeux

concernant la nature de ce consentement sont présents. Par exemple, le consentement doit-il être explicite pour être considéré comme légal? Qui est-ce qui a le pouvoir de donner ce consentement? Peut-il être révoqué? Comment peut-on savoir si le consentement est réel ou qu'il n'a pas été obtenu par la contrainte? Bref, l'absence générale de transparence concernant le déploiement des armes commandées à distance peut compliquer le problème du consentement.

Alors qu'elle pourrait être contreproductive par le « recrutement » d'insurgés additionnels, l'utilisation de drones armés dans des pays tels que le Pakistan, la Somalie et le Yémen pourrait compromettre vivement l'évolution et l'application des lois, ainsi que la capacité de gouvernance dans ces régions.

Il existe un grand nombre de problèmes ayant des implications éthiques:

- Le basculement des notions de légitime défense et de menace. Est-ce un comportement éthique que de tuer si l'on n'est pas menacé personnellement?
- Les conséquences psychologiques pour ceux qui vivent dans un climat de menace et d'attaques.
- Les conséquences psychologiques pour les opérateurs, leur famille et leur société.
- La fragilisation des valeurs traditionnelles d'honneur et de courage militaires.

Enfin, il existe aussi des problèmes sérieux quant à la sécurité de ces systèmes. Comme toute machine, elle peut avoir des défaillances mécaniques (internes ou causées par des phénomènes externes), qu'elle soit contrôlée à distance ou non, qui peuvent causer des morts et d'autres dommages. Contrairement à d'autres systèmes, les systèmes télécommandés peuvent éventuellement être piratés; l'identité de l'utilisateur peut être usurpée par l'adversaire, ce qui peut même permettre à ce dernier de « détourner » le système télécommandé contre l'utilisateur.

3. ARMES AUTONOMES

Les mêmes problèmes de *distinction*, de *proportionnalité*, d'*imputabilité* et de *transparence* des véhicules armés télécommandés se retrouvent

dans le système d'armes autonomes. Le fait de retirer complètement l'être humain de la boucle décisionnelle amplifie plusieurs phénomènes et soulève des questions éthiques et légales additionnelles. Plusieurs de ces problèmes sont reliés à la capacité des logiciels de prendre des décisions normalement prises par des êtres humains, notamment celui du choix de vie ou de mort, et celui de la moralité de déléguer ce choix à une machine.

En ce qui concerne le problème de la *distinction*, les systèmes de discrimination sont nécessairement basés sur les paramètres d'une recherche donnée. Si ces paramètres sont bien définis, il se peut qu'un logiciel de reconnaissance d'images sophistiqué puisse, en effet, identifier un objet correctement, mais il ne pourra pas agir de façon fiable en dehors de sa stricte définition. Par exemple, lorsqu'un combattant obtient une nouvelle série de droits – c'est-à-dire comme prisonnier de guerre ou comme soldat blessé – comment le logiciel peut-il décider que ces nouveaux droits ont été acquis ? Il y a aussi toute une série de directives du Comité international de la Croix-Rouge qui concernent les civils directement impliqués dans les combats, et qui doivent être interprétées par des êtres humains. Comment un logiciel peut-il distinguer un système de défense d'un système d'attaque ? Bien que les systèmes avancés de reconnaissance d'images pourront être capables dans le futur de reconnaître un homme armé, seront-ils capables de faire la distinction entre un combattant et un policier civil ? Le problème de la capitulation d'un ennemi face à un robot télécommandé a déjà été examiné plus haut. Dans le contexte de systèmes autonomes, puisqu'il n'y a aucun critère précis quant à la méthode de capitulation, une réponse à celle-ci ne peut, par définition, être programmée. Une tentative de reddition *exige* donc, encore par définition, la participation du jugement humain.

Le critère de la *proportionnalité* pose encore plus de défis pour un système autonome. Une machine peut-elle évaluer la nécessité militaire qui exige de prendre en considération toute une gamme d'enjeux dont un grand nombre sont impossibles à quantifier ? Un système numérique peut-il évaluer et infliger le degré de dommage nécessaire pour riposter face à une situation particulière ?

Comme discuté plus haut dans le contexte des systèmes téléguidés, l'évaluation du critère de proportionnalité ne consiste pas seulement en une estimation des dommages collatéraux dans la zone de destruction d'un projectile. Une machine autonome est-elle en mesure

d'évaluer le nombre de morts collatéraux qui serait justifié? Comment pourrait-elle évaluer la perte de la vie humaine liée à la destruction des infrastructures ou la perte de moyens assurant la survie, ou même considérer dans cette équation toutes les autres conséquences sociales et psychologiques mentionnées précédemment dans la section 3? Pour évaluer la proportionnalité, il faut être capable de jugement, du type de jugement attendu d'un « commandant militaire raisonnable ». Ceci est un art, et non une science, qui nécessite une conscience humaine capable d'apprécier le contexte. On ne peut que conclure à la nécessité du jugement humain dans ces circonstances, car il ne peut être délégué sur le plan éthique à un système autonome.

Un des enjeux principaux de la robotique, quant aux éléments mentionnés ci-haut, concerne le fonctionnement de ce type d'équipements dans des environnements « ouverts » plutôt que « structurés ». Un robot est conçu pour accomplir une ou plusieurs fonctions à l'intérieur d'un environnement déterminé (« structuré »), et il fonctionne bien quand l'environnement lui est adapté. La machine devient imprévisible quand elle fonctionne dans un environnement différent et plus ouvert. Comme nous l'avons constaté, même un logiciel sophistiqué de reconnaissance d'images qui fait face à de petits changements dans l'environnement (telle une journée ensoleillée) peut avoir un fonctionnement défectueux. Puisqu'il n'est pas possible de programmer dans un logiciel tous les environnements possibles ou toutes les probabilités, le comportement du robot en dehors de sa zone de confort pourrait être catastrophique.

Les conséquences éventuelles d'une réaction imprévisible deviennent encore plus inquiétantes lorsque l'on considère deux groupes de véhicules autonomes programmés selon des règles de fonctionnement différents et s'ignorant mutuellement. Un krach de la Bourse causé par l'imprévisibilité de l'interaction de systèmes informatiques complexes est déjà assez inquiétant; une défaillance imprévue qui implique des machines mortelles le serait encore plus.

Dans ce contexte, il faut aussi considérer que des systèmes relativement simples peuvent avoir des défaillances et que l'intervention humaine peut alors être requise pour résoudre ces problèmes techniques (un exemple évident est celui du papier bloqué dans une imprimante). Les systèmes plus complexes peuvent devenir imprévisibles lorsque, par exemple, une carte de circuit imprimé est endommagée par une balle, ou qu'un Cheval de Troie s'est introduit dans leur système pendant la fabrication. Les conséquences peuvent être catastrophiques sur le

champ de bataille et entraîner des événements imprévus en cascade dont la résultante serait une escalade militaire sans intervention humaine.

Les conséquences de l'usurpation de l'identité de l'utilisateur et de la prise de contrôle des machines autonomes seraient encore plus problématiques que pour celle des véhicules téléguidés.

4. L'AUTORISATION DE TUER

Pour tuer légalement, une machine doit pouvoir agir selon les règlements du DIH. On a beaucoup parlé de la possibilité de créer un «gouverneur éthique» programmé pour suivre ces règlements; mais pour que cela soit possible, il faudrait réduire ces règlements en des termes très simples et adéquats pour programmer un robot. Le problème rencontré ici est que le DIH comprend des *directives* qui tentent d'engager des processus de raisonnement. Il serait donc non seulement inapproprié, mais aussi dangereux d'essayer de les réduire ou de les simplifier.

Un autre problème important concerne l'autorisation de tuer ou la délégation de cette autorisation. Or, les DIH et DDH exigent que la décision soit prise par un être humain. Une des prémisses sur laquelle se basent la plupart des codes juridiques et moraux est que la décision d'ôter la vie doit être prise par un être humain. Par exemple, la Convention de La Haye exige qu'un combattant «soit commandé par une personne». La clause de Martens, qui est une règle ancienne et fondamentale du DIH, requiert l'application du «principe d'humanité» dans le combat armé. C'est un être humain qui peut passer outre mon droit à la vie, et ce droit ne peut être laissé au calcul d'une machine. Même s'il était techniquement possible de programmer un robot selon les règlements du DIH, il ne serait pas acceptable de lui déléguer cette décision sans faire des changements juridiques importants.

Déléguer à une machine l'autorisation de tuer, même si sa programmation est très complexe, peut être considérée comme un déni d'un élément essentiel de la valeur intrinsèque de la personne humaine, soit la dignité humaine. La dignité humaine est fragile et doit être respectée et protégée, car elle est le fondement éthique et juridique de la Déclaration universelle des droits de l'homme, de la Charte des droits fondamentaux de l'UE et de la Constitution de l'Allemagne. La

Convention des DDH devrait pouvoir protéger la dignité humaine, mais elle reste faible à cet égard et de nouvelles prémisses pour la renforcer devraient être établies.

5. IMPLICATIONS STRATÉGIQUES

Les armes robotisées ont aussi des *implications stratégiques* capables de changer les règles du jeu. Leur utilisation diminue la barrière d'activation des conflits armés de façon importante, ce qui atténue l'importance et l'exigence du *jus ad bellum*, et déplace ainsi l'évaluation de la nécessité militaire. Les exécutions ciblées ne sont pas récentes, mais les systèmes robotisés armés les rendent maintenant plus faciles et plus fréquentes. Nous faisons face au danger d'un tournant stratégique majeur, selon lequel nous pourrions conduire des «guerres à l'abri de tout risque», de façon continue et à long terme. Nous pouvons commencer à imaginer un monde où la guerre de faible intensité serait la règle. Basé sur des exécutions ciblées, ce monde serait très différent, puisque la possibilité d'exécutions continues remplacerait le développement à long terme d'une coexistence stable. Une fois ce seuil franchi, le retour en arrière semble bien difficile.

Une autre implication stratégique de l'utilisation des armes robotisées concerne le fait qu'elle risque de miner l'équilibre stratégique et de compliquer les avancées positives vers la réduction des armes nucléaires.

6. RESPONSABILITÉS PERSONNELLES

L'état de développement des logiciels a été grandement exagéré par certains opérateurs qui affirment que certains systèmes disponibles ont atteint un tel degré de perfection qu'ils peuvent accomplir des tâches complexes dans des environnements complexes avec une fiabilité incontestable. Ces affirmations peuvent influencer les décideurs et faire aller au-delà des zones justifiées par l'état d'avancement ou le potentiel réaliste de la technologie. Il est donc important que les scientifiques développent un code d'éthique qui ne s'applique pas seulement à l'usage potentiel de leur travail, mais qui s'étend à la façon dont ils communiquent avec le public à propos de leur travail et de ses potentialités.

Une personne peut devoir faire face à de décisions éthiques impossibles à prendre qui sont imposées par le design du système dans lequel elle travaille – l'expérience classique et bien aimée des philosophes, le dilemme moral du Trolley, illustre bien ce point. Il est donc important de concevoir l'imputabilité en fonction de ceux qui utilisent les systèmes, mais aussi en fonction de ceux qui les subissent. Il faut donc s'assurer de l'inclusion des normes juridiques et éthiques pour permettre aux ingénieurs de «concevoir de façon responsable» (*design with responsibility*), et de travailler à une «conception qui tienne compte des valeurs» (*value sensitive design*).

CONCLUSION

Pour évaluer si et comment ces systèmes d'armes sont à la hauteur des exigences du DIH, nous serions bien avisés de :

- contrôler la correspondance entre les applications de ces nouvelles technologies et les lois et coutumes de la guerre ;
- essayer de comprendre comment ces lois et coutumes peuvent ou non être suivies, en fonction des développements technologiques actuels et futurs ;
- établir de façon urgente un ensemble de principes éthiques liés au développement et à l'utilisation de la robotique moderne dans le cadre de la guerre.

Plutôt que de dissiper «le brouillard de la guerre», le développement de systèmes armés robotisés l'a épaissi. Nous avons développé des systèmes complexes qui opèrent dans un nouveau cadre épistémique. Il est adéquat et essentiel d'explorer leurs conséquences éthiques et juridiques. Des machines à tuer, complètement autonomes, n'existent pas encore. Cependant, nous devons affronter ces enjeux aux stades initiaux de leur éventuel développement si nous ne voulons pas que notre capacité d'utiliser les systèmes robotiques de façon éthique se retrouve dépassée par les événements.

Traduction : Carroll Guérin-Lafèche

Révision : Marie-Hélène Parizeau

CHAPITRE 11

Le Kamikaze et le drone

Éthique de l'assassinat

Soheil KASH

Si pour Hobbes (1990 [1668]) l'idée de la « guerre de tous contre tous » est un danger à éviter, car elle fait basculer la société à l'état de nature (Hobbes, 2000 [1651]), son invention la plus originale et la plus moderne fut de créer un monstre d'État qui serait capable, en monopolisant la violence, d'obtenir l'obéissance absolue des citoyens tout en assurant leur protection de toute guerre éventuelle. Le paradoxe de son invention est de fonder la modernité de son État sur une image mythique empruntée au Livre de Job de l'Ancien Testament : le couple, dont le premier (Léviathan) représente l'ordre et la sécurité et s'aliène toutes les volontés individuelles, et le second (Béhémoth), les forces du chaos de la guerre de tous contre tous.

Cette image mythique comporte plusieurs connotations : la première signification (le « Grand Homme ») affiche sa suprématie dès le titre du livre ; la deuxième, celle du « Grand Animal », découle du titre du livre (Léviathan) et évoque l'animal marin du Livre de Job ; la troisième signification est celle d'une énorme « machine » qui inspire la peur par son pouvoir de réprimer et de tuer ; la quatrième évoque un « contrat social » qui consiste à soumettre la liberté des citoyens à ce « Monstre

d'État» qui donne naissance à une personne souveraine représentative qui a tous les qualificatifs d'un «Dieu mortel».

Dans sa lecture du «Léviathan» de Hobbes, Carl Schmitt (2002 [1938]) constate qu'avec ces quatre significations de l'État moderne, on atteint une totalité mythique comprenant un dieu non transcendantal, l'homme, l'animal et la machine. Toutefois, il ne reste de ce «dieu mortel» que la machine qui a échoué à surmonter la crainte de la guerre de tous contre tous dans l'état chaotique de la nature; d'autant plus que ce Léviathan, avec son État absolu à l'intérieur de la société civile, n'a pu freiner la montée néolibérale qui oppose au maximum d'État à la Hobbes un minimum d'État à la Locke. Elle n'a pu, non plus, mettre de l'ordre dans le chaos qui caractérise les relations internationales entre les grandes puissances; les organisations internationales ne font que refléter le rapport de forces entre ces mêmes puissances à l'aune de leur capacité à faire la guerre.

Pour Carl Schmitt (1972 [1932]), au commencement de la politique, il y a la guerre. Le principe de la politique consiste donc pour lui à affronter la mort contre un ennemi qui est toujours déjà là. En d'autres termes, la notion de la politique qui consiste à déterminer l'ennemi et l'ami est un acte de guerre en dehors duquel toute politique de souveraineté est un mot vide (Schmitt, 2002 [1938]).

Avec la suprématie de la guerre telle que formulée par Carl Schmitt et reprise par Michel Foucault (1997), on peut effectivement renverser la conception de Clausewitz en disant que ce n'est pas la guerre qui est le prolongement de la politique par d'autres moyens, mais c'est plutôt la politique qui n'est qu'une activité guerrière au départ, exactement comme Hobbes (1990 [1668]) le craignait et comme Hegel l'avait déjà pressenti dans «la phénoménologie de l'esprit». Avec la dialectique du maître et de l'esclave, Hegel (1941 [1807]) n'a fait que développer la confrontation entre ces deux protagonistes, confrontation guerrière qui lui semblait inévitable comme jeu avec la mort.

Ce renversement de la maxime de Clausewitz (1955 [1832]) peut-il nous aider à voir dans quelles mesures les technologies de la guerre (y compris les nanotechnologies, l'intelligence artificielle et les drones militaires en vogue actuellement) constituent la base qui détermine le rapport de forces entre les nations et qui génère les écarts et l'inégalité entre elles?

Dans quelles mesures ce principe (la politique est un acte de guerre) peut-il nous aider à analyser, non seulement l'ordre international tel

qu'il fonctionne à la métropole, mais aussi le rapport de forces entre la métropole et ses périphéries régionales non modernes ?

J'essaie dans ce texte de susciter une réflexion philosophique sur les différentes formes de guerre et de voir quelles sont les limites éthiques que les guerriers ont inventées pour gérer leurs façons de s'entretuer ou de faire de la politique. Il s'agit donc dans ce texte de deux armes non classiques : l'attentat-suicide que j'ai appelé opération kamikaze qui est différente du phénomène kamikaze japonais de la fin de la Seconde Guerre mondiale ; ainsi que le drone militaire qui est censé relever le défi terroriste du kamikaze moderne qui est devenu en Occident synonyme de toutes les formes du fondamentalisme islamique.

Si l'on admet donc avec Schmitt et Foucault que la politique est d'abord un acte de guerre, les hommes se sont toujours employés à mettre des limites éthiques à leur façon de s'entretuer. Le courage et l'honneur étaient des valeurs prémodernes qui méritaient d'être défendues au prix de sa vie dans un monde technique où la guerre impliquait l'engagement personnel des adversaires pour défendre leur honneur. D'ailleurs, toute la théorie de la Ghalaba (vaincre), développée par Ibn Khaldoun dans ses Prolégomènes, est fondée sur une *açabya* (esprit de corps) capable par sa seule combattivité physique de fonder un État en risquant sa vie dans une guerre sans merci.

C'est la disparition de ce monde prémoderne que déplore le Don Quichotte de Cervantès avec les progrès de la technique guerrière moderne ; les temps modernes sont venus remplacer le courage du chevalier par la supériorité technique d'un nouvel adversaire, si lâche soit-il, muni d'un pistolet ou d'un canon pour battre le chevalier le plus courageux et le plus aguerri.

En effet, l'histoire de l'Occident moderne dans son rapport à son extérieur non moderne ne fait que retracer les moments décisifs qui ont caractérisé les technologies de la guerre : si la venue de la modernité déplorée par Don Quichotte a remplacé l'épée du brave chevalier par le pistolet et le canon, elle n'a pas fait basculer pour autant les règles de s'entretuer ; mais cette modernité a plutôt ouvert la voie au non-moderne pour rattraper son retard technologique et garder sa place dans un combat, si inégal soit-il. À la guerre classique de Clausewitz, le non-moderne a pu opposer la guerre révolutionnaire de la guérilla qui a caractérisé les mouvements de libération nationale après la Seconde Guerre mondiale. Sur un autre registre, toute la littérature

du Western américain du XIX^e siècle est régie par ce code d'honneur de pouvoir tuer pour se défendre.

Ce n'est qu'après la Seconde Guerre mondiale qu'on assiste à la naissance de ce que Jean-François Lyotard appelle un Occident nord-américain postmoderne. Dans sa description de cette postmodernité, il retient trois phénomènes qui font la démarcation avec notre modernité : en premier lieu figure la bombe atomique qui est venue bousculer toute notre conception d'une guerre classique, si sophistiquée soit-elle ; il s'agit d'une nouvelle arme qui est censée dissuader l'adversaire de ne plus faire la guerre ; Hiroshima n'était qu'un argument tragique pour éviter toute tentative de guerre. Mais, depuis Hiroshima, les guerres n'ont fait que se multiplier à travers le monde. Les autres phénomènes associés à la postmodernité sont : la révolution cybernétique et le changement du rapport entre la science et la technique. Lyotard ne pouvait pas prévoir dans les années 1970 la nouvelle arme (le drone militaire) qui annonce l'ère de tuer sans faire la guerre, et qui était conçue, au départ, comme réponse à l'attentat-suicide dit « terroriste » depuis le 11 septembre 2001.

En effet, la suprématie absolue de la bombe atomique occidentale n'a pas empêché les Vietnamiens de déjouer cette suprématie en cultivant les anciennes vertus guerrières, et en allant donc jusqu'à sacrifier leur vie dans une guerre révolutionnaire de libération qui n'excluait pas les attentats-suicides contre des objectifs militaires au départ. C'est justement, de l'échec de la guerre classique menée par les États-Unis contre la guerre révolutionnaire vietnamienne que va se développer une logique robotique de guerre à distance ; cette orientation vers l'application de la guerre à distance va commencer à prendre sa place dans l'armée américaine après le 11 septembre 2001, et connaître un regain d'intérêt sous le mandat d'Obama qui a décidé de retirer son armée de l'Afghanistan et de l'Irak, tout en intensifiant le recours aux drones comme nouvelle arme antiterroriste.

Une des meilleures approches des drones militaires et de la guerre à distance est faite par le chercheur français Grégoire Chamayou dans son livre *Théorie du drone* (2013). Dans ce livre, Chamayou fait une critique éthico-juridique des partisans de cette arme ; il se réfère à Walter Benjamin pour faire la comparaison entre deux logiques d'une même dialectique : celle du Kamikaze (terroriste) face à l'avion sans pilote dirigé à distance. Mais il a omis de dire que, même inscrits dans la dialectique du maître et de l'esclave de Hegel, ces deux protagonistes

viennent remettre en question les fondements mêmes de cette dialectique, à savoir l'attitude vis-à-vis de la mort qui est le principe même de la définition et du maître et de l'esclave.

En effet, Hegel (1941 [1807]) part justement du jeu avec la mort pour dégager sa définition du maître comme étant celui qui ne craint pas la mort et qui est prêt à risquer sa vie pour imposer sa volonté à un esclave qui a fondamentalement peur d'affronter la force de l'adversaire pour préserver sa vie. Entre deux êtres humains s'impose la guerre qui va différencier le vainqueur du vaincu, le maître de l'esclave : le maître est la conscience autonome, pour laquelle l'être-pour-soi est la réalité essentielle. L'autre est la conscience dépendante, pour laquelle la réalité essentielle est la survie animale. Celle-là est le Maître, celle-ci l'Esclave (Kojève, 1947).

Cet Esclave est l'adversaire vaincu, qui n'est pas allé jusqu'au bout dans le risque de sa vie, et qui n'a pas adopté le principe des Maîtres : vaincre ou mourir. Il a accepté la vie accordée par un autre. Il dépend donc de cet autre. Il a préféré l'esclavage à la mort, et c'est pourquoi, en restant en vie, il vit en esclave (Hegel, 1941 [1807]). Dans cette lutte à mort qui oppose le Maître à l'Esclave, Hegel insiste sur le fait que le Maître vainqueur ne supprime son adversaire que dialectiquement (supprimer dialectiquement veut dire ici : supprimer en conservant le supprimé), car il ne sert à rien au Maître vainqueur de tuer son adversaire physiquement, étant donné qu'il cherche plutôt à l'asservir et à être reconnu par lui (Kojève, 1947).

Or, avec le couple Kamikaze/drone, c'est toute la logique de Hegel qui se présente à l'envers : du côté kamikaze c'est l'engagement intégral, alors que du côté de l'opérateur du drone c'est le désengagement total ; alors que le kamikaze implique la fusion complète du corps du combattant avec son arme, le drone assure leur séparation radicale. L'attentat-suicide implique la mort du kamikaze, alors que le drone l'exclut de façon absolue, et l'opérateur du drone n'a plus besoin d'avoir le profil d'un combattant prêt à risquer sa vie sur le terrain, car il n'a qu'à appuyer sur des boutons dans une salle climatisée du Nevada pour faire disparaître sa cible en Afghanistan ou au Yémen ; sa philosophie est devenue le contraire du maître chez Hegel qui cherche la mort pour imposer sa volonté ; plus que ça, il ne cherche plus, comme dans la dialectique de Hegel, à sauvegarder la vie de l'esclave parce qu'il a besoin de sa reconnaissance, il cherche plutôt à l'anéantir. Entre le kamikaze et le drone, les définitions mêmes de maître et d'esclave sont piégées :

le maître de la haute technologie fuit la mort pour faire au kamikaze une chasse à l'homme à distance qui lui épargne l'affrontement guerrier comme au beau vieux temps des guerres classiques.

À ce constat fondamental s'ajoutent les arguments éthico-juridiques que je partage avec Chamayou en ce qui concerne l'arme du drone.

Avec les drones militaires, on assiste à une technique guerrière où l'adversaire le mieux loti se cache loin de la scène du combat et délègue sa supériorité à une machine d'assassinat ciblé. Cette nouvelle façon de mener la guerre nous impose une nouvelle éthique, celle de l'assassinat secret d'individus opéré de façon impersonnelle.

Le drone militaire prolonge et radicalise les procédés existants de la guerre à distance, aboutissant à supprimer le combat. Mais par-là, c'est la notion même de la « guerre » qui entre en crise. Un problème central se pose alors : si la « guerre des drones » n'est plus exactement la guerre, à quel « état de violence » correspond-elle ?

Cette tentative d'éradication de toute réciprocité dans l'exposition à la violence dans l'hostilité reconfigure non seulement la conduite matérielle de la violence armée, techniquement, tactiquement, psychologiquement, mais aussi les principes traditionnels d'un ethos militaire officiellement fondé sur la bravoure et l'esprit de sacrifice. À l'aune des catégories classiques, le drone apparaît comme l'arme du lâche.

Cela n'empêche pas les partisans du drone militaire de le proclamer être l'arme la plus éthique que l'humanité ait jamais connue. L'œuvre des philosophes de l'éthique militaire partisans de la technologie militaire la plus avancée trouve son point d'appui sur le terrain de la théorie du droit. La « guerre sans risque », dont le drone constitue sans doute l'arme la plus accomplie, met en crise les principes juridiques constitutifs du droit de tuer à la guerre. Avec la guerre à distance, on remet en question la notion même de guerre et de combats entre partenaires qui s'affrontent selon des règles de droit établies, pour les remplacer par un nouveau droit « d'assassinat ciblé », quitte à passer outre le droit des conflits armés.

Ce qui est en train de se produire avec les drones, c'est un glissement sémantique : on passe de la contre-insurrection délimitée dans un territoire et dont il s'agit de gagner la population pour isoler politiquement les insurgés, à une chasse à l'homme qui change le paradigme de la lutte contre-insurrectionnelle et la remplace par la lutte anti-terroriste mondiale.

Ceux qui considèrent le drone militaire comme étant l'arme privilégiée de la lutte antiterroriste nous promettent une guerre *clean* sans perte ni défaite. Mais cette guerre ne fait que perpétuer la violence. Au lieu d'en finir avec les kamikazes, cette machine à tuer ne fait que les multiplier.

Les cibles sont présumées coupables parce qu'elles se trouvaient dans le champ visuel du drone militaire, leur innocence ne pouvant être prouvée qu'à titre posthume. Contrairement à la légende si répandue, le drone s'apparente en réalité à une arme non discriminante d'un nouveau genre: en supprimant la possibilité du combat, il sape la possibilité même d'une différenciation entre combattants et non combattants; il ne cherche pas à gagner une population qui peut le soutenir contre des terroristes isolés. Il n'a besoin que de sa propre technique de tuer; en supprimant le combat, il supprime en même temps l'action politique auprès d'un ennemi devenu collectif.

Dans leur étude *La machine à tuer: la guerre des drones*, Jeremy Scahill et l'équipe de *The Intercept* (2017) expliquent l'intuition éclairante de leur livre:

Le drone est un instrument, pas une mesure politique. La politique, en l'occurrence, c'est l'assassinat. Alors que tous les présidents depuis Gerald Ford ont prorogé un décret présidentiel interdisant aux militaires américains de commettre des assassinats, le Congrès a évité de légiférer sur ce point jusqu'à ne pas définir le mot «assassinat». Les partisans de la guerre des drones ont ainsi pu redorer l'image des assassinats, notamment en les rebaptisant «assassinats ciblés».

Dans le même livre, Edward Snowden (2017) nous informe que la série de révélations sur le programme d'assassinats du gouvernement Obama dont il est question dans ce livre montre qu'une part de l'âme des États-Unis est fascinée par un exercice du pouvoir sans contrôles ni limites. En effet, selon Jeremy Scahill (2016): «Dès ses premiers jours à la tête des États-Unis, l'arme de choix du président Barack Obama a été le drone utilisé par la CIA et l'armée pour traquer et éliminer les personnes que le gouvernement a donné l'ordre d'exécuter, par des procédures secrètes, sans mise en examen ni procès.»

La question centrale soulevée par cette étude concerne les critères adoptés par les services américains pour construire leurs bases de données sur les «terroristes» à exécuter à travers la planète. La technologie de l'assassinat est fondée sur une technologie de l'information

et de l'analyse qui ne fait que refléter l'ignorance catastrophique de ce savoir américain de l'histoire des peuples de la planète à traquer comme terroristes présumés.

Entre le kamikaze et le drone la question ne consiste donc pas à choisir entre le courage suicidaire de l'un qui choisit de se faire tuer avec des innocents civils, ou la lâcheté de l'autre qui nous promet de perfectionner son arme pour éviter les dommages collatéraux des civils qui sont nés aux mauvais endroits : la question revient plutôt à réfléchir sur les raisons qui poussent les hommes à mourir et nous faire mourir. Il se peut que la réponse à cette question ne soit pas dans le perfectionnement de la technologie de la guerre de plus en plus sophistiquée ; elle se trouve peut-être du côté des injustices subies par des hommes qui ont décidé de nous tuer.

BIBLIOGRAPHIE

- Brzezinski, Zbigniew (1971). *La révolution technétronique*. Paris, Calmann-Lévy.
- Chamayou, Grégoire (2013). *Théorie du drone*. Paris, La Fabrique.
- Clausewitz, Carl Von (1955 [1832]). *De la guerre*. Paris, Minuit.
- Clausewitz, Carl Von (2014). *Théorie du combat*. Paris, Éditions Economica.
- COMEST, 2017, Rapport sur l'éthique de la robotique, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000253952_fre
- Dumouchel, Paul et Damiano, Luisa (2016). *Vivre avec les robots : Essai sur l'empathie artificielle*. Paris, Seuil.
- Foucault, Michel (1997). « *Il faut défendre la société* » *Cours au Collège de France de 1975-1976*. Paris, Gallimard.
- Habermas, Jürgen (1992 [1991]). *De l'Éthique de la discussion*. Paris, Le Cerf.
- Hegel, Georg Wilhelm Friedrich (1941 [1807]), *La Phénoménologie de l'esprit*. Paris, Aubier Montaigne.
- Heidegger, Martin (1984 [1931]). *La «Phénoménologie de l'esprit» de Hegel*. Paris, Gallimard.
- Heidegger, Martin (1954 [1935]). « La Question de la technique », dans *Essais et conférences*, Paris, Gallimard.
- Hobbes, Thomas (2000 [1651]), *Léviathan ou matière, forme et puissance de l'État chrétien et civil*, Paris, Gallimard.
- Hobbes, Thomas (1990 [1668]). *Béhémot ou le Long Parlement*, Paris, Vrin.
- Ibn Khaldoun (2002). *Al-Muqaddima. Le livre des exemples*. Paris, La Pléiade, Gallimard, 2002.
- Kash, Soheil (2012). « La guerre comme essence du politique », dans Peter Kemp et Noriko Hashimoto (dir.), *Eco-Ethica, volume 2, Le Nouvel Ordre du Monde*. Tomonobu Imamichi Institute for Eco-Ethica, Copenhagen, Tokyo.
- Kojève, Alexandre (1947). *Introduction à la lecture de Hegel*, Paris, Gallimard.

- Liotard, Jean-François (1979). *La condition postmoderne*. Paris, Les éditions de Minuit.
- Liotard, Jean-François (1988). *La postmodernité expliquée aux enfants*. Paris, Galilée.
- Liotard, Jean-François (1993). *Moralités postmodernes*. Paris, Galilée.
- Rawls, John (1987 [1971]). *Théorie de la justice*. Paris, Seuil.
- Rawls, John (1996 [1993]). *Le droit des gens*. Paris, Éditions Esprit.
- Scahill, Jeremy et l'équipe de *The Intercept* (2017). *La machine à tuer: La guerre des drones*. Montréal, Lux Éditeur.
- Schmitt, Carl (1972 [1932]). *La notion de politique*. Paris, Calmann-Lévy.
- Schmitt, Carl (2002 [1938]). *Le Léviathan dans la doctrine de l'État de Thomas Hobbes. Sens et échec d'un symbole*. Paris, Seuil.
- Snowden, Edward (2017). « Avant-propos. Choisis des circonstances », dans Scahill, Jeremy et l'équipe de *The Intercept* (2017). *La machine à tuer: La guerre des drones*, Montréal, Lux Éditeur.
- Spengler, Oswald (1948 [1923]). *Le déclin de l'Occident*, 2 vol. Paris, Gallimard.
- Todd, Emmanuel (2002). *Après l'empire. Essai sur la décomposition du système américain*. Paris, Gallimard.
- United Nations-Human Rights Council (2013). *Report of the Special Rapporteur on extrajudicial summary or arbitrary executions*, Christof Heyns, A/HRC/23/47.
- Vilmer, Jean-Baptiste (2012). *La Guerre au nom de l'humanité: tuer ou laisser mourir*. Paris, PUF.
- Vilmer, Jean-Baptiste (2013). « Légalité et légitimité des drones armés », *Politique étrangère*, vol. 3.

CHAPITRE 12

Les robots : une expérimentation sociale en mal d'encadrement éthique

Marie-Hélène PARIZEAU

1. MISE EN SITUATION

Aéroport Trudeau à Montréal fin d'après-midi, en correspondance, je prends un café et je m'installe à cinq mètres d'un petit robot humanoïde de type Peper, qui interpelle les gens qui passent devant lui par un claironnant « bonjour ! ». J'ai observé pendant plus d'une heure les interactions entre ce robot et les nombreux voyageurs qui passaient dans ce coin de restauration de l'aéroport. Ce petit robot d'environ 120 cm, avait un écran d'ordinateur à la place de la poitrine. Après vous avoir entretenu par quelques phrases anodines en français ou en anglais, il vous entraînait à regarder sur l'écran les photos des plats d'un menu, puis vous invitait à venir les consommer au restaurant derrière lui. Autant vous le dire tout de suite, il s'agissait d'un robot publicitaire.

En une heure d'observation, à peine une douzaine de personnes en majorité des hommes et deux familles, se sont arrêtés pour interagir avec ce robot. La curiosité était semble-t-il, la principale motivation

des adultes et des enfants. Les interactions ne durèrent guère plus de deux à trois minutes car rapidement les gens découvraient la finalité publicitaire du robot. D'ailleurs, personne avec ou sans l'interaction avec le robot n'est allé à ce restaurant. Les traits enfantins de Peper ont attiré l'attention des enfants de moins de 10 ans. Trois enfants d'environ 4 ans voulaient le toucher comme si le robot était un jouet. L'un d'entre eux, dans les bras de son père, a commencé à palper avec ses deux mains la tête du robot – ses capteurs visuels et auditifs. Aussitôt dans une tonalité nettement moins aimable, le robot a dit : « Attention ! Ma tête est fragile ». Le père et l'enfant s'en sont allés.

Il est frappant quand on observe ce type de robot et l'usage publicitaire qui en est fait, de constater le peu d'interactions humain-machine et la pauvreté des échanges. D'ailleurs, la grande majorité des personnes à proximité du robot, soit n'a pas remarqué sa présence, ou encore, l'a simplement regardé et continué son chemin, quelque fois avec un sourire ou une mine dégoûtée ou même une figure effrayée.

Les robots dans notre vie quotidienne ne nous laissent pas indifférents et peuvent donc induire une variété de réactions, d'émotions et de comportements. Interagir avec des robots ne va pas de soi car ce sont des machines sophistiquées. Elles nous sollicitent sur différents registres que nous n'avons pas l'habitude de côtoyer ensemble au sein d'une seule machine : le langage, le mouvement, l'apparence humanoïde ainsi que la captation visuelle et auditive. Or actuellement, on peut dire que les robots sont en quelque sorte « lâchés » dans la société où ils sont expérimentés par les travailleurs, les consommateurs, les patients, les élèves, les enfants et les gens de tout âge. Ceux-ci servent comme autant de cobayes dans une vaste expérimentation sociale dont nous allons tenter d'examiner le fonctionnement et les enjeux éthiques. Nous examinerons d'abord le cas des robots industriels puis celui des robots sociaux. Nous verrons comment se déploie cette différence en nous appuyant sur les analyses de Gunther Anders et de la théorie du *care*. Nous expliquerons ensuite pourquoi cette expérimentation sociale demande à être encadrée du point de vue éthique.

La petite histoire de ce robot publicitaire constitue une bonne illustration de cette expérimentation sociale qui a cours en particulier avec les robots sociaux. Cette expérimentation contraste avec l'utilisation progressive mais non linéaire, calculée techniquement et économiquement des robots industriels dans les usines de production.

2. LES ROBOTS INDUSTRIELS : UNE LOGIQUE DE PRODUCTIVITÉ ÉCONOMIQUE

Dans les ouvrages techniques, le critère d'utilisation des robots industriels est *l'efficacité* et l'augmentation de la *productivité* dans la production de biens de consommation en raison de leur fiabilité et de leur précision technique par rapport aux ouvriers. Les robots sont utilisés essentiellement pour des raisons économiques au sein du capitalisme mondialisé. Ainsi, la production de certains objets techniques complexes (les avions par exemple) exige une très grande quantité de pièces dont la qualité technique doit être quasi-parfaite, quel que soit, le lieu et le pays de leur provenance.

Globalement, l'utilisation des robots est devenue moins coûteuse que la main d'œuvre humaine, en tout cas dans les pays développés. Certains pays comme l'Allemagne, se sont tournés sur ce modèle d'industrialisation permettant ainsi de maintenir au pays des emplois de qualité et de minimiser les délocalisations dans les pays du Sud ou émergents, en particulier en Asie. Les robots industriels peuvent fonctionner tout le temps :

Les robots, contrairement aux ouvriers, ne se fatiguent pas (ils tombent cependant en panne) : ils ne se plaignent pas ; ils ne sont jamais distraits dans leur travail ; ils ne font pas la grève ; ils n'ont pas la gueule de bois le lundi matin... Les robots coûtent moins cher. Ils sont souvent plus efficaces et précis que les travailleurs humains. Ils n'ont pas de retraite, ni assurance santé, ni droits légaux. Nous voulons que les robots aient toutes les qualités que les maîtres recherchent dans leurs esclaves, les patrons dans leurs employés, les commandants dans leurs soldats, et nous voulons qu'ils n'aient ni leurs faiblesses, ni leurs manques, ni surtout cette tendance irrépressible à l'insubordination, à l'indépendance et à n'en faire qu'à leur tête qui caractérise les travailleurs humains. (Dumouchel, Damiano, 2016, p. 10).

Les robots sont donc des machines programmées à faire des tâches très précises et répétitives. Cependant d'autres critères sociaux confortent le développement des robots en particulier la question de la sécurité en milieu hostile, c'est-à-dire là où des risques d'accident sont importants à cause de la manipulation de pièces (la fonderie, la peinture, le découpage du métal à froid), de l'utilisation de produits chimiques ou d'agression physique (bruit intense, rayonnement nucléaire), ou encore

de conditions de travail difficiles (charges physiques, monotonie du poste) (Jean-Paul Vautrin, 1987). Cependant, la présence de robots dans les usines entraîne d'autres types d'accidents entre ouvrier et robot (collision avec le robot, blessure suite à la projection d'un objet par le robot ou de la reprise manuelle du robot, coincement dans le robot, etc).

Il reste que les critères de la pénibilité du travail et de répétitivité de la tâche encouragent l'utilisation des robots dans un contexte de production où les tâches et les emplois sont susceptibles d'automatisation. La définition précise des tâches, leur morcellement et le fait qu'elles soient répétitives, sont également des conditions essentielles de l'organisation du travail qui favorisent l'utilisation des robots dans un contexte industriel et de services.

Au plan historique, il est intéressant de noter que la robotisation des usines s'est développée dans les années 1970 en particulier dans la construction automobile sur les chaînes de montage. Déjà en 1980, de nombreuses publications se préoccupaient des impacts négatifs de la robotisation sur l'emploi (perte d'emploi dans les secteurs industriels, augmentation du chômage, etc). Cette préoccupation s'accroît aujourd'hui à cause, d'une part, des effets de la mondialisation et de désindustrialisation des pays développés en faveur des pays asiatiques, et d'autre part, des enjeux complexes de la requalification des travailleurs et de la transformation de l'emploi (Nicolas Le Ru, 2016).

Or, le paysage des robots industriels a changé considérablement depuis 40 ans. Quelques chiffres illustrent ce phénomène : en 2004, il s'est vendu 97 000 robots industriels, en 2017, 387 000¹. Le coût moyen actuel est de 46 000\$ l'unité². Les acheteurs principaux sont la Chine, la Corée du Sud, les États-Unis et l'Allemagne. Ces robots sont vendus en particulier dans le secteur de l'automobile, l'industrie électrique et électronique, l'industrie du caoutchouc et des plastiques, l'industrie pharmaceutique et des cosmétiques, l'industrie de l'alimentation et des breuvages, l'industrie du métal et de la machinerie (Executive Summary World Robotics, 2018). Le critère utilisé maintenant est celui de la densité de robots par 10 000 employés : Europe 106, Amérique 91, Asie 75. Mais ces chiffres cachent de grandes disparités, la Corée du Sud a un ratio de 710/10 000, l'Allemagne de 322, le Japon 308. Cette tendance lourde à la robotisation industrielle répond à des

1. <https://www.statista.com/statistics/264084/worldwide-sales-of-industrial-robots/>

2. <https://www.statista.com/statistics/830578/average-selling-price-of-industrial-robots/>

changements technologiques (efficacité énergétique, programmation facilitée et optimisation autonome des robots) et économiques (expansion du marché, produits sur mesure et individualisés, rapidité de la production et de la livraison, etc). La fabrication industrielle utilise donc de plus en plus les robots et la tendance pousse vers l'interaction robot-ouvrier nécessitant une « collaboration » dans le travail, l'un assistant l'autre et vice-versa. Car la prochaine étape technologique est celle des objets connectés où les machines et les robots communiquent entre eux mais également avec les travailleurs (M.A. Kamarul Bahrin et coll., 2016). Cette nouvelle étape de la robotisation industrielle inclut des tâches plus complexes dans la production ou liées à l'optimisation : augmentation de la rapidité de la production, vérification de la qualité des pièces par des capteurs et les données massives, planification de la maintenance et de la logistique, etc.

Il reste que plus les robots interagissent avec les humains et les travailleurs, plus, du point de vue du robot et de ses concepteurs, les enjeux de sécurité, fiabilité, utilité, performance, coût, (éventuellement empreinte écologique), deviennent déterminants. Du côté des humains et des travailleurs, la littérature spécialisée focalise sur « l'acceptabilité », c'est-à-dire l'acceptation psychologique et culturelle d'une personne d'interagir avec un robot. Il s'agit d'établir ou de mesurer, la volonté démontrable d'un groupe d'utilisateurs d'utiliser un robot (ou un système d'information) pour la tâche pour laquelle le robot a été conçu et construit (Salvini P. et al, 2010). Cette acceptabilité est complexe puisqu'elle englobe tant la présence physique du robot avec sa matérialité (taille, forme ou non humanoïde, type de mobilité et de senseurs, matériaux utilisés, etc), que la qualité des interactions robot-utilisateur (sécurité, utilité, facilité d'usage et d'interactions, capacité cognitive exigée, modes d'interactions favorisés physique, vocal, visuel). Un autre niveau de complexité s'ajoute lorsque le robot agit dans l'environnement social, et qu'il est présent au-delà d'un espace confiné (une pièce, un hôpital, un immeuble) et se déplace dans des environnements non confinés comme la rue ou l'aéroport. Il existe au sein des populations, des personnes et des groupes réfractaires à ce type de technologie et ce, pour de multiples raisons qui peuvent aller de la crainte, à la méfiance, ou encore qui préfèrent garder le contrôle ou qui ont une conception stricte de leur vie privée ou qui conçoivent autrement les relations humaines, sociales, culturelles et économiques, etc.

Par ailleurs, l'intrusion de robots dans des environnements non confinés pose des problèmes supplémentaires en termes d'organisation

sociale, de règles éthiques, sociales et juridiques. Jusqu'où et pour quelles raisons, devrait-on changer l'organisation sociale et spatiale d'une ville par exemple, pour introduire de façon significative ces robots dans les activités usuelles des gens? Les débats sur la voiture autonome sont assez exemplaires des problèmes liés à l'acceptabilité sociale d'une technologie (T. Litman, 2019) alors que l'automatisation d'une ligne de métro – un métro sans chauffeur – passe relativement inaperçu du grand public, ce qui n'est évidemment pas le cas du côté des chauffeurs de métro qui risquent de perdre leur emploi.

On saisit ici que le passage du robot au laboratoire entouré de scientifiques à l'usine où il effectue des manœuvres précises avec des travailleurs constitue une première étape de la robotisation des activités de travail en contexte industriel. Cette étape a nécessité plus de 30 ans avec des adaptations matérielles et humaines. Elle répondait essentiellement à des objectifs de productivité et d'efficacité économiques dans un contexte de capitalisme avancé. Dans cette perspective, penser que toutes les usines du monde deviendraient entièrement robotisées, constitue probablement une utopie. Le transfert de technologies des pays les possédant vers les pays ne les ayant pas, ne va pas de soi pour des raisons de domination du marché. Cependant la puissance du mouvement actuel de robotisation accentuée au plan social, une étrangeté croissante pour le travailleur des objets produits, une dévalorisation tant du travail que des travailleurs dans le processus de production industrialisé, comme Marx l'avait déjà montré au XIX^e siècle. Cette dévalorisation du travailleur s'accroîtrait par «l'obsolescence de l'homme» face à la machine, comme l'a expliqué Günther Anders, au point où l'homme ordinaire et limité dans ses capacités, qui travaille en usine ou effectue des tâches spécialisées, se doit d'imiter la machine efficace ou le robot parfait : *« ce n'est plus l'homme en tant qu'instrument parmi les instruments, mais l'homme en tant qu'instrument pour les instruments »* (Anders, 2002, p 48).

Cette extension de la robotisation industrielle vers les robots sociaux aura-t-elle aussi des effets sur la transformation des institutions et des rapports sociaux? Comment alors les sentiments et les représentations du monde de l'homme ordinaire pourraient-ils se mettre en phase avec les activités de la machine ou du robot? Anders met en garde contre une robotisation de la société toute entière :

Il serait tout à fait concevable que la transformation des instruments soit trop *rapide*, bien trop *rapide*; que les produits nous demandent quelque chose d'excessif, quelque chose d'impossible; et que nous nous enfonçons vraiment, à cause de leurs exigences, dans un état de pathologie collective. Ou bien dit autrement, du point de vue des producteurs: il n'est pas complètement impossible que nous, qui fabriquons ces produits, soyons sur le point de construire un monde au pas duquel nous serions incapables de marcher et qu'il serait absolument au-dessus de nos forces de «comprendre», un monde qui excéderait absolument notre force de compréhension, la capacité de notre imagination et de nos émotions, tout comme notre responsabilité. (Anders, 2002, p 32).

La rapidité de la robotisation qui s'est mise en place comme modèle de productivité dans les usines pourrait selon l'analyse d'Anders induire des pathologies sociales vécues par les travailleurs ordinaires dépassés par ce phénomène de la robotisation du travail et de la perfection d'exécution des robots. Dans cette logique robotique, Anders souligne que le travailleur est soumis à «l'ingénierie humaine» (human engineering), bref à une transformation expérimentale de son corps et de son esprit pour le soumettre à ces nouvelles contraintes de la robotisation. Pour Anders, la robotisation des travailleurs et des rapports sociaux induit la perte de la liberté humaine donc ultimement, la déshumanisation. Et il ajoute: «Or pour le moraliste qui ne peut se résoudre à faire son deuil de l'idée de l'homme, il est évident que cela mène tout droit à la catastrophe» (Anders, 2002, p59).

Ce diagnostic pessimiste d'Anders sur le phénomène de la robotisation de l'homme au travail et qui, face à des robots, l'obligent à travailler suivant les paramètres définis par les machines, peut-il être étendu aux robots sociaux? Est-ce que les robots sociaux transformeraient les relations sociales en les standardisant du point de vue robotique? Un peu à la manière dont nous avons appris à répondre de façon minimale et univoque aux questions des machines pour certains services téléphoniques. C'est ce que Jean-Michel Besnier a appelé «l'homme simplifié». Besnier réclame par ailleurs le retour de l'intelligence dans les communications humaines afin d'éviter l'appauvrissement du langage et la délégation de l'existence humaine aux machines (Besnier, 2012, p14).

Que cherche-t-on à faire lorsque les chercheurs, les ingénieurs et les industriels développent des robots sociaux? Dans cette vaste

expérimentation à l'échelle planétaire, assisterait-on à une sorte d'ingénierie sociale, à une robotisation des structures et des rapports sociaux? Celle-ci viserait-elle un des objectifs de Norbert Wiener (1952), le fondateur de la cybernétique, soit la pacification et le contrôle des comportements humains pour éviter une autre guerre mondiale et ce, au prix d'une réduction essentielle de la liberté individuelle par une surveillance généralisée plus ou moins contraignante? Un tel développement de la robotisation d'une société pourrait alors se traduire par des formes politiques de type totalitaire comme l'illustre depuis quelques décennies la littérature de science-fiction (Gladden, 2014).

3. LES ROBOTS SOCIAUX À L'HÔPITAL, À L'ÉCOLE ET À LA MAISON: UNE EXPÉRIMENTATION SOCIALE À LARGE ÉCHELLE

Le robot qui interagit dans des environnements confinés ou non, pour des fonctions sociales précises (robots sociaux) avec des utilisateurs ou des personnes en présence, constitue donc une nouvelle étape de la robotisation au sein de la société. Un robot social se définit :

«comme un robot autonome qui interagit et communique avec des êtres humains ou d'autres agents physiques autonomes en respectant les normes sociales et les comportements sociaux attachés à ses fonctions. Certains chercheurs soulignent à propos de cette catégorie, les dimensions fonctionnelles et affectives des robots sociaux qui caractérisent tout particulièrement les appareils robotiques que l'on appelle 'robots compagnons'» (COMEST, 2017, p 37).

Nous allons maintenant examiner le contexte d'expérimentation sociale à large échelle généré par la multiplication des robots sociaux et questionner certaines finalités.

La question des robots sociaux peut être analysée du point de vue de leur utilisation comme dans le rapport de la COMEST sur l'éthique de la robotique ou comme dans le livre de Royakkers et van Est qui font état de toutes les utilisations possibles et actuelles des robots.

Notre classification sera un peu différente dans la mesure où nous allons examiner la pénétration de certains robots dans trois sphères: les institutions sociales que sont l'école et l'hôpital, l'oïkos (la maison) et l'intimité personnelle. Ces trois sphères sont inter-reliées,

cependant, une présence accrue souvent non-sollicitée des robots induirait-elle un phénomène concomitant de robotisation des relations sociales, familiales, et personnelle? La robotisation de ces relations conduirait alors à la perte partielle voire complète, de l'autonomie des personnes par des schémas comportementaux et langagiers dictés par l'interaction robot-humain, entraînant, du même coup, au plan éthique, une dilution de la responsabilité individuelle et collective dans les relations sociales. Nous utilisons ici les critères moraux développés par Anders. En d'autres termes, la question de l'utilisation des robots sociaux ne peut se réduire à une question de coût/efficacité ou de risques/bénéfices, elle va au-delà d'un conséquentialisme économique ou éthique de type utilitariste. L'enjeu se situe au niveau déontologique par rapport à une série de questions : dans les interactions humains-robots, quel genre d'humanité voulons-nous favoriser collectivement pour nous, nos enfants et les générations futures? Sur quels principes ou valeurs voulons-nous faire reposer nos choix sociaux et politiques dans l'utilisation des robots sociaux? Quels sont nos critères éthiques pour faire face à l'expérimentation sociale en cours?

Examinons brièvement la situation actuelle de façon globale dans les pays développés. Au sein des institutions sociales, les robots sont expérimentés en particulier dans deux grandes institutions, l'hôpital (les lieux de soins) et l'école.

Dans le domaine des soins de santé, les robots ont différents types de finalité. Des robots qui fournissent une aide physique pour soulever des patients de leur lit ou de leur chaise roulante sont testés à l'hôpital au Japon afin de faciliter les soins infirmiers. Ce pays comme d'autres pays développés, fait face à une pénurie d'infirmières et d'aides-soignants avec une augmentation de la population des personnes de plus de 65 ans (environ 28 % de la population). D'autres robots ont pour vocation les activités de soins et de surveillance des patients en particulier des personnes âgées en perte progressive d'autonomie dans le but de les maintenir à domicile. Ces robots peuvent donner des médicaments et visent dans le futur à aider la personne dans des tâches telles le nettoyage de la maison, la préparation des repas ou l'hygiène personnelle. Actuellement, ces robots servent surtout à assurer la sécurité de la personne (mobilité, chutes, problèmes cardiaques) en la surveillant. Ce type de robot est aussi utilisé en contexte de résidence pour personnes âgées ayant une autonomie réduite. Des tests sont en cours dans des pays comme l'Allemagne ou en Grande-Bretagne (Jenkins, Draper,

2015) et la question éthique est posée de façon répétitive: quand et comment utiliser les robots et pour quels types de soins? (Santoni de Sio, van Wynsberghe, 2015; Sharkey, Sharkey, 2012).

Dans le contexte des soins, le robot compagnon est nettement plus répandu car son rôle est de tenir compagnie à la manière d'un animal de compagnie (Paro, le robot-phoque ou Aibo le chien-robot). Ces robots compagnons sont utilisés dans des résidences pour personnes âgées dépendantes ou ayant des troubles cognitifs dans le but de les occuper et de rompre leur isolement. Cependant ces robots compagnons pénètrent aussi la sphère familiale en remplaçant les animaux de compagnie. Ainsi Aibo, transformé en un objet-jouet connecté, est devenu le compagnon des enfants au Japon. Plus de 20 000 Aibo se sont vendus en 2018 à un coût d'environ 3000\$US. Cet exemple montre comment d'un usage à visée de soins, l'utilisation du robot peut s'étendre à des usages relationnels différents qui s'intègrent à la vie familiale (des personnes âgées aux enfants).

L'école est aussi un lieu d'expérimentation des robots. Sur une base encore expérimentale au Japon, des robots humanoïdes sont utilisés pour enseigner une deuxième langue à des élèves du primaire. En stimulant l'interaction en classe et en répétant inlassablement les mots et les phrases à apprendre, ce robot soutiendrait la motivation des élèves par son côté ludique. Des robots sont aussi utilisés pour faciliter l'apprentissage des mathématiques et des processus scientifiques. Enfin, au Japon, des robots interagissent avec des élèves pour les aider à s'exprimer et à créer de façon collaborative des récits et des histoires (Tanaka, Kimura, 2010). Pour l'instant, les résultats de ces expérimentations ne permettent pas au plan empirique de saisir les impacts réels de l'utilisation de ces robots dans les processus d'apprentissage. Cependant, des études montrent l'influence du robot sur certains enfants qui lui attribuent des sentiments ou encore qui croient ce que le robot dit sans aucun sens critique. L'enjeu de l'attachement affectif de certains enfants au robot plutôt qu'à l'enseignant n'est pas sans poser des questions sur les conséquences d'une relation affective simulée par le robot dans le développement psychique de l'enfant. Comme le constate le rapport de la COMEST :

«L'éducation des enfants est profondément enracinée dans une culture et engage à la fois la transmission et la construction de connaissances et de valeurs sociales. Les attitudes des parents et d'autres acteurs sociaux à l'égard de l'utilisation de robots

« dans l'éducation et les soins de santé varient d'un pays à l'autre » (COMEST, 2016, p39).

Un sondage européen de 2012, indique que 34 % des répondants sont favorables à l'interdiction des robots dans l'éducation alors qu'au Japon, les robots ayant une image positive, sont largement expérimentés dans les secteurs de la santé et de l'éducation.

Les robots sont de plus en plus présents au sein de la maisonnée. Les robots ménagers ou de services – robots-aspirateurs (Roomba), robots nettoyeur de plancher, de vitres, de piscine, etc. – font partie maintenant de l'équipement ménager. Cependant ces robots de services sont souvent équipés de capteurs, de caméra et connectés à internet, devenant ainsi des objets connectés. Ils interagissent de façon quotidienne dans la vie privée des gens et des familles. Ils peuvent potentiellement enregistrer des quantités de données sur les habitudes de vie, les conversations, etc. Cette présence de robots pour des tâches précises peut aussi se transformer en surveillance indirecte dont les finalités peuvent être commerciales ou dans certains contextes, étatiques.

Enfin, les robots peuvent aussi s'immiscer dans la vie personnelle. Les robots sexuels sont généralement des poupées plus ou moins mobiles (Harmony), couplées à l'intelligence artificielle coûtant environ 14 000\$US. Certains chercheurs s'interrogent sur les impacts de ces pratiques sur les formes de la sexualité humaine ainsi que la relation à l'autre qu'il soit humain ou machine (Whitby, 2012).

4. LES ROBOTS, UNE RÉPONSE TECHNIQUE À DES PROBLÈMES SOCIAUX ?

Dans leur diversité, les robots sociaux ont en commun le fait de s'immiscer dans les lieux (maison, hôpital, école) où s'exprime la relation de la sollicitude, du « prendre soin » (le *care*), que ce soit envers la personne âgée malade ou non, l'enfant, les membres de la famille ou l'animal de compagnie. Rappelons que le *care* se définit comme le fait de « se soucier de la personne » (l'attention à l'autre), de « prendre en charge cette personne » (la responsabilité), de « prendre soin de cette personne » (la compétence) et pour cette personne, de « recevoir le soin » (la capacité de réponse) (J. Tronto, 2009). Cette incursion des robots jusqu'alors limitée aux usines vers l'espace social et les lieux de l'intimité familiale, a de quoi étonner : comment les robots peuvent-ils répondre aux exigences pratiques et éthiques du *care*, de cette

sollicitude humaine envers l'autre ? Pourquoi au-delà des considérations économiques ou ludiques, introduire les robots dans l'espace et la relation du *care* et de l'attention à l'autre ? Comment justifier leur utilité, voire leur nécessité pour imposer leur présence dans ces espaces relationnels si particuliers au plan éthique, que sont, par exemple, les soins aux personnes âgées et l'éducation aux enfants ?

Prenons l'exemple du soin aux personnes âgées. Si dans beaucoup de pays, la femme s'est vue attribuée socialement de gré ou de force, le rôle de « care-giver » de ses parents ou beaux-parents âgés, dans d'autres pays, avec l'emploi des femmes et l'allongement de la vie, les femmes sont moins disponibles pour ce rôle social. Des lieux spécialisés (hébergement pour personnes âgées, autonome, semi-autonome ou dépendante) ont remplacé certaines dimensions de l'entre-aide inter-générationnelle qui reposait généralement sur la femme. Cette entre-aide, encore largement à l'œuvre dans les familles, est devenue « un travail », c'est-à-dire un emploi, souvent faiblement rémunéré et féminin, peu technique et peu valorisé socialement comme le montre Joan Tronto. On peut donc saisir ici que l'utilisation de robots de soins devient une réponse technique à un problème social complexe, celui du prendre soin des personnes âgées alors que dans les pays développés, une majorité des femmes ont un emploi et du temps limité pour s'occuper et prendre soin de leurs enfants, conjoint, parents, voire petits-enfants. À cela s'ajoute un contexte socio-économique particulier qui considère les personnes âgées, d'une part, comme peu productives puisqu'elles n'ont plus d'emploi, et d'autre part, comme des consommateurs de biens et de soins coûteux lorsqu'elles sont en perte d'autonomie. La logique économique et technicienne impose les robots pour pallier non seulement à un manque de personnel soignant ou un manque d'enseignants (le Japon étant un cas paradigmatique à cet égard) mais, plus globalement, à une dévalorisation sociale des emplois du soin. Les critères d'efficacité et de productivité similaires à ceux des robots industriels sont alors appliqués à ce type d'emploi qui semble, comme les autres emplois, réductible à une série de tâches décomposables et donc programmables pour un robot. On fait comme si la relation de soins était équivalente à une fonctionnalité, ce que toute la littérature sur l'éthique du care a démenti (Bennaroyo L *et al.*, 2010 ; V. Nurock, 2010). Au lieu de valoriser la relation humaine dans le prendre soin de l'autre, la réponse sociale actuelle est une valorisation technique qui tend alors vers la déshumanisation de la relation à l'autre. C'est-à-dire que la relation humaine quelles que soient ses

insuffisances, est disqualifiée au profit de gestes techniques et d'une gestion technique, ici une robotique qui surveille, mesure ou divertit.

L'utilisation des robots en contexte scolaire répond à un autre problème social. La crise plus ou moins latente des systèmes d'éducation à travers le monde, émerge avec le développement non seulement des robots mais plus largement du numérique à l'école qui posent deux questions fondamentales : que doit-on apprendre à l'école ? Et qui doit l'enseigner (le professeur ou le robot, la tablette, l'ordinateur, le téléphone portable, etc.) ? (Holmes, Bialik, Fadel, 2019).

Il faut donc souligner qu'à ces deux questions importantes auxquels font face avec plus ou moins d'intensité, les pays développés, la réponse actuelle est à nouveau d'ordre technique. Dans cette logique, l'enfant apprend a priori mieux s'il est entouré d'objets techniques numériques dont les robots, pour le stimuler de façon ludique, maintenir son attention et le surveiller afin d'obtenir des résultats scolaires correspondants aux objectifs d'apprentissage. Dans ses apprentissages, l'enfant doit donc mimer le robot. Cette logique met l'accent sur la conformité et la discipline qui vont de pair avec « la culture ludique » actuelle qui offre des produits culturels (dont les robots) par lesquels l'enfant apprend à maîtriser un univers symbolique particulier. C'est-à-dire ce que la société attend de lui. (G. Brougère, 2002). Être obéissant et apprenant comme un robot ? Nous sommes ici plutôt loin des visées éducatives du développement affectif et cognitif de l'enfant ainsi que de son épanouissement personnel.

Les robots pour les soins et pour l'éducation participent donc à cette ingénierie sociale expérimentale du numérique. En ce sens, les robots pour les soins, éducatifs, compagnons, ou de services ne peuvent être isolés de ces nouvelles médiations sociales générées par le numérique (Internet, réseaux sociaux, intelligence artificielle, objets connectés, données massives, etc) et le nouveau modèle économique qui y est attaché.

Les principales questions éthiques concernant les robots se résument ainsi :

- À quoi et à qui sont-ils utiles ?
- En quoi constituent-ils un vrai progrès social et non un gadget (au sens où une prothèse auditive permet à une personne âgée d'entendre distinctement à nouveau et ainsi d'être *en relation* avec les autres) ?

- Quelles sont les limites et les finalités des interactions entre l'humain et le robot ?
- Comment doit-on protéger les groupes plus vulnérables – enfants, personnes âgées- et les gens en général, d'interactions humain-robot non sollicitées et non-consenties ?
- Dans les usages des robots, comment valoriser les relations humaines pour éviter la déshumanisation, c'est-à-dire être traité comme un robot ou une machine fonctionnelle ?

5. CRITÈRES ET PRINCIPES ÉTHIQUES D'UNE EXPÉRIMENTATION SOCIALE AVEC LES ROBOTS SOCIAUX

Depuis peu, certaines réponses à ces questions viennent des universitaires, mais aussi des ONG et surtout des milieux professionnels du numérique (ingénieurs, informaticiens, concepteurs et producteurs de robots) et des institutions internationales. Les milieux des ingénieurs se soucient de plus en plus des impacts éthiques de la production des robots et des autres objets numériques utilisant l'intelligence artificielle et les données massives. Certaines initiatives comme celle de l'Institute of Electrical and Electronics Engineers (IEEE) qui a élaboré en 2018 un document de référence intitulé, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, témoignent d'une volonté d'éduquer les ingénieurs aux enjeux éthiques liés au développement de la robotique et de l'intelligence artificielle. L'OCDE de son côté, a énoncé certains principes pour aider les gouvernements à développer l'intelligence artificielle de manière responsable, fondée sur la confiance et les Droits de l'homme. Dans sa « recommandation sur l'intelligence artificielle » adopté en mai 2019, elle retient les valeurs centrées sur l'humain et l'équité, la transparence et l'explicabilité des systèmes, la robustesse et la sécurité ainsi que la responsabilité des acteurs.

Dans son rapport de 2017, la COMEST de son côté, a proposé un cadre éthique s'appuyant sur le développement technique des robots en faisant une distinction entre les robots déterministes programmables de part en part et les robots cognitifs utilisant l'intelligence artificielle et ayant une fonction décisionnelle autonome. En collant au plus près de la technique, la COMEST a montré que plus un robot avait une autonomie décisionnelle, plus le principe de précaution s'appliquait et plus il devenait nécessaire que l'humain puisse maintenir son contrôle sur

la machine. *C'est donc le principe de la responsabilité humaine qui doit guider le développement de la robotique et de l'intelligence artificielle.* Le robot reste une machine, un outil pour des finalités déterminées et à évaluer en fonction d'un certain nombre de principes: la dignité humaine, l'autonomie de la personne dans sa dimension d'inter-relation avec les autres humains et avec l'environnement, le respect de la vie privée, l'innocuité, le principe de précaution, la bienfaisance et la justice. La COMEST a dessiné à grands traits ce qui pourrait fournir un cadre éthique pour évaluer tout nouveau robot avant de le rendre disponible sur le marché économique.

Il reste ici un effort théorique encore important à faire pour fournir une éthique de la robotique plus normative et qui puisse répondre aux problèmes éthiques concrets que nous avons soulevés à propos des robots sociaux et des populations vulnérables comme les enfants et les personnes malades ou âgées. Un parallèle peut être effectué ici avec le développement des règles d'éthique concernant la recherche biomédicale au début des années 1970 et qui a construit le champ de la bioéthique américaine.³

Ne faudrait-il pas faire le même effort normatif, multidisciplinaire et interdisciplinaire pour encadrer, voire règlementer le développement des robots (et plus globalement du numérique, de l'intelligence artificielle et des données massives) et, de façon plus urgente, le domaine des robots sociaux ?

La nouveauté et la complexité des robots sociaux exigent des règles éthiques et de justice, qui s'appliquent autant dans leur conception que dans leur utilisation ou leurs finalités. Il n'est pas dit que l'utilisation des robots soit une solution technique adéquate et durable aux problèmes sociaux multidimensionnels et relationnels, en particulier dans le domaine de l'éducation et des soins. Parce qu'elle vise surtout des populations vulnérables, les enfants et les personnes malades, âgées ou non, l'utilisation des robots devrait non seulement être encadrée, mais soumise à des règles d'évaluation scientifique et éthique tout à fait

3. Au Canada, ce sont les organismes subventionnaires qui ont mis en place des règles précises en ce sens. L'Énoncé de politique des trois Conseils révisé en 2014, intitulé « Éthique de la recherche avec des êtres » met l'accent sur les conditions du consentement à la recherche, l'importance de la confidentialité et du respect de la vie privée, l'équité, l'évaluation par les comités d'éthique de la recherche. Ce sont des bases éthiques minimales à partir desquelles la recherche en robotique pourrait être encadrée. Cependant une éthique de la robotique exige davantage afin de répondre de façon plus précise aux enjeux de la robotique et du numérique.

similaires à celles de la recherche biomédicale qui se sont développées depuis plus de quarante ans dans le champ de la bioéthique.

Avant d'introduire les robots sociaux auprès des différents groupes de la population, des recherches scientifiques indépendantes sur les impacts relationnels des robots devraient être prérequis avant leur mise en marché afin de prouver leur innocuité relationnelle, leur utilité thérapeutique, de soins ou éducatif, et plus globalement, d'évaluer leurs impacts sociaux et environnementaux. Ces populations plus vulnérables et la population en général, n'ont pas à être les cobayes non consentants de l'expérimentation sociale qui a cours et dont les ressorts sont avant tout d'ordre économique et d'ingénierie techno-sociale.

BIBLIOGRAPHIE

- Anders Günther (2002). *L'Obsolescence de l'homme. Sur l'âme à l'époque de la deuxième révolution industrielle* (1956). Paris, Éditions de l'Encyclopédie des nuisances, Éditions Ivesa.
- Anderson Michael, Anderson Susan Leigh, (eds) (2011). *Machine Ethics*, Cambridge University Press.
- Benaroyo L., Lefèvre C., Mino J-C, Worms F., (eds) (2010). *La philosophie du soin. Éthique, médecine et société*, Paris, PUF.
- Besnier Jean-Michel (2002). *L'homme simplifié. Le syndrome de la touche étoile*, Paris, Fayard.
- Bontems Vincent (2018). «La machine respectueuse. L'éthique des techniques de Simondon à l'ère des robots», *Revue Française d'éthique appliquée*, (1), 22-33.
- Brogère Gilles (2002). «L'enfant et la culture ludique», *Spirale*, 24, 25-38.
- COMEST, (2017). Rapport sur l'éthique de la robotique, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000253952_fre
- Executive Summary World Robotics (2018). *Industrial Robots*. https://ifr.org/downloads/press2018/Executive_Summary_WR_2018_Industrial_Robots.pdf
- Jenkins S., Draper H. (2015). «Care, Monitoring, and Companionship: Views on Care Robots from Older People and Their Carers», *International Journal of Social Robotics*, 7, p. 673-683.
- Gunkel David (2012). *The Machine Question. Critical perspectives on AI, robots, and Ethics*, Cambridge, The MIT Press.
- Gladden M. (2014). «The Social Robot as 'Charismatic leader': a Phenomenology of Human Submission to Nonhuman Power», *Sociable Robots and the Future of Social Relations: Proceedings of Robot-Philosophy 2014*, vol. 273, IOS Press, p. 329-339.
- Kamarul Bahrin M.A., Othman M. F., Nor Azli N.H., Farihin Talib M. (2016). «Industry 4.0: A review on industrial automation and robotic», *Jurnal teknologi*, 78, p. 6-13.
- Holmes W., Bialik M., Fadel C. (2019). *Artificial Intelligence in Education*, The Center for Curriculum Redesign, USA.

- Hottois G., Missa J-N., Perbal L. (dir.) (2015). *Encyclopédie du trans/posthumanisme*, Paris, Vrin.
- Le Ru Nicolas, (2016). «L'effet de l'automatisation sur l'emploi: ce que l'on sait et ce qu'on ignore», La note d'analyse, *France Stratégie*, juillet 2016, n° 49.
- Litman Todd, (2019). *Autonomous Vehicle Implementation Predictions. Implications for Transport Planning*, Victoria Transport Policy Institut.
- Lin Patrick, Abney Keith, Bekey George, (eds) (2012). *Robot Ethics. The ethical and social implications of robotics*, Cambridge, The MIT Press.
- Nurock Vanessa, (2010). *Carol Gilligan et l'éthique du care*, Paris, PUF.
- OCDE, (2019). Recommandation du Conseil sur l'intelligence artificielle, <https://legalinstruments.oecd.org/fr/instruments/OECD-LEGAL-0449>
- Royakkers L., van Est R. (2016). *Just Ordinary Robots. Automation from Love to War*, New York CRC Press.
- Sharkey A., Sharkey N. (2012). «Granny and the robots: ethical issues in robot care for elderly», *Ethics and Information Technology*, 14, p. 27-40.
- Salvini P., Laschi C., Dario P., (2010). «Design for Acceptability: Improving Robots' Coexistence in Human Society», *Int J Soc Robot*, 2, p. 451-460.
- Santoni de Sio F., van Wynsberghe A. (2016). «When Should we use Care Robots? The Nature-of-Activities Approach», *Science and Engineering Ethics*, 22, p. 1745-1760.
- Sharkey, Noel, Sharkey, Amanda (2012). «The rights and wrongs of robot care» in: Lin Patrick, Abney Keith, Bekey George, (eds), (2012), *Robot Ethics. The ethical and social implications of robotics*, Cambridge, The MIT Press, p. 267-282.
- Tanaka F., Kimura T. (2010). «Care-receiving robot as a tool for teacher in child education», *Interactions Studies*, 11(2), p. 263-268.
- Tronto Joan (2009). *Un monde vulnérable. Pour une politique du care*, Paris, La Découverte.
- Vautrin Jean-Paul, (1987). *Robotique industrielle et sécurité*, Centre de recherche sur la culture technique, Neuilly-sur-Seine.
- Wiener Norbert (1952). *Cybernétique et société*, Paris, Édition des deux Rives.
- Whitby Blay (2012). «Do you want a Robot Lover? The ethics of caring technologies» in: Lin Patrick, Abney Keith, Bekey George, (eds), (2012), *Robot Ethics. The ethical and social implications of robotics*, Cambridge, The MIT Press, p. 233-248.

CHAPITRE 13

Les robots humanoïdes peuvent-ils venir en aide aux enfants atteints du trouble du spectre de l'autisme ?

Josée Anne GAGNON

De plus en plus d'enfants souffrent du trouble du spectre de l'autisme (TSA); on parle d'une incidence aussi élevée qu'un enfant sur 152 (Elsabbagh *et al.*, 2012). Au cours des dernières décennies, le nombre d'enfants atteints n'a cessé d'augmenter pour des raisons encore nébuleuses qui vont au-delà des meilleures performances diagnostiques. Il n'existe pas de présentation unique de cette condition, puisque les symptômes peuvent varier d'un enfant à l'autre, ce qui explique la terminologie de «spectre de l'autisme» pour désigner cette pathologie. Cette réalité complique malheureusement le diagnostic et, conséquemment, en sous-estime la prévalence. Cependant, des traits cliniques communs à tous ces enfants sont répertoriés, notamment des difficultés dans leurs interactions sociales et leur communication, des stéréotypies dans leurs mouvements et attitudes, et des intérêts limités spécifiquement pour les visages humains, surtout chez les garçons. Ces particularités de leurs habiletés sociales ont des impacts dans leurs activités de la vie quotidienne, et ultérieurement, dans leurs activités professionnelles (APA, 2013). Malgré des différences dans la sévérité de leurs symptômes, la socialisation de chacun de ces enfants représente une difficulté. Leurs relations avec les pairs sembleront absentes,

inadéquates ou étranges. En l'occurrence, ils s'adaptent difficilement au contexte en raison d'un manque de flexibilité cognitive et d'une psychorigidité.

Pour mieux comprendre nos questionnements quant à l'utilisation des robots dans la thérapie du TSA chez les enfants, il apparaît utile dans un premier temps de s'attarder de manière plus précise à la symptomatologie clinique, aux défis diagnostiques et aux thérapies reconnues du TSA. En général, les parents préoccupés par l'état de santé de ces enfants consulteront les professionnels pour l'une des raisons suivantes. Leur enfant sollicite rarement leur attention, il ne pointe pas les objets, il ne répond pas à son nom, il ne recherche pas le contact visuel, il ne parle pas ou peu, ou il n'exprime pas de sentiment de bonheur. Lorsque l'enfant s'exprime, son ton apparaît souvent uniforme ou robotique, il est possible qu'il répète des mots et des phrases de manière inappropriée, ou encore il utilise des mots non adaptés aux situations ou accole des syllabes sans signification. Parfois, ce sont des gestes répétés de manière compulsive qui peuvent inquiéter les parents, des jeux mécaniques comme l'alignement d'objets ou des intérêts pour les objets plus que pour les personnes. Les jeux d'alignement d'objets ou les répétitions de gestes ou de phrases peuvent se présenter comme de véritables obsessions. Il arrive qu'un tout petit changement de leur routine, comme un léger déplacement d'objet les contrarie de manière totalement inhabituelle, voire entraîne leur colère (Masi *et al.*, 2017). Pour les parents, ces accès de colère sont souvent très impressionnants et difficiles à vivre.

Le diagnostic du trouble du spectre de l'autisme demeure un défi et ne peut être fait avant l'âge de deux ans. Auparavant, d'autres diagnostics devront être éliminés, comme l'anxiété, la timidité, le mutisme, le retard de langage, et le déficit intellectuel auxquels le TSA pourra être associé. Par ailleurs, le TSA peut accompagner certains syndromes génétiques comme le syndrome 22q11, le syndrome du X-Fragile, l'agénésie du corps calleux, et certaines autres anomalies du système nerveux central. Le TSA peut aussi être associé au trouble déficitaire de l'attention. Les étiologies du TSA, quant à elles, sont multifactorielles; cependant, tous les enfants atteints du TSA présentent soit des anomalies anatomiques, soit des anomalies fonctionnelles cérébrales (Doshi-Velez *et al.*, 2014). Dans les études, on parlera d'anomalies fonctionnelles au niveau des synapses neuronales puisque plusieurs gènes identifiés chez les enfants avec TSA sont responsables de protéines impliquées dans la transmission synaptique. Notamment, la cause

génétique la plus fréquente de TSA est le syndrome du X Fragile, lui-même causé par l'absence d'une protéine encodée par le gène FMR1, laquelle anomalie entraîne des dysfonctionnements au niveau des synapses (Penzes *et al.*, 2011). Certaines études s'intéressent également au mauvais fonctionnement du système sérotoninergique, puisque, normalement, la sérotonine contribue au bon fonctionnement de nombreux processus neurobiologiques et au développement cérébral (Cook et Leventhal, 1996). Il apparaît également que l'ocytocine, une hormone qui module plusieurs comportements sociaux dont l'instinct maternel, l'attachement avec les pairs, la mémoire sociale et la reconnaissance des visages pourrait jouer un certain rôle dans cette pathologie (Ross et Young, 2009). D'autres études scrutent le métabolisme de la vitamine D et constatent que l'administration de cette vitamine à haute dose peut augmenter les taux de sérotonine et de GABA et ainsi améliorer certains symptômes du TSA (Feiyong *et al.*, 2018).

Ainsi, puisque les étiologies et les mécanismes causals évoqués sont multiples, aucun test biologique ne peut confirmer ou infirmer ce diagnostic : pour le moment, le diagnostic du TSA demeure un diagnostic clinique. Ce diagnostic clinique commande le travail d'une équipe pluridisciplinaire composée de psychologues, psychiatres, pédiatres, neurologues, orthophonistes, audiologistes, psychoéducateurs et travailleurs sociaux. Cette évaluation implique le recensement de l'histoire du développement psychomoteur de l'enfant, ainsi que des observations cliniques de son comportement et de sa communication et socialisation au moyen d'instruments validés pour le faire¹. L'ADOS (*Autism diagnostic observation schedule*) est un outil validé pour confirmer le diagnostic du TSA chez les enfants de plus de deux ans et chez les adolescents. Il s'agit d'un outil d'observation semi-structuré. Il comporte quatre modules dont chacun permet d'évaluer le développement langagier à des étapes différentes du développement de l'enfant. Chaque module possède son propre protocole avec des activités adaptées selon l'âge chronologique et le stade de développement. Lors de ces tests d'évaluation, l'observateur notera le jeu de l'enfant en scrutant des aspects précis de son développement. Comment sont

1. ADOS-2 Autism Diagnosis Observation Schedule-2. Les activités standardisées de l'ADOS-2 permettent d'évaluer la communication, l'interaction sociale réciproque, le jeu et/ou l'utilisation créative d'un matériel, le comportement stéréotypé, les intérêts restreints et d'autres comportements atypiques, chez des participants avec autisme ou présentant un trouble du spectre de l'autisme (TSA) (allant d'enfants de 12 mois à des adultes possédant un langage fluide [5 modules]).

les interactions de l'enfant, interagit-il avec son évaluateur, établit-il un contact visuel avec celui-ci? Arrive-t-il à formuler des demandes? Parle-t-il de ses intérêts? L'enfant a-t-il un jeu symbolique? Peut-il faire semblant de faire manger les poupées par exemple? Peut-il imiter son évaluateur en répétant la même tâche? Utilise-t-il le jouet de la manière attendue pour un enfant de son âge? Aligne-t-il les objets au lieu de répondre aux consignes? Répète-t-il des gestes de manière non adaptée au contexte du jeu? À la fin du test, l'évaluateur devrait être en mesure de documenter les habiletés de communication de l'enfant, ses interactions sociales réciproques, son jeu, ses comportements stéréotypés et ses intérêts restreints. Il faut comprendre que tous les enfants avec TSA ne performeront pas de manière identique au test de l'ADOS, puisque chaque enfant présente une variante du tableau clinique correspondant au spectre de cette pathologie.

Devant la diversité des causes et des manifestations cliniques, aucun traitement pharmacologique définitif ou génétique n'existe pour le TSA. On parlera de thérapie comportementale qui devra être débutée le plus précocement possible et appliquée le plus intensivement possible dans le but d'obtenir les meilleurs résultats, soit améliorer le pronostic fonctionnel et la qualité de vie actuelle et future de ces enfants. Sachant que le TSA peut induire plusieurs problèmes psychiatriques et médicaux sérieux à l'âge adulte, une thérapie précoce efficace est impérative. Les études sur les interventions parents-enfants dans le cas des enfants de bas âge démontrent un effet positif sur plusieurs aspects du développement social de ces enfants; cependant, l'évidence d'un gain persistant à long terme reste à démontrer (Oono *et al.*, 2013).

À la question «pourquoi ne pas utiliser des robots à la rescousse des enfants autistiques?» il est pertinent de répondre que, considérant qu'aucune thérapie parfaitement efficace n'existe, cette stratégie technologique peut sembler justifiée. Il pourrait également apparaître logique de déduire que l'attrait des enfants avec TSA pour les objets, et parallèlement l'anxiété qu'engendrent chez eux les contacts sociaux, légitimise cette solution en apparence prometteuse. Ainsi, puisque le robot est un objet, son utilité thérapeutique reposerait dans le fait précis qu'il est non humain, bien que conçu pour lui ressembler. Avec un comportement programmé pour être d'humeur égale et prévisible, le robot constitue un atout indéniable pour capter l'attention de ces enfants avec TSA et serait un avantage (Scasellati, 2012). L'hypothèse de leur utilisation repose donc sur le fait que le robot, en étant plus semblable à l'objet qu'à l'humain, est moins menaçant et que, de ce

fait, il captivera davantage l'enfant avec TSA que le visage humain et sera susceptible de susciter davantage de réactions et d'interactions. Encore faut-il cependant que cet effet bénéfique soit reproductible à plus ou moins long terme dans les interactions de l'enfant TSA avec de vrais humains.

Mais qu'en est-il, dans les faits, de l'efficacité des robots dans les interventions ciblant les habiletés sociales des enfants présentant un TSA ? Les résumés d'articles de recherches qui rapportent l'utilisation thérapeutique expérimentale de robots auprès de ces jeunes patients révèlent que les robots sont utilisés tantôt comme partenaires de jeux, tantôt comme initiative de bon comportement ou comme médiateur ou acteur dans la relation (Diehl *et al.*, 2012). On divise donc leur application selon 4 catégories : 1- Agent initiateur de comportements ; 2- Modèle d'habiletés relationnelles à imiter pour l'enfant TSA ; 3- Agent de rétroaction de performance ; 4- Répondant à l'interaction à titre d'humain. Même si ces initiatives semblent globalement prometteuses dans la rééducation des enfants autistes, plusieurs auteurs nous mettent en garde quant au fait que certains enfants ne répondront pas de manière favorable aux robots. En fait, l'intervention des robots pourrait aider certains enfants pour certains aspects de leur interaction avec le robot, mais les bénéfices réalisés dans les interactions réelles avec de vrais humains ne sont pas concluants. À l'heure actuelle, ces robots ne sont pas en mesure d'offrir des soins personnalisés aux enfants avec TSA. De plus, l'absence d'études cliniques randomisées valides ne permet pas d'arriver à des recommandations solides ou consensuelles pour leur utilisation (Institut universitaire en déficience intellectuelle et en trouble du spectre de l'autisme, 2015).

D'un point de vue clinique, il peut effectivement sembler stratégique à première vue d'utiliser le robot comme objet d'intérêt ou de fascination pour amener ces enfants à entrer en relation, d'abord avec un pseudo-humain, puis progressivement avec un autre humain à part entière. Si la preuve devient irréfutable que le robot humanoïde représente un tremplin vers l'épanouissement d'une relation entre un enfant avec TSA et un autre humain, les avantages ne seront pas négligeables. On devra malgré tout faire la preuve d'une absence d'inconvénients graves pour l'enfant. Cependant, il faudrait démontrer auparavant que cette thérapie dyade enfant avec TSA-robot est plus bénéfique que la dyade enfant avec TSA-autre humain, ce qui jusqu'à présent n'a pas été démontré. N'y a-t-il pas un danger à présumer sans preuve que toute utilisation des robots est à priori inoffensive ?

Il nous semble primordial de nous interroger d'abord sur les conséquences négatives potentielles d'une approche impliquant le robot dans une thérapie de rééducation des enfants TSA. Cette intervention nous renvoie au concept même de la relation thérapeutique, et par extension à la relation humaine qui est la base de cette relation d'aide. Cette dimension représente la finalité de la thérapeutique impliquée, soit de permettre à cet enfant de devenir un être de relation avec ses semblables. Entrer en relation demande effectivement aux interlocuteurs de communiquer, laquelle communication dépend de leur capacité à percevoir leurs émotions réciproques. Il existe une distinction pratique entre la communication et la création d'une relation vivante et dynamique. La communication est l'échange d'un message entre des interlocuteurs. Ce que peuvent faire les robots par programmation de phrases circonstancielles joyeuses ou neutres. Cependant, la relation est un lien qui se crée entre des personnes qui communiquent de manière signifiante pour l'une et pour l'autre. Cette relation, pour être significative pour chaque interlocuteur, implique un référent émotif et une transmission d'émotions par rapport à ce référent. Une relation implique la démonstration d'une variété d'émotions. Dans le cas du robot, elles sont possiblement neutres, ce qui entre en contradiction avec l'essence même de l'humain, qui est un être d'émotions, et qui s'implique dans toute relation avec ces émotions. La difficulté relationnelle de l'enfant avec TSA et le but de la thérapie réfèrent directement à cette composante émotive de la relation humain-humain. Les auteurs mentionnent précisément que la constance des interventions du robot et la palette neutre d'émotions qu'il transmet réconfortent l'enfant avec TSA. Cette uniformité dans la communication du robot, provenant du fait que la composante émotive en est absente, évacue la notion d'imprévisibilité et celle d'unicité qui sont propres à la relation humaine. Le robot n'entre pas en relation au sens humain du terme. Il est un objet animé et programmé qui parle à un enfant, et cet enfant lui parle justement parce qu'il n'est pas humain.

Certains n'y verront que du bien, prétextant que l'utilisation de cet objet thérapeutique peut être considérée comme une modalité pouvant ultimement conduire à une véritable relation enfant-humain. Mais il faudrait prouver d'abord que cette modalité d'intervention est supérieure à la thérapie enfant-autre humain, et exiger sans concession que cette thérapie robot-humain ne soit jamais offerte seule, sans la thérapie d'intervention humain-humain, afin que l'enfant puisse faire la différence entre les deux types de relations, puisque l'une est

«fausse» et l'autre est «vraie». Finalement, il faut se rappeler que cette relation enfant-robot ne constitue pas une finalité en soi, puisqu'il ne s'agit pas d'une relation humaine, mais bien d'un faux-semblant. Un robot humanoïde, aussi perfectionné qu'il soit ou qu'il puisse le devenir, ne sera jamais aussi imparfait et imprévisible dans ses réactions que l'humain, dont les réactions et les interactions ne sont pas toujours raisonnables ni logiques, et qui sont même parfois totalement irrationnelles. Malheureusement, ce sont ces relations-là, imparfaites, que nous devons affronter au quotidien et auxquelles nous devrions préparer les enfants avec TSA afin qu'ils puissent prendre leur place dans la société telle que nous la connaissons. Il n'existe aucune évidence de la transférabilité des gains thérapeutiques, si gains il y a, dans des situations relationnelles «naturelles» au quotidien. Cela dit à moins que nous désirions évacuer l'imprévisibilité comme dimension implicite dans les relations humaines, et à moins que la relation humain-robot devienne la nouvelle norme relationnelle à atteindre. Éventuellement, à force de ne pas nous méfier de ce que représente l'utilisation des robots dans les interactions, nous allons entrer de manière progressive dans une nouvelle ère posthumanité. Une nouvelle ère dont nous ne pouvons mesurer pleinement les effets au moment présent et dont peu de chercheurs semblent se préoccuper. Un futur de science-fiction où la notion de relation entre humains irait bien au-delà de la réalité espérée par le transhumanisme, celle où seul le corps biologique devrait être amélioré et non pas l'essence même de ce qui constitue l'homme social. Et c'est selon moi cette toute dernière question, lorsqu'on évalue les effets de l'intervention des robots dans les relations d'aide auprès d'enfants, qui devrait précéder toutes les autres : dans quel monde désire-t-on que nos enfants vivent ? Et si l'homme social est une valeur que nous désirons préserver, alors est-il éthique ou même moral de faire intervenir le robot comme figure représentative de l'homme social ?

BIBLIOGRAPHIE

- APA. (2013). *Diagnostic and statistical manual of mental disorders* (5^e éd.). Washington, DC, American Psychiatric Publishing.
- Cook, E.H., Jr, Leventhal, B.L. (1996). «The serotonin system in autism». *Curr Opin Pediatr*, 8, p. 348-354.
- Diehl, J-J. Crowell, CR, Villano M, *et al.* (2012). «The clinical Use of Robots for individuals with Autism Spectrum Disorders ; A Critical Review». *Res Autism Spectr Disord*, 6(1), p. 249-262.

- Doshi-Velez F, Ge Yaorong, Kohane I., (2014). « Comorbidity Clusters in Autism Spectrum Disorders: An Electronic Health Record Time-Series Analysis », *Pediatrics*, 133 (1), p. e54-e63.
- Elsabbagh, Mayada, Gauri Divan, Yun-Joo Koh, Young Shin Kim, Shuaib Kauchali, Carlos Marcín, Cecilia Montiel-Nava, Vikram Patel, Cristiane S. Paula, Chongying Wang, Mohammad Taghi Yasamy et Eric Fombonne, (2012). « Global Prevalence of Autism and Other Pervasive Developmental Disorders », *Autism Research*, 5 (3), p. 160-179.
- Feiyong, Jia, Ling Shan, Bing Wnag *et al.*, (2018). « Bench to bedside review: Possible role fo vitamin D in autism spectrum disorder ». *Psychiatry Research*, 260, p. 360-365.
- Institut universitaire en déficience intellectuelle et en trouble du spectre de l'autisme, (2015). *L'efficacité du robot dans les interventions ciblant les habiletés sociales des enfants présentant un TSA*. Les collections de l'Institut universitaire en DI et en TSA. Gouvernement du Québec.
- Masi Anne, Marilena M, De Mayo, Glozier Nicholas, Guastella Adam J., (2017). « An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options ». *Neurosci. Bull.* Avril, 33(2), p. 183-193.
- Oono, I.P., Honey, E.J., McConachie, H., (2013). *Parent-mediated early intervention for young children with autism spectrum disorder (ASD) (Review)* Cochrane Library.
- Penzes, P, Cahill, M.E, Jones, K.A., VanLeeuwen J.E., Woolfrey K.M., (2011). « Dendritic spine pathology in neuropsychiatric disorders ». *Nat Neurosci*, 14, p. 285-293.
- Ross, H.E., Young, L.J., (2009). « Oxytocin and the neural mechanisms regulating social cognition and affiliative behavior ». *Front Neuroendocrinol*, 30, p. 534-547.
- Scasellati, B., (2012). « Robots for use in autism research ». *Annual Review of Biomedical Engineering*, 14, p. 275-294.

CHAPITRE 14

Robots soignants en gériatrie : réalité ou illusion ?

Félix PAGEAU

INTRODUCTION

Depuis l'apparition d'abord controversée des robots en usine, leur présence dans nos vies est grandissante et de moins en moins critiquée. Après le remplacement des militaires par des drones, la prochaine limite, que l'industrie semble vouloir franchir, est le robot qui prend la place des soignants - particulièrement auprès des personnes âgées. Mais pourquoi cette population précise est-elle ciblée? Les soins des gens vieillissants sont ardues physiquement pour les employés et peuvent être perçus comme répétitifs à certains égards. Les tâches de donner un bain par semaine, d'aider lors des transferts et des déplacements sont assez difficiles. Or, le remplacement ou l'aide aux soignants par les robots mènent à plusieurs questionnements éthiques.

1. DES ROBOTS POUR LE SOIN DES PERSONNES ÂGÉES : MISE EN CONTEXTE

Reprenons pour commencer les facteurs qui influencent l'intégration des robots dans les soins gériatriques et quelques éléments majeurs qui justifient leur emploi.

1.1 Les données sociodémographiques

Il existe plusieurs influences sociétales qui participent à la montée de l'utilisation des robots dits « soignants » en gériatrie. Ces facteurs sont multiples et souvent reliés. Les plus importants selon un certain nombre d'auteurs sont : le taux de natalité qui baisse dans plusieurs pays industriels ; le fait que de plus en plus de femmes sont sur le marché du travail et que la population mondiale vieillit (Chan, 2015 ; Sharkey et Sharkey, 2012 ; Sparrow, 2006 ; Vandemeulebroucke *et al.*, 2017). En effet, les femmes ont actuellement librement accès aux méthodes de contraception, font moins d'enfants ou les ont plus tardivement (Sharkey et Sharkey, 2012). La diminution des naissances entraîne une diminution du nombre d'enfants pouvant éventuellement s'occuper de leurs parents vieillissants. Les descendants, moins nombreux donc, ont eu moins de contacts privilégiés avec leur mère ou leur père, surtout, s'ils travaillaient à l'extérieur du foyer. Aussi, les enfants du capitalisme veulent travailler davantage pour bien gagner leur vie. Un cercle vicieux de la valorisation sans borne du travail (même au coût d'une famille unie) est engendré dans lequel les enfants ne sentent guère le devoir de s'occuper de leurs parents après s'être sentis eux-mêmes délaissés. De plus, les femmes qui traditionnellement s'occupaient de leurs aïeux n'ont plus le temps de le faire, puisqu'elles doivent travailler à l'extérieur et s'occuper de leurs propres enfants (Sharkey et Sharkey, 2012). Et malheureusement, les hommes ne semblent pas avoir pris la relève de leurs compagnes. Ce point est important, car la faute n'est pas imputable aux femmes. L'entrée de celles-ci sur le marché du travail n'a pas créé de sortie équivalente des hommes prenant leur relève au foyer. Ces derniers auraient dû aider davantage aux tâches traditionnellement laissées aux femmes, mais ils ne l'ont pas nécessairement fait partout.

Par ailleurs, la population mondiale vieillit, et cela, même dans certains pays en voie de développement, ce que note l'Organisation mondiale

de la santé de façon presque alarmante dans son rapport de 2015 sur le vieillissement mondial (Chan, 2015, p.3). Un fort déséquilibre advient entre les besoins de prises en charge des aînés et la capacité des familles de les combler. Un problème culturel et démographique émerge et la solution semble, pour certains, l'utilisation de robots.

1.2 *Hype*, travail et transhumanisme

La réponse rapide des constructeurs de robots se traduit par la mise en marché de machines automatisées qui semblent pouvoir combler les besoins de soins des populations vieillissantes. Évidemment, le *hype* mondial que crée la robotique n'épargne aucun pays industrialisé et dit « moderne ». Le gouvernement de Justin Trudeau annonçait en 2017 un investissement de 125 millions de dollars pour l'intelligence artificielle, dont un montant de 40 millions sur cinq ans pour le Québec (Benessaïeh, 24 mars 2017), ceci afin de contrer un retard technologique perçu par rapport aux États-Unis. Et pour faire face au défi de recrutement des experts du domaine, « [l]'axe le plus important de la stratégie est donc "d'attirer et retenir les meilleurs", explique Denis Therrien, de l'ICRA [International Conference on Robotics and Automation]. Comment faire ? Il n'y a pas de secret : en donnant à ces experts les budgets, les ressources et des salaires compétitifs par rapport aux Google et Facebook de ce monde. » (Benessaïeh, 24 mars 2017, p. 3) Étant donné que la compétition est très forte dans le domaine de la robotique et de l'intelligence artificielle, il nous faudrait donc, selon cette stratégie, suivre le modèle états-unien dominant. Il est intéressant de noter que dans les centres hospitaliers de soins de longue durée (CHSLD) du Québec, la compétition est au contraire plutôt faible pour l'obtention d'un emploi. Les quelques employés recrutés sont épuisés et leurs capacités de travail poussées à l'extrême (Moisan, 17 mars 2017). Outre le phénomène du *hype* ainsi que la réalité du travail accablant pour les soignants en gériatrie, la vague de l'idéologie transhumaniste participe à l'essor de la robotique. Pour Julian Huxley, le « transhumanisme » soutient l'idée selon laquelle un humain doit dépasser ses capacités biologiques par la technique (Huxley, 2015 p. 14). Les exosquelettes, les robots soignants et la maison intelligente sont vus comme des extensions du corps humain devant aider l'humain à maîtriser son environnement.

1.3 Les raisons médicales en faveur de l'emploi du robot

Malgré ce contexte favorable, comment les robots pourraient-ils prodiguer des soins? En gériatrie, il est évident que surtout les personnes âgées qui souffrent de problèmes de mobilité et de démence pourraient bénéficier du soutien des robots.

Trouble de mobilité et chutes

La fréquence des troubles de mobilité avec ou sans chute augmente avec l'âge, car les causes possibles s'accumulent avec le temps chez une même personne. Mentionnons entre autres : l'arthrose, les troubles neurologiques (accident vasculaire cérébral [AVC], atteintes nerveuses, vertige et maladie de Parkinson), la baisse de tension artérielle debout (hypotension orthostatique), les médicaments (surdose, prise inadéquate ou insuffisante), les pertes de muscles, la perte de vision, l'alcoolisme ou les troubles reliés à la démence. Cette longue énumération de causes de chutes démontre bien le nombre important de maladies multiples qui peuvent atteindre la population gériatrique.

Ces différentes pathologies peuvent aussi entraîner une grande quantité de problèmes avec des effets variés néfastes et menant même, à la mort. Or, certains considèrent que la mort peut devenir une issue favorable à un âge avancé. Il faut donc considérer d'autres complications graves des chutes pour comprendre à quel point ce problème peut être délétère. Mentionnons, d'abord, les fractures, dont celle de la hanche, qui sont très souffrantes et limitatives. Les saignements intracrâniens, l'écrasement des muscles, la peur de retomber, les vertiges sont d'autres complications qui, sans causer la mort, peuvent amener un lot important de souffrance physique ou psychologique. Ces problèmes prédisposent aussi à d'autres chutes. L'utilisation de robots peut-elle traiter et prévenir de tels problèmes? La réponse n'est pas évidente.

Trouble neurocognitif majeur : critères du DSM 5

La fréquence des troubles neurocognitifs¹ augmente avec l'âge de la population étudiée (Mordoch, 2013, p. 15). Un robot qui stimule la cognition ou aide à exécuter des tâches rendues difficiles par les pertes de mémoire pourrait aider la personne âgée, selon certains auteurs (Vandemeulebroucke, 2017, p. 1 ; Mordoch, 2013, p. 18). Il est important

1. Nouvelle dénomination qui remplace le terme de «démence» dans les textes médicaux.

de clarifier ce que sont les critères diagnostiques des troubles neuro-cognitifs (TNC) majeurs pour le comprendre. Il existe quatre critères et quelques spécifications pour diagnostiquer une démence selon le *Manuel diagnostique et statistique des troubles mentaux* (DSM) récemment mis à jour par l'Association américaine de psychiatrie. La perte d'autonomie fonctionnelle est constatée lorsque la personne est incapable de faire une des activités de la vie domestique (AVD), soit : faire seul(e) le ménage, la lessive, les repas, les courses, prendre ses médicaments, gérer son argent, conduire un véhicule et utiliser des appareils électroniques (dont un ordinateur ou un robot!). Les activités de la vie quotidienne (AVQ) regroupent : se nourrir, s'habiller, se laver, se déplacer et changer seul(e) ses protections en cas d'incontinence.

2. AVANTAGES ET INCONVÉNIENTS DES ROBOTS DANS LES SOINS AUX PERSONNES ÂGÉES

Il faut mieux comprendre quels sont les principaux arguments du débat actuel concernant l'utilisation des robots dits « soignants » chez les gens atteints de démence.

2.1 La sécurité et l'autonomie fonctionnelle (AVD, AVQ)

La sécurité est un élément primordial en faveur de l'intégration des robots dans les soins en gériatrie étant donné que le fonctionnement de la personne atteinte se détériore progressivement (critère B du TNC majeur). Par exemple, certains malades atteints de démence oublient de fermer les ronds de la cuisinière après les avoir utilisés ou programment trente minutes au lieu de trois sur la minuterie du four à micro-ondes. Ces oublis et ces erreurs placent les gens malades, ainsi que leurs voisins, à risque d'incendie et de décès par le feu. Une cuisinière robotisée ou une maison équipée de détecteurs automatisés pourraient réduire grandement le risque de feu. Les auteurs d'une intéressante étude ont analysé le comportement des participants souffrant de démence qui devaient faire du thé sous la supervision d'un robot. Pour certains, le robot a été très utile, mais pas pour tous (Begum, 2013). Les équipes de Ya-Hue Wu et de Sandra Bedaf ont d'ailleurs écrit sur les besoins ressentis par les gens âgés dans leur étude qualitative. L'appel automatique au centre d'aide, un support cognitif et un système de sécurité sûr sont des éléments prioritaires chez les robots selon ce qu'ont déclaré les personnes âgées questionnées

(Wu *et al.*, 2010; Bedaf, 2013). L'équipe d'Alan Pearce dans une revue de la littérature de 2012 (Pearce, 2012) démontre une certaine utilité des robots pour l'aide dans les tâches ménagères.

Les troubles de mobilité et les chutes peuvent être diminués par les robots. Ils peuvent alors augmenter la sécurité des gens atteints. Selon Rajani Sadasivam et son équipe, un robot commandé à distance envoyé au domicile des gens atteint de troubles de mobilité détecte plus d'éléments qui mettent une personne à risque de chute qu'un observateur humain seul (Sadasivam, 2014). Les robots étudiés peuvent aussi améliorer la mobilité des survivants d'accidents cérébraux vasculaires (Buesing, 2015; Schück, 2012). Par ailleurs, des lits robotisés se sont révélés utiles, selon la compagnie Reysone qui propose un lit se transformant en fauteuil roulant. La mobilité est améliorée par cette technique. Il existe ainsi de nombreuses applications possibles des robots pour aider la mobilité selon toutes ces études. Par contre, il ne faut pas oublier que ces machines demeurent au stade expérimental, pour la plupart, et que plus de données probantes seront nécessaires pour conclure à leur pertinence et à leur sécurité réelle.

La domotique prend aussi de l'ampleur et permet un contrôle à distance de plusieurs domaines de la maison (éclairage, alarme de feu, chauffage, etc.). En gériatrie, la surveillance avec la domotique par caméras, microphones ou capteurs peut permettre aux familles d'avoir un œil sur leurs proches atteints de troubles de mobilité, de chutes répétées, d'incontinences ou d'oublis fréquents. Un service d'urgence peut être rapidement appelé au besoin. Le tout se fait automatiquement ou à distance et permet aux gens qui ont perdu la capacité cognitive d'utiliser un téléphone de rester chez eux plus longtemps. Tom Sorell, un des auteurs qui prônent l'utilisation de la robotique dans les soins aux personnes âgées, écrit : « la sécurité est suprême, dépassant même l'autonomie [trad. libre] » (Sorell et Draper, 2014, p. 192). Pour lui ainsi que pour Linda et Robert Sparrow, une maison dite « intelligente » permet aux personnes qui souffrent de démence ou de troubles de mobilité de rester seules à leur domicile plus longtemps (Sparrow, 2006, p. 141 ; Sorell et Draper, 2014, p. 183). Les AVD (faire seul(e) le ménage, la lessive, les repas, les courses, prendre sa médication, gérer son argent, conduire un véhicule et utiliser des appareils électroniques) et AVQ (se nourrir, s'habiller, se laver, se déplacer et changer seul(e) ses protections) seraient prises en charge par les robots. Pour Sorell, au fur et à mesure que l'intelligence artificielle se développera, les commandes données aux robots seront de plus

en plus centrées sur les besoins de l'utilisateur. Il pourra choisir un horaire plus adapté à ses besoins sans dépendre de soignants ou de sa famille (Sorell et Draper, 2014, p. 189). Les auteurs y voient un rôle d'apprentissages interdépendants (Sorell et Draper, 2014, p. 189). Dans cette programmation, le robot reconnaît de plus en plus les besoins et les limites de l'utilisateur, et ce dernier s'adapte tranquillement au robot. Différents robots ont d'ailleurs été étudiés à cette fin. Le robot de l'étude de Momotaz Begum, que nous avons mentionné plus haut, aidait les gens atteints de démence à faire du thé (Begum, 2013); les conclusions de l'étude stipulent que les robots n'ont pas bien fonctionné pour cette application par contre. Le robot FRASIER étudié par l'équipe de Taskin Padir semblait de son côté pouvoir aider à effectuer des tâches domestiques tel un servante robotisée (Padir, 2015). Pour une autre équipe de chercheurs qui a répertorié les études sur le sujet, les robots apparaissent utiles, mais on note que les recherches analysées étaient petites et peu robustes (Pearce, 2012). L'étude menée sur Robadom, un projet européen d'incorporation des robots dans les soins, a démontré que plus de travail était à faire pour une intégration harmonieuse de robots dans les soins, selon les personnes âgées interrogées (Wu, 2013). Plusieurs modèles sont à l'étude actuellement. Mentionnons le Robear, un robot avec l'apparence d'un ours, qui aide lors des déplacements et des transferts de la position assise à la position debout. ACCOMPANY est un autre exemple de robot-serviteur présenté par Sorell (Sorell, 2014). L'utilisation d'une tablette électronique étant cependant nécessaire, l'utilisateur ne doit donc pas avoir d'atteinte cognitive pour pouvoir se servir du robot. Des limitations importantes apparaissent dans les études au point de vue technique et la recherche actuellement est plutôt décevante. Les vendeurs de robots tentent toutefois de promouvoir leurs produits avec la désapprobation de certains.

2.2 Sécurité et liberté

D'ailleurs, de nombreuses critiques se sont élevées en réaction à l'éloge de la sécurité que font les auteurs favorisant l'usage des robots. Selon Linda et Robert Sparrow, plusieurs raisons historiques et pratiques rappellent que ces machines préprogrammées peuvent être dangereuses. Ces auteurs mettent en lumière le fait que dans les usines de montage, les robots sont isolés parce qu'ils sont jugés trop risqués pour les travailleurs. Pourquoi en serait-il autrement avec les robots soignants? Les bénéficiaires de soins pourraient faire des chutes en

se prenant le pied dans un robot ou en étant renversés par ce dernier. Les utilisateurs pourraient aussi être écrasés ou frappés par le robot «soignant» en raison d'une erreur de programmation ou par malchance (Sparrow, 2006, p. 146). Est-ce que les robots obéiront si leur propriétaire leur demande de l'assistance pour leur suicide? Quelles sont les limites acceptables qu'un robot pourrait imposer à son utilisateur? Le contrôle des apports alimentaires caloriques pour éviter l'obésité, le retrait de tout alcool et l'empêchement de sortir sont des limitations qui semblent exagérées et qui donnent même froid dans le dos lorsqu'on y pense bien. Un robot paternaliste, est-il souhaitable? Amanda et Robert Sharkey écrivent que cela revient à un «emprisonnement dans sa maison sans procès» (Sharkey, 2012, p. 32). Est-ce que le robot doit respecter l'autonomie ou la sécurité? Une maison équipée de capteurs, de senseurs, d'un système de caméra crée pour la personne malade une perte évidente sur le plan de la vie privée, une autre atteinte à son autonomie.

2.3 La vie privée

Tom Sorell et Heather Draper répondent à cette épineuse notion de la protection de la vie privée par le consentement éclairé (Sorell, 2014, p. 187). La personne qui accepte d'être ainsi surveillée le fait de son plein gré et après avoir été adéquatement informée. Les bénéfices et les désavantages sont initialement bien expliqués et ils sont acceptés. Un des avantages est l'absence physique d'étranger dans la résidence.

Au lieu de relations sociales tout court, des relations sociales choisies apparaissent plus désirables. C'est l'utilisateur qui décide, comme la majorité des adultes le font quotidiennement, qui inclure et qui ne pas inclure dans son cercle privé. [trad. libre] (Sorell et Draper, 2014, p. 192)

Les contacts seront choisis et limités aux personnes significatives. De plus, les soins d'hygiène pourront être plus privés, car, si ce sont les robots qui les effectuent, aucun être humain ne verra le bénéficiaire de soins nu et dans des positions de vulnérabilité. En outre, la surveillance peut prendre les techniques des jeux vidéo qui ne détectent que les mouvements sans enregistrer le visage de l'utilisateur ni aucune caractéristique physique reconnaissable.

2.4 Les relations sociales

En plus de la vie privée préservée par différentes techniques de programmation informatique, la robotique et la domotique pourraient augmenter les relations significatives. Effectivement, plusieurs auteurs mentionnent l'augmentation de l'accès aux médias sociaux, au web et à des moyens de communication rapides pour rester en contact avec ses proches (Sparrow, 2006, p. 147 ; Sorell et Draper, 2014, p. 184). Ainsi, seulement les moments agréables seraient partagés entre un malade et son entourage; les soins étant réservés aux robots. La robotisation des soins semble contrer la fatigue et le manque d'empathie des proches qui se sont épuisés à soigner leurs parents vieillissants (Sorell et Draper, 2014, p. 185). Tom Sorell et Heather Draper mentionnent même que l'utilisation des robots soignants pourrait amener une diminution des mauvais traitements et de la négligence (Sorell, 2014, p. 185). Par un maintien du « masque empathique » (Sorell et Draper, 2014, p. 189), les robots seraient capables d'interagir avec les personnes âgées et de les rendre plus sociables ou, du moins, de maintenir leurs acquis relationnels. Les exosquelettes actuellement utilisés en réadaptation pourraient aussi aider les gens âgés à se déplacer vers leurs proches. La méthode de Siena (Sorell et Draper, 2014, p. 192), qui consiste à ajuster le ton de la voix du robot à celui du patient, pourrait même améliorer les habiletés sociales des patients.

Quelques robots animaux ont été étudiés pour aider à réduire la solitude des gens âgés ou à diminuer les problèmes de comportements associés à la démence. PARO (bébé phoque), Aibo (chien) et NeCoRo (chat) sont parmi les robots qui ont été les plus utilisés. PARO est d'ailleurs le sujet de plusieurs petites études positives en gériatrie. L'une d'entre elles démontre une augmentation des interactions sociales des participants de l'étude en contact avec PARO (Moyle *et al*, 2013). Une petite étude randomisée contrôlée a aussi démontré la diminution de la sensation de solitude et même celle des symptômes dépressifs chez les personnes qui avaient plus de liens avec PARO (Robinson *et al*, 2013). Néanmoins, l'étude de Melanie Birks et ses collègues demeure critique par rapport aux effets réels de PARO (Birks *et al*, 2016). Roger Bemelmans et son équipe ont révisé la littérature pour répertorier les études sur les robots animaux. Leurs résultats ont démontré une tendance positive pour l'utilisation d'Aibo et de Paro, mais une tentative infructueuse pour NeCoRo (Bemelmans R. G. *et al*, 2012). Une étude plus récente du même auteur remet par contre en doute cette idée

(Bemelmans R. G. *et al*, 2016). Chez des patients âgés sans démence, les robots semblent aider, selon une étude de 2008 de Marian R Banks (Banks, 2008). L'équipe d'Élaine Mordoch mentionne dans son étude de 2013 que l'utilisation d'un robot en soins de longue durée peut augmenter l'attention et la collaboration des patients avec démence (Mordoch *et al*, 2013). D'autres chercheurs ont révélé que les robots avaient aussi démontré une certaine utilité pour garder les patients éveillés en soins de longue durée (Peri *et al*, 2015). Inversement, les chercheurs de l'équipe d'Elizabeth Broadbent remettent en doute leur utilité dans leurs études des patients en gériatrie (Broadbent *et al*, 2016). La littérature scientifique ne semble pas pouvoir actuellement trancher la question, malgré qu'une tendance à l'utilisation des robots semble émerger en gériatrie dans les pays industrialisés. Bien que ces données soient préliminaires, on voit apparaître dans certains CHSLD le robot Paro, ce qui est inquiétant, considérant le manque de données probantes.

2.5 La fiabilité

Bien que les robots semblent utiles, selon ces études, certains problèmes bien connus quant à la fiabilité des ordinateurs peuvent les toucher, comme c'est le cas de tout type de machines informatisées. Des problèmes de fonctionnement peuvent survenir ; cela fait partie des inquiétudes des participants âgés de l'étude de Vandemeulebroucke *et al*. (2017, p. 16). Le piratage, les problèmes de virus, et même l'espionnage des propriétaires de robots, pourraient se produire et réduire grandement la sécurité des gens qui en possèdent un. Les caméras, les microphones et les capteurs sont autant de moyens, pour des esprits mal intentionnés, de s'infiltrer dans les maisons des gens âgés vulnérables. Par ailleurs, l'emploi de techniciens et la réparation de ces nombreux problèmes ne sont pas sans coût. Enfin, trouver des mesures de sécurité de plus en plus renforcées demeure un obstacle à l'utilisation à grande échelle de telles technologies. En Allemagne, Cayla, une poupée destinée aux enfants et qui était connectée à Internet, a ainsi été interdite de vente par crainte d'espionnage ou de sévices par des pédophiles (TVA Nouvelles, 2017).

2.6 Adaptation aux personnes avec démence

Une autre critique importante de la robotisation des soins réside dans l'importance clinique du critère A de la démence, soit le « déclin cognitif dans au moins une sphère [trad. libre] » (Association américaine de psychiatrie, 2013). Les pertes progressives qui adviennent chez les personnes atteintes de démence sont un défi de taille pour la robotique actuelle. Cette atteinte lente et graduelle requiert l'usage de robots qui font preuve d'adaptations constantes. Il faut à la fois qu'ils évitent l'infantilisation des aînés, mais aussi la négligence de leurs besoins de soins. Actuellement, les systèmes d'exploitation et les algorithmes informatiques ne sont absolument pas en mesure de répondre à de tels besoins. Un simple appel téléphonique auprès d'une grande compagnie de services révèle rapidement l'inefficacité des répondeurs automatiques et leur champ d'application restreint, et ce, auprès de gens sans déficience cognitive ; il suffit d'imaginer dans cette situation la confusion des gens souffrant de démence pour comprendre les limites actuelles de la robotique.

En médecine, un facteur majeur d'incertitude et de variabilité est « le facteur humain ». Chaque être humain porte avec lui un bagage d'émotions, de problèmes, de maladies pas encore diagnostiquées ni traitées. Le facteur humain est cette valeur unique, complexe et incalculable que chacun amène dans la relation médecin-patient et qui rend le tout impossible à standardiser en algorithme. Par exemple, un patient qui consulte pour une sinusite, lors d'un rendez-vous planifié de 10 minutes, révèle à son médecin qu'il est déprimé. Le médecin a besoin d'au moins une heure pour bien comprendre de façon empathique les symptômes dépressifs et les idées suicidaires du patient, dresser un plan de traitement et de suivi. Le temps planifié initialement était franchement insuffisant par rapport au temps que le médecin a finalement consacré au patient. Le médecin est inévitablement en retard, ce qui était imprévisible, car le patient n'a révélé ces symptômes qu'à la toute dernière minute. La médecine demande une grande adaptabilité de la part du clinicien, surtout en gériatrie. Actuellement, la puissance des robots est insuffisante pour rendre compte des réalités des besoins des patients âgés ayant de multiples maladies complexes, ce qu'on appelle en médecine « le facteur humain ».

2.7 Les coûts

Les robots se vendent aussi à des prix prohibitifs pour la majorité des systèmes de soins de santé actuels (Peri *et al*, 2015). Les prix devront considérablement baisser pour que ces machines deviennent attrayantes. C'est une considération importante pour les gouvernements et pour les personnes qui prennent soin des gens âgés qui souffrent de démence ou de troubles de mobilité. Les employés du milieu de la santé qui s'occupent des gens âgés sont souvent parmi les travailleurs les moins bien payés. Sparrow cite la Fédération australienne des infirmières: «Workers in the aged-care sector are already by-and-large paid very low wages.» (Australian Nurses Federation, 2005, citée par Sparrow, 2006, p. 149). Il existe donc une différence importante entre le salaire présentement versé aux infirmières et le montant que devront déboursier les acheteurs de ces robots soignants.

3. REPRISE DES CRITIQUES DES ROBOTS DITS «SOIGNANTS»

Il existe donc divers avantages, mais aussi plusieurs inconvénients qui viennent contredire l'utilisation des robots dits «soignants». La littérature scientifique récente sur les robots énonce la sécurité, l'autonomie fonctionnelle, la vie privée et les relations sociales pour justifier l'utilisation des robots. Ceux qui critiquent l'emploi de machines automatisées mentionnent le risque du bris de la sécurité (bogues, virus, espionnage), une atteinte de la liberté, un manque d'adaptabilité et des coûts non compétitifs. Or, il semble que pour les défenseurs des robots (aidés du *hype*), les critiques peuvent être résolues par le développement de l'intelligence artificielle, l'augmentation de la puissance computationnelle des machines et le financement de la recherche. Il faut pourtant bien comprendre les autres problèmes possibles, soulevés par les critiques, qui ne se résolvent pas par le développement de telles technologies.

Pour les producteurs de robots, un développement plus avancé de leur science pourrait permettre une grande adaptabilité des automates informatisés, une compréhension des émotions et même une production (externe ou interne) d'émotions par la machine, ce que les auteurs anglophones nomment «*emotionality*» et que nous traduisons par émotivité. Il est important de revenir sur les éléments de l'émotivité externe et interne pour comprendre comment une telle fonction mène à une dérive dangereuse. L'émotivité est le fait de générer une

émotion. Elle peut être externe au robot, c'est-à-dire qu'elle est vécue seulement par l'utilisateur. Cette dernière peut, aussi, émaner du robot (interne). La machine n'est plus qu'un algorithme complexe qui calcule, elle aurait de véritables émotions, comme l'humain. Or, pour l'instant en robotique, l'émotivité demeure externe. L'émotion est uniquement vécue par l'utilisateur et non par la machine. Le robot fait croire en une émotion qui suscite à son tour l'émotivité de la personne humaine qui lui fait face. Par exemple, un robot pourrait être programmé à dire : « Je t'aime ». Or, il n'aime pas vraiment, mais l'être humain qui entend une telle phrase chargée d'émotion peut développer un sentiment qui s'approche de l'amour du robot. C'est une relation à sens unique. Le robot ne procure qu'un reflet factice dans la relation. L'illusion qu'amènerait un robot dit « social » par l'émotivité externe est néfaste. Ces robots sont conçus pour tenir compagnie, pour comprendre les émotions de la personne et pour lui faire vivre une émotion. Pour Linda et Robert Sparrow, vouloir aller si loin avec le robot est grave. Ces humanoïdes programmés sont toujours un mensonge et un pauvre simulacre de relations parce qu'ils ne ressentent rien ; ils tentent de faire vivre l'émotion à l'utilisateur sans rien réellement ressentir quoi que ce soit (émotivité externe) (Sparrow, 2006, p. 155). La relation est unidirectionnelle. L'humain se voit dans le robot, comme Narcisse à la surface de la marre. La relation se trouve entre l'utilisateur et son image miroitée par le robot. Une telle relation à soi-même par l'entremise du robot se fait toujours au détriment des vraies relations humaines, comme pour le cas des robots animaux (Paro, NeCoRo, Aibo, etc.). La personne âgée isolée qui est en présence du robot n'est plus considérée comme une fin en soi, car on la place devant ce miroir robotisé pour ne plus avoir à en prendre soin. Il est alors évident que la personne et sa dignité sont ignorées, qu'on ne considère plus cette personne comme notre semblable avec qui il faut avoir une relation humaine véritable. Le patient n'est donc plus traité en tant qu'être humain dans toute sa dignité, mais plutôt comme un fardeau pour lequel on touche un salaire. Et le robot devant lequel on place la personne âgée facilite cette tâche. La personne devient ainsi un « moyen » pour obtenir une fin (salaire). Pour Amanda et Noel Sharkey, le patient âgé est victime d'une certaine « objectification » (Sharkey, 2012, p. 30) ; il est détenu dans cette relation fautive et c'est un crime grave. Ils écrivent :

Un droit tel est peut-être inscrit dans l'article 5 de la Déclaration universelle des droits de la personne : « Nul ne sera soumis à la torture, ni à des peines ou traitements cruels, inhumains ou »

dégradants»; ou dans l'article 9: «Nul ne peut être arbitrairement arrêté, détenu ou exilé» [Trad. libre] (Sharkey et Sharkey, 2012, p. 29).

De plus, le robot soignant pourrait être utilisé par la famille pour diminuer un sentiment de culpabilité. Dans les mots des auteurs, cela peut s'exprimer ainsi: «Ne t'inquiète pas pour Mamie, elle a le robot pour lui tenir compagnie [trad. libre]» (Sharkey et Sharkey, 2012, p. 35). Ainsi, un robot-compagnon risque d'isoler indûment la personne âgée, ce qui va contre les droits de cette dernière en tant que personne légale. C'est d'ailleurs ce que des participants à l'étude de Tijs Vandermeulebroucke *et al.* (2017, p.16) ont mentionné. La crainte de se retrouver seul avec son robot est grande pour les principaux intéressés.

Nous avons mentionné ci-haut quelques éléments problématiques de l'utilisation des robots: isolement, faux-semblant, perte sur le plan de la dignité humaine. Un autre aspect important touche le manque d'une vraie relation humaine. La relation sincère avec une entité qui ressent réellement des émotions se forme au profit des deux parties de la relation. Deux individus qui se réalisent dans une relation humaine développent une connexion et une satisfaction à aider un autre être sensible. La valeur de cette relation se retrouve inhérente aux émotions partagées. Les vraies relations ont une valeur parce qu'elles demandent un effort constant pour être maintenues. Elles sont exigeantes, imparfaites et renferment une part de contingence. Elles sont aussi véritables. La valeur de vérité est essentielle. Ces caractéristiques sont-elles programmables? Le *deep-learning* pourrait-il être la solution de la valeur de la relation humain-robot?

Peut-être, mais il faut aussi considérer dans la relation la mort comme fin de la vie. Elle est impossible à prévoir avec certitude et donne une valeur forte à la vie qui la précède. Il existe toujours un risque de mort d'une des parties de la relation. La fin abrupte possible fait valoir que chaque instant partagé compte dans cette relation éphémère. Elle place une valeur dans la personne que l'on peut perdre subitement et sans qu'il y ait possibilité de retour en arrière. Cette vision est théorisée par la philosophe Gaëlle Fiasse avec le concept de fragilité. Elle écrit: «La fragilité provient d'une tension entre deux traits spécifiques: ce qui est précieux et ce qui est cassable. Elle renvoie à la possibilité d'une perte de quelque chose qui suscite l'admiration et le respect.» (Fiasse, 2012, p. 9) Le robot n'est pas vivant, il est remplaçable, réparable, reprogrammable. C'est un objet sans vie. Avoir de l'affection pour un robot

relève du fétichisme, car l'aimer est identique à l'excitation ressentie par certains à la vue d'un soulier, d'un godemichet ou d'un objet en cuir. C'est confondre la partie pour le tout ou l'outil pour l'humain. Une forme certaine d'anthropomorphisme accompagne cette idée. Les gens qui « aiment » leur voiture, l'appellent « leur bébé » et la dorlotent comme la prunelle de leurs yeux sont des exemples frappants de l'anthropomorphisme fréquent chez l'homme.

Les concepteurs informatiques diront, peut-être, que si les robots ont des émotions inhérentes (émotivité interne), il sera donc acceptable et possible d'avoir une relation vraie avec des entités humanoïdes, robotisées, rationnelles et sensibles. Une question morale encore plus forte se pose alors : « Est-ce éthiquement acceptable d'utiliser un être raisonnable et sensible pour notre bon plaisir et pour diminuer nos souffrances ? » D'abord, la robotique est très loin d'être arrivée à créer de tels êtres. Il faut faire preuve de beaucoup d'imagination et de foi en la technique pour croire que les roboticiens y arriveront un jour. Cela dit, même si la science ne nous a pas encore menés aussi loin, la morale nous somme d'arrêter avant. Si créer de tels individus est possible, il est dangereux de laisser les roboticiens le faire. Kant, dans sa *Métaphysique des mœurs*, a énoncé qu'il faut traiter tout humain « comme une fin, et jamais simplement comme moyen » (Kant, 1993, p. 105). Les prérequis de l'impératif catégorique sont que l'homme est un être rationnel et sensible. Si les robots deviennent aussi sensibles et rationnels, devons-nous les traiter comme des fins en soi ? Il est dangereux de créer de tels êtres qui nous seraient égaux par leurs caractéristiques (raison et sensibilité). Nous avons la possibilité d'arrêter avant de nous rendre à une telle finalité, qui demeure spéculative au plus haut point.

3.1 Le jugement clinique : aspect critique

Un autre aspect à considérer dans la suite des propos sur les émotions externes et internes concerne le jugement clinique. En médecine, le remplacement des médecins par des robots soignants est quasi impossible, malgré le développement du *deep-learning*. En effet, le jugement clinique, qui est souvent effrayant pour un étudiant en médecine, est impossible à programmer. Il intervient lorsque le clinicien d'expérience arrive aux limites des lignes directrices cliniques des grandes associations de sa spécialité. Souvent, on énonce le jugement clinique quand il y a un manque d'évidences scientifiques pour appuyer une conduite

thérapeutique ou lorsque le patient ne correspond pas parfaitement à la population étudiée lors de l'élaboration des lignes directrices. Le jugement clinique se forme à partir de l'addition d'expériences favorables et défavorables avec les patients. C'est un concept près de la *phronésis* d'Aristote² (*Éthique à Nicomaque*). Il était peut-être lui-même, inspiré par son père médecin, qui sait ! Il n'y a évidemment pas d'algorithme possible pour vraiment rendre justice au jugement clinique, car il ne correspond pas à la sagesse (*sophia*) qui s'enseigne, mais à autre chose de plus concret et variable, ajusté à chaque situation unique.

3.2 Un argument majeur : l'analyse sociale

Bien que les aspects techniques et cliniques semblent suffisants pour éviter l'utilisation de robots soignants autonome, une raison encore plus fondamentale remet en question leur utilité. Il faut comprendre pourquoi les robots ont été créés. Paul Dumouchel reprend les raisons qui ont mené à l'essor des machines robotisées dans *Vivre avec les robots – Essai sur l'empathie artificielle*. Il rappelle comment le travail en usine est abrutissant et épuisant. Les robots ont été mis en place pour remplacer l'homme dans des tâches automatiques et nécessitant un travail éreintant (Dumouchel, 2016, p. 39-43). Ils participent à la production de biens et sont de loin plus efficaces que l'homme dans les tâches répétitives. Dans notre société capitaliste, les robots sont de fiers représentants des paradigmes de productivité et d'efficacité.

Aussi, puisque le travail est si valorisé, la majorité des gens travaillent pour trouver une place, leur valeur dans la société. Nous demandons souvent à un étranger : « Qu'est-ce que vous faites dans la vie ? » Un certain malaise est ressenti si la réponse n'est « rien de spécial ». Avoir un travail est un premier moyen de rentrer en contact avec l'autre, se faire valoir. La pression sociale est forte pour que des parents travaillent toujours plus pour subvenir aux besoins de leurs enfants. Dans une société du travail et de la productivité, les parents sont poussés à les délaisser (souvent contre leur gré) pour travailler plus et avoir un meilleur niveau de vie. Alors, ces fils et ces filles ne développent pas

2. La *phronésis* est : « un état vrai, accompagné de raison, qui porte à l'action quand sont en jeu les choses bonnes ou mauvaises pour l'homme » (*Éthique à Nicomaque*, 1140b, 5), c'est une sagesse pratique bâtie par l'expérience et qui nous permet d'atteindre un équilibre.

une relation significative suffisante pour vouloir prendre soin de leurs parents qui vieillissent.

Les robots soignants semblent être une bonne solution pour éviter cette tâche perçue comme un fardeau non mérité. De plus, les robots participent à l'idée de productivité et d'efficacité. Si un individu veut travailler plus et s'occuper moins de ses parents vieillissants, le robot soignant est un choix évident, mais ô combien illusoire ! Le robot ne sait pas prendre soin comme nous l'avons démontré plus haut. Il imite, il reflète et il trompe. Les soins ne peuvent être prodigués que par un être sensible et rationnel, un être humain. La robotique est loin d'être rendue à ce point de perfectionnement, et elle ne devrait pas s'y rendre non plus. La société devra changer pour donner au robot le rôle qui lui revient ontologiquement, c'est-à-dire faire des tâches répétitives, éreintantes et abrutissantes, ce qui est l'inverse de celui de procurer des soins de qualités adaptés et sagaces (*phronèsis*).

Si les recherches en intelligence artificielle aboutissent à de tels robots, ils seront soit traités comme des sous-humains ou des personnes (au sens juridique). C'est insensé, car nous avons besoin des robots pour les tâches exténuantes, répétitives et dangereuses et que nous n'aurons plus le droit de leur en faire exécuter s'ils deviennent des personnes juridiques. De plus, les robots sociaux sont inutiles, puisque les êtres humains sont déjà les meilleurs pour se tenir compagnie entre eux, pour se soigner ou pour diminuer mutuellement leur solitude.

3.3 L'environnement

Enfin, il est impossible de parler de robots et d'éviter la question de l'environnement. L'environnement ne peut qu'être affecté par une production effrénée de robots, même les plus utiles. Il faut demeurer conscient des répercussions de leur construction sur l'environnement. La production de robots requiert une quantité importante d'énergie et de métaux rares. L'absence de limitation du point de vue de l'environnement est dangereuse. La production devrait se limiter aux robots franchement utiles, ce que ne sont pas les robots soignants autonomes et sensibles. Seul le robot-outil demeure pertinent pour réduire la charge de travail des professionnels de la santé.

CONCLUSION

Les changements culturels, sociaux et démographiques ont motivé un développement accru de nouvelles technologies en robotique pour remplacer certains travailleurs par des robots. La prévalence de la démence et des troubles de mobilité qui augmente avec l'âge inquiète les différentes associations, dont l'Organisation mondiale de la santé (Chan, 2015, p. 3). La perte d'autonomie fonctionnelle, le désengagement des familles et l'importance du travail mènent à la tentative d'insertion de robots dits «soignants» en gériatrie. De nombreuses études tentent de démontrer l'importance qu'on devrait leur accorder. Toutefois, la science demeure assez limitée pour répondre de façon certaine sur leur réelle utilité. D'un point de vue médical gériatrique, il faut rester prudents, avant de les utiliser, car trop peu de données probantes existent à l'heure actuelle pour assurer leur utilisation sécuritaire.

Sur le plan de l'éthique, certains éléments sont amenés pour justifier leur emploi, soit la sécurité, la dénatalité et le maintien à domicile des gens en perte d'autonomie fonctionnelle (AVQ et AVD). Toutefois, des contre-arguments tout aussi majeurs émergent de l'analyse de l'utilisation des robots. La protection de la vie privée est mise à risque par les bogues informatiques, le piratage et l'espionnage qui mettent au défi la fiabilité des robots - dont nous devons douter. Les relations interpersonnelles vraies sont mises à l'épreuve par l'utilisation de robots dits «sociaux» qui isolent plus qu'ils ne connectent les gens entre eux. De plus, les coûts sont prohibitifs et les robots sont peu adaptés à la réalité clinique des patients âgés. La robotique semble répondre à ces critiques par la mise de l'avant d'un développement plus avancé de la machine. Le progrès des nouvelles technologies semble la réponse aux ratées en robotique. Certains défenseurs de la robotique sociale supposent qu'ils parviendront à créer des robots sociaux qui remplaceront les soignants humains, ce que nous désapprouvons d'un point de vue médical, éthique et ontologique.

Selon les arguments ontologiques, que nous avons mis en lumière, il apparaît évident que les robots peuvent être des outils utiles, mais ils ne doivent jamais devenir des compagnons ni des êtres sensibles. Ils perdraient leurs sens comme outils. Le mensonge lié à l'*émotivité* externe justifie aussi l'absence de telles programmations, pour éviter l'anthropomorphisme et le fétichisme envers les robots dont souffrirait beaucoup l'humain-utilisateur. De tels robots seront ce que l'homme

a créé de plus similaire à lui-même – un être rationnel et sensible. Au mieux, ce seront des sous-humains esclaves parce qu'ils seront des créations humaines. Au pire, ils deviendront des personnes (au sens juridique), des personnes qui seraient faites de silicone, de métal et d'huile, mais qui ressentiraient les mêmes peines, les mêmes souffrances et les mêmes difficultés émotives que l'homme de chair, d'os et de sang. Ces deux possibilités sont tout aussi inutiles qu'insensées. C'est futile, car des créatures sensibles et rationnelles existent déjà sur Terre – les êtres humains (cela ne vaut pas la peine d'en inventer d'autres). Et c'est insensé, puisque les robots sont utiles à l'homme pour les tâches exténuantes, répétitives et dangereuses qu'il est impossible d'imposer à une personne juridique; or, c'est ce que deviendraient des robots sociaux sensibles et rationnels (cela contreviendrait alors aux droits et libertés de la personne). L'être humain est suffisant pour soigner ses semblables et le robot n'est que son outil, jamais son remplaçant.

Il faut valoriser et mieux payer les travailleurs du domaine des soins auprès de la population gériatrique, plutôt que l'industrie des robots dits « sociaux ». La proche-aide et le travail des professionnels en gériatrie doivent être revalorisés pour attirer plus de gens à aider les personnes âgées. C'est véritablement la voie à emprunter. Nous sommes convaincus qu'il faut éviter la robotique sociale, du moins en gériatrie.

Nous avons donc démontré que le but des robots et leur finalité sont ceux de l'outil, des moyens pour une fin. Pour nous, il est évident que les robots, même dits « soignants », ne devraient pas devenir des êtres sensibles ni des compagnons. De telles créatures, qui produisent volontairement des émotions chez l'humain-utilisateur, augmentent les faux-semblants et relèvent de la supercherie. Les robots ont, selon nous, un sens en tant qu'outil et non comme compagnon, encore moins comme amis. L'humanité risque des dérives majeures s'ils deviennent des fins en soi en devenant capables de ressentir des émotions (encore que la robotique n'y soit pas du tout parvenue). Il est important également de prendre en considération les relations vraies et la protection de l'environnement pour éviter la productivité déchaînée; n'oublions pas que les robots sont les émissaires puissants de la productivité et de l'efficacité, lesquelles entrent souvent en conflit avec les préoccupations écologiques de maintien des écosystèmes et de protection de la Nature et de relations humaines véritables.

BIBLIOGRAPHIE

- Aristote (2014). *Éthique à Nicomaque*, (R. Bodéüs, Trad.) Paris, France, GF Flammarion.
- Association américaine de psychiatrie (2013). *DSM-5, Manuel diagnostique et statistique des troubles mentaux*, Arlington, VA, États-Unis.
- Banks, M. R. *et al.* (2008). «Animal-Assisted Therapy and Loneliness in Nursing Homes: Use of Robotic versus Living Dogs», *Original Studies*, mars, vol. 9, p. 173-177.
- Banks, W. A. (2013). «Artificial Emotions: Robots Caring for the Elderly», *JAMDA*, vol. 14, p. 635-636.
- Bedaf, S. *et al.* (2013). «Selecting Services for a Service Robot Evaluating the problematic activities threatening the independence of elderly persons». *IEEE International Conference on Rehabilitation Robotics du 24-26 juin 2013*, Seattle: IEEE, p. 1-6.
- Begum, M. *et al.* (2013). «Performance of Daily Activities by Older Adults with Dementia: The Role of an Assistive Robot», *IEEE International Conference on Rehabilitation Robotics du 24-26 juin 2013*, Seattle: IEEE, p. 1-8.
- Bemelmans, R. *et al.* (2012). «Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness», *JAMDA*, vol. 13, p.114-120.
- Bemelmans, R. *et al.* (2016). «How to use robot interventions in intramural psycho-geriatric care; A feasibility study», *Applied Nursing Research*, vol. 30, juillet, p. 154-157.
- Benessaïeh, K. (2017). «Intelligence artificielle: Montréal recevra 40 millions d'Ottawa», *La Presse*, 24 mars. Consulté en ligne.
- Birks, M. *et al.* (2016). «Robotic Seals as Therapeutic Tools in an Aged Care Facility: A Qualitative Study. *Journal of Aging Research*, octobre, p. 1-7.
- Bolduc, A. *et al.* (2013). «Enquête auprès des responsables des Unités de courte durée gériatriques (UCDG) sur leurs ressources professionnelles: ratios actuels et niveau de satisfaction», *Rapport du RUSHGQ*, Québec, Librairie nationale.
- Broadbent, E. *et al.* (2016). «Research Benefits and problems of health-care robots in aged care settings: A comparison trial», *Australasian Journal on Ageing*, vol. 35, n° 1, mars, p. 23-29.
- Buesing, C. *et al.* (2015). «Effects of a wearable exoskeleton stride management assist system (SMA®) on spatiotemporal gait characteristics in individuals after stroke: a randomized controlled trial», *Journal of NeuroEngineering and Rehabilitation*, vol. 12, n° 69, août, p.1-14.
- Chan, M. (2015). «Préface» dans *Résumé: rapport mondial sur le vieillissement et la santé*, de l'Organisation mondiale de la santé, Luxembourg, Organisation mondiale de la santé.
- Dumouchel, P. et L. Damiano (2016). *Vivre avec les robots – Essai sur l'empathie artificielle*. Paris, Seuil.
- Fiasse, G. (2015). *Amour et fragilité – Regards philosophiques au cœur de l'humain*. Québec, Canada, Presse de l'Université Laval.
- Huxley, J. (2015). «Transhumanism», *Ethics In Progress*, vol. 6, n° 1, p. 12-16.
- Kant, E. (1993). *Fondement de la métaphysique des mœurs*. (V. Delbos, Trad.) Paris, Librairie Générale Française.

- Moisan, M., 2017, « *Recherche préposés désespérément* », *Le Soleil*, 16 mars 2017. Consulté en ligne.
- Mordoch, E. *et al.* (2013). « Use of social commitment robots in the care of elderly people with dementia: A literature review », *Maturitas*, vol. 74, octobre, p. 14-20.
- Moyle, W. *et al.* (2013). « Exploring the Effect of Companion Robots on Emotional Expression in a Pilot Randomized Controlled Trial », *Journal of Gerontological Nursing*, vol. 39, n° 5, p. 46-53.
- Mugueta-Aguinaga, I. *et al.* (2017). « Is Technology Present in Frailty? Technology a Back-up Tool for Dealing with Frailty in the Elderly: A Systematic Review », *Aging and Disease*, vol. 8, n° 2, p. 176-195.
- Padir, T., J. Skorinko et V. Dimitrov (2015). « User-Centric Design of a Personal Assistance Robot (FRASIER) for Active Aging », *IEEE International Conference on Rehabilitation Robotics du 25-29 août 2015*, Milan: IEEE, p. 5020-5023.
- Pearce, A. J. *et al.* (2012). « Review Article – Robotics to Enable Older Adults to Remain Living at Home », *Journal of Aging*, vol. 2012, octobre, p. 1-10.
- Peri, K. *et al.* (2015). « Lounging with robots – social spaces of residents in care: A comparison trial. Australasian Journal on Ageing », *Australasian Journal on Ageing*, vol. 35, n° 1, mars, p. 1-6.
- Robinson, H. *et al.* (2013). « The Psychosocial Effects of a Companion Robot: A Randomized Controlled Trial », *JAMDA*, vol. 14, p. 661-667.
- Sadasivam, R. *et al.* (2014). « Robot-assisted home hazard assessment for fall prevention: a feasibility study », *Journal of Telemedicine and Telecare*, vol. 20, n° 1, p. 3-10.
- Schück, A. *et al.* (2012). « Feasibility and effects of patient-cooperative robot-aided gait training applied in a 4-week pilot trial », *Journal of NeuroEngineering and Rehabilitation*, vol. 9, n° 31, mai, p. 1-14.
- Sharkey, A. et Noel Sharkey, 2012, “Granny and the robots: Ethical issues in robot care for the elderly”, *Journal Ethics and Information Technology*, vol. 14, n° 1, juillet 2010 (publié en ligne), pages 27-40.
- Sorell, T. et Heather Draper (2014). « Robot carers, ethics, and older people », *Ethics Information and Technology*, vol. 16, n° 3, mars, p. 183-195.
- Sparrow, S. L. (2006). « In the hands of machines? The future of aged care ». *Mind Mach*, 16(2), p. 141-161.
- TVA Nouvelle, 2017, « *Une poupée parlante... espionne?* », *TVA Nouvelle*, 18 février. Consulté en ligne.
- Vandemeulebroucke, T. *et al.* (2017). « How do older adults experience and perceive socially assistive robots in aged care: a systematic review of qualitative evidence », *Aging and Mental Health*, février, p. 1-19.
- Wu, Y.-H. *et al.* (2013). « Le projet Robadom: conception d'un robot d'assistance pour les personnes âgées », *Revue de Gériatrie*, vol. 38, n° 5, mai 2013, p. 349-353.
- Wu, Y.-H. *et al.* (2010). « Robotic agents for supporting community-dwelling elderly people with memory complaints: Perceived needs and preferences ». *Health Informatics Journal*, vol. 17, n° 1, p. 33-40.

CHAPITRE 15

Un amour de robot : robot émotionnel et travail d'aidant

Monique LANOIX

«Et pourtant, pourtant, Je n'aime que toi...»

Charles Aznavour, «Et pourtant».

Dans ce texte, nous nous interrogeons sur l'utilisation de robots émotionnels destinés aux personnes âgées en milieu institutionnel. La croissance de la population vieillissante a déjà des répercussions importantes sur les soins de santé. Les médias nous avertissent de cette transformation sociale et de ses conséquences depuis plusieurs années. Les personnes âgées ont très souvent besoin d'aide pour accomplir les tâches de tous les jours et cette aide, si elle provient en partie de la famille, doit souvent être complétée par des travailleurs rémunérés. Pour cette raison, les sociétés qui font face à ces changements démographiques s'inquiètent du manque de main-d'œuvre. Plusieurs pays ont recours à des travailleurs d'autres pays, mais cette main-d'œuvre demeure limitée. Une des solutions envisagées est de se tourner vers les robots. Le Japon est un des chefs de file dans ce domaine et a déjà créé plusieurs modèles de robots aidants.

D'une façon générale, la technologie fait déjà partie des services de santé. Elle peut être assez simple, comme dans le cas des lève-patients, ou plus complexe, par exemple les aides médicales robotisées. La

recherche scientifique propose encore d'autres avenues dans le secteur de la robotisation, dont les robots émotionnels qui sont conçus pour aider des enfants dans leur apprentissage ou encore pour tenir compagnie à des personnes âgées. Le robot qui fera l'objet de notre analyse se nomme PARO, un phoque robotique utilisé avec les personnes qui souffrent de démence.

Nous avançons que l'utilisation d'aides robotiques émotionnelles est le signe d'un malaise profond face au travail rémunéré des aidants qui interviennent auprès des personnes âgées. Plus précisément, l'aide robotique illustre deux tendances qui sont caractéristiques de nos sociétés industrielles. La première est d'isoler la personne âgée, même si le but premier du robot émotionnel est d'encourager l'engagement de la personne âgée dans son milieu de vie. Deuxièmement, l'apport du robot fait en sorte que le travail de l'aidant devient un travail encore plus marchandisé. En effet, en nous basant sur les travaux de Plessner et d'Agamben, nous allons montrer comment le travailleur se trouve réduit au statut d'outil plutôt que de travailleur responsable de prendre soin des malades.

1. POURQUOI PARO ?

PARO, un phoque robotique, est décrit comme suit : « recouvert d'une fourrure bactéricide, PARO est équipé de 7 moteurs, qui lui permettent de bouger la tête (haut, bas, droite, gauche), cligner des yeux, remuer la queue et actionner ses deux nageoires latérales. Le son de sa voix provient d'un réel enregistrement de bébé phoque et la tête de PARO est renforcée en polycarbonate (utilisé pour les vitres pare-balles) afin d'assurer un niveau de robustesse en adéquation avec l'environnement d'utilisation. Une douzaine de capteurs (toucher, positionnement, lumière) et 3 microphones (détection de la provenance du son par triangulation) renvoient des informations sur l'interaction avec le malade à un logiciel d'intelligence artificielle, qui adapte en conséquence les mouvements et l'intonation de PARO afin de fournir à chaque malade la meilleure stimulation cognitive possible. PARO peut donc communiquer au patient des émotions telles que la joie, la surprise ou le mécontentement » (<http://www.phoque-paro.fr/>).

Sur le marché depuis 2005, PARO est conçu pour les personnes âgées atteintes de démence, car ces personnes ont plus de difficulté à rester en contact avec le monde qui les entoure. On craint, entre

autres, que ces personnes se retirent complètement du monde et n'aient plus aucune interaction avec les autres. De plus, lorsque la maladie progresse, les personnes atteintes de démence n'ont souvent plus la capacité de parler, ce qui fait qu'elles peuvent facilement être délaissées. Les milieux de soins de longue durée ont quand même la vocation de prendre soin de toute la personne ou de la personne dans son ensemble, ce qui veut dire, à un certain niveau, donner des soins de nature sociale. Ainsi, comme le souligne le centre d'hébergement Grace Dart à Montréal :

[l]a mission d'un centre d'hébergement et de soins de longue durée est d'offrir, de façon temporaire ou permanente, un milieu de vie substitut, des services d'hébergement, d'assistance, de soutien et de surveillance ainsi que des services de réadaptation, psychosociaux, infirmiers, pharmaceutiques et médicaux aux adultes qui, en raison de leur perte d'autonomie fonctionnelle ou psychosociale, ne peuvent plus demeurer dans leur milieu de vie naturel, malgré le support de leur entourage. (CHSLD Grace Dart)

Pour cette raison, les centres de soins de longue durée organisent des activités qui visent à faciliter l'engagement de ces personnes¹.

Cependant, avec les compressions budgétaires, ces activités sont souvent éliminées afin de privilégier les soins plus fondamentaux tels que l'alimentation, les soins médicaux et corporels. Les soins de nature sociale ne sont pas considérés comme des soins aussi essentiels que les soins médicaux. Puisque le personnel est souvent surchargé, un outil comme PARO apparaît tout à fait approprié. Il semblerait que PARO peut faire le travail d'un aidant, car il peut rester dans les bras de la résidente et la garder sinon heureuse du moins calme. D'ailleurs, plusieurs études montrent que l'utilisation de PARO réduit le stress et l'anxiété chez ces personnes (Petersen *et al.*, 2017) et même diminue la consommation de médicaments (Jøranson *et al.*, 2014). Ces interventions, qui ont lieu quelques fois par semaine, sont d'une durée de 20 minutes et se font en présence d'une thérapeute qui encourage les résidents à interagir avec PARO. Les études mentionnent aussi que l'utilisation des robots est préférable aux animaux à cause des allergies, des risques de blessures, etc.

1. Au Québec, les personnes qui requièrent plus de trois heures de soins par jour sont hébergées dans des centres de soins de longue durée ou CHSLD (Santé Montréal, <https://www.santemontreal.qc.ca/reseau-de-la-sante/ressources/ressources-dhebergement/chsld/>)

La question qui se pose ici ne porte pas tant sur les effets de PARO, considéré comme une aide à la communication gérée par un intervenant. Nous voulons plutôt réfléchir à un autre aspect, celui de la substitution d'un humain par un robot émotionnel. L'expression même de robot émotionnel nous amène à nous demander d'où proviennent ces émotions. Remarquons, toutefois, qu'il ne s'agit pas des émotions que le robot pourrait avoir (ou semble avoir), mais plutôt des émotions que ce dernier suscite chez une personne. En effet, le robot peut sembler avoir des émotions, car il bouge, interagit et a même un air de vulnérabilité, mais il ne s'agit que de mouvements, de semblants d'émotions, que la personne regardant PARO interprète comme étant authentiques. Dans la mesure où le robot produit des émotions factices, il fait surgir par mimétisme des émotions chez la personne qui interagit avec lui².

Nous avançons que le risque lié à l'utilisation de PARO est de substituer l'interaction du robot, qui est toujours docile, à l'interaction humaine. Nous proposons d'analyser l'interaction de madame Körner avec son robot PARO, comme on peut l'observer dans le documentaire *Mechanical Love*, de Phie Ambo (2007).

Le documentaire s'intéresse à madame Körner, une personne autonome qui réside dans un centre d'accueil en Allemagne. On la voit interagir avec PARO: elle lui parle, et il semble lui répondre à sa manière. Lorsqu'elle entre dans sa chambre, il crie; il semble la reconnaître et paraît heureux qu'elle soit de retour. Il est évident qu'elle a un lien d'affection avec PARO; il est son compagnon. Elle décide même de l'amener à une soirée de chants dans une des salles communes où sont installées d'autres résidentes pour célébrer la fête de Noël. Puisque PARO réagit au son, il est bruyant durant les chansons et interrompt la soirée constamment. Cela pousse les autres résidentes à chuchoter à propos de madame Körner et de PARO. Ces résidentes la traitent de folle et, de son côté, madame Körner s'enorgueillit et rassure PARO en lui disant qu'ils sont bien tous les deux ensemble. Madame Körner trouve refuge dans son rapport avec PARO, qui est en vérité un rapport à sens unique, mais qui la rassure tout de même.

2. Ces effets ont été reconnus récemment dans un article paru dans le *Washington Post* (Geoffrey Fowler, 2017), qui discute du sentiment d'attachement que les robots émotionnels provoquent, comme si nous devenions programmés pour l'amour.

Ce petit scénario illustre bien les dangers du robot émotionnel, qui a pour effet d'isoler davantage la personne. Dans ce cas-ci, puisque le robot devient la cible de critiques, madame Körner désire protéger son être cher. Le rapport avec PARO devient plus important que celui qu'elle entretient avec les autres résidentes. L'attachement de madame Körner pour PARO est au centre de sa vie. Puisqu'elle vit en milieu institutionnel, les rapports qu'elle peut entretenir avec d'autres personnes sont limités, d'où l'importance de la garder engagée dans son milieu. Sauf qu'en compagnie de son robot, elle se referme sur elle-même. Elle sait bien que PARO est un robot. Elle ne souffre pas de démence ; cependant un dérapage se produit, car le robot offre un semblant d'attachement qui, en retour, isole madame Körner. La question qui se pose est donc de savoir si ce rapport est nocif.

Le rapport pourrait être nocif s'il se substitue à tout autre rapport. Les parents d'enfants qui utilisent des robots préfèrent que les robots soient des compagnons de jeu, car il n'y a pas « de prise de responsabilité du robot, ni une trop grande intimité avec l'enfant » (Johal *et al.*, 2015, p. 133). En effet, selon cette étude, les parents ne veulent pas être remplacés par un robot dans des moments qu'ils considèrent comme privilégiés avec leurs enfants et souhaitent aussi conserver leur rôle de parent (p. 133). Mais dans le cas d'une personne en hébergement, très souvent les liens intimes ont disparu. Plus personne ne craint d'être remplacé par un robot. L'intervenant qui aide madame Körner à regagner sa chambre n'a pas cette inquiétude. Il est là simplement pour la guider et pour s'assurer qu'elle est bien installée pour la nuit. On voit dans le film que madame Körner apprécie l'aide du préposé, mais que son souci premier est de veiller à parler avec PARO. Il est facile de voir pourquoi elle est attirée par PARO, car il semble attentif à ces mouvements. Madame Körner est dans un centre d'accueil et est par conséquent déjà isolée en quelque sorte du reste de la société. Substituer PARO à une interaction avec d'autres êtres humains semble un pas de plus vers la ségrégation totale de madame Körner.

Le danger de tisser des liens avec un robot émotionnel est aussi discuté par Turkle *et al.* (2006). Ces auteurs analysent le rapport du robot émotionnel avec les enfants et les personnes âgées et ils reconnaissent que le cyberaccompagnement est complexe. Un des robots utilisés dans cette étude est My Real Baby. Les auteurs qualifient l'interaction de cette poupée robotique comme étant une fenêtre sur le passé pour les personnes âgées (« Seniors : robots as a prism for the

past», 2006, p. 354-357). Un des résidents, Andy, traite My Real Baby comme si celle-ci était consciente et il veut aussi en prendre soin. Pour les auteurs, Andy substitue le robot aux êtres humains qu'il aimerait avoir dans sa vie (2006, p. 356). Les auteurs poursuivent leur étude en analysant l'effet que PARO produit sur un groupe de personnes âgées. PARO suscite de l'admiration, de la curiosité et même un désir d'attachement. Les résidents demandent à PARO s'il a faim ou s'il s'endort et le robot semble leur répondre (2006, p. 357). Dans un livre paru en 2011, Turkle écrit que nous sommes programmés pour vouloir prendre soin de ce que l'on aime, mais aussi pour aimer ces êtres dont nous prenons soin³. Turkle déplore le fait que la société ne semble pas être troublée par cet attachement au robot⁴.

Ces auteurs ne veulent pas bannir complètement tout robot émotionnel. Cependant, ils avancent qu'un des dangers est celui de la liminalité dans le cas des robots émotionnels. Nous expliquons : pour ces auteurs, l'être humain crée et entretient un rapport avec un robot qui semble être engagé auprès de cet humain. Selon les auteurs, ce rapport n'est pas un vrai rapport, mais on ne peut pas non plus affirmer qu'il n'y a aucun rapport entre madame Körner et PARO, par exemple. En ce sens, nous pouvons dire que le rapport est liminal, car il est placé à la frontière d'un rapport authentique et d'un rapport fictif. De plus, pour Turkle *et al.*, ces robots sont des artefacts relationnels (2006, p. 359); ils ne sont pas vivants, mais ils sont quand même animés. Ils remarquent que les résidents de centre de soins de longue durée sont aussi dans un espace liminal, car le centre est un endroit qui est situé entre la vie et la mort (2006, p. 359)⁵. C'est ce qui ajoute au problème ou du moins devrait susciter un questionnement quant à l'usage de robots dans les centres d'hébergement.

Dans leur étude sur les lieux de résidence, la notion de chez soi et la perception des centres d'hébergement chez les personnes âgées,

3. «We are psychologically programmed not only to nurture what we love but also to love what we nurture.» (Turkle, 2011, p. 272)

4. Dans un documentaire, Turkle affirme que PARO est un usage inapproprié de la technologie (cité dans De Falco, 2016).

5. Turkle *et al.* écrivent : «Relational artifacts, as objects between the living and not living (Ramey 2005), may have some special resonance for their frail nursing home playmates. The residents themselves experience their existence in the home as a placement on the boundary between their own lives and their deaths. Like relational artifacts, they are liminal creatures.» (2006, p. 359)

Leibing, Guberman et Wiles (2016) discutent aussi de la liminalité. Pour ces auteurs, l'idée centrale est celle d'incertitude, de zone grise et d'ambiguïté. La perception des centres d'hébergement est que ces lieux sont liminaux, c'est-à-dire que la personne qui réside dans ce type de centre vit en marge de la société. Elle n'est pas tout à fait exclue, mais elle ne fait plus partie de la société, comme lorsqu'elle était plus jeune. Elle se situe dans une zone grise. Et ce qui ajoute à cette incertitude est que l'établissement est considéré comme le dernier jalon d'un parcours qui mène à la mort. Ce qui veut dire que PARO, un être liminal, semble interagir avec des personnes en position liminale dans un endroit également liminal.

Si PARO peut avoir un effet négatif sur le statut des personnes en hébergement, la robotisation a aussi des conséquences pour la personne elle-même. D'autres questions se posent : l'interaction avec PARO est-elle basée sur un véritable rapport ou simplement sur un rapport illusoire ? Et que nous dit ce rapport sur la condition de madame Körner ? Afin d'étudier ces effets, nous nous tournons vers les écrits du philosophe allemand Plessner, qui a élaboré une anthropologie philosophique et qui s'est intéressé à l'existence corporelle.

2. L'IMPOSSIBILITÉ DE SE DÉCENTRER

Plessner a écrit à la même époque qu'Heidegger. Un des textes les plus importants de Plessner porte sur l'être et l'espace (*Die Stufen des Organischen und der Mensch. Einleitung in die philosophische Anthropologie*, parut en 1928). Si Heidegger discute de la finitude humaine centrée sur le temps, pour Plessner la dimension essentielle est celle de l'espace. Certains de ses travaux ont été traduits en français et en anglais, mais plus récemment l'anthropologie philosophique de Plessner a connu un regain de popularité dans le monde anglophone grâce aux travaux de Jos de Mul (2003 ; 2014). L'anthropologie philosophique de Plessner se centre sur la positionalité (*positionality*).

Certains des concepts que présente Plessner sont d'intérêt ici, en particulier l'attention que porte Plessner à la manière dont l'espace circonscrit l'être humain. Pour Plessner, la vie est une unité psychique, et le corps vécu peut apparaître corps ou esprit, dépendamment de la perspective. De plus, un être humain vivant est dans son corps et

aussi à l'extérieur de son corps⁶. Pour Plessner, nous possédons un corps (*Körper*) et nous sommes aussi un corps (*Lieb*).

Nous pourrions dire que l'humain navigue à sa frontière corporelle en se sachant un être pensant, mais aussi en constatant qu'il est en lien avec le monde extérieur. Comme l'explique de Mul, pour Plessner, l'être humain vit dans trois mondes : le monde extérieur, la vie intérieure et le monde de la culture qui est partagé avec les autres (de Mul, 2003, p. 252). Plessner définit l'humain comme un être qui est ex-centrique plutôt que décentré (de Mul, 2003, p. 252)⁷.

Dans l'anthropologie philosophique de Plessner, la positionalité sert de guide pour penser la différence entre le monde végétal, animal et humain (de Mul, 2003). L'humain est le seul être qui peut être conscient de la frontière entre lui-même et le monde et qui peut en quelque sorte franchir cette frontière. Si les plantes et les animaux sont déterminés par leurs limites physiques, l'humain a la possibilité d'avoir un rapport réflexif à la vie. Parce qu'il peut se voir ou être conscient de lui-même, il est décentré ; il a une positionalité excentrique. En raison de son corps physique, l'humain est un être fini dans l'espace, mais il est conscient de cette frontière et en étant conscient de son agir dans le monde, il est capable d'être à l'extérieur de lui-même, d'où la possibilité d'être décentré. L'être humain peut se décentrer et cette capacité est essentielle à son être.

Avec les avancées technologiques et la robotisation, la pensée de Plessner devient plus pertinente encore. De Mul reprend les travaux de Plessner pour étudier le phénomène des avancées numériques, la réalité virtuelle et la téléprésence. Ce qui nous intéresse dans les travaux de Plessner est l'idée centrale de positionalité, qui peut être utilisée de deux façons dans le cas de PARO. La première est la façon dont madame Körner projette ses émotions sur cet objet animé et les conséquences de son rapport avec PARO ; la deuxième sera abordée lors de notre discussion sur le travail d'aidant.

En reprenant le cas de madame Körner et de ses liens avec PARO, le risque évoqué par Turkle, dans son livre *Alone together and why we expect more from technology and less from each other* (2011), est

6. « A living person is a body, is in his body and at the same time outside his body as the perspective, from which he is both » (cité dans de Mul, 2003, p. 252).

7. De Mul explique que, selon Plessner, « as excentric beings we are not where we experience, and we don't experience where we are » (de Mul, 2003, p. 252).

que madame Körner soit vulnérable à un attachement technologique. Comme nous le montre le documentaire, l'attachement de madame Körner pour PARO devient un obstacle à ses interactions avec les autres résidentes. Nous pourrions affirmer que ce genre d'obstacle peut être une des conséquences de n'importe quel attachement, qu'il soit robotique ou autre. Cependant, il y a un effet qui vient s'ajouter pour ce qui est du rapport que madame Körner entretient avec le robot. Le robot réagit à la voix de madame Körner, donc il semble s'intéresser à elle. Mais c'est madame Körner qui donne une vie intérieure au robot. Ainsi, puisque le robot réagit, elle suppose qu'il possède une certaine agencéité, une vie intérieure, et qu'il réagit parce qu'il a de la sollicitude à son égard. L'illusion qui se glisse de plus en plus fortement dans le rapport avec le robot est que le lien avec celui-ci soit de nature affective. Le lien est assurément affectif de la part de madame Körner, mais il ne l'est pas pour PARO. Il faut rappeler que madame Körner ne souffre pas de démence et que, par conséquent, elle est très consciente de ce fait. Malgré cela, le glissement se produit petit à petit et le risque est que cet attachement, qui semble vrai et réconfortant, finisse par isoler davantage madame Körner. Le danger est non seulement sur le plan de l'isolement social, il se situe aussi à un niveau encore plus profond qui touche l'être même de madame Körner. En effet, l'interaction avec PARO pose le risque d'appauvrir la vie intérieure de madame Körner. PARO lui fait perdre la capacité de se décentrer, du moins lors de leurs interactions. Si madame Körner semble se projeter vers un autre être, ce qu'elle fait en réalité n'est pas un mouvement vers l'autre, mais un retour sur elle-même. Elle est dans l'illusion qu'elle est en rapport avec un être qui lui est extérieur.

L'humain peut franchir la frontière de sa vie intérieure grâce non seulement à ses capacités cognitives, mais aussi au fait qu'il se trouve face à des vies qui sont différentes de la sienne, face à la diversité humaine et à la diversité du monde en général. Donc, l'être humain est mis en présence de manières d'agir qui ne sont pas les siennes ; il devient conscient du monde culturel dans lequel il vit, que ce soit celui qu'il partage ou celui qu'il observe. Il en est de même pour ce qui est des interactions avec les animaux. L'humain doit être en contact avec une diversité de l'agir qui n'est pas simplement le reflet de sa propre vie intérieure pour se rendre compte qu'il partage un monde et que ce monde est extérieur, indépendant de lui. Madame Körner interagit avec un objet animé, mais elle ne fait que projeter ses propres émotions et sentiments sur le robot. Cependant, PARO n'a aucune vie intérieure,

de sorte que les rapports qu'elle entretient avec cet être robotique ne peuvent l'aider à apprécier le monde extérieur parce que ce monde extérieur (PARO) est simplement un dédoublement de son propre monde intérieur. Elle perd donc un volet important de sa vie, car elle perd la possibilité de se concevoir comme un être vivant avec une vie intérieure et un accès à une culture qui lui est extérieure. Ici aussi, nous voyons encore cet effet de liminalité. Madame Körner se trouve dans une zone indéterminée : elle se croit en rapport avec un être, mais cet être est tout simplement constitué du for intérieur de madame Körner. Elle est au seuil d'une frontière (sa vie intérieure et le monde extérieur) qu'elle ne franchit jamais.

Nous pourrions avancer que cet aspect, s'il est significatif pour une analyse de la condition des personnes âgées et indicatif de la ségrégation des personnes âgées, n'a pas de conséquence négative si le robot permet à la personne de se sentir moins seule. C'est un moyen de pallier les circonstances malheureuses de nos sociétés postindustrielles. Évidemment, lorsqu'une personne souffre de démence, PARO devient un être animé qui est là, une présence. Nous pourrions concéder que, dans le cas de personnes atteintes de démence, la compagnie de PARO puisse être bénéfique. Comme le montrent maintes analyses déjà citées, cette interaction réduit le stress et l'anxiété. Si les signes vitaux sont positifs, qu'importe la situation de la personne dans son vécu quotidien ?

Si les répercussions négatives sur la personne âgée peuvent être atténuées par des considérations concernant leur bien-être, il en est tout autrement du travail des aidants dans ces centres. La présence de PARO aura des répercussions sur les relations des personnes qui font partie de l'entourage de la résidente. Ces effets ne seront pas toujours négatifs ; il se peut que PARO soit un lien entre la personne âgée et les autres. Mais l'utilisation d'un robot émotionnel peut avoir des conséquences sur les conditions de travail des aidants en milieu institutionnel. La question que nous nous posons est de savoir ce que l'utilisation d'un robot pour faire un travail émotionnel auprès du résident implique pour le travail d'aidant. Dans la prochaine section, nous allons analyser la question de la robotisation émotionnelle en milieu de travail par l'entremise des écrits d'Agamben.

3. LE TRAVAIL D'AIDANT

Le travail d'aidant en milieu institutionnel a fait l'objet de plusieurs études sociologiques, mais il a été moins étudié sous l'angle philosophique⁸. Les centres de soins de longue durée pour personnes dites dépendantes sont des endroits basés sur un modèle médical de soins (voir Lanoix 2011 ; 2013a ; 2013b). Les résidents de ces centres sont des personnes en perte d'autonomie, qui très souvent souffrent de démence⁹. Même si ces centres se veulent des milieux de vie substitués, il n'en demeure pas moins que les soins sociaux n'y sont pas valorisés. La hiérarchie médicale fait en sorte que les préposés aux bénéficiaires doivent s'occuper avant tout des besoins corporels de la personne hébergée. Les études sur les centres d'hébergement montrent que les soins qui priment sont les soins corporels : le résident doit être habillé, avoir pris ses médicaments, avoir bien mangé. Cependant, les travailleurs auprès des personnes âgées en milieu institutionnel déplorent qu'il leur manque de temps pour parler avec elles et qu'il soit difficile, voire impossible de tisser un rapport autre que simplement instrumental (voir Armstrong *et al.*, 2009 ; Armstrong et Braedley, 2013 ; Diamond, 1992).

Lorsqu'une personne a besoin d'aide pour vaquer aux activités de la vie quotidienne, elle bénéficie d'une aide technique. Par exemple, aider la personne à se lever est une tâche qui demande que l'aidant possède une façon de faire particulière pour ne pas se blesser. Il en est de même pour d'autres activités plus complexes, comme aider une personne dysphagique à manger. Mais il y a toujours une composante d'interaction qui ne vise pas simplement l'activité comme telle. Ainsi, le préposé peut demander à une dame qui va manger comment elle se sent ou si sa fille est venue la voir. Cette composante est nommée travail affectif ou émotionnel, ce dernier terme étant plus en usage dans les études anglo-américaines¹⁰.

8. Le terme *aidant* est un des termes en usage, comme le terme *préposé aux bénéficiaires*, lequel est utilisé pour désigner les aidants rémunérés en milieu institutionnel.

9. Le terme *personne en perte d'autonomie* est le terme utilisé dans les politiques qui gèrent ces centres. Cependant nous aimerions noter que ce terme est problématique et qu'il serait préférable d'utiliser *personne qui requiert de l'aide*, terme qui, malheureusement, n'est pas très élégant.

10. Le concept de *travail émotionnel* a été introduit par Arlie Hochschild (1983) lors de ses études sur les agents de bord. Il a été repris par plusieurs théoriciennes-infirmières pour indiquer le travail de soutien qu'elles font auprès des patients (voir Hyde *et al.*, 2005).

Dans une de nos études sur le travail d'aidant, nous avançons que ce travail est composé d'actions instrumentales, par exemple les actions que l'aidant effectue pour laver une personne. Mais une des composantes importantes du travail d'aidant est l'action communicative. Nous démontrons que le rapport aidant-aidé requiert une dimension communicationnelle, s'il se veut éthique (Lanoix, 2013a)¹¹. Cependant, le travail émotionnel n'est pas reconnu comme faisant partie du travail des aides aux personnes âgées. De fait, un préposé peut s'entretenir avec une résidente lors du repas, mais il n'y aura pas beaucoup d'autres interactions de ce genre qui se produiront en dehors des activités qui ciblent les besoins corporels. De plus, en raison des régimes d'austérité qui font en sorte que les centres doivent mettre en place des mesures d'efficacité, le travail des aidants devient de plus en plus réglementé et les préposés doivent accomplir leurs tâches dans une période de temps précise (c'est aussi le cas dans le secteur de l'aide à domicile : voir Müller, 2016 ; Grant *et al.*, 2004 ; Parks, 2003).

Il existe déjà beaucoup d'aide technique qui vise à alléger le travail du préposé. En effet, dans son introduction au volume *Robot Ethics*, Lin discute de l'utilisation des robots pour exécuter des travaux qui sont dangereux, sales et répétitifs (« *The 3 Ds, Dull, Dirty or Dangerous* », p. 4). Si les robots pouvaient intervenir dans les tâches qui visent seulement les soins corporels, l'utilisation de ces robots aidants pourrait libérer les préposés de tâches fastidieuses tout en leur donnant plus de temps pour s'occuper des résidents. Cependant, la tendance est alors d'ajouter plus de tâches en augmentant le nombre de résidents auxquels les préposés doivent fournir des services¹². Le travail demeure centré sur les activités qui répondent aux besoins physiques des résidents. On ne peut pas s'attendre à un autre résultat, puisque

11. Notre conclusion dans cette étude est que : « Within the institutional framework of ancillary care, meeting the physical and medical needs of the care receivers is the primary goal of caregiving. The result is that the physical needs of an individual become separated from the emotional ones that may be associated with her physical needs. This splits the care receiver in two parts: one that is being washed and fed and the other, the emotional part, which then becomes insignificant. Put differently, the care receiver's body becomes the locus of care but her agency and her emotional responses are disregarded. This dissociative process is morally problematic as it reduces the agency of the person being assisted. This can easily lead to the objectification of the person who is being assisted. » (2013a, p. 111)

12. En effet les compressions budgétaires au Québec impliquent des réductions dans les soins, Ariane Lacoursière, « Nouvelles compressions de 242 millions en santé », *La Presse*, 18 juin 2016.

le travail de préposé est conçu selon le modèle du travail à la chaîne (voir Lanoix, 2011 ; 2013b).

Pour ce qui est des robots émotionnels, leur apport pourrait non pas alléger la tâche des préposés, mais la compléter. Il semblerait donc qu'un robot tel que PARO permettrait de pallier ce manque d'interactions sociales. Cependant, la fonction de PARO ne sera pas de s'insérer dans la trame des soins prodigués par les préposés et dans tout l'appareil institutionnel et médical des centres d'hébergement. Nous avançons que l'apport de PARO est de faire en sorte que le travail d'aidant soit davantage axé sur les soins corporels et que toute la dimension du travail émotionnel soit progressivement éliminée, afin que le travailleur devienne une sorte d'instrument de l'appareil institutionnel.

Dans les prochaines lignes, nous allons tout d'abord nous intéresser à l'analyse d'Agamben sur l'usage des corps pour montrer que, dans son travail, l'aidant sera de plus en plus considéré comme un instrument de travail et non comme une personne qui possède une agencéité. Nous verrons que les travaux de Plessner sur l'espace, c'est-à-dire la positionalité, seront aussi d'intérêt pour cette question.

4. L'AIDANT COMME OUTIL

Dans ses nombreux écrits, Agamben explore la question du pouvoir politique et oriente particulièrement son étude sur l'État d'exception. Pour Agamben, cet État n'est pas « une dictature (constitutionnelle ou inconstitutionnelle, de commissaire ou souveraine), mais un espace vide de droit, une zone d'anomie où toutes les déterminations juridiques – et avant tout la distinction même entre public et privé – sont désactivées » (2003, p. 86). La vie nue (*zoè*) est constituée par et dans cet espace. Toujours en poursuivant son but d'éclairer comment est constitué le pouvoir, dans son dernier livre *L'Usage des corps*, Agamben reprend les travaux d'Aristote et de Thomas d'Aquin sur l'instrumentalité afin d'explorer le lien constitutif entre esclavage et technique.

Le but de l'ouvrage d'Agamben est d'explorer comment l'évolution du travail et le rapport de l'humain avec la technique régissent l'humain et impliquent que celui-ci s'apparente plus à une forme de vie qu'à un sujet. Pour notre analyse, l'étude d'Agamben se révèle pertinente, car elle jette un regard critique sur le rapport de l'humain avec les outils

dont il se sert pour travailler. Le chapitre qui nous intéresse est celui sur «l'instrument animé et la technique». Nous relevons dans l'analyse d'Agamben un élément qui rejoint notre étude sur la robotisation et le travail d'aidant, celui de l'humain comme instrument de travail. Nous avançons, dans un premier temps, que l'utilisation de robots émotionnels a pour conséquence de rendre encore plus invisible le travail émotionnel et relationnel de l'aidant et, dans un deuxième temps, que le risque est de réduire sa condition à celle de simple instrument de travail.

Agamben remarque que l'esclave chez Aristote est un instrument animé : l'esclave n'est pas un agent, mais un instrument (2015, p. 118). C'est un être «dont l'œuvre est l'usage du corps» (2015, p. 27). Pour Agamben, la définition aristotélicienne de l'esclave montre que chez Aristote «l'idée dominante était celle d'une vie humaine se développant entièrement dans la sphère de l'usage (et non dans celle de la production)» (2015, p. 124). L'esclave accomplit la tâche qui est exigée par le maître ; il est comme la hache qui sert à façonner le lit. Ce qui est significatif, selon Agamben, est que l'esclave «constitue donc un double seuil : par lui la vie animale passe dans l'humain, de même que le vivant passe dans l'inorganique (dans l'instrument) et vice versa» (2015, p. 124). Comme on vient de le voir, dans l'Antiquité, une vie pouvait être considérée comme instrumentale, mais qu'en est-il pour le travail d'aidant, le travail d'aider ou de prendre soin d'une autre personne ?

Le travail d'aidant, de soins auprès des proches, était avant la fin du XX^e siècle un travail non rémunéré et surtout effectué par les femmes au foyer. Ce travail n'était pas digne d'analyse pour les grands philosophes, qui le considéraient comme une activité relevant de l'instinct. Cependant, au début des années 1980, avec l'ouvrage de Carol Gilligan, *In a different voice*, les philosophes féministes se sont intéressées au *care* et au travail du «prendre soin». Depuis les dernières décennies, certaines philosophes féministes arguent que prendre soin d'une autre personne est un travail qui requiert de la réflexion et n'est pas simplement instinctuel. Ce travail est tout aussi légitime que le travail de l'ouvrier. Si nous admettons que prendre soin de quelqu'un est un travail, comment l'analyse d'Agamben peut-elle s'appliquer ?

Il faut rappeler que le travail d'aidant est un travail non productif, car, selon Marx, le travail de reproduction ne s'inscrit pas dans le cadre du travail de production. Ce travail est essentiel, mais il n'est pas sous la

responsabilité de ceux qui détiennent le pouvoir de production. Donc, le travail d'aidant, comme travail de *care*, pourrait bien être pensé comme une activité qui requiert une agencéité¹³. Et ce point de vue est soutenu par les théoriciennes du *care*, telles Joan Tronto, Virginia Held et Pascale Molinier. Cependant, le travail d'aidant en tant que travail rémunéré et régi dans un cadre médical, de même que sujet à une logique d'efficacité qui prend en considération les limites budgétaires, devient un travail qui s'apparente très peu aux soins que l'on offre aux membres de sa famille. Nous avons discuté ailleurs (Lanoix, 2011) de l'effet de la marchandisation sur le travail d'aidant qui fait de ce dernier un travail à la chaîne, comme en usine, et que cette marchandisation a des effets nocifs non seulement sur l'aidant, mais aussi sur la personne aidée (Lanoix, 2009).

Puisque la marchandisation agit sur la manière dont est structuré le travail d'aidant, celui-ci s'apparente plus à un travail productif même s'il ne produit pas de plus-value (*surplus value*). Soumis à ce régime de travail, le travailleur devient un outil plus qu'un être travaillant. Il est l'outil par lequel s'effectue le travail qui est déterminé par l'institution ou le centre de soins de longue durée¹⁴. Nous expliquons.

PARO, comme aide au préposé, pourrait devenir un outil qui soit une aide pour la communication avec la personne aidée ou un moyen d'alléger une tâche qui peut être difficile¹⁵. Cependant, même lorsque PARO est utilisé en groupe avec une thérapeute, le but est de combler un besoin émotionnel. Le travail émotionnel est donc relégué au robot : le préposé, l'être humain n'est plus le seul locus de soins. On retire au travailleur les soins qui ont trait aux émotions et on les resitue dans le personnage du robot. Ce qui implique qu'avec PARO, qui s'occupe en quelque sorte des besoins émotionnels des personnes hébergées, le travail de l'aidant devient davantage un travail instrumental, car

13. Dans la présentation du livre *Le Souci des autres*, Patricia Paperman et Sandra Laugier expliquent que la traduction du terme *care* est problématique. Comme elles voulaient garder l'aspect pratique et la notion d'attention, elles ont décidé de conserver le terme *care* au lieu de le traduire par *sollicitude* (2005, p. 11).

14. Il faut comprendre que l'institution a le devoir de gérer la structure par laquelle les soins seront prodigués et que cette structure est hiérarchique. Ainsi, le médecin prescrit les médicaments, ensuite les infirmières prennent la responsabilité du plan de soin qui sera effectué par les préposés.

15. Nous pensons aux activités qui pourraient être déplaisantes pour une personne atteinte de démence, comme prendre un bain.

toute communication et tout échange seront alors ciblés sur les besoins corporels.

Lors des soins, le préposé va communiquer avec la personne qu'il assiste, mais cette communication concernera principalement la tâche qui doit être accomplie, c'est-à-dire que l'aidant exécutera les soins requis simplement comme un instrument. Aucun travail émotionnel ne fera partie de ses activités auprès des résidents. L'aidant devient la hache citée dans l'exemple d'Agamben. Puisque le travail de l'aidant est un travail physique, son corps devient un instrument du système. L'aidant perd encore plus d'agencéité lorsque le travail de communication qu'il doit effectuer est surtout centré sur les soins corporels. Si les besoins émotionnels et de communication sont comblés par d'autres personnes que le préposé, celui-ci risque de devenir un simple automate qui doit se charger des besoins physiques.

5. L'EFFET PARO

Afin de poursuivre l'analyse agambienne de PARO, il faut examiner les raisons pour lesquelles PARO devient un élément constitutif des soins aux personnes âgées. Si PARO est utilisé comme une aide technique, c'est l'intervenant qui dirige les activités, et le robot sert alors d'outil afin d'aider les personnes âgées atteintes de démence à communiquer. Ainsi, le thérapeute a une intention, celle de garder ces personnes en relation avec le monde qui les entoure. L'intervenant exécute cette tâche par l'entremise du robot. Ici, PARO est vraiment un instrument animé.

Mais si le robot est seul avec la personne pour lui tenir compagnie, qui manie cet instrument ? PARO devient-il alors un agent ? Pour répondre à cette question, il est intéressant de voir ce que propose Agamben lorsqu'il se penche sur la question de la cause instrumentale. Selon l'interprétation d'Agamben, pour Thomas d'Aquin la cause instrumentale prend tout son sens dans la mesure où elle est employée dans le contexte d'une technique (2015, p. 118). Pour Agamben, « la technique est la dimension qui s'ouvre quand l'opération de l'instrument s'est rendue autonome et en même temps s'est scindée en deux opérations distinctes et liées » (2015, p. 118). L'analyse de Thomas d'Aquin porte sur les sacrements et, comme l'explique Agamben, « le caractère de cause instrumentale n'appartient pas seulement à l'élément matériel (l'eau, les huiles consacrées, etc.) : il concerne avant tout le célébrant lui-même » (2015, p. 119). Mais le célébrant n'est pas seulement mû par

l'agent principal, il est un instrument animé, car il peut se mouvoir par sa volonté propre (2015, p. 119). Cela signifie que «le sacerdoce sacramental est lié à l'esclavage de façon généalogique et non pas seulement terminologique» (2015, p. 120). Ce qu'il faut noter ici est que le ministre est «l'instrument animé d'une opération dont l'agent principal est le Christ» (2015, p. 120). Ce qui implique pour Agamben que le célébrant, en devenant un instrument animé, est séparé de son engagement personnel et de sa responsabilité morale (2015, p. 121). Agamben souligne que cet engagement et cette responsabilité sont relégués à l'intériorité de la pratique sacramentelle (2015, p. 121). Il doit en être ainsi, car le Christ est l'agent principal.

La discussion de Thomas d'Aquin sur la distinction entre l'œuvre opérante et l'œuvre opérée qu'il propose afin d'assurer «la validité du sacrement» est pour Agamben une «sorte de prophétie du machinisme» (2015, p. 121). Ce machinisme sera réalisé plusieurs siècles plus tard quand l'artisan se transformera en un instrument de la machine.

En somme, si nous transposons ces réflexions au cas de PARO, un instrument animé qui permet d'apaiser les personnes âgées, nous pouvons affirmer que PARO est la cause instrumentale. PARO comme instrument animé n'a aucune responsabilité morale ni aucun engagement personnel, car il est évidemment un robot sans vie intérieure. Cependant, nous voyons que ce n'est pas essentiel dans le cas d'un instrument animé, robot ou humain. Le but ultime de PARO est de procurer le calme. En effet, comme nous l'avons déjà mentionné, les études montrent que les signes médicaux qui indiquent qu'une personne est calme (par exemple la baisse de la pression artérielle et de l'anxiété, qui se traduit par la diminution de la médication) sont présents lorsqu'un résident passe quelques heures avec PARO lors de sessions qui prennent place trois fois par semaine.¹⁶

Puisque PARO semble pallier le manque d'interaction avec les humains, l'effet qu'il produit est de libérer les humains de toute interaction et, surtout, de toute obligation d'interaction avec les personnes âgées atteintes de démence. Cette conséquence est-elle voulue de ceux qui ont développé PARO? Sans doute que non, cependant, même si

16. La recherche de Peterson *et al.* (2017) a pris place sur une période de trois mois. Les personnes atteintes de démence passaient vingt minutes trois fois semaine en compagnie d'une infirmière et de PARO. Les signes vitaux ont été mesurés avant et après chaque session et l'usage de divers médicaments a été évalué à la fin de l'étude.

cet effet n'est pas souhaité par les ingénieurs ni par les directeurs des établissements de soins de longue durée, il est quand même présent. Dans son étude portant sur le dispositif, Agamben reconnaît que «le dispositif est donc, avant tout, une machine qui produit des subjectivations et c'est pourquoi il est aussi une machine de gouvernement» (2007, p. 42). Par sa présence, PARO change la dynamique des comportements des travailleurs et aussi de leur travail.

L'instrument, l'inorganique, remplace l'humain et en remplaçant l'humain, dans notre exemple l'aidant auprès de la personne avec démence, l'instrument place cette dernière encore plus en marge de l'humain. En même temps, il place aussi l'aidant en marge de l'agencéité. Nous pourrions objecter que le préposé est déjà en marge et que l'effet de PARO est assez minime. Cependant, l'apport de PARO en tant que substitut de l'humain démontre la volonté de penser le travail de «prendre soin» comme requérant seulement une attention au corps de la personne à aider, le travailleur étant lui aussi réduit à son corps. Cette manière de penser le travail d'aidant s'inscrit dans la foulée d'une pensée cartésienne qui réifie la scission du corps et de l'esprit. Ce qui est problématique dans cette vision est l'absence de toute dimension affective et incarnée du travail d'aidant. En effet, l'aidant devient non seulement l'instrument du système institutionnel des centres de longue durée, mais il risque aussi de devenir l'instrument du robot émotionnel, car son travail sera déterminé en fonction de PARO et non en fonction de ce que l'aidant, lui-même, pourrait déterminer comme besoin à combler.

Avec PARO comme instrument de travail émotionnel, le travail affectif devient ex-centrique, pour reprendre le langage de Plessner. Cependant ici l'ex-centricité sert à éliminer le travail émotionnel du travail d'aidant au détriment de ce travail. Au lieu de soutenir les travailleurs dans leurs activités de «prendre soin», l'apport robotique constitue plutôt une entrave.

CONCLUSION

Nous pouvons voir que PARO, qui est une aide à la communication et a pour but de faciliter l'engagement de la personne, n'atteint pas nécessairement ce but. Dans le cas de madame Körner, elle devient plus isolée des autres personnes de la résidence. Certes, elle communique

et semble avoir un lien affectif avec PARO, mais le phoque n'a pas vraiment aidé madame Körner à avoir une vie sociale plus riche.

Le cas de madame Körner est un peu différent de celui des personnes atteintes de démence, car elle possède une certaine autonomie. Elle sait où se rendre pour ses activités, elle a une chambre qu'elle reconnaît. Dans le film, nous voyons que la personne qui l'aide interagit assez minimalement avec elle, puisque madame Körner a la capacité d'accomplir beaucoup de tâches par elle-même. Il est possible que PARO amplifie en quelque sorte le côté conflictuel de madame Körner, car, lors des l'activité musicale, elle défend son PARO et refuse de voir comment le robot aurait pu déranger les chansons. Nous pourrions conclure que, puisque madame Körner possède encore une agencéité, c'est elle qui détermine le rôle que prendra PARO. Néanmoins, nous avons l'impression qu'elle se replie sur elle-même. Nous la voyons glisser vers un comportement qui devient manipulé en quelque sorte par le robot ou plutôt par les attentes qu'elle porte envers le robot.

Pour conclure, il est essentiel de faire appel à une autre dimension de l'analyse de Plessner. Dans le développement de sa philosophie anthropologique, Plessner propose trois lois. La première est que l'humain est toujours constitué de l'artificiel. Donc, l'instrument passe par l'humain. La deuxième est la prise de conscience des frontières qui sera de plus en plus relayée par les technologies. Enfin, la troisième est que les créations gagnent une certaine indépendance et peuvent à leur tour déterminer la vie humaine (de Mul, 2014, p. 471- 473)¹⁷. Nous le voyons avec les avancées technologiques qui illustrent la troisième loi de Plessner reconnaissant que la culture peut évoluer d'une façon autre que celle qui était prévue au départ. Dans le cas que nous avons étudié, l'instrument animé sert à inscrire les personnes âgées atteintes de démence dans un espace liminal. Ensuite, les robots émotionnels peuvent limiter l'accès à la vie extérieure par laquelle tout être humain se voit en relation avec une diversité d'êtres qui ont une vie différente de la sienne. Enfin l'utilisation de PARO aura aussi des conséquences sur le travail d'aidant. Ces constats illustrent bien la conclusion d'Agamben au sujet des dispositifs : « ce qui définit les dispositifs auquel nous avons à faire [sic] dans la phase actuelle du capitalisme est qu'ils n'agissent plus par la production d'un sujet, mais bien par des processus que nous pouvons appeler des processus de désubjectivation » (2007, p. 43).

17. Les trois lois selon de Mul : *The law of natural artificiality; the law of mediated immediacy; the law of utopian standpoint.*

Les robots émotionnels ne sont pas des avancées neutres. En effet, comme on le voit dans les discussions sur la télé-médecine, tout dispositif technique n'est jamais neutre. Si PARO est un animal avec une peluche douce et un air attachant, ce qu'il représente doit être analysé dans le contexte d'une humanité vulnérable.

BIBLIOGRAPHIE

- Agamben, Giorgio (2003). État d'exception. *Homo Sacer*, 2. Joël Gayraud, trad. Paris, Seuil.
- Agamben, Giorgio (2007). *Qu'est-ce qu'un dispositif?* Martin Rueff, trad. Paris, Rivages.
- Agamben, Giorgio (2015). *L'Usage des corps. Homo Sacer IV*, 2. Joël Gayraud, trad. Paris, Seuil.
- Ambo, Phie (2007). *Mechanical Love*. (Documentaire.)
- Armstrong, Pat et Braedley, Susan (dir.) (2013). *Troubling Care: Critical Perspectives on research and Practice*. Toronto, Canadian Scholar's Press.
- Armstrong, Pat, Bannerjee, Albert, Szebehely, Marta, Armstrong, Hugh, Daly, Tamara, et Lafrance, Stirling (2009). *They deserve better: The long-term care experience in Canada and Scandinavia*. Ottawa, Canadian Centre for Policy Alternatives.
- DeFalco, Amelia (2016). Beyond prosthetic memory: Posthumanism, embodiment, and caregiving robots. *Age Culture Humanities*. Consulté en ligne.
- de Mul, Jos (2003). «Digitally mediated (dis)embodiment: Plessner's concept of excentric positionality explained for cyborgs». *Information, Communication & Society*, vol. 6, n° 2, p. 247-266.
- de Mul, Jos (dir.) (2014). *Plessner's philosophical anthropology*. Amsterdam, Amsterdam University Press.
- Diamond, Timothy (1992). *Making Grey Gold: Narrative of Nursing Home Care*. Chicago, University of Chicago Press.
- Fowler, Geoffrey A. (17 novembre 2017). «These robots don't want your love». They want your job. *Washington Post*. Consulté en ligne.
- Gilligan, Carol. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, Massachusetts, Harvard University Press.
- Grant, Karen R., Amaratunga, Carol, Armstrong, Pat, Boscoe, Madeline, Perderson, Ann, et Wilson, Kay (dir.) (2004). *Caring For/Caring About: Women, home Care and Unpaid caregiving*, Aurora, Ontario, Garamond Press.
- Held, Virginia (2006). *The Ethics of Care*. Oxford Scholarship Online Monographs.
- Hochschild, Arlie Russell (1983). *The Managed heart: Commercialization of human feelings*. Berkeley, University of California Press.
- Hyde, Abbey, Margaret Treacy, P. Anne Scott, Michelle Butler, Jonathan Drennan, Kate Irving, Anne Byrne, B. Pdraig, MacNeea, et Marian Hanrahan (2005). «Modes of rationality in nursing documentation: Biology, biography and the voice of "nursing"». *Nursing Inquiry*, vol. 12, n° 2, p. 66-77.
- Johal, Wafa, Adam, Carole, Fiorino Humbert, Pesty, Sylvie, Jost, Céline, et Duhaut, Dominique (2015). «Acceptabilité d'un robot compagnon dans les situations de

- la vie quotidienne» dans Marine Grandgeorge, Céline Jost, Brigitte Le Pévédic et Frédéric Puignière-Saavedra (dir.), *Vers un communication Homme-animal-Machine?* Bruxelles, EME Éditions.
- Jøranson N., Pedersen I., Rokstad A.M.M. et Ihlebæk C. (2016). «Change in quality of life in older people with dementia participating in Paro-activity: a cluster-randomized controlled trial». *Journal of Advanced Nursing* vol. 72, n° 12, p. 3020-3033.
- Klein, Barbara et Cook, Glenda (2012). «Emotional Robotics in Elder Care – A Comparison of Findings in the UK and Germany» dans Goebel, Randy, Tanaka Yuzura, et Wahlster Wolfgang (dir.), *Social Robotics. Lecture Notes in Artificial Intelligence*, vol. 7621. Berlin Heidelberg, Springer-Verlag, p. 108-117.
- Lacoursière, Ariane (18 juin 2016). «Nouvelles compressions de 242 millions en santé». *La Presse*.
- Lanoix, Monique (2013a). «Caring for Money: Communicative and Strategic Action in Ancillary Care», *International Journal of Feminist Approaches to Bioethics*, special issue on Aging and Long-Term Care, vol. 6, n° 2, p. 94-117.
- Lanoix, Monique (2013b). «Labour as Embodied Practice: The Lessons of Care Work», *Hypatia*, vol. 28, n° 1, p. 85-100.
- Lanoix, Monique (2011). «Assembly Line Care: Paid Ancillary Care in post-Fordist Economies», *Work: A Journal of Prevention, Assessment and Rehabilitation*, special issue on care labour, Septembre, vol. 40, n° 1, p. 41-50.
- Lanoix, Monique (2009). «A Body No Longer of One's Own», dans Sue Campbell, Letitia Meynell et Susan Sherwin (dir.), *Agency and Embodiment*, Philadelphia, University of Pennsylvania Press, p. 164-183.
- Leibing, Annette, Guberman, Nancy et Wiles, James (2016). «Liminal homes: Older people, loss of capacities, and the present future of living space». *Journal of Aging Studies*, vol. 37, p. 10-19.
- Lin, Patrick, Abney, Keith, Bekey, George A. (dir.) (2012). «Robot ethics: The Ethical and Social implications of robotic». Cambridge, Massachusetts: MIT Press.
- Molinier, Pascale (2013). *Le travail du care*. Paris, La Dispute.
- Müller, Beatrice (2016). *Work-Abjektion*. Westfälisches Dampfboot.
- Paperman, Patricia, et Laugier, Sandra (dir.) (2005). *Le souci des autres: Éthique et politique du care*. Paris, Édition de l'École des hautes études commerciales.
- Parks, Jennifer (2003). *No place like Home? Feminist ethics and home health care*. Bloomington Indiana, Indiana University Press.
- Petersen, Sandra, Houston, Susan, Qin, Huanying, Tague, Corey et Studley, Jill (2017). «The utilization of robotic pets in dementia care». *Journal of Alzheimer's Disease*, vol. 55, p. 569-574.
- Plesner Helmut (2017). *Les degrés de l'organique et l'homme*, Paris, Gallimard.
- Turkle, Sherry (2011). *Alone together: Why we expect more from technology and less from each other*. New York, Basic Books.
- Turkle, Sherry, Taggart, Will, Kidd, Cory D. et Dasté Olivia (2006). «Relational artifacts with children and elders: the complexities of cybercompanionship». *Connection Science*, vol. 18, n° 4, p. 347-361.
- Tronto, Joan (2009). *Un monde vulnérable: pour une politique du care*. Hervé Maury, trad. Paris, La Découverte.

CHAPITRE 16

Enjeux et perspectives de l'apprentissage avec des robots pédagogiques

Raoul KAMGA
et Margarida ROMERO

INTRODUCTION

L'approche par compétences s'est développée au cours des dernières décennies dans de nombreux pays (Halász et Michel, 2011). Du Canada au Cameroun, en passant par la France, les différents systèmes éducatifs ont visé le dépassement d'une approche centrée principalement sur les connaissances pour intégrer de manière plus importante l'approche par compétences. Cette approche, sans s'opposer à l'apprentissage des connaissances spécifiques (Tremblay, 2014), vise à placer l'apprenant au centre de son apprentissage et à développer chez ce dernier des compétences qui lui permettront de mieux appréhender le monde dans lequel il vit et où il est appelé à demeurer, son environnement social et économique de même que son futur environnement professionnel. Le développement des compétences dites du XXI^e siècle pourrait contribuer à l'épanouissement de l'apprenant tant sur le plan personnel que sur les plans social et professionnel. En effet, sur le plan professionnel, le World Economic Forum (WEF, 2016)

souligne que dans les prochaines années se produira une migration des emplois nécessitant des compétences liées à la capacité physique vers ceux impliquant des compétences liées à la résolution des problèmes complexes. Cette migration pourrait s'expliquer par le fait que plusieurs emplois comprenant l'exécution de tâches routinières ou qui sollicitent principalement la capacité physique de l'individu pourraient être effectués par les robots. Bien que la question de l'utilisation des robots dans la société ne soit pas nouvelle, celle de la contribution de l'usage de ces derniers dans l'éducation est encore récente. En fait, étant donné la transformation de nos environnements par la présence des technologies robotiques, il est important non seulement de mettre en question le type d'éducation à donner aux apprenants du XXI^e siècle, mais aussi d'analyser le potentiel pédagogique et l'apport de l'usage des robots dans l'éducation.

1. USAGE PÉDAGOGIQUE DES ROBOTS ET PERSPECTIVES D'APPRENTISSAGE

La robotique pédagogique est l'usage pédagogique des technologies robotiques (robots et trousseaux robotiques) pour développer une compétence ou pour atteindre un objectif d'apprentissage dans un contexte d'apprentissage formel ou informel (Romero et Kamga, 2016). Dans l'actualité, aussi bien les technologies robotiques que les usages pédagogiques de robots en éducation sont nombreux et variés. Des automates BeeBot à l'usage des robots humanoïdes de type Nao, il existe une panoplie de robots couramment utilisés en éducation (Dupont *et al.*, 2010; Reverd, 2016). Dans la littérature scientifique, les études portant sur l'intégration pédagogique des robots en éducation sont de plus en plus nombreuses, que ce soit celles qui abordent les contenus disciplinaires (Chambers *et al.*, 2008; Chih-Wei Chang et Gwo-Dong Chen, 2010; Kanda *et al.*, 2009) ou les compétences (Fleming, 2014; Jungho, 2015).

En effet, sur le plan disciplinaire, la littérature scientifique présente plusieurs études abordant l'usage de robots pédagogiques pour l'apprentissage de notions en mathématique, en science et technologie, ou encore en anglais langue seconde (Romero et Kamga, 2016). Nombreux sont les auteurs qui soulignent les effets positifs de l'intégration pédagogique des robots dans l'enseignement de la science et sur la motivation des élèves pour l'apprentissage des sciences (Park, 2015) et

la compréhension de plusieurs concepts abstraits tels que la relation entre les engrenages, la vitesse relative des roues et leur nombre de tours (Chambers *et al.*, 2008).

Un des avantages de l'intégration pédagogique des robots en éducation est l'interdisciplinarité qu'ils permettent d'aborder. En effet, les activités de robotique pédagogique offrent la possibilité à l'apprenant de mobiliser des notions issues de plusieurs champs disciplinaires. Par exemple, dans l'étude menée par Varney et ses collaborateurs (2012), les apprenants de l'enseignement primaire sont engagés dans une activité nécessitant la mobilisation de notions de mathématiques, de science et technologie, etc.

En ce qui concerne le développement des compétences des apprenants par l'usage des technologies robotiques, des études se sont intéressées à des compétences comme la collaboration, la résolution de problèmes, la communication, la créativité, la compétence en technologies de l'information et de la communication (Romero et Kamga, 2016). Selon Voogt et Roblin (2012), ces compétences sont essentielles pour le développement personnel et professionnel des citoyens au XX^e siècle.

Les usages de robots en éducation ne se limitent pas seulement à l'atteinte d'objectifs d'apprentissage ou au développement des compétences des apprenants. On distingue aussi l'usage des robots de téléprésence, tels Kubi ou Beam. Ces robots peuvent être utilisés pour améliorer l'interaction entre des personnes participant à une même activité, mais ne se trouvant pas dans la même enceinte. En effet, les études de Fels et ses collaborateurs (2001) et de Yeung et Fels (2005) soulignent l'usage des robots de téléprésence pour faciliter la participation de certains apprenants aux activités d'apprentissage. Par ailleurs, ils peuvent constituer une option de rechange pour les apprenants qui, pour une raison ou pour une autre, ne peuvent se déplacer jusqu'à l'école.

Plusieurs universités ont intégré ces usages de robots dans leurs activités. Dans le cas de l'Université Laval, des robots de téléprésence sont utilisés lors de la présentation des examens de doctorat, des projets de recherche et des soutenances de thèse de certains étudiants. Ces robots permettent de solliciter la participation des examinateurs externes ou des enseignants ayant accepté de faire partie du jury mais qui, pour une raison quelconque, ne peuvent pas faire le déplacement. Cet usage de robot donne l'impression aux participants à distance d'être assez proches de ceux qui sont présents. De plus, sur le plan économique, cet

usage de robot peut contribuer à réduire les coûts liés au déplacement, à l'hébergement et même aux repas. Il est un facteur d'économie en temps, puisque les déplacements et les séjours des acteurs sont alors moins fréquents.

2. DÉFIS DE ROBOTIQUE PÉDAGOGIQUE ET RAPPORT À L'ERREUR

Au-delà de la diversité des technologies robotiques, nous devons considérer la diversité des activités pédagogiques pouvant être réalisées en classe. Si certains usages de la robotique pédagogique engagent l'apprenant dans une démarche procédurale où il suffit de suivre des étapes pas à pas, d'autres activités l'engagent dans des défis en équipe d'une certaine complexité.

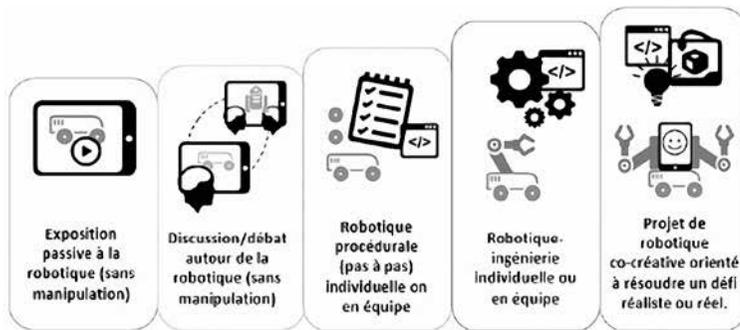


FIGURE 1 Diversité d'activités de robotique pédagogique (Romero et Dupont, 2016).

Dans la figure 1, nous observons cinq types d'activités. Tout d'abord, les activités transmissives dans lesquelles les apprenants sont exposés à des explications ou des ressources sur la robotique, mais qui ne leur permettent pas de manipuler ; en deuxième lieu, les activités de discussion sans manipulation ; en troisième lieu, les activités très guidées, de type procédural ; en quatrième lieu, les activités d'ingénierie qui engagent l'apprenant dans un processus créatif de résolution d'une situation-problème ; en dernière instance, les projets de robotique cocréative visent à résoudre un défi réaliste ou réel et engagent les apprenants dans une démarche participative d'apprentissage qui dépasse le contexte scolaire, comme dans le cas des compétitions

de Robotique Zone 01 ou First Lego League (Romero, 2016). Dans le cadre des activités de robotique pédagogique qui engagent l'apprenant dans un défi d'une certaine complexité, les apprenants font face à des démarches de résolution de problèmes. Pendant la résolution de ces problèmes qui sont souvent caractérisés par une certaine complexité, les apprenants sont amenés à commettre des erreurs de temps en temps. Lorsqu'ils identifient et traitent ces erreurs, les apprenants s'engagent dans une démarche d'autorégulation ou de corégulation. En effet, lorsque la réponse proposée par l'apprenant pour résoudre un problème ne fonctionne pas, celui-ci s'en rend compte tout seul, sans avoir à recourir à un enseignant pour valider ou invalider sa réponse. Ainsi, l'erreur est davantage perçue comme une richesse et une occasion de relever un nouveau défi qu'une source de pénalité. Un exemple de ce rapport à l'erreur a été observé dans les activités de robotique pédagogique organisées dans un laboratoire créatif numérique, dénommé EspaceLab, avec des apprenants de 7 à 14 ans de la ville de Québec. Pendant ces activités, nous avons observé des apprenants qui n'avaient pas peur de se tromper lorsqu'ils proposaient une solution à un problème, qui utilisaient leurs erreurs pour améliorer la résolution de problèmes, et qui sollicitaient de moins en moins l'enseignant ou le facilitateur pour savoir si leur réponse fonctionnait.

3. LES DÉFIS LIÉS À L'INTÉGRATION PÉDAGOGIQUE DES ROBOTS À L'ÉCOLE

L'intégration pédagogique des robots à l'école a certes progressé depuis les premières expériences de Pepper, mais il reste de nombreux défis à relever. L'un d'entre eux est lié à la formation du personnel enseignant ou d'un personnel qualifié dans l'intégration pédagogique des robots en éducation. Si des enseignantes et enseignants, comme ceux ayant reçu une formation à l'université Laval ces trois dernières années, ont eu la possibilité pendant leur formation initiale de se familiariser avec l'intégration pédagogique des technologies robotiques, ce n'est pas toujours le cas pour leurs collègues formés dans les autres régions du monde. De même, les membres du personnel enseignant et les conseillères et conseillers pédagogiques qui travaillent depuis plusieurs années ne bénéficient pas tous d'une formation continue axée sur l'intégration pédagogique des technologies robotiques. Ce facteur à lui seul ne justifie pas les résistances observées lors de l'intégration pédagogique des robots à l'école. Un autre facteur est le coût

des technologies robotiques. Celles utilisées pour l'apprentissage sont parfois très coûteuses, ce qui limite leur accès à des écoles ayant les ressources financières suffisantes. En effet, certaines écoles ayant la volonté d'intégrer ces technologies pour l'ensemble de leurs apprenants font souvent face à des coûts très élevés.

Ces différents freins à l'intégration de la robotique pédagogique dans les écoles soulèvent la question de l'égalité des chances liées à l'accès à une éducation de qualité pour tous les apprenants. Est-ce une pratique réservée ou des outils destinés aux pays « riches » ? Précisément aux écoles « riches » ?

Pour favoriser l'accès de tous les apprenants aux technologies robotiques, il faut former le personnel enseignant à leur intégration pédagogique, développer des ressources moins coûteuses ou alors créer des laboratoires de fabrication numérique (FabLab) dotés des technologies robotiques. Ces laboratoires seront des lieux ouverts à toute personne intéressée par les usages pédagogiques de ces technologies.

BIBLIOGRAPHIE

- Chambers, J.M., Carbonaro, M., Murray, H. (2008). «Developing conceptual understanding of mechanical advantage through the use of Lego robotic technology», *Australasian Journal of Educational Technology*, 24, p. 387-401.
- Chih-Wei Chang, Gwo-Dong Chen (2010). «Using a Humanoid Robot to Develop a Dialogue-based Interactive Learning Environment for Elementary Foreign Language Classrooms», *Journal of Interactive Learning Research*, 21, p. 55-74.
- Dupont, Y., Sauvé, E., Touchette, P. (2010). *La robotique pédagogique en classe*. Québec, Éditions Didactic.
- Fels, D.I., Waalen, J.K., Zhai, S., Weiss, P.L. (2001). «Telepresence under Exceptional Circumstances: Enriching the Connection to School for Sick Children», *INTERACT*, p. 617-624.
- Fleming, V.F. (2014). «Developing 21st century skills through a virtual robotics course», *The Children's Technology & Engineering*, 18, p. 28-31.
- Halász, G., Michel, A. (2011). «Key Competences in Europe: interpretation, policy formulation and implementation», *European Journal of Education*, 46, p. 289-306.
- Jungho, P. (2015). «Effect of Robotics-Enhanced Inquiry-Based Learning in Elementary Science Education», *Journal of Computers in Mathematics & Science Teaching*, 34, p. 71-95.
- Kanda, T., Nishio, S., Ishiguro, H., Hagita, N. (2009). «Interactive Humanoid Robots and Androids in Children's Lives», *Children, Youth & Environments*, 19, p. 12-33.
- Park, J. (2015). «Effect of Robotics enhanced inquiry based learning in elementary Science education in South Korea», *Journal of Computers in Mathematics and Science Teaching*, 34, p. 71-95.

- Reverd, C. (2016). «Quelle est la place de la robotique pédagogique au sein de l'éducation ?», site web Vitrine technologie-éducation.
- Romero, M. (2016). «La robotique pédagogique, un outil privilégié pour le développement des compétences du 21^e siècle», *Zone01 Magazine*, 2, p. 28-32.
- Romero, M., Dupont, Y. (2016). Educational robotics: from procedural learning to co-creative project oriented challenges with lego wedo (p. 61596163). Présenté à International Conference on Education and New Learning Technologies, Barcelona, Spain.
- Romero, M., Kamga, R. (2016). « Usages de la robotique pédagogique en éducation primaire selon son intégration disciplinaire et le développement des compétences du 21^e siècle ». Présenté dans le cadre d' Intelligences numériques 2016, Québec.
- Tremblay, R.R. (2008). « Au-delà de l'opposition entre les connaissances et les compétences : réflexions sur l'approche par compétences et la formation générale [au collégial] ». *Pédagogie collégiale*, vol. 22, n^o 1, automne 2008, p. 23-29.
- Varney, M.W., Janoudi, A., Aslam, D.M., Graham, D. (2012). «Building Young Engineers: TASEM for Third Graders in Woodcreek Magnet Elementary School», *IEEE Transactions on Education*, 55, p. 78-82.
- Voogt, J., Roblin, N.P. (2012). «A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies». *Journal of Curriculum Studies*, 44, p. 299-321. doi:10.1080/00220272.2012.668938
- World Economic Forum (WEF) (2016). «The Future of Jobs Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution».
- Yeung, J., Fels, D.I. (2005). «A remote telepresence system for high school classrooms, in: Electrical and Computer Engineering», Canadian Conference on Electrical and Computer Engineering. *IEEE Xplore Digital Library*, p. 1465-1468.

CHAPITRE 17

Les animaux robots : quelles conséquences pour les enfants ?

Sophie CLOUTIER

L'entrée des robots dans nos vies et dans l'intimité de nos maisons aura des conséquences sur notre condition humaine et sur notre rapport aux autres et à la nature. Avant d'introduire ces nouvelles technologies dans notre environnement, il convient, pour reprendre l'injonction de Hannah Arendt, de penser ce que nous faisons. C'est dans cet esprit que nous aborderons la problématique des robots émotionnels, et plus spécifiquement, des animaux robots destinés aux enfants. Les enfants sont des êtres vulnérables, impressionnables et en formation, il faut donc agir avec beaucoup de prudence lorsque nous développons de nouvelles technologies destinées à leur usage. Nous proposons de réfléchir à certaines conséquences que pourraient avoir les animaux robots sur les enfants, et particulièrement sur leur développement moral. Qu'est-ce qu'ils pourraient gagner ou perdre au contact de ces technologies ? Qu'est-ce que nous voulons transmettre et développer comme potentiel chez les enfants ?

Le point de départ de notre réflexion sera la critique de Hannah Arendt de la science moderne qui contribue à l'aliénation par rapport au monde ; aliénation qui ouvre la voie au totalitarisme et contre laquelle il faut donc résister. Nous analyserons ensuite les animaux

robots comme une forme de technologie symptomatique de cette aliénation au monde. Cette analyse s'appuiera notamment sur les recherches de Sherry Turkle sur l'authenticité, la réciprocité et les interactions émotionnelles entre les humains et les robots. L'inquiétude de Turkle sur le narcissisme moderne, qui risque d'être exalté par les robots, permettra de poursuivre la réflexion sur la question de l'aliénation en abordant sa dimension psychologique, et d'ouvrir sur celle de la domination. La perspective qui se dégagera de cette analyse est celle de l'individu moderne aliéné au monde, rejeté en lui-même, ayant perdu la capacité de supporter l'étrangeté, qui ne trouve de réconfort émotionnel que dans la machine qu'il peut contrôler. Il se dessinera en filigrane une opposition entre une relation avec un être vivant, où il faut s'ouvrir à l'autre pour établir une réciprocité, et une relation avec un objet qui n'est qu'un miroir de nous-mêmes. Cette opposition sera approfondie dans la troisième section par le biais des apprentissages moraux que les enfants peuvent faire au contact des animaux biologiques. Notre hypothèse est que les relations de qualité avec les animaux contribuent au développement et à l'actualisation du potentiel éthique des enfants et pourraient ainsi servir de rempart à l'aliénation, contrairement à ce que pourraient offrir les robots, qui risquent de reconduire et d'alimenter l'aliénation et le mépris pour notre condition humaine vulnérable et mortelle.

1. REMARQUES PRÉLIMINAIRES SUR L'ALIÉNATION AU MONDE

Hannah Arendt, philosophe juive allemande, publie *The Human Condition* en 1958. Ce livre se voulait sa réponse au totalitarisme, c'est-à-dire qu'Arendt voulait repenser les conditions de possibilité d'existence dans un monde non totalitaire. Dans *The Origins of Totalitarianism*, publié en 1951, elle avait montré que le totalitarisme avait ses germes dans des problèmes du vivre-ensemble qui n'avaient pas trouvé de solutions politiques. Le totalitarisme était bel et bien né dans nos sociétés libérales modernes atomisées. Un de ces problèmes, qui est au cœur de *Condition de l'homme moderne*, est l'aliénation par rapport au monde, entendue comme le fait de devenir étranger à la terre et aux autres. Arendt évoque trois événements qui marquent le seuil de la modernité et «en fixent le caractère: la découverte de l'Amérique suivie de l'exploration du globe tout entier; la Réforme qui, en expropriant les biens ecclésiastiques et monastiques, commença le

double processus de l'expropriation individuelle et de l'accumulation de la richesse sociale; l'invention du télescope et l'avènement d'une science nouvelle qui considère la nature terrestre du point de vue de l'univers» (Arendt, 1983, p. 315).

Ces trois événements contribuent tous à leur manière à l'aliénation par rapport au monde. L'exploration a permis de mesurer et de cartographier la terre, ce qui a eu pour conséquence de rétrécir les distances et, du coup, la surface terrestre elle-même. Ce processus de rétrécissement nécessitait de mettre une distance entre l'humain et la Terre, d'aliéner l'humain de son milieu terrestre naturel (Arendt, 1983, p. 319). De son côté, la Réforme et le processus d'expropriation ont contribué à la perte de la condition politique élémentaire qui est l'appartenance-au-monde; certaines couches de la population ont été aliénées du monde. Arendt avance que la laïcité et la perte de foi et de transcendance ne correspondent pas à une passion renouvelée pour les choses de ce monde :

L'Histoire, au contraire, montre que les modernes n'ont pas été rejetés dans le monde : ils ont été rejetés en eux-mêmes. Une tendance persistante de la philosophie moderne depuis Descartes, sa contribution la plus originale peut-être à la philosophie, est le souci exclusif du moi, par opposition à l'âme, à la personne, à l'homme en général, la tentative de réduction totale des expériences, vécues par rapport au monde ou par rapport aux humains, à des expériences qui se passent entre l'homme et son moi. (Arendt, 1983, p. 322)

Arendt reprend la formule de Whitehead pour décrire la découverte du télescope : « jamais peut-être si grande chose n'est arrivée avec si peu de bruit » (1983, p. 326). Cet instrument, qui permettait de dévoiler ce qui autrement devait échapper aux sens humains, allait déterminer le cours d'autres événements qui donneraient naissance à l'époque moderne et sa science nouvelle. Arendt explique que la conception astrophysique moderne remet en question l'aptitude de nos sens à percevoir le réel. L'utilisation d'instruments fait en sorte que ce que nous connaissons de l'univers se réduit à la manière dont ses qualités affectent nos instruments de mesure. « En d'autres termes, au lieu de qualités objectives, nous trouvons des appareils, et au lieu de la nature de l'univers – pour citer Heisenberg – l'homme ne rencontre que soi. » (Arendt, 1983, p. 330-331) La science moderne est ainsi marquée par une aliénation par rapport à la Terre. L'algèbre moderne joue un

rôle important dans ce processus d'aliénation en ce qu'elle permet de réduire en symboles mathématiques les données des sens et les mouvements terrestres. Avec Descartes, la connaissance certaine, calquée sur le modèle de la certitude mathématique, est donnée par l'entendement qui joue seul avec ses formes et ses formules. Les mathématiques modernes offraient à la science une nouvelle possibilité d'appréhender des concepts et des dimensions infiniment petites ou infiniment grandes qui autrement demeuraient inaccessibles aux sens, fournissant du coup des moyens inouïs d'aborder la nature et d'entrer en contact avec elle. Comme l'ajoute Arendt :

C'est dans l'expérimentation que l'homme mit en pratique la liberté qu'il venait de gagner en brisant les chaînes de l'expérience terrestre; au lieu d'observer les phénomènes naturels tels qu'ils lui étaient donnés, il plaça la nature dans les conditions de son entendement, c'est-à-dire dans des conditions obtenues à partir d'un point de vue astrophysique universel, d'un point d'appui situé hors de la nature. (1983, p. 335)

À la différence de la géométrie platonicienne, dont les formes étaient données, les mathématiques modernes sont le produit de l'entendement humain, elles deviennent une science de la structure de l'entendement humain. Ou pour reprendre la formule d'Arendt, «les mathématiques réussirent à réduire et traduire tout ce que l'homme n'est pas en des schémas identiques aux structures mentales de l'homme» (1983, p. 336). Arendt voit le début de ce processus dans la géométrie analytique de Descartes qui fait en sorte que l'homme peut enfin se mouvoir dans la nature avec la sécurité de ne rencontrer que ses propres schémas, de ne finalement rencontrer que son propre moi. Les phénomènes qui ne pourraient pas se réduire à un ordre mathématique seront disqualifiés de toute façon. Nous sommes ainsi très loin de la démarche platonicienne qui recommandait l'étude de la géométrie pour préparer l'esprit à la révélation de l'Être véritable. Au contraire, tout est maintenant réduit à la mesure de l'esprit humain¹. Nous avons perdu le contact avec l'altérité de la nature, voire la science récuse le témoignage de la nature observée par les sens.

1. Dans une note, Arendt reprend la distinction de Whitehead entre la méthode scientifique traditionnelle de classification et la méthode moderne des mensurations: «la première suit les réalités objectives dont le principe se trouve dans l'altérité de la nature; la seconde est entièrement subjective, ne dépend pas des qualités et ne requiert autre chose qu'une multitude d'objets donnés.» (1983, p. 337, note 1)

Les inquiétudes que nous soulevons concernant les animaux robots destinés aux enfants sont ancrées dans cette lecture arendtienne de l'aliénation par rapport à la nature. Le contact avec les robots risque de creuser davantage ce fossé entre l'humain et son milieu de vie naturel. La création d'animaux robots nécessite de mathématiser leurs comportements ; l'animal robot n'est pas un animal, mais une représentation mentale de ce que nous pensons qu'il est. En ce sens, le contact des enfants avec l'animal robot risque aussi d'avoir des conséquences sur leurs capacités d'interactions avec ce qui est étrange et différent. Il favorisera l'enfermement en soi plutôt que la rencontre et l'ouverture à l'autre et à l'étrangeté. Ceci risque en retour d'avoir des conséquences sur le développement moral et la capacité d'empathie des enfants.

2. LES ANIMAUX ROBOTS ET L'ATTACHEMENT NARCISSIQUE

Le processus d'aliénation par rapport au monde semble aujourd'hui prendre de nouvelles proportions. Comme le notait Carol Gigliotti, il y a quelque chose de paradoxal à une quête de vie artificielle, alors même que la dévastation de la réelle vie non humaine s'accélère (2005, p. 72). La vie artificielle, comme désir d'une « vie-comme-elle-pourrait-l'être » (Gigliotti, 2005, p. 74), permet de constater à quel point l'humain cherche à échapper aux conditions dans lesquelles la vie sur terre est donnée. On peut y lire le désir ultime de maîtrise et de domination sur la nature. Les animaux robots pourraient aussi apparaître comme une nouvelle forme de domination sur les animaux, au sens où nous en serions rendus à les accepter dans nos foyers seulement si leur comportement est contrôlé. L'animal robot pourra effectivement être programmé pour ne pas faire de dégâts dans la maison et ne pas avoir de comportements agressifs ou imprévisibles. En ce sens, l'animal robot soulève des enjeux pédagogiques importants, notamment sur le plan du développement moral, sur lesquels nous reviendrons dans la troisième section.

Sherry Turkle, psychologue et professeure au MIT, s'intéresse depuis le début des années 1990 aux relations que les personnes entretiennent avec les technologies. Ses recherches montrent que les gens se sentent de plus en plus seuls et isolés et que le recours aux technologies est à la fois symptôme et cause d'aggravation de ce sentiment. L'isolement social et l'impression de ne pas être écouté par les autres poussent les

gens à chercher des connexions en ligne ou avec des robots, ce qui en retour a pour conséquence qu'ils perdent leurs capacités relationnelles avec les autres. Comme elle l'observe :

Students' trust in Eliza did not speak to what they thought Eliza would understand but to their lack of trust in the people who would understand. [...] A graduate student confides that she would trade in her boyfriend for a "sophisticated Japanese robot", if the robot would produce "caring behavior". The graduate student says she relies on a "feeling of civility" in the house. If the robot could "provide the environment", she would be "happy to produce the illusion that there is somebody really with me". (2007, p. 502)

Une machine qui est conçue pour simuler un intérêt stimulera ce que Turkle appelle nos « boutons darwiniens », *Darwinian buttons* (Turkle, 2007, p. 503), qui envoient le signal que nous sommes devant une entité ayant un but relationnel. Les robots émotionnels ou les animaux robots, tels que Furby, Aibo, My Real Baby et Paro, sont justement conçus pour afficher des comportements qui font en sorte que les gens sentent qu'ils sont des créatures sensibles qui se soucient de leur présence. Ils ont des caractéristiques comme le contact visuel, le suivi de la localisation de la personne dans une pièce et des gestes pour mimer la reconnaissance de la présence de la personne. Toutes ces caractéristiques stimulent le désir de prendre soin et aussi le désir de réciprocité. Les gens veulent prendre soin de leur robot et désirent que le robot prenne soin d'eux en retour. La stimulation de ce désir de prendre soin est aussi devenue une des applications les plus recherchées par les concepteurs, puisqu'elle procure un avantage sur le marché et permet de se distancer de la compétition (Turkle, 2007, p. 506).

Selon Turkle, le problème central de la relation avec les robots ne concerne pas tant les sentiments des humains envers les machines, mais les sentiments des machines envers les humains, ou pour être plus précis, leur manque de sentiments. Les robots ne peuvent pas réciproquer l'attachement, ils ne peuvent que le feindre. Avec l'avènement des robots émotionnels, il faudra donc trouver des critères pour départir les relations authentiques des relations inauthentiques, ce qui s'avère une question particulièrement urgente pour les enfants grandissant entourés de ces technologies. Comme Turkle l'explique :

With robots, people are acting out “both halves” of complex relationships, projecting the robot’s side as well as their own. Of course, we can also behave this way when interacting with people who refuse to engage with us, but people are at least capable of reciprocation. We can be disappointed in people, but at least we are disappointed about genuine potential. For robots, the issue is not disappointment, because the idea of reciprocation is pure fantasy. (2007, p. 504-505)

Les recherches avec les enfants montrent en effet qu’ils s’attachent aux animaux robots parce qu’ils doivent en prendre soin, ils deviennent comme leurs parents. Si les robots émotionnels sont pourvus de caractéristiques qui stimulent cette pseudorelation de parentalité, ils sont cependant dépourvus de la capacité de réciproquer l’attachement. La personne doit donc construire elle-même la réponse émotionnelle supposée du robot. Par exemple, l’enfant prend soin de son animal robot et croit que son robot l’aime en retour. L’artefact relationnel est ainsi construit comme une partie de soi. En ce sens, Turkle soulève l’inquiétude suivante :

Someday, if relational artifacts can give the impression of aliveness and not disappoint, they may have a “comparative advantage” over people as self objects and open up new possibilities for narcissistic experience. For some, predictable relational artifacts are a welcome substitute for the always-resistant human material. What are the implications of such substitutions? Do we want to shore up people’s narcissistic possibilities? (2007, p. 510)

En effet, qu’advierait-il de notre monde si les humains préféraient interagir avec des robots plutôt qu’avec d’autres êtres vivants? Étant donné que notre expérience avec les robots émotionnels est fondée sur une fiction – ils nous connaissent et se soucient de nous –, est-ce que ce type d’attachement est une bonne chose pour nous en tant qu’êtres moraux? Est-ce que nous voulons vraiment d’une humanité qui développe une plus grande intimité avec des machines? Les enfants qui grandissent avec ces nouvelles technologies et qui vivent dans un monde où les relations humaines sont déplacées dans le monde virtuel ont de plus en plus de difficulté à faire la distinction entre ce qui est en vie et ce qui a l’apparence d’être en vie. Leur animal robot a autant, sinon plus, de réalité que l’oiseau qui vole dans le ciel là-bas. Pourront-ils encore se soucier du fait que certaines espèces animales disparaissent? Comme le signale Turkle :

Ancient cultural axioms that govern our concepts about aliveness and emotion are at stake. Robots have already shown the ability to give people the illusion of relationship: Paro convinced an elderly woman that it empathized with her emotional pain; students ignored the fact that Eliza was a parrot-like computer program, choosing instead to accept its artificial concern. Meanwhile, examples of children and the elderly exchanging tenderness with robotic pets bring science fiction and technophilosophy into everyday life. Ultimately, the question is not whether children will love their robotic pets more than their animal pets, but rather, what loving will come to mean. Going back to the young woman who was ready to turn in her boyfriend for a “sophisticated Japanese robot” is there a chance that human relationships will just seem too hard? (2007, p. 514)

Le narcissisme que Turkle remarque dans la relation avec les animaux robots prolonge dans le champ de la psychologie l'analyse arendtienne de l'aliénation par rapport au monde. Les humains, rejetés en eux-mêmes, n'ont plus qu'eux-mêmes à aimer, ils ne peuvent aimer que leurs propres productions comme miroir d'eux-mêmes. À un moment où nous assistons à la disparition de plusieurs espèces animales, pourquoi créer des animaux robots? N'est-ce pas un transfert émotionnel qui prive les animaux de cette ressource même qui pourrait les préserver? Quelles conséquences les relations d'attachement avec les robots auront-elles sur les relations avec les autres êtres humains et les animaux? Allons-nous perdre la capacité de surmonter le silence de l'autre, sa non-réaction, sa réaction tardive? Par contraste, la machine répondra toujours et rapidement, à moins qu'elle soit brisée, et elle devra alors être réparée ou remplacée. Nous risquons de devenir de plus en plus allergiques à ce qui n'est pas nous; les autres seront réduits à des facteurs d'agression qui nous dérangent et nous contrarient.

Peter H. Kahn Jr., Heather E. Gary et Solace Shen (2013) se sont aussi penchés sur les relations que les enfants peuvent entretenir avec les robots sociaux et les conséquences sur leur développement moral. Ils ont comparé le comportement des enfants avec un chien en peluche et avec Aibo (un chien robot) et ont remarqué que les enfants avaient une tendance beaucoup plus grande à croire qu'Aibo était capable de réciprocité que le chien en peluche. Bien que les trois quarts des enfants fussent conscients qu'Aibo était un artefact technologique, ce savoir ne les empêchait pas de concevoir et de traiter le chien robot

comme s'il était un chien biologique (Kahn *et al.*, 2013, p. 33). Leur recherche soulève la question fondamentale de la confusion des catégories ontologiques causée par les robots sociaux. Les personnes sont généralement capables de catégoriser les entités dans le monde, c'est-à-dire de distinguer entre les objets animés et les objets inanimés, et elles n'attribueront pas les mêmes capacités aux deux types d'entités. À moins de troubles graves, même les enfants en bas âge ont déjà développé ces capacités de distinction. Le problème, c'est que les animaux robots viennent défier ces distinctions parce qu'ils brouillent l'appareil de catégories qui s'applique aux humains, aux animaux et aux artefacts en incarnant certains aspects de chacune de ces trois catégories. Les chercheurs posent ainsi l'hypothèse de l'émergence d'une nouvelle catégorie ontologique liée à la production des animaux robots :

Just as young children experience the color orange as a unique color, not merely as a combination of red and yellow, the NOC [new ontological category] hypothesis leads to the prediction that children growing up with social robots of the near future will likely see, conceptualize, and interact with them as a unified entity, and not merely a combinatorial set of its constituent properties. Granted, we might continue to ask such questions as “are these robots alive,” “are they fundamentally an artifact,” and “are they a social being.” Yet children might find such questions puzzling and answer each of them “yes and no”. (Kahn *et al.*, 2013, p. 34)

Cette confusion entre les catégories causant l'émergence d'une nouvelle catégorie ontologique risque d'avoir de graves conséquences sur le développement moral des enfants. Plusieurs psychologues, dont Jean Piaget et Lawrence Kohlberg, ont montré que le développement de la réciprocité s'effectuait principalement dans la sphère des relations entre pairs. La capacité de réciprocité est centrale pour le développement moral de l'enfant puisqu'elle permet de développer le souci pour le bien-être des autres ainsi que les notions d'égalité, d'équité et de justice. Les relations affectives et sociales que les enfants entretiennent avec leur robot soulèvent ainsi plusieurs inquiétudes :

What will happen if children grow up interacting with robots as peers and even friends because of the robots' sociality, but also objectifying if not dominating the robots because the children understand that the robots are a human technological creation? Imagine, for example, if an 8-year-old's best friend is a social

robot. The child plays with the robot every day and goes to it for comfort and companionship. The robot always does what the child wants and the child never needs to accommodate to the social interests and needs of the robot. Does that situation put into motion a master-servant relationship that we would not want to reify? (Kahn *et al.*, 2013, p. 35)

Le fait que les robots sont considérés par les enfants à la fois comme entité sociale et comme objet ouvre la voie à une relation potentielle de domination de type maître-esclave. L'enfant utilise le robot pour assouvir certains besoins, dont des besoins d'ordre affectif, sans avoir pour autant à se soucier des besoins de son robot, qui n'est en somme qu'un artefact. Cette situation d'objectivation partielle des robots sociaux pourrait avoir des conséquences néfastes sur le développement moral de l'enfant puisque la manière dont les enfants traitent une catégorie ontologique peut se transposer sur une autre catégorie (Kahn *et al.*, 2013, p. 35). Plusieurs recherches montrent en effet que la violence et la cruauté perpétrées envers les animaux sont des précurseurs de la violence faite à l'encontre des humains. Par exemple, Carol J. Adams (1995) relate de manière troublante comment les hommes qui battent leur conjointe ont aussi souvent maltraité ou tué des animaux, voire ont forcé la femme à avoir des rapports sexuels avec des animaux. Adams analyse ce phénomène en termes de système de domination dans une lecture intersectionnelle entre le sexisme et la violence faite aux animaux. La domination sur les femmes et sur les animaux participerait d'un même type de logique d'objectivation (Adams, 2000).

Il est donc préférable d'appliquer le principe de précaution et de réfléchir aux conséquences potentielles des interactions entre les enfants et les robots sociaux. Si ces derniers peuvent avoir certains avantages sur le plan pédagogique, ils pourraient aussi avoir des effets néfastes sur le développement moral. Nous avons soulevé les enjeux du narcissisme et d'une potentielle réification de la relation maître-esclave, qui influencent négativement la capacité de réciprocité. À défaut d'un apprentissage des relations intersubjectives, nous risquons de condamner les enfants à l'aliénation au monde. Les nouvelles technologies sont souvent introduites en pédagogie sans qu'elles aient nécessairement fait leurs preuves ou sans que leurs conséquences aient été bien évaluées². En

2. Nous allons ici dans le sens de la thèse d'Ulrich Beck (2008) sur la société du risque. La modernité, avec son idéal de progrès, a engendré toute une série de maux, de menaces et de dangers que Beck range sous la catégorie des risques. Selon lui, le principal problème

revanche, plusieurs recherches mettent en évidence le rôle que les relations avec les animaux peuvent avoir sur le développement moral, notamment sur le développement de la capacité d'empathie.

3. L'APPRENTISSAGE MORAL PAR LES RELATIONS AVEC LES ANIMAUX BIOLOGIQUES

Afin d'approfondir la question du développement moral, nous procéderons dans cette section sous forme de comparaison animal/machine, c'est-à-dire que nous considérerons certaines études portant sur le rôle des animaux biologiques dans le développement moral des enfants en nous demandant si ces apprentissages seraient possibles avec les robots. Les enfants sont en formation, il importe donc de se demander ce que nous voulons leur transmettre et comment développer leurs capacités. De plus, les robots sont très dispendieux et leur utilisation massive en éducation entraînerait d'importants coûts, tandis que certaines initiatives plus abordables ont fait leurs preuves, comme nous le verrons avec les animaux, mais ne sont pas très répandues. Il y a fort à parier que l'influence d'une certaine idéologie du progrès fait pencher la balance en faveur des nouvelles technologies et qu'un certain préjugé à l'égard des animaux nous empêche de considérer sérieusement les bénéfices qu'apportent les relations avec eux.

Nous allons dans le sens de Maurice Hamington (2008) qui fait valoir que les relations de qualité avec les animaux contribuent au développement de nos facultés morales. Hamington argumente en effet que les humains ont beaucoup à apprendre des animaux, et cela non seulement sur le plan scientifique par l'observation de leur comportement, mais surtout sur le plan éthique, puisque les interactions avec les animaux stimulent l'imagination morale qui est à la base du sentiment d'empathie (2008, p. 177). Il s'inscrit dans la perspective de l'éthique du *care*³ et avance que l'empathie et le souci sont des

est justement l'action dans la nature, c'est-à-dire le fait que des nouvelles technologies sont introduites dans la nature sans que nous connaissions la portée de leurs conséquences, mobilisant ainsi la critique arendtienne de la modernité (1983).

3. Nous référons à l'ouvrage de Marie Garrau et Alice Le Goff (2010) pour une définition et une mise en contexte de l'émergence de l'éthique du *care*. L'ouvrage de Joan Tronto (2009) offre aussi une lecture de la portée politique de l'éthique du *care* et en présente les quatre phases.

capacités morales fondamentales. Selon lui, l'éthique du *care* a aussi l'avantage de considérer les animaux comme des êtres particuliers qui font partie de notre monde et avec lesquels nous entretenons des relations particulières, plutôt que de les considérer comme des êtres abstraits ou de simples données biologiques⁴. L'éthique du *care* défend aussi une anthropologie qui contraste avec la vision moderne du sujet comme être autonome. Elle considère plutôt l'humain comme un être fondamentalement vulnérable, un être social inscrit dans un réseau de relations – qui implique plusieurs relations de dépendance qui font de chacun de nous à la fois des bénéficiaires de soins (enfance, maladie, vieillesse) et des pourvoyeurs de soins. Les animaux font partie de ces réseaux et auraient même avantage à en faire partie dès le plus jeune âge puisque, selon Hamington, les relations de qualité avec les animaux contribuent à l'actualisation du potentiel éthique des humains.

Hamington définit l'empathie comme la capacité à utiliser notre imagination pour comprendre les sentiments et les actions des autres ou la situation dans laquelle ils se trouvent. Les humains ont une grande capacité d'empathie, même pour ce qui ne leur est pas familier; ils peuvent en effet éprouver de l'empathie pour des étrangers, des animaux ou des personnages fictifs en littérature ou au cinéma. Selon Hamington, l'empathie est vécue dans un rapport au corps, elle est incarnée et en ce sens, « [i]magination is a crucial link between what the body knows, empathy, and care » (Hamington, 2008, p. 181). La proximité physique avec les animaux joue ainsi un rôle dans le développement de l'empathie, comme l'avance Hamington :

If proximal bodily relations provide experiential resources for empathy between humans, it follows that such experiences also occur with our interactions with other embodied beings. Certain species of animals such as cats and dogs have a history of strong embodied relations with humans. The term “natural” is not one that philosophers invoke lightly, but the contact between companion animals and humans often appears natural in terms of ease, frequency, and mutual enjoyment. The medical, hospice, and convalescent communities have acknowledged the human-animal connection through Animal-Assisted Activities and

4. À propos de la question de la défense des animaux, Hamington défend la pertinence de l'éthique du *care* et la distingue de l'approche de Tom Regan, centrée sur la question des droits des animaux, et de celle de Peter Singer, basée sur l'utilitarisme. Sur cette question, nous référons aussi à l'ouvrage collectif dirigé par Sandra Laugier (2012), où plusieurs chapitres abordent l'apport de l'éthique du *care* dans le débat sur la défense des animaux.

Therapy (AAA/T) Programs. The stories of animals mitigating the distress of patients through their presence and contact are legion. Such animals play an affective role in a manner that no material object can in improving the quality of care for patients. There is a familiar comfort between the bodies of humans and animals that does not require articulation to initiate. (2008, p. 181)

Ainsi, le rapport avec les animaux implique la corporéité, il nous rappelle à notre corps, à notre animalité, et se passe entre deux corps vivants différents. Cette relation diffère aussi de celle entre les humains, puisqu'elle n'implique pas la même relation au langage. Il faut en effet apprendre à écouter un autre type de langage. Cet apprentissage peut s'avérer très bénéfique pour la relation avec les étrangers qui ne parlent pas notre langue ou avec les personnes avec certains handicaps qui ne peuvent pas ou ne peuvent plus employer le langage verbal usuel. Les enfants qui sont en contact de proximité avec les animaux peuvent ainsi apprendre à sentir les besoins de l'autre par-delà un langage commun, voire développer des outils de communication entre espèces. Cet apprentissage ne nous semble cependant pas possible au contact des robots, puisqu'ils ne partagent pas notre condition corporelle d'êtres vivants ayant des besoins particuliers et capables de ressentir la douleur. L'entretien d'un robot demande des connaissances techniques et engage des coûts qui dépassent les compétences d'un enfant. L'enfant devra certes apprendre à ne pas briser son robot, mais non pas à en prendre soin de la même manière qu'il peut apprendre à prendre soin d'un animal ayant des besoins particuliers. Par le biais des relations avec des corps vivants non humains, les enfants peuvent ainsi développer leur capacité de « prendre soin » et de souci (*care*) pour les animaux, qui pourra se transposer aux humains non familiers. Au contraire, le contact avec les robots risque d'éloigner les enfants de leur condition corporelle et mortelle, voire les encourager à entretenir des préjugés à l'égard de la vulnérabilité. À cet égard, l'étude de Turkle *et al.* (2006) donne l'exemple de Jimmy, un jeune garçon atteint d'une maladie congénitale, qui croit qu'Aibo entretient envers lui des sentiments similaires à ceux de son chien biologique, qu'il peut s'ennuyer de lui quand il n'est pas là. Cependant, Aibo possède un avantage que son chien biologique n'a pas, il ne peut pas vraiment mourir :

During several sessions with Aibo, Jimmy talks about Aibo as a super dog that shows up his own dog as a limited creature. Jimmy says: "Aibo is probably as smart as Sam and at least he isn't as scared as my dog [is]". When we ask Jimmy if there are things that

his dog can do that Aibo cannot do, Jimmy answers not in terms of his dog's strengths but in terms of his deficiencies: "There are some things that Sam can't do and Aibo can. Sam can't fetch a ball. Aibo can. And Sam definitely can't kick a ball". On several other occasions, when Aibo completed a trick, Jimmy commented "My dog couldn't do that!" Aibo is the "better" dog. Aibo is immortal, invincible. Aibo cannot get sick or die. In sum, Aibo represents what Jimmy wants to be. (Turkle *et al.*, 2006, p. 353)

Le robot Aibo vient ici éloigner l'enfant malade de sa condition mortelle et de sa vulnérabilité. À l'inverse, l'animal nous ramène à notre corporéité, notre vulnérabilité et notre mortalité. Il s'agit bien là d'aspects fondamentaux de notre condition humaine que les enfants ont avantage à découvrir et apprivoiser plutôt qu'à dénier. Les relations que les enfants entretiennent avec les robots risquent au contraire d'alimenter le fantasme de l'autonomie, de la puissance et de l'invincibilité et le mépris concomitant pour la vulnérabilité, la souffrance et la mortalité. Ce mépris aura en retour des conséquences politiques et sociales et contribuera à l'atomisation sociale et l'aliénation au monde. L'éthique du *care* a abondamment montré qu'afin de mettre en place des services pour aider, il faut d'abord reconnaître le besoin⁵. La reconnaissance de la vulnérabilité serait ainsi à la base de la socialité humaine, l'humain a besoin des autres et d'un certain environnement pour survivre. En tissant des relations avec les animaux, les enfants peuvent développer leurs capacités d'empathie et d'attention aux autres nécessaires à la prise en charge des besoins. Différentes études ont en effet montré que les défenseurs des droits des animaux et de l'environnement avaient eu des expériences significatives avec des animaux dans leur jeunesse, démontrant ainsi le potentiel des relations avec les animaux dans le développement de la sensibilité éthique des enfants (Rice et Rud, 2016; Renck Jalongo, 2013; Hamington, 2008; Pattnaik, 2004).

Selon Hamington, des initiatives aussi simples que le jeu de l'anthropomorphisme peuvent contribuer au développement de l'imagination morale, entendue comme la capacité à se mettre à la place des autres et à adopter une mentalité élargie. En effet, par le biais de jeux éducatifs où les enfants donnent des caractéristiques humaines aux animaux, ils peuvent apprendre à se familiariser avec ce qui est d'abord étranger, voire menaçant. Ces jeux stimulent leur imagination et leur permettent de développer une plus grande sensibilité pour le

5. Voir par exemple Tronto, 2009.

sort des animaux, sensibilité qui se transpose ensuite plus facilement aux humains (Hamington, 2008, p. 185). Si les sciences sociales s'intéressent de plus en plus au lien entre la cruauté envers les animaux et celle envers les humains, Hamington veut montrer le lien inverse, c'est-à-dire que le souci et le « prendre soin » des animaux favorisent le souci pour les humains – ce qui fait écho à la thèse de Kahn *et al.* (2013) sur le transfert du traitement d'une catégorie ontologique à une autre. Hamington conclut :

Empathizing with unfamiliar others may be the moral challenge of the twenty-first century. Technology has made the world smaller than ever but the newfound intimacy has not translated into widespread sympathetic understanding. Violent conflict continues to permeate a world separated into “us” and “them.” [...] If we can see our way to care for non-human creatures, not as property, but as extensions of ourselves, perhaps we can also come to care for and about one another. [...] if I have the moral will not to participate in violence perpetuated against animals, the challenge to care for bipedal animals that are a lot like me does not seem so great. (2008, p. 186)

Plusieurs recherches en éducation vont dans le sens de Hamington et montrent le potentiel pédagogique des relations avec les animaux. Les évaluations de différents projets éducatifs avec des animaux présentent des résultats très positifs pour des initiatives qui sont souvent très simples et peu coûteuses, surtout si on les compare aux coûts associés aux robots. L'ouvrage collectif de Rice et Rud (2016) relate différentes initiatives, dont les Scouts et les Girl Scouts qui permettent aux enfants d'apprendre que le bien-être des humains est lié à celui des animaux et où les enfants ont l'occasion de nouer des liens d'amitié avec des êtres d'autres espèces. Cet apprentissage permet en retour aux enfants de développer une sensibilité pour la justice sociale, la paix et l'écologie. L'ouvrage rapporte aussi l'initiative d'une professeure d'anglais qui a décidé d'avoir un lapin dans sa classe, procurant ainsi aux enfants l'expérience d'un apprivoisement mutuel. Les enfants ont en effet dû apprendre à bien s'occuper du lapin, lui donner les meilleures conditions de vie et lui construire un habitat approprié afin que celui-ci puisse en retour apprendre à socialiser avec eux. Ils ont développé leur sens des responsabilités et leur souci en apprenant à être moins turbulents en classe afin de ne pas effrayer leur ami lapin. L'ouvrage donne aussi en exemple un projet dans les prisons américaines avec des lévriers. L'initiative est relativement simple, il s'agit de faire appel

à des détenus pour réhabiliter les lévriers qui ne sont plus aptes aux courses à réintégrer la vie sociale afin de pouvoir être adoptés. Ces chiens ont été dressés pour les courses et n'ont jamais vécu en proximité avec des humains; sans cette «réhabilitation», ils ne peuvent vivre dans des familles et sont condamnés à être tués. Les détenus reçoivent des formations de dressage qui interdisent les réprimandes et favorisent plutôt la récompense des bons comportements. L'aspect intéressant de cette initiative est l'apprentissage mutuel, la coformation entre espèces. Les détenus apprennent le contrôle de leurs émotions, la patience et comment donner et recevoir de l'amour (Rice et Rud, 2016, p. 93). Les détenus témoignent par exemple qu'ils ont appris à devenir de meilleures personnes, à être plus compatissants et généreux, plus attentifs aux besoins des autres grâce aux relations avec leur lévrier. Les recherches de Jyotsna Pattnaik (2004) montrent que le contact avec les animaux procure aux jeunes enfants plusieurs avantages pour le développement: la stimulation tactile est favorisée par le fait de toucher l'animal; le développement physique par le fait de suivre l'animal et le développement émotionnel par la relation d'amour réciproque entre l'enfant et l'animal. Les relations de proximité avec les animaux permettent aux enfants de tester leurs capacités de gentillesse, d'empathie, de cruauté et d'indifférence (Pattnaik, 2004, p. 94). La compréhension des besoins particuliers de leur animal de compagnie permet en outre aux enfants de quitter graduellement le stade de l'égoïsme et du narcissisme propre à l'enfance. Ils apprennent à décoder les différents états d'un autre être à travers des indices non verbaux.

Nous pourrions multiplier les exemples d'initiatives et la liste des bienfaits puisque l'apprentissage avec les animaux n'a finalement rien de nouveau. Il est au cœur des pratiques éducatives autochtones et va de soi pour les personnes habitant dans des régions agricoles, particulièrement pour les enfants grandissant sur de petites fermes en proximité avec les animaux. Comme le remarque Susan Laird: «The idea of befriending animals with young people as an educational practice has early and late modern roots, today commonly neglected, and now requires further critical and experimental thought, especially concerning its possible value for learning to live responsibly in this postmodern era of global corporatist climate crisis» (citée par Rice et Rud, 2016, p. 169). Et il s'agit bien ici d'un enjeu de fond: en dépit de sa longue tradition, la promotion de l'apprentissage moral par le biais des relations avec les animaux souffre aujourd'hui de l'aliénation

au monde et du préjugé que la modernité entretient à l'égard des animaux.

Le philosophe Jacques Derrida a analysé ce préjugé comme une guerre contre les animaux⁶. Bien que la question animale fasse partie des réflexions tardives et inachevées de Derrida, elle se situe néanmoins dans l'ensemble de son projet de déconstruction de l'idée de la rationalité occidentale par laquelle l'homme s'est octroyé le privilège de dominer l'ensemble des vivants et la nature. Derrida (2006) avance que le privilège sur et la domination des animaux s'élabore comme une tentative de se distinguer du reste des êtres vivants. Tout se passe comme si les hommes ne voulaient pas faire partie du règne animal, tentaient de le fuir pour échapper à leur propre finitude. Pour le dire avec Jean Grondin : « Le soupçon de Derrida est que l'autopromotion de la différence qu'incarnerait la raison humaine s'érige sur un oubli, une répression, voire une violence originaire, celle qui consiste à nier l'animalité de l'homme et, par conséquent, sa mortalité. Tout son travail se veut une lutte contre cette violence, car elle équivaudrait, littéralement, à une "dénaturation" de l'homme. » (2007, p. 34) Cet oubli de l'animalité fait écho à l'analyse arendtienne de l'aliénation au monde : rendus étrangers à la nature, les humains peuvent plus aisément ignorer leur appartenance corporelle à l'environnement.

Il faut souligner que dans *L'animal que donc je suis*, Derrida réfléchit à partir de l'expérience de la relation qu'il entretient avec son chat – et de sa honte d'avoir honte d'être nu devant son chat (2006, p. 18) – afin de confronter une tradition philosophique qui méprise et fait violence aux animaux. Il exemplifie en ce sens l'argument de Hamington sur la corporéité, l'imagination et l'empathie. Son expérience vécue lui permet d'être attentif à la violence qui est faite aux animaux lorsqu'on nie leur particularité pour les dissoudre dans une catégorie, comme le règne animal. Dans la relation avec son chat, Derrida expérimente sa propre vulnérabilité et sa propre mortalité ; ce chat particulier, qui s'adresse à lui, est d'une irremplaçable singularité :

il vient à moi comme ce vivant irremplaçable qui entre un jour dans mon espace, en ce lieu où il a pu me rencontrer, me voir,

6. Il écrit en effet : « je crois que le cartésianisme appartient, sous cette indifférence mécaniste, à la tradition judéo-christiano-islamique d'une guerre contre l'animal, d'une guerre sacrificielle aussi vieille que la genèse. [...] cette violence ou cette guerre ont été jusqu'ici, constitutives du projet ou de la possibilité même du savoir technoscientifique dans le processus d'humanisation ou d'appropriation de l'homme par l'homme » (2006, p. 140).

voire me voir nu. Rien ne pourra jamais lever en moi la certitude qu'il s'agit là d'une existence rebelle à tout concept. Et d'une existence mortelle, car dès lors qu'il a un nom, son nom lui survit déjà. Il signe sa disparition possible. La mienne aussi – et cette disparition, d'ici là, *fort/da*, s'annonce chaque fois que, nudité ou non, l'un de nous quitte la pièce. (Derrida, 2006, p. 26)

Derrida expérimente l'être-avec l'animal et s'étonne que les philosophes modernes aient pu observer les animaux sans jamais se sentir observés par eux, regarder un animal sans reconnaître que l'animal pouvait aussi les regarder et s'adresser à eux. Ce déni de mutualité possible serait au fondement de la justification de la domination sur les animaux – et nous pourrions ajouter qu'il trahit un manque de relation de proximité de qualité avec les animaux. Les philosophes ont dénié le statut de « sujet » aux animaux parce qu'ils les ont considérés comme étant dépourvus de certaines caractéristiques, comme la raison, le langage ou la conscience. Derrida met en question ce déni et joue sur la polysémie du mot « suis » qui peut évoquer une modalité de l'être (être un animal) ou l'action de suivre (traquer l'animal/venir après l'animal⁷):

En quel sens du « prochain » (qui n'est pas forcément celui d'une tradition biblique ou gréco-latine) devrais-je dire que je suis proche ou près de l'animal, et que je le suis, et dans quel ordre de pression ? l'être-avec-lui comme être-près-de-lui ? être-auprès-de-lui ? être-après-lui ? *être-après-lui* au sens de la chasse, du dressage, du domptage ou *être-après-lui* au sens de la succession et de l'héritage ? [...] Et depuis cet être-là-devant-moi, il peut se laisser regarder, sans doute, mais aussi, la philosophie l'oublie peut-être, elle serait même cet oubli calculé, il peut, lui, me regarder. Il a son point de vue sur moi. Le point de vue de l'autre absolu, et rien ne m'aura jamais tant donné à penser cette altérité absolue du voisin ou du prochain que dans les moments où je me vois vu nu sous le regard d'un chat. (2006, p. 28)

Il nous semble qu'être auprès d'un robot n'aurait pas la même portée, n'équivaldrait pas à la même expérience qu'être auprès d'un animal vivant particulier. Il s'agirait de se reconnaître soi-même dans son

7. Sur le sens de la succession et de l'héritage, nous référons aux travaux du biologiste et philosophe Georges Chapouthier (2009a; 2004) qui démontrent la continuité biologique et comportementale entre les animaux non humains et les humains. Chapouthier défend, entre autres, l'idée que l'animalité serait la racine de la culture.

produit ou de constater les prouesses techniques de l'humain plutôt que de s'ouvrir à l'altérité et l'étrangeté. L'animal robot peut certes interagir avec l'humain, mais il ne peut l'interpeler comme peut le faire le regard de l'animal biologique en ouvrant la question de la responsabilité. Cet animal vivant particulier qui s'avance vers moi, établit un lien, une relation, il peut communiquer un besoin et me demander d'y répondre. Il se rappelle à moi et son corps m'éveille à mon corps, sa mortalité me renvoie à la mienne. Cet animal particulier peut aussi me surprendre, me contrarier, s'opposer à moi. La relation avec l'animal biologique est ainsi une expérience très différente de celle avec un animal robot où, comme nous l'avons vu précédemment, l'enfant construit la réponse du robot et ne joue finalement qu'avec lui-même.

En somme, Derrida, ainsi que plusieurs autres penseurs que nous avons abordés, cherche à nous responsabiliser et à nous rappeler nos obligations à l'égard du vivant en général. Nous savons que les animaux peuvent souffrir, alors la violence et la cruauté envers les animaux reviennent à violer notre sens de la compassion et de la pitié. L'animal robot risque ainsi de devenir une manière d'éviter la question fondamentale posée par Derrida, à savoir la « guerre au sujet de la pitié » (Derrida, 2006, p. 50) qui se traduit dans un manque de compassion envers certains animaux, voire certains êtres humains. L'animal robot peut bien alimenter un sentiment d'attachement, il ne pourra jamais complètement nous faire pitié, il ne sera qu'un ersatz de pitié. L'animal robot destiné aux enfants pose ainsi le risque de nuire à leur développement affectif puisqu'ils cultiveront un attachement faible et facilement révoquant : il s'agit après tout d'un artefact technologique qui deviendra sans doute obsolète rapidement. Les enfants qui grandissent avec des animaux robots plutôt qu'avec des vivants risquent aussi de reconduire la conception cartésienne de l'animal-machine.

Le raisonnement de Descartes sur l'animal-machine est en effet lourd de conséquences pour la modernité, puisqu'il justifie la domination humaine sur le règne animal. Descartes avance que si l'on pouvait fabriquer des automates – aujourd'hui des animaux-robots –, qui ressembleraient en tout et pour tout à des animaux, il serait impossible de les distinguer. Cependant, si on faisait de même avec des machines qui ressembleraient à des humains, il y aurait deux moyens infaillibles de les distinguer des hommes : jamais les machines ne pourraient user de la parole pour répondre à tout ce qui serait dit en leur présence et jamais ces machines ne pourraient faire tout ce que les hommes peuvent faire. Un manque les trahirait toujours et c'est aussi par

un manque que Descartes distingue les humains des animaux : ces derniers sont dépourvus de raison. Mais, insiste Derrida, « il s'agit de l'animal théorétique » (2006, p. 116), celui qui est supposé vu sans voir. C'est sur la base de ce manque, de cette déficience présumée des animaux, qui ne sont que des automates en chair et en os, que l'humain s'est arrogé le droit de les dominer sans mouvement de compassion. Ce qui recoupe ce que Georges Chapouthier nomme le « "solipsisme d'espèce", proche de ce que certains auteurs ont appelé le "spécisme", selon lequel la pensée humaine constitue la seule pensée » (2009, p. 210), idée qu'il déconstruit par ailleurs. Il n'y a ensuite qu'un pas à faire, qu'à transposer ce manque d'empathie d'une catégorie ontologique à l'autre pour que le pire arrive. L'histoire a sinistrement montré la violence et l'horreur qui naissent quand les humains considèrent leurs semblables comme des êtres déficients, des sous-humains, des animaux – et traiter certains humains comme des robots n'augure rien de bon. Face au déferlement de nouvelles technologies, il importe de penser ce que nous faisons. Il faut se demander quelles capacités nous voulons que les enfants développent et quel monde nous voulons leur léguer. Ultiment, et pour le dire avec Arendt :

L'éducation est le point où se décide si nous aimons assez le monde pour en assumer la responsabilité [...] C'est également avec l'éducation que nous décidons si nous aimons assez nos enfants pour ne pas les rejeter de notre monde, ni les abandonner à eux-mêmes, ni leur enlever leur chance d'entreprendre quelque chose de neuf, quelque chose que nous n'avions pas prévu, mais les préparer à l'avance à la tâche de renouveler un monde commun. (1972, p. 251-252)

BIBLIOGRAPHIE

- Adams, Carol J., (2000). *The Sexual Politics of Meat: a Feminist-Vegetarian Critical Theory*, New York, Continuum.
- Adams, Carol J. (1995). « Woman-Battering and Harm to Animals », dans Adams, C. J. et Donovan, J. (dir.), *Animals and Women Feminist Theoretical Explorations*, Durham, Duke University Press, p. 55-84.
- Arendt, Hannah (1983). *Condition de l'homme moderne*, trad. G. Fradier, Paris, éd. Calmann-Lévy.
- Arendt, Hannah (1972). *La crise de la culture. Huit exercices de pensée politique*, trad. P. Lévy, Paris, Gallimard, « Folio/Essais ».
- Beck, Ulrich (2008). *La société du risque. Sur la voie d'une autre modernité*, trad. L. Bernardi, Paris, Flammarion, « Champs essais ».

- Chapouthier, Georges, (2009a). *Kant et le chimpanzé. Essai sur l'être humain, la morale et l'art*, Paris, Belin-Pour la science.
- Chapouthier, Georges (2009b). « Réflexions sur l'altérité et l'animalité », *Revue philosophique de la France et de l'étranger*, Tome 134, n° 2, p. 207-216.
- Chapouthier, Georges (2004). « L'animalité », *Revue Philosophique de la France et de l'Étranger*, Tome 194, n° 3, p. 299-305.
- Derrida, Jacques (2006). *L'animal que donc je suis*, Paris, Galilée, « La philosophie en effet ».
- Garrau, Marie et Le Goff, Alice (2010). *Care justice et dépendance. Introduction aux théories du Care*, Paris, Presses universitaires de France.
- Gigliotti, Carol (2005). « Kenneth Rinaldo and France Cadet. Artificial Life and the Lives of the Non-Human », *Parachute 119*, p. 68-83.
- Grondin, Jean (2007). « Derrida et la question de l'animal », *Cités* (PUF), 30, p. 31-39.
- Hamington, Maurice (2008). « Learning Ethics From Our Relationships with Animals: Moral Imagination », *International Journal of Applied Philosophy*, volume 22, n° 2, p. 177-188.
- Kahn Jr., Peter H.; Gary, Heather E. et Shen, Solace (2013). « Children's Social Relationships With Current and Near-Future Robots », *Child Development Perspectives*, volume 7, n° 1, p. 32-37.
- Laugier, Sandra (dir.) (2012). *Tous vulnérables ? Le care, les animaux et l'environnement*, Paris, Payot, « Petite bibliothèque Payot ».
- Pattnaik, Jyotsna (2004). « On Behalf of Their Animal Friends: Involving Children in Animal Advocacy », *Childhood Education*, 81 :2, p. 95-100.
- Renck Jalongo, Mary (dir.) (2013). *Teaching Compassion: Humane Education in Early Childhood*, Dordrecht, Springer.
- Rice, Suzanne et Rud, A.G. (dir.) (2016). *The Educational Significance of Human and Non-Human Animal Interactions. Blurring the Species Line*, New York, Palgrave Macmillan.
- Tronto, Joan (2012). *Le risque ou le care ?*, trad. F. Brugère, Paris, Presses universitaires de France, « Care studies ».
- Turkle, Sherry (2007). « Authenticity in the Age of Digital Companions », *Interaction Studies*, 8:3, p. 501-517.
- Turkle, Sherry; Taggart, Will; Kidd, Corry D. et Dasté, Olivia (2006). « Relational Artifacts with Children and Elders: the Complexities of Cybercompanionship », *Connection Science*, vol. 18, n° 4, décembre, p. 347-361.

CHAPITRE 18

Ferme laitière et robot de traite : Que reste-t-il de la relation entre l'humain et l'animal ?

Sandra BLOUIN

La robotisation s'incruste dans de plus en plus de milieux et les fermes laitières n'y font pas exception. Beaucoup d'agriculteurs sont attirés par les promesses de nouvelles technologies d'élevage de précision qui permettent l'automatisation des fermes tout en augmentant les performances des animaux. Une panoplie d'équipements s'offre aujourd'hui aux agriculteurs pour réduire la lourdeur de leur travail. Chaque équipement automatisé remplace une tâche de l'éleveur. Nous retrouvons d'abord le principal outil relié à l'automatisation, le robot de traite, qui nécessite une restructuration de l'étable afin de laisser les vaches s'y promener à leur aise. D'autres robots servent aussi à remplacer les tâches reliées à l'alimentation. Il existe par exemple un robot qui distribue le fourrage et un robot qui nourrit les veaux en compilant de l'information sur chaque individu. De plus, pour le nettoyage, un robot peut se charger d'écurer les planchers. À ces machines s'ajoute un robot pour nettoyer les pattes et une brosse à dos qui peut servir à détendre l'animal. Il existe également une application mobile qui permet de contrôler les paramètres de la ferme à distance.

Si ces équipements peuvent faire miroiter un travail plus facile pour les agriculteurs qui se font remplacer par des robots, ils soulèvent

également beaucoup de questions. En effet, lors de l'implantation de ces machines sur une ferme, l'impact sur la relation entre l'humain et ses vaches ne semble pas être d'emblée considéré. Pour nous, il est important de réfléchir au fait que l'introduction de machines réduit le contact direct entre l'éleveur et son troupeau, et c'est ce que nous tenterons de faire dans le présent article. Notre réflexion sera davantage centrée sur le robot de traite, car il remplace un moment privilégié de contact interspécifique quotidien et il accumule de l'information sur les vaches qui est destinée à l'éleveur. Ce robot est donc celui qui entraîne le plus de changements dans la dynamique d'une ferme.

Pour bien comprendre cette problématique, nous décrirons le fonctionnement de fermes automatisées et les motivations qui ont conduit à ce changement. Puis, nous expliquerons les inconvénients liés à la robotisation des fermes. Ensuite, en nous appuyant principalement sur Christian Nicourt et Dominique Lestel, nous pourrions réfléchir sur les relations entre les humains et les animaux. Finalement, nous aborderons l'idée d'inclure la relation à l'humain comme un critère de bien-être animal. Ce cheminement nous permettra de nous questionner sur l'impact relationnel des robots d'élevage.

1. POURQUOI SE PROCURER UN ROBOT DE TRAITE ?

Les robots de traite trouvent leur principal attrait dans le nouveau mode de vie qu'il propose aux agriculteurs. En effet, ils permettent de passer d'une routine demandant beaucoup de travail physique et une régularité rigide à un horaire flexible et un effort moins éreintant. Son ancêtre, la traite mécanique, nécessite de prendre le temps de mettre une trayeuse mécanique sur chaque vache. Il faut aussi pratiquer cette activité deux fois par jour à des heures régulières et elle prend beaucoup de temps dans une journée. En revanche, elle permet de partager un moment avec l'animal qui demande sa bonne collaboration. Généralement, la traite mécanique se déroule plus aisément si l'animal est habitué à la présence de l'éleveur et qu'il entretient une bonne relation avec ce dernier. Cette activité, qui exige du temps, est également un moment privilégié qui offre une proximité avec les vaches et qui permet à l'humain de communiquer avec l'animal et de vérifier qu'il se porte bien.

De son côté, la traite robotisée se déroule tout autrement, car le rôle de l'éleveur est complètement bouleversé, tout comme la nature de

son travail. Lors de l'installation d'un robot de traite, les vaches sont détachées de leurs chaines et sont désormais en stabulation libre dans l'étable. Motivées par la distribution de moulée par le robot, elles sont chargées de se rendre elles-mêmes à la traite tour à tour. L'éleveur se contente dorénavant de superviser le déroulement de l'activité à distance. La vision qu'il a de ses vaches change également, car elle est désormais médiatisée par le robot. En effet, cette machine arrive à identifier chaque vache individuellement et recueille une panoplie de données sur chacune d'entre elles lors de la traite. Elle peut noter leur poids, leur température, la quantité de lait produite, etc. Bref, l'éleveur est donc invité à évaluer la santé et la performance de ses vaches par le biais des profils individuels produits par le robot.

Plusieurs compagnies dans le secteur agricole recourent à des arguments de vente plutôt convaincants pour inciter les éleveurs à se tourner vers la traite robotisée. Parmi elles, nous retrouvons la compagnie Lely qui a été fondée en 1948 aux Pays-Bas (Lely, sans date d). Chef de file dans le domaine de la robotique agricole, elle présenta le premier robot de traite en 1992 (Lely, sans date a). Elle affirme avoir vendu son vingt-millième robot de traite en 2014 (Lely, 2014) et possède aujourd'hui des réseaux de vente dans plus de 40 pays, dont le Canada (Lely, sans date d). Même si de grosses fermes se procurent des robots de traite, le produit est surtout destiné aux fermes de petites et moyennes tailles (moins de 500 têtes) (Schewe et Stuart, 2015, p. 203), ce qui correspond au marché québécois où le nombre moyen de vaches par ferme est d'environ 60, bien que cette moyenne augmente quelque peu d'année en année (Gr. Agéco, 2017).

Lorsqu'ils pensent à l'achat d'un robot de traite, les agriculteurs se laissent également convaincre par l'augmentation promise de la production de lait des vaches qui, bien qu'elle ne soit pas la même pour toute, serait en moyenne de 3 % par vache et pour une période d'un an (Veysset, Wallet, Prugnard, 2001, p. 55). Le robot augmenterait la productivité, car les vaches en liberté peuvent être traitées jusqu'à trois fois par jour plutôt que deux comme c'est le cas avec la traite mécanique. De plus, au sein d'une ferme automatisée, des compagnies comme Lely affirment que tout est conçu pour le bien-être et la stabilité de l'animal dans le but d'augmenter sa productivité; une visite sur leur site Internet permet de le constater rapidement.

Cependant, les producteurs laitiers sont avant tout attirés par l'allègement du travail que proposent les machines (Veysset, Wallet,

Prugnard, 2001, p. 54). Celles-ci permettent aussi aux producteurs de gérer leur temps de manière moins stricte et avec un horaire flexible. La proposition de ce nouveau mode de vie est très attrayante surtout dans le milieu agricole québécois, où les fermes sont essentiellement des entreprises familiales. En effet, il est plus motivant pour un jeune adulte de reprendre la ferme de ses parents en pensant qu'il n'aura pas la vie de dur labeur vécue par ces derniers (Schewe et Stuart, 2015, p. 205). De plus, le robot de traite permet de réduire le personnel au sein de la ferme, ce qui est un argument supplémentaire pour envisager son achat (Veysset, Wallet et Prugnard, 2001, p. 53). En quelque sorte, pour certains éleveurs, il est plus aisé de se fier au robot que de trouver la bonne personne à qui confier ses vaches (Schewe et Stuart, 2015, p. 206).

En fin de compte, les agriculteurs peuvent percevoir dans leur quotidien plusieurs avantages au robot de traite, ce qui les incite à se tourner vers ce produit. Il est cependant nécessaire de s'attarder aux inconvénients qu'ils entraînent pour mieux comprendre les conséquences de cette décision pour les producteurs laitiers. C'est ce que nous ferons dans la prochaine section.

2. LES INCONVÉNIENTS DU ROBOT DE TRAITE

L'implantation de l'automatisation au sein d'une ferme demande une réflexion approfondie pour un agriculteur, car les robots de traite et la stabulation libre nécessitent de repenser et souvent de reconstruire les infrastructures de la ferme. Il doit d'abord considérer le coût que représente un tel changement. En effet, le prix de chaque installation comportant un robot va de 150 000 à 200 000 dollars américains (Jacobs, 2011, p. 6), et c'est sans compter les coûts importants de réaménagement ou de maintenance. Il faut aussi considérer les frais liés à la mise en marche et au temps d'adaptation au robot, car pendant un certain temps, les vaches n'auront pas encore appris à l'utiliser de manière optimale (Beauregard, 2008, p. 3).

De plus, il est compliqué d'évaluer la rentabilité économique de la machine, et ce, en raison de plusieurs facteurs. D'abord, le robot fonctionne plus efficacement avec 60 vaches. Par conséquent, si l'éleveur ne possède par exemple qu'une quarantaine de vaches, l'augmentation de la production ne couvrira peut-être pas le coût de la machine et des rénovations. En outre, si un agriculteur avait une production assez

optimale avant l'installation du robot, il est peu probable qu'il augmente significativement la quantité de lait produite par vache. Pour que son investissement en vaille la peine, l'agriculteur doit répartir efficacement le temps qu'il gagne grâce au robot. De surcroît, des facteurs externes, comme le prix du lait sur les marchés, peuvent aussi influencer la rentabilité. Ainsi, l'augmentation de la production laitière au sein d'une ferme n'est pas toujours suffisante pour rentabiliser le robot. Il est nécessaire par ailleurs que l'agriculteur sache l'utiliser de manière optimale.

Ensuite, en se procurant un robot, le fermier doit être prêt à faire face à un autre inconvénient, soit l'étape de l'adaptation à la nouvelle routine. Elle peut être déstabilisante et demander beaucoup de temps, autant pour lui que pour son troupeau. En fait, un délai d'environ un mois est nécessaire pour que 80 à 90 % des vaches utilisent le système de traite volontairement (Jacobs, 2011, p. 8). Les vaches qui n'auront pu s'adapter au nouveau système seront quant à elles souvent réformées (engraissées pour être abattues) ou vendues.

Pour l'agriculteur, l'adaptation comprend également l'acquisition de nouvelles compétences. Il doit, entre autres, apprendre à faire la maintenance du robot, comprendre le fonctionnement de l'ordinateur et de l'application mobile, et analyser les données fournies par le robot. En contrepartie, la question de l'automatisation de l'élevage nous oblige à nous demander si l'éleveur ne risque pas aussi de perdre certaines compétences ou connaissances. En effet, la machine peut lui fournir des informations dont il ne pouvait prendre connaissance autrement, comme les taux hormonaux (Tiers, 2014), ce qui peut le pousser à se défaire de certaines pratiques qu'il avait auparavant. Pour donner un exemple concret : sans un robot de traite, un éleveur peut remarquer quelques signes qui indiquent le moment où la vache est féconde, comme une vulve rosée, une vache qui se laisse monter, etc. De plus, certains éleveurs peuvent percevoir des signes plus subtils : « les éleveurs qui connaissent bien leurs animaux remarquent qu'une vache vient en chaleur parce qu'elle est plus alerte et a une apparence nerveuse. Elle peut changer son comportement de façon plus évidente avec sa voisine d'étable ou l'opérateur de traite » (Lacerte, 2003, p. 3). Par contre, avec un robot de traite, l'éleveur n'a plus qu'à se fier au taux d'hormones donné par la machine pour connaître le moment où ses vaches sont fécondes. Le robot est aussi plus compétent que l'humain pour détecter les mastites. Comme il ne sera plus nécessaire pour l'éleveur d'observer cet aspect de l'animal, la préservation des

connaissances évoquées n'est pas garantie, ce qui peut engendrer avec le temps une dépendance au robot de traite. Nous pouvons également supposer que la dépendance risque de s'accroître d'une génération à l'autre, car les plus jeunes éleveurs n'auront pas connu l'étable sans la robotisation.

D'autres éléments liés aux contraintes du robot doivent être considérés. Par exemple, une fois qu'un robot de traite conçu pour une soixantaine d'individus aura été installé, les possibilités d'agrandissement du troupeau sont réduites (Beauregard, 2008, p. 12). En outre, comme les robots constituent un équipement particulier, leur réparation nécessite parfois la venue d'un technicien. En plus des honoraires, cette pratique engendre des coûts pour l'agriculteur, puisqu'un bris marque un temps lors duquel le robot n'est pas en fonction et ne peut permettre l'accumulation de litres de lait (Kassler, 2001, p. 238). L'importance d'une présence humaine en permanence sur la ferme est une autre contrainte qui s'ajoute aux précédentes. En effet, le robot peut déclencher des alarmes jour et nuit en cas de problème (Parent, 2016), mais il le fait même pour des cas peu urgents, ce qui constitue un inconvénient pour certains agriculteurs. Ils doivent être en mesure d'aller répondre à ces alarmes à tout moment, car elles ne se résolvent pas à distance. Enfin, à moins qu'il y ait au moins deux personnes au sein d'une même entreprise qui apprennent le fonctionnement du robot, il sera difficile pour un éleveur de se faire remplacer (Veysset, Wallet, Prugnard, 2001, p. 55).

De surcroît, le travail sur une ferme laitière munie d'un robot de traite implique d'effectuer de nouvelles tâches au quotidien. D'abord, en fin de journée, l'éleveur doit emmener lui-même au robot les vaches qui ne s'y sont pas encore rendues. Il doit faire l'entretien des robots tout en assurant leur programmation et leur paramétrage. Le travail comprend beaucoup d'observation des animaux, car l'éleveur doit vérifier ce que les machines ne peuvent pas mesurer. Par exemple, même s'il arrive à détecter des problèmes de santé chez la vache, le robot ne peut voir les blessures ou constater les dynamiques entre les vaches au sein du troupeau. L'éleveur doit donc prendre ce temps d'observation pour vérifier des éléments que la traite lui permettait de voir auparavant. Néanmoins, comme cette vérification ne s'inscrit plus dans un échange et un contact direct avec l'animal, l'agriculteur risque donc de ne pas être en mesure de porter attention aux mêmes détails qu'auparavant. Enfin, il doit aussi prendre le temps de consulter les données produites par le robot de traite, ce qui s'ajoute à sa nouvelle routine.

3. COMPRENDRE LES RELATIONS ENTRE LES HUMAINS ET LES ANIMAUX

En plus des nouvelles tâches qu'ils apportent, les robots de traite imposent un autre modèle de gestion de ferme qui vise à optimiser la production et aider les agriculteurs à prendre les meilleures décisions qui soient pour leur entreprise. Cette nouvelle gestion passe par l'analyse des données que procure le robot. Bien que l'idée d'une meilleure gestion soit attirante pour les producteurs agricoles, cette nouvelle façon de faire entraîne chez nous quelques appréhensions, car elle modifie le regard que les éleveurs portent sur leurs animaux et sur leur travail. Cette vision nouvelle peut affecter le lien qu'a l'éleveur avec ses vaches. Nous tenterons dans cette section de comprendre de quelle façon cela est susceptible de se produire. Puis, nous verrons ce qui constitue fondamentalement les relations interspécifiques.

D'abord, il faut noter que le rôle d'un éleveur change lorsque l'automatisation est intégrée à sa ferme. Le robot permet d'accentuer la position d'entrepreneur d'un agriculteur, mais surtout sa position de gestionnaire. Au lieu d'être concentré sur le soin de l'animal, le travail de l'éleveur est désormais orienté vers des outils de gestion, qui servent à la prise de décisions favorisant la productivité et qui impliquent des tâches s'effectuant automatiquement. Par exemple, l'éleveur peut programmer le robot pour qu'il rejette le lait d'une vache faisant une mastite.

Les données et les indicateurs de performance deviennent centraux dans le travail de l'éleveur, car il y en a une panoplie à gérer quotidiennement. Il peut être ardu pour certaines personnes de savoir quelles informations cibler pour augmenter la productivité. Si elles n'apprennent pas à le faire, elles risquent de se sentir surchargées (Tiers, 2014) et de perdre beaucoup de temps à analyser des données sans savoir quoi en faire. Par exemple, certains éleveurs témoignent qu'ils n'arrivent à utiliser qu'une portion des données fournies (Schewe et Stuart, 2015, p. 205). Nous pouvons donc remarquer que tous les agriculteurs ne se servent pas des données de la même manière. Ils peuvent les évaluer selon les différents objectifs qu'ils choisissent d'améliorer. Sur son site Internet, l'entreprise Lely leur suggère d'ailleurs de cibler certains objectifs précis, comme mieux contrôler la reproduction, augmenter la production laitière, avoir plus de contrôle sur la santé des vaches, etc. (Lely, sans date c). Ces suggestions éclairent les agriculteurs pour savoir quelles données utiliser, ce qui permet de concevoir ce modèle de gestion comme un atout pour la ferme.

Une fois que les objectifs de l'entreprise sont fixés, les éleveurs peuvent utiliser les données en fonction de leurs nouveaux besoins. Il faut noter que l'utilisation des données change le sens qu'un agriculteur va donner à son travail. En effet, il sera dorénavant porté à évaluer son travail à partir d'indicateurs chiffrés qui lui permettront de définir en quoi consiste une bonne performance. Avec autant de données, il est donc peu probable que le travail soit évalué à partir de critères liés au bien-être animal ou au plaisir qu'on peut tirer du travail. De plus, le fait de travailler avec cet objectif de performance comporte des risques, puisque « le but est d'améliorer en permanence » (Nicourt, 2013, p. 131). Ainsi, la quête d'amélioration et de performance ne peut jamais être pleinement satisfaite, puisqu'elle est inachevable.

Comme Nicourt le fait remarquer, cette médiatisation du travail par les chiffres engendre de nouveaux rapports entre les agriculteurs et leurs tâches :

Les chiffres construisent un langage qui marque une distance avec la réalité vécue du travail [...]. Focalisés sur des indicateurs chiffrés, les débats professionnels restreignent en effet le contenu des échanges et transforment leur signification. Le langage des chiffres enferme dans un monde qui distingue d'abord ceux qui sont capables d'apprécier toute leur signification, c'est-à-dire les éleveurs les plus performants. L'usage des chiffres, pour décrire et débattre des situations, constitue ensuite une prise de distance envers la complexité du travail d'élevage ; celle qui porte sur la gestion des multiples aléas qui caractérisent les phénomènes vivants. En les ramenant à des performances, l'homme et l'animal disparaissent. Les chiffres appauvrissent le travail et le transforment en une représentation qui occulte ses épreuves et ses plaisirs, ils érodent ses dimensions sensibles (Nicourt, 2013, p. 130-131, s'appuyant sur Schwartz, 2000)

Nous pouvons remarquer que les données modifient ce sur quoi les agriculteurs se concentrent lorsqu'ils réfléchissent à leur métier. Ils s'évaluent par des chiffres, ce qui les éloigne de la réalité vécue de leur travail. Ainsi, les éleveurs perçoivent leur travail de manière à s'en distancer.

De plus, les indicateurs créent des normes de performance qui permettent aux agriculteurs de comparer leur efficacité, puisqu'ils peuvent mettre en parallèle les mêmes références chiffrées. Ces indicateurs normalisent les échanges en lien avec le travail : « Pour les dirigeants

agricoles, les chiffres sont des outils d'échange socialement situés ; ils fournissent un langage partagé [...]» (Nicourt, 2013, p. 133) Dans un tel modèle, les agriculteurs discutent de leur élevage par le biais de chiffres. C'est par la comparaison aux autres qu'un éleveur peut se classer et ainsi développer son identité professionnelle (p. 133), car des indicateurs performants indiquent la « bonne manière » de travailler pour un agriculteur.

Le robot de traite change aussi la dynamique du travail pour l'éleveur dans la mesure où les données qu'il perçoit orientent ses décisions. Lely veut d'ailleurs montrer aux éleveurs qu'avec son système, ils ne prendront pas de meilleures décisions, mais bien les bonnes décisions : « Vous contrôlez mieux la santé de vos vaches et prenez les bonnes décisions grâce aux capteurs high-tech et aux informations en temps réel. » (Lely, sans date c) Cependant, ces outils influencent le jugement de l'éleveur et c'est en ce sens que Nicourt les qualifie même « d'artefact cognitif » (2013, p. 127). Cette manière de procéder comporte un danger, car elle se substitue aux anciennes méthodes de réflexion et de discussion qui amenaient l'éleveur à faire des choix, et il peut en devenir dépendant. En effet, comme la machine peut capter des informations invisibles à l'œil humain, l'éleveur peut croire qu'il doit absolument baser sa réflexion sur ces données et, ainsi, perdre de l'autonomie décisionnelle.

Il était pertinent de s'attarder aux propos de Nicourt pour montrer que le recours aux données pour gérer la ferme fait en sorte que l'agriculteur concentre son travail sur l'évaluation de données, ce qui médiatise sa relation avec les vaches, car il est alors plus porté à prendre des décisions sur la base des chiffres. L'automatisation d'une ferme entraîne une distance physique entre l'éleveur et ses vaches. En effet, il se trouve davantage devant son ordinateur, alors qu'il consacrait auparavant beaucoup de temps en contact avec les animaux. Pour l'agriculteur, la référence à des données chiffrées pour comprendre son travail et orienter ses décisions ne date pas d'hier. Elle est cependant fortement accentuée par un outil comme le robot de traite qui procure de nouveaux éléments chiffrables tous les jours. Donc, avec le robot, cette méthode de travail est encore plus présente.

L'introduction du robot a donc un impact significatif sur le travail de l'agriculteur. Pour comprendre si le robot de traite affecte même les fondements de la relation entre l'humain et les animaux, il est important de s'arrêter à ce qui constitue la base de ces relations. Nous croyons qu'il est opportun de se questionner sur les rapports entre

les humains et les animaux sous l'angle du concept de communauté hybride décrit par Dominique Lestel (2004). En effet, partant du constat que toutes les sociétés humaines ont côtoyé des animaux (p. 15), l'auteur explique de manière détaillée les éléments qui composent les relations entre les humains et les animaux. Nous pouvons d'abord nous attarder à la définition qu'il attribue à la communauté hybride : « une association d'hommes et d'animaux, dans une culture donnée, qui constitue un espace de vie pour les uns et pour les autres, dans lequel sont partagés des intérêts, des affects et du sens » (p. 19). Cette définition contient plusieurs éléments, dont l'idée qu'au sein d'une même communauté interspécifique, les différentes espèces partagent un espace commun dans lequel elles s'influencent et se transforment les unes les autres. Pour favoriser le bon vivre-ensemble, les espèces n'ont donc d'autres choix que de se considérer.

Selon Lestel (2004), les communautés hybrides constituent des communautés d'intérêts. En effet, elles correspondent aux intérêts de chaque espèce qui y participent, car elles sont un échange de services. Comme plusieurs animaux domestiques sont des proies, il est dans leur intérêt que les humains leur offrent de la nourriture et une protection pour qu'ils ne soient plus à la merci de leurs prédateurs. En échange, les animaux peuvent fournir à l'humain un service, comme le fait le chien guide, ou un produit, comme de la laine ou du lait. Pour expliquer ces propos, Lestel aborde le cas d'un berger et de son troupeau de moutons. Ceux-ci sont ainsi protégés des loups et le berger peut vendre leur laine (2004, p. 22).

Le service attendu dans la relation peut être simplement affectif, comme dans le cas des petits animaux de compagnie. Il faut mentionner que cet échange affectif est présent dans toutes les relations entre humains et animaux qui impliquent une véritable coopération (2004, p. 22). La part d'affectivité est essentielle à la communauté hybride. En effet, si la relation interspécifique n'est qu'utilitaire, fonctionnelle ou d'ordre exclusivement économique, nous ne sommes pas en présence d'une communauté hybride.

En outre, la présence d'affect est impérative à la construction d'une relation interspécifique, car l'observation et l'interprétation des comportements de l'animal permettent la communication entre les espèces. Sans que la communication passe par le dialogue comme chez les humains, les relations interspécifiques sont composées d'interactions qui engendrent une histoire commune entre les espèces dans un même

espace (Lestel, 2004, p. 21). Ces interactions répétées servent également à construire une complicité et une familiarité entre les membres de ces relations. Chaque relation interspécifique est unique et s'appuie sur les habitudes de comportements et les actions de chacun (p. 28). Enfin, les deux partis d'une relation en viennent à agir avec l'autre en fonction des éléments mutuellement connus et présents à répétition dans l'échange.

Le dernier élément de la définition de Lestel est celui du sens qui est construit par les membres d'une communauté. Pour développer ce sens commun, il est nécessaire que la relation se déroule dans une continuité temporelle et une proximité spatiale (Lestel, 2004, p. 21). C'est de cette manière que des habitudes et des références communes apparaissent en lien avec les choses vécues ensemble. Néanmoins, il faut mentionner que, malgré une histoire commune, les humains et les animaux accordent des représentations différentes aux choses, ce qui fait en sorte que les humains ne peuvent développer de communautés hybrides avec des espèces animales qui sont trop éloignées d'eux : « Personne n'a jamais établi de relations privilégiées [...] avec une baleine bleue, dont la taille et le mode de vie excluent l'*intimité* avec l'humain » (p. 21). Il faut donc être en mesure d'échanger avec une espèce pour avoir une relation avec elle.

De surcroît, les communautés hybrides comprennent d'autres caractéristiques, dont l'unicité. L'individualité des êtres qui la composent rend effectivement chaque relation particulière. L'unicité de chaque individu prend plus d'importance dans la relation que l'espèce à laquelle il appartient (Lestel, 2004, p. 20). Plus précisément, une relation interspécifique n'est pas seulement composée d'un chien et d'un humain. Elle comprend deux individus uniques qui se considèrent l'un l'autre parce qu'ils connaissent leurs particularités. L'individu animal de la relation porte un nom et possède sa propre personnalité qui le distingue des autres individus de son espèce. Enfin, le même animal ne se comporte pas de la même manière avec tous les humains qu'il côtoie, ce qui témoigne bien de l'unicité de chaque relation.

Lestel suggère aussi de concevoir les communautés hybrides comme une sorte de contrat de confiance qui s'appuie sur la familiarité développée entre les individus d'une communauté (Lestel, 2004, p. 26). Même s'ils ne sont pas en mesure de comprendre le concept de contrat, les animaux participent et négocient dans la relation. Nous pouvons ajouter qu'ils s'investissent affectivement dans la relation, ce qui est

une façon d'adhérer au contrat (p. 26). Cette idée de contrat est également abordée par Catherine et Raphaël Larrère qui voient dans la cohabitation avec l'animal domestique « une sorte de négociation, d'où se dégage, par apprentissage mutuel – l'attitude de l'un s'adaptant aux attentes de l'autre – une forme d'arrangement, comme s'il y avait eu une entente, un accord » (Catherine et Raphaël Larrère, 1997, p. 15). Des situations peuvent montrer que les animaux ont la possibilité de négocier lors d'un échange pour faire changer une pratique, par exemple une vache qui défèque pendant la traite (p. 15).

Un dernier aspect digne de mention est que les communautés hybrides sont fondamentalement asymétriques et hiérarchiques, et ce, sur plusieurs points. Tout d'abord, l'humain a un pouvoir de décision sur son animal qui n'est en rien réciproque. Ce dernier ne possède qu'un faible pouvoir d'initiative dans son rapport à l'humain (Lestel, 2004, p. 29). Par exemple, si l'humain déménage et amène ses animaux avec lui, il décide de les changer d'environnement sans les consulter. Le contraire est évidemment impossible. Aussi, les humains ont des attentes plus grandes envers les animaux que l'inverse (p. 29).

Nous venons de survoler la question des rapports coopératifs entre les humains et les animaux, une coopération dont les humains n'ont pas toujours conscience lorsqu'ils interagissent avec un animal. L'aspect relationnel est important, parce qu'il nous permet de nous interroger sur la façon dont nous voulons agir avec les animaux. Nous avons vu qu'une bonne relation se fonde dans l'échange et dans la routine au quotidien et, comme les robots de traite changent cette routine, nous croyons qu'ils ont un impact sur la manière dont les espèces se perçoivent. Les prochains paragraphes nous permettront d'explorer cette idée.

Les élevages de plus en plus mécanisés peuvent entraîner la déshabitation des animaux à l'humain, phénomène qui est réciproque (Sigaut, 1988, p. 63). Nous pouvons facilement imaginer ce phénomène s'amplifier avec la robotisation et l'automatisation, car le contact entre les deux espèces est alors d'autant plus réduit et qu'il est médiatisé par la machine. Nous pouvons donc nous demander si cette perte de familiarité entre les espèces peut affecter leur relation, non seulement sur le plan affectif, mais aussi sur le plan du bon déroulement du travail. Les robots actuellement commercialisés sur les fermes laitières remplacent des activités qui peuvent s'inscrire dans un échange agréable entre l'humain et ses animaux, comme la traite ou l'alimentation. Il

faut rappeler que Lely vend aussi des équipements qui contribuent au confort de l'animal, comme un bain de pied ou une brosse pour la peau automatisée. Ces activités peuvent reconforter l'animal, mais avec des machines automatisées, les vaches peuvent se faire reconforter sans contact avec l'humain. Nous croyons que cette situation ne favorise pas pour l'animal une association positive à l'humain, puisque même si, avec l'automatisation, rien ne l'empêche de cajoler ses vaches, l'éleveur n'est plus nécessairement impliqué dans ce type d'activités reconfortantes.

De surcroît, les tâches de l'humain en interaction avec l'animal qui demeurent à ce jour non robotisées peuvent être associées à un désagrément pour la vache. Il s'agit par exemple de la vaccination, de l'insémination artificielle ou de la mise bas. Même si elle n'a pas été planifiée ainsi, cette dynamique de relayer les tâches avec un potentiel de développement relationnel positif aux robots et de préserver les tâches plutôt désagréables à l'humain pourrait faire en sorte que les vaches associent les agriculteurs aux situations stressantes et qu'elles en développent une certaine peur (Tiers, 2014). Plus concrètement, une relation interspécifique dans laquelle les individus animaux ne sont pas habitués à l'humain rend le travail plus difficile pour l'humain et engendre plus de stress chez l'animal. La simple présence d'un humain non familier peut être une source de peur chez l'animal (Bouissou, 1992, p. 303). En revanche, une relation positive rend l'animal plus collaboratif lors des tâches (p. 306). Nous pouvons citer un exemple de vaches en salle de traite : « La comparaison de troupeaux laitiers à forte ou au contraire à faible production laitière montre que dans ces derniers la relation Homme-Animal est moins bonne : les animaux entrent moins facilement en salle de traite et y défèquent plus, ont une distance de fuite supérieure, approchent moins facilement volontairement un humain, le trayeur leur parle moins et les touche moins » (Seabrook, 1984, cité par Bouissou, 1992, p. 308). Cette comparaison montre bien l'importance de la relation entre les espèces : quand elle est bonne, les vaches collaborent mieux, ce qui rend le travail plus agréable et plus facile.

4. L'IMPORTANCE DE LA RELATION POUR L'ANIMAL

Dans cette partie nous tenterons de montrer qu'il est important de considérer l'aspect relationnel lors de l'évaluation des robots de traite. En effet, nous estimons, tout comme l'auteure que nous présenterons

dans cette partie, que la relation entre l'humain et l'animal participe au bien-être de ce dernier et qu'elle devrait en être un critère. C'est pour nous un fait : le robot ne peut remplacer l'humain pour cerner les besoins ou la personnalité d'un animal. Toutefois, avant de revenir à la question de la relation avec l'animal, nous voulons mieux comprendre le concept de bien-être animal.

Il est pratiquement impossible de définir exhaustivement le bien-être animal, parce que la définition qu'on en donne est changeante et dépend des critères qu'on cherche à y intégrer. Il faut d'abord relever que ce terme ne peut pas se réduire à l'absence de souffrance ou de stress. Nous ne pouvons pas non plus nous limiter à évaluer l'intégrité physique de l'animal, car ce critère ne permet pas de prendre en compte sa subjectivité. Par ailleurs, en agriculture, on a souvent la perception que les critères du bien-être animal sont assurément respectés si l'animal est performant : « Pour l'éleveur, les indicateurs de performance zootechnique sont le garant de la bonne santé du troupeau. [...] L'idée implicite est que l'animal ne peut produire que s'il est en bonne santé, au sens physique et mental du terme. » (Dantzer, 2001, p. 94) Cependant, les critères de performance ont été établis au départ pour évaluer la santé économique de la ferme, et non celle des animaux (p. 94). Il n'est donc pas facile ni pertinent de les utiliser pour évaluer le bien-être de l'animal. Finalement, nous voyons que le concept de bien-être animal est complexe et que sa définition dépend de plusieurs facteurs. Nous ne croyons pas être en mesure de le définir précisément dans cette étude (ce qui n'en est d'ailleurs pas le but), mais nous suggérerons dans cette section des aspects qui devraient faire partie de la définition du bien-être animal.

La perception du bien-être animal varie selon les intérêts en jeu. Par exemple, la compagnie Lely ne semble pas posséder de définition propre du bien-être animal, même si elle utilise ce terme abondamment sur son site Internet. Elle semble l'associer au confort, mais rapporte aussi que le bien-être favorise la productivité des vaches (Lely, sans date b). En outre, des compagnies comme Lely défendent l'idée que les vaches ont un meilleur bien-être avec un robot de traite, car il leur permet de se déplacer de manière autonome dans l'étable, ce qui favorise la liberté individuelle des vaches et respectent leur subjectivité (Holloway, 2007, p. 1049). La dynamique engendrée par le robot de traite serait donc intéressante pour la vache, car elle lui permettrait de suivre une routine plus « naturelle » qui dépend de ses désirs, vu que c'est elle qui choisit quand manger, se reposer ou se faire

traire (p. 1049). De plus, le robot de traite individualise les vaches en créant des profils qui compilent les données recueillies sur chacune d'elles, ce qui permet de soigner chaque vache plus adéquatement. Le robot peut par exemple détecter des problèmes de santé comme des mastites. La machine adapte également l'alimentation de chaque animal à son âge et à son niveau de production.

Pour répondre à cet argument de vente de Lely, des auteurs tels que Driessen et Heutnick qui se sont penchés sur le cas des robots de traite nous permettent de remarquer que la définition du bien-être est multiple et évolutive selon les points de vue et les contextes (2015, p. 7). Ils notent par ailleurs que l'idée d'inclure la liberté individuelle comme critère de bien-être est récente et qu'elle est associée aux comportements naturels des animaux d'élevage (2015, p. 10). Cette idée est confortée par les compagnies qui souhaitent vendre des robots de traite. Cependant, les auteurs remettent en question le fait que les vaches aiment se faire traire quand elles le veulent, puisqu'au départ (avant que de la moulée ne soit distribuée lors de la traite), elles ne s'y rendaient pas. Ils mentionnent au contraire que le robot exerce un pouvoir sur les vaches, justement parce qu'elles ne peuvent obtenir de moulée sans se faire traire. Cette liberté suppose donc pour les vaches une discipline et la responsabilité d'aller elles-mêmes à la traite (2015, p. 11).

Maintenant que nous avons exploré le bien-être animal et compris son équivocité, nous pouvons ouvrir les horizons de ce dernier pour y inclure l'aspect relationnel avec l'humain, car il aurait un rôle à y jouer. La sociologue Jocelyne Porcher (2002) est d'avis que la qualité de la relation entre l'humain et l'animal favorise son bien-être et sa qualité de vie. L'automatisation des fermes permettrait d'individualiser chaque vache parce que le robot lui crée un profil personnel et parce que la vache répond elle-même à ses besoins, mais pour Porcher c'est le lien et l'échange qui permettent véritablement d'individualiser les animaux, car l'éleveur connaît la personnalité de chaque animal et reconnaît les besoins particuliers auxquels il peut répondre.

Porcher cible deux composantes nécessaires de la relation humain-animal, ce qui recoupe d'ailleurs le concept de communauté hybride de Lestel. Il s'agit d'abord de la communication qui sous-entend la pertinence d'être attentif aux messages envoyés par les animaux. Puisqu'elle ne passe pas par le langage, la communication avec l'animal s'effectue à travers les gestes, les attitudes, les regards, etc.

(Porcher, 2002, p. 198). Elle permet d'accéder aux désirs de l'autre et demande une connaissance de cet autre. Un éleveur qui connaît ses animaux pourra cibler plus rapidement une situation anormale en remarquant qu'une vache adopte momentanément des comportements qui diffèrent de ses habitudes.

Le second élément de la relation interspécifique qui participe selon Porcher au bien-être est l'affectivité. Il y a des réticences à l'inclure comme critère servant à définir le bien-être animal, car, en plus d'être une source potentielle d'anthropomorphisme, l'affection est habituellement considérée comme un sentiment empêchant de traiter les animaux objectivement et sur une base scientifique (Porcher, 2002, p. 65-66). Toutefois, selon Porcher, aimer l'animal signifie qu'on se soucie de son bien-être et cela devrait être une compétence de base chez un éleveur (p. 59). En tout état de cause, même lorsqu'on tente d'en faire abstraction, une part d'affection demeure chez l'humain en contact avec l'animal. Le métier d'éleveur implique un rapport sincère à l'autre qui va au-delà des tâches quotidiennes et de gestion (p. 68), et ce, même en exerçant son travail à travers une logique productiviste. Porcher mentionne que les deux caractéristiques du bien-être sont liées, puisque l'affection envers un individu suscite une meilleure écoute, ce qui favorise la communication (p. 91). Finalement, nous pouvons noter que les robots de traite ne sont pas aptes à communiquer ou à aimer les vaches qui interagissent avec eux.

CONCLUSION

Nous avons d'abord expliqué pourquoi les agriculteurs se munissent de robots de traite. L'argument principal est qu'il allège le travail et les horaires. Il est donc tout à fait compréhensible qu'ils soient attirés par ce produit. Cependant, il est important de réfléchir également à l'impact qu'aura le robot dans la dynamique d'une ferme et, plus particulièrement, à la façon dont il influencera la connaissance qu'a l'éleveur de ses animaux.

Certains éleveurs sont très à l'aise d'avoir une ferme automatisée, car ils adhèrent à l'idée que le robot donne un mode de vie plus « naturel » aux vaches. Ils sont convaincus que, puisque les vaches décident elles-mêmes de leur horaire de la journée, leurs comportements sont ainsi plus naturels. Cependant, il arrive qu'ils croient également qu'ils doivent s'effacer le plus possible du quotidien des vaches pour ne pas

déranger leur routine plus « naturelle » (Holloway, 2014, p. 191). Ce genre d'agriculteur favorise donc l'éloignement par rapport à ses animaux, au profit du robot. De plus, le fait de voir leurs vaches à travers des données peut faire perdre certaines compétences aux agriculteurs, comme évaluer la santé des vaches.

En ce qui concerne la vache, celle-ci perd son pouvoir de négociation lorsqu'elle est en interaction avec les robots. En effet, si on doute de la possibilité qu'elle adhère à un contrat avec l'humain, il est évident que cette éventualité ne se présente pas avec un robot. Ainsi, la vache ne peut pas construire avec le robot un partage de sens qui permettrait au robot de mieux la connaître et de comprendre plus adéquatement ses besoins.

Pour terminer sur une note plus positive, nous croyons que la dégradation de la relation entre l'humain et l'animal lors de l'acquisition d'un robot de traite n'est pas une fatalité. Notre objectif en exposant comment se construisent les relations interspécifiques et la dynamique qu'entraîne le robot est de prendre conscience de ces phénomènes pour mieux y réfléchir. Nous souhaitons que le développement de bonnes relations entre les vaches et les humains se perpétuent. Pour y arriver, l'éleveur peut se rappeler que la relation à l'animal n'a pas à être qu'utilitaire, et favoriser l'importance de donner de l'affection aux vaches. Malgré la robotisation dans les fermes, il demeurera donc toujours important de préserver une bonne relation avec l'animal pour qu'il se sente en confiance et collabore bien.

BIBLIOGRAPHIE

- Beauregard, Guy (30 novembre 2008). « Les robots de traite : la réalité? », Conférence pour Le Centre de référence en agriculture et agroalimentaire du Québec, Drummondville, Canada.
- Bouissou, Marie-France (1992). « La relation homme-animal, conséquences et possibilités d'amélioration », INRA Productions animales, vol. 5(5), p. 303-318.
- Dantzer, Robert (2001). « Comment les recherches sur la biologie du bien-être animal se sont-elles construites? », dans Florence Burgat, et Robert Dantzer (dir.), *Les animaux d'élevage ont-ils droit au bien-être?*, Paris, Inra, p. 85-104.
- Driessen, Clement, F. M. Heutinck, Léonie (2015). « Cows desiring to be milked? Milking robots and the co-evolution of ethics and technology on Dutch dairy farms », *Agriculture and Human Values*, vol. 32, n° 1, p. 3-20.
- Groupe Agéco (2017). <http://www.groupeageco.ca/fsl/> « Nombre moyen de vaches laitières par ferme et par province au 1er juillet, Canada ». Consulté en ligne.

- Holloway, Lewis (2007). «Subjecting cows to robots: farming technologies and the making of animal subjects» *Environment and Planning D: Society and Space*, vol. 25, p. 1041-1060.
- Holloway, Lewis, Bear Christopher, Wilkinson Katy (2014). «Robotic milking technologies and renegotiating situated ethical relationships on UK dairy farms» *Agriculture and Human values*, vol. 31, p. 185-199.
- Jacobs, Jacquelyn Ann (2011). «Dairy cow adaptation to an interaction with an automatic milking system», Maitrise en science animale soumise à la Michigan state University, 103 pages.
- Kassler, Michael (2001). «Agricultural automation in the new millenium», *Computers and Electronics in Agriculture*, vol. 30, p. 237-240.
- Lacerte, Guy (30 octobre 2003). «La détection des chaleurs et le moment de l'insémination» Conférence pour Symposium sur les bovins laitiers, Centre de référence en agriculture et agroalimentaire du Québec, Saint-Hyacinthe, Canada.
- Larrère, Catherine, Larrère, Raphaël (1997). «Le contrat domestique», *Courrier de l'environnement de l'Inra*, n° 30, p. 5-17.
- Lely (sans date a). «25 ans de traite automatisée». Consulté en ligne.
- Lely (sans date b). «Lely Luna. Pour le confort optimal des vaches». Consulté en ligne.
- Lely (sans date c). «Lely Astronaut A5. Une nouvelle étape dans la traite robotisée». Consulté en ligne.
- Lely (sans date d). «Qui nous sommes et ce que nous faisons». Consulté en ligne.
- Lely (27 octobre 2014). «Lely installe le 20 000^e robot de traite Lely Astronaut cette année». Consulté en ligne.
- Lestel, Dominique (2004). *L'animal singulier*, Paris, Seuil.
- Nicourt, Christian (2013). *Être agriculteur aujourd'hui, l'individualisation du travail des agriculteurs*, Paris, Éditions Quae.
- Parent, Marie-Josée (16 mars 2016). «Bilan de l'utilisation du robot de traite», *Le bulletin des agriculteurs*. Consulté en ligne.
- Porcher, Jocelyne (2002). *Éleveurs et animaux, réinventer le lien*, Paris, Presses universitaires de France.
- Schewe Rebecca L., Stuart Diana (2015). «Diversity in agricultural technology adoption: How are automatic milking systems used and to what end?», *Agriculture and Human Values*, vol. 32, n° 2, p. 199-213.
- Sigaut, François (1988). «Critique de la notion de domestication», *L'Homme*, vol. 28, n° 108, p. 59-71.
- Tiers, Nathalie (25 avril 2014). «Robot : 4 questions à se poser avant d'investir», Mon-Cultivar élevage. Consulté en ligne.
- Veysset, P., Wallet, P., Prugnard, E. (2001). «Le robot de traite: pour qui? pourquoi? Caractérisation des exploitations équipées, simulations économiques et éléments de réflexion avant investissement», *INRA Productions animales*, vol. 14, p. 51-61.

CHAPITRE 19

Face à la culture : animaux, robots et êtres humains

Georges CHAPOUTHIER

Une culture, c'est un ensemble de traits de comportement qui se transmettent entre les individus, par imitation ou par enseignement, sans passer par les mécanismes génétiques. Longtemps l'être humain s'est cru le seul être terrestre doué de culture. Cette conviction s'est trouvée battue en brèche dans deux domaines. En éthologie, le récent progrès des connaissances sur le comportement des animaux a montré, chez nombre d'entre eux, l'existence de traits culturels (Lestel, 2001, Chapouthier, 2009). D'autre part, en robotique, des traits culturels ont pu être démontrés chez des êtres construits par les êtres humains, les robots (Chapouthier, Kaplan, 2011). Après avoir développé la question des traits culturels dans ces deux domaines, nous essaierons de définir comment se place, face à eux, la culture humaine.

1. L'ANIMAL, ÊTRE DE CULTURE

Comme nous venons de le rappeler, de nombreux travaux récents d'éthologie ont montré l'existence de traits culturels chez les animaux. Il s'agit surtout d'animaux fortement « céphalisés », comme les vertébrés, les mollusques céphalopodes ou certains insectes. En outre, l'aptitude à vivre en sociétés (ou en familles, qui constituent de petites

sociétés) favorise les traits culturels, dans la mesure où les animaux sociaux ont plus de facilité à échanger des pratiques comportementales. C'est vrai pour les vertébrés, mais aussi pour les insectes sociaux.

Les traits culturels peuvent être des actions relativement simples, comme l'utilisation d'outils permettant de prolonger l'action des organes du corps. Telles sont, par exemple les techniques de fracture de noix par certaines populations de chimpanzés, qui se transmettent, de parents à enfants, à l'intérieur de la population. Mais beaucoup de traits culturels ont des rapports complexes avec des attitudes sociales et avec l'intelligence, qu'ils contribuent à moduler. Tels sont le rire, l'humour, le langage, la morale ou l'esthétique.

Le rire n'est pas, à proprement parler, un trait culturel, puisqu'il repose sur des bases génétiques, mais il est généralement associé à des phénomènes sociaux, comme le jeu, qui contribuent à la mise en place des schémas culturels du sujet. Il en va probablement de même de l'humour, qui est une sorte de rire différé ou médiateur. Le rire, avec des manifestations sonores et comportementales identiques à celles que l'on connaît chez les sujets humains, existe chez la plupart des primates. On peut aussi en trouver des aspects similaires chez les chiens. Lorsqu'un chien est heureux, il peut le manifester par une sorte de halètement qui ressemble beaucoup au rire humain. Plus original, le groupe de Panksepp a démontré l'existence du rire chez les rats (Panksepp, Burgdorf, 2003). Mais les vocalisations se situent alors dans la gamme des ultrasons et ne sont pas directement perceptibles par l'oreille humaine. On peut enregistrer de tels « rires » aussi bien quand on chatouille des rats que lorsque des ratons jouent entre eux. Le groupe de Panksepp a même pu montrer que les ratons qui rient le plus sont préférés comme camarades de jeu.

L'humour, quant à lui, a été peu étudié, même si beaucoup de propriétaires de chats et de chiens affirment que leur animal de compagnie fait preuve d'un certain humour. Je voudrais citer ici une aventure personnelle. Durant l'année que j'ai passée avec des chimpanzés, j'étais responsable d'un groupe de quatre individus adolescents. J'avais notamment un rapport particulièrement amical avec le chimpanzé François, mais quand, une fois par jour, nous mettions nos animaux en récréation pour observer leur comportement social, il se précipitait vers moi avec une posture de menace que je ne comprenais pas du tout. Jusqu'au jour où, par inadvertance, j'ai eu un mouvement de retrait,

comme si j'avais peur. François est parti en rigolant à gorge déployée : sa menace était une blague !

Comme on l'a vu, le rire, et sans doute l'humour, sont liés au jeu. Le jeu est un comportement très commun chez les mammifères et les oiseaux qui, lorsqu'ils sont jeunes, forgent ainsi leur futur comportement d'adultes. Ils miment des combats, des actes sexuels ou des comportements parentaux de la même manière que les enfants humains jouent à être adultes. Chez la plupart des animaux, le jeu cesse chez l'adulte dont les comportements sont alors assez rigides. Mais certains animaux, qui restent « juvéniles » (on dit en science « néoténiques » (Chapouthier, Policar, 2015)), peuvent jouer toute leur vie. C'est le cas de l'homme et de deux de ses animaux de compagnie les plus communs, le chat et le chien¹. Cette propension au jeu donne sans doute une place prépondérante au rire, voire à l'humour, dans l'espèce humaine.

2. LANGAGE, MORALE, ESTHÉTIQUE CHEZ LES ANIMAUX

Le langage, la morale ou l'esthétique entrent plus directement dans les traits culturels, dans la mesure où leur transmission fait intervenir des échanges non génétiques importants (Chapouthier, 2009). En éthologie, on distingue le langage de la communication simple. La communication simple, qui est extrêmement répandue dans le règne animal, fait référence à des éléments présents dans l'environnement quand l'animal communique. Elle pourrait se traduire par : « ceci est mon territoire », « attention, prédateur en vue ! », « j'ai faim », « je vous copuler », etc. Le langage, en revanche, fait référence à des éléments qui ne sont plus présents dans l'environnement immédiat quand l'animal s'exprime. Par exemple, quand une abeille trouve une source de nourriture, elle revient à la ruche et, par une sorte de danse, transmet à ses congénères la distance et la direction par rapport au soleil de la source qu'elle a trouvée. Mais la source de nourriture n'est pas dans la ruche. Il s'agit bien, non d'une communication simple, mais d'un langage, ou plutôt d'un « protolangage », puisque l'animal n'utilise que quelques « mots ».

1. Indépendamment des aptitudes au jeu, on peut trouver aussi des caractères anatomiques. Les oreilles pendantes de certains chiens et le ronronnement du chat sont des traits de carnivores « bébés ».

Dans l'état actuel des connaissances, on doit considérer que les langages, tels qu'ils viennent d'être définis, sont rares dans le règne animal. Il y a, bien sûr, les langages humains, dans leur double articulation et leur extrême complexité. Intéressants sont aussi les langages que l'être humain a pu enseigner à des anthropoïdes, chimpanzés ou gorilles. Comme ces animaux proches de l'homme ne possèdent pas les structures vocales adaptées, on a utilisé le langage par geste (langage des sourds-muets) ou des langages formés par l'affichage de signes graphiques arbitraires comme un carré rouge pour une banane, un triangle noir pour une pomme, etc. Les anthropoïdes peuvent ainsi composer des phrases simples comme « J'entends (la) voiture (du) lait(ier) » ou « Donne (la) pomme couleur orange », et maîtriser l'usage de quelques centaines de mots, plus quelques règles simples de grammaire. Intéressante est aussi la maîtrise du langage par les chiens. Les chiens ne parlent pas, mais ils peuvent comprendre certains éléments du langage humain. Certains chiens peuvent apprendre les noms de plus d'une centaine d'objets différents. Ils peuvent aussi apprendre la différence entre le nom d'un objet (un os, un ballon ...) et l'action à effectuer sur cet objet (aller le chercher, rester en signallement près de l'objet sans le toucher ...), bref la distinction entre un nom et un mot d'action (c'est-à-dire un verbe).

La morale, ou, tout au moins, des règles « proto-morales » sont nécessaires chez tous les animaux qui vivent en groupe ou en famille, afin que les individus ne puissent pas faire n'importe quelle action avec leurs congénères et afin de maintenir une certaine harmonie du groupe. Ce sujet est en pleine évolution, mais, d'ores et déjà, des règles (proto)morales nombreuses ont été mises en évidence par Frans De Waal chez les chimpanzés (De Waal, 1989). Sympathie, attachement, intérêt pour les jeunes, aide aux handicapés, punitions, négociations, coopérations, réconciliations..., autant d'aptitudes que De Waal a pu décrire dans les troupes de chimpanzés en semi-liberté et que nous interpréterions comme des comportements moraux, même si les chimpanzés ne formulent pas des morales discursives, même si leurs comportements moraux restent du domaine de l'action pratique.

Beaucoup d'animaux sont capables d'effectuer des choix esthétiques (Michel Kreutzer, Verena Aebischer, 2015), de préférer certaines couleurs, certaines formes, certains motifs de chant (Watanabe, 2013). Pour les couleurs, ce sont, en général, les couleurs du ou de la partenaire sexuel(le), mais ensuite l'animal peut aimer les mêmes couleurs de façon générale et abstraite. On sait d'ailleurs que, dans

l'espèce humaine aussi, les goûts esthétiques peuvent avoir une origine sexuelle: qu'on en juge par le nombre de femmes nues qui peuplent nos peintures ou nos sculptures.

3. L'INTELLIGENCE AU-DELÀ DE LA CULTURE

Nous avons déjà vu, à propos du rire et de l'humour, que des éléments non directement culturels peuvent être très liés à la culture. Le langage, la morale et l'esthétique sont, eux-aussi, très fortement liés à des aspects de l'intelligence sans lesquels ils ne pourraient exister sous leur forme complexe. Ce sont notamment les aptitudes à la mémoire et à la conscience.

Plutôt que de «la» mémoire, il vaut mieux parler «des» mémoires, qui sont des aptitudes assez différentes les unes des autres, apparues au fur et à mesure de l'évolution du cerveau de nos ancêtres animaux (Chapouthier, 2006). Il est clair que tout trait culturel un peu complexe suppose l'existence de mémoires sans lesquelles il ne peut se conserver chez un individu. Ce n'est pas un hasard si les traits culturels les plus marquants se rencontrent chez les vertébrés, animaux fortement céphalisés et dotés de mémoires cognitives très complexes. Liée aux mémoires se situe la conscience. Tout trait culturel n'est pas nécessairement conscient et on peut imaginer que la production des champignons par les fourmis, ou diverses autres formes d'utilisations d'outils par les insectes sociaux, reposent largement sur des processus génétiques inconscients, mais, ici encore, ce n'est pas un hasard si l'essentiel des phénomènes culturels ont été rapportés chez les vertébrés qui, du fait de l'importance de leur cortex cérébral, possèdent des aptitudes de conscience.

On distingue classiquement deux types de consciences (Chapouthier, 2009). La «conscience d'accès» est la «conscience de l'environnement» où l'on vit et tous les animaux fortement céphalisés, vertébrés et pieuvres notamment, possèdent une telle conscience. La «conscience réflexive ou phénoménale», c'est la «conscience d'être conscient» et, dans l'état actuel des connaissances, seuls quelques mammifères et oiseaux, comme le chimpanzé, l'éléphant ou la pie, ont pu en faire la démonstration. Les aptitudes culturelles développées sont liées à une intelligence capable de conserver ses souvenirs et consciente de ses possibilités.

Mémoires et conscience sont des aptitudes mentales, apparues progressivement dans l'évolution, par la complexification du cerveau des animaux, et particulièrement lors de la céphalisation des vertébrés. Vertébré lui-même, l'homme est finalement un être de culture parce qu'il est issu de l'animalité, porteuse déjà de potentialités culturelles. Nous verrons plus loin les importantes conséquences que cela peut avoir sur la spécificité de l'espèce humaine. Toutes ces considérations amènent à parler d'une intelligence animale (Vauclair, 1995), même si, du fait des différences comportementales entre les espèces, cette notion n'est pas facile à définir de manière universelle. Elle se manifeste par l'aptitude à une forte adaptation intellectuelle à des processus complexes ou abstraits, comme le classement d'objets ou l'acquisition de règles abstraites, des processus particulièrement développés dans certains groupes de mammifères (anthropoïdes, éléphants, dauphins, carnivores...) ou d'oiseaux (corvidés, geais, perroquets...).

4. LES ROBOTS ET LA CULTURE

Construits par l'homme, il ne fait cependant pas de doute que les robots sont complexes (Chapouthier, Kaplan, 2011 ; Oudeyer, 2015). Il ne fait pas non plus de doute qu'ils ont des capacités de mémoire, certes fondées sur d'autres bases que les mémoires physiologiques animales, mais susceptibles de traiter de grandes quantités d'information. Cette complexité et ces aptitudes de mémoire font que les robots possèdent des formes d'intelligence, qui se manifestent notamment par la résolution de problèmes complexes, logiques, mathématiques ou autres, par l'aptitude à jouer, souvent mieux que les humains, à des jeux complexes comme le jeu d'échecs ou le jeu de go, ou encore, même si elle n'en est qu'à ses premiers balbutiements, par la capacité à utiliser des (proto) langages. Pour des psycholinguistes comme Jerry Fodor, notre langage humain serait l'externalisation d'une sorte de langage interne, qui serait celui de la pensée (Fodor, 1975). Les ordinateurs disposent d'un ensemble de processus logiques internes qui pourraient être comparés à ce langage interne fodorien et, un peu comme lors de l'élaboration de la pensée humaine, ces processus logiques internes simulerait le monde extérieur. Ils permettraient ainsi aux machines une prévision des résultats de certaines de leurs actions, voire la possibilité de tenir compte de certaines de leurs erreurs pour modifier leur comportement ultérieur, donc d'apprendre.

Forts de ces aptitudes complexes d'intelligence, les robots peuvent posséder des traits culturels. L'exemple le plus remarquable en est l'expérience qui a permis à des groupes de robots de créer leur propre vocabulaire. « Un des robots désignait un des objets présents par une séquence de sons. A partir des sons entendus, le deuxième robot indiquait (...) l'objet qu'il pensait être désigné » (Chapouthier, Kaplan, 2011, p. 85). Il s'ensuivait un dialogue entre les robots et des comportements de confirmation ou d'infirmité des termes (sonores) utilisés. Par de tels échanges entre les robots se constituait peu à peu un vocabulaire sonore commun pour désigner les objets, alors que « sans communication, les robots auraient tous construits des systèmes différents » (Chapouthier, Kaplan, 2011, p. 85). Si on ajoutait ensuite au groupe de nouvelles machines, elles adoptaient le vocabulaire commun tout en contribuant à son évolution éventuelle. Une sorte de sélection culturelle s'opérait pour optimiser le langage commun.

Plus délicates sont les possibilités d'imaginer des normes morales ou esthétiques chez les robots. Certes on peut donner aux machines des contraintes « extrinsèques », on peut les programmer pour qu'elles ne fassent pas certaines actions comme : « ne pas porter atteinte à une autre machine ». Mais il est difficile d'imaginer, de nos jours, que le « langage intérieur » de la machine, évoqué plus haut, la mène à des motivations internes de comportements moraux conscients, même si on ne peut exclure qu'à terme une sorte de morale « aveugle » « émerge spontanément entre machines apprenantes » (Chapouthier, Kaplan, 2011, p. 95). Pour le moment les machines se comportent « comme si » elles suivaient certaines règles morales, sans que l'on puisse imaginer qu'elles en aient vraiment conscience. Les mêmes considérations peuvent être développées pour les choix esthétiques. Les machines pourraient être programmées pour faire certains choix de formes ou de couleurs, voire faire émerger certains choix « aveugles » par leur fonctionnement même, sans que l'on puisse dire qu'il s'agirait de choix « conscients ». La conscience est, nous semble-t-il, le facteur limitant, dans l'état actuel des robots, pour une vraie comparaison avec les performances humaines.

5. LES LIMITES DES ROBOTS ACTUELS

On peut élargir cette notion de conscience à une des caractéristiques des animaux sur laquelle nous aimerions insister ; la notion de douleur. Ici encore, on peut programmer des machines pour qu'elles affichent :

«j'ai mal», quand un objet extérieur porte atteinte à leur intégrité, mais il paraît difficile, dans l'état actuel des connaissances en robotique, de penser que le processus d'affichage correspond à une douleur effective dans le vécu existentiel de la machine. Nous sommes toujours dans le «comme si». Le vécu existentiel de la douleur, qui, chez les animaux les plus céphalisés, se manifeste par des émotions, ne peut pas raisonnablement être attribué aux machines actuelles. Héritée de sa nature animale, la conscience humaine inclut, au contraire, une composante émotionnelle et affective essentielle.

Une autre différence nous paraît essentielle entre le comportement des robots actuels et celui des animaux. Il s'agit de l'aptitude à un apprentissage systématique. Certes les robots ont des capacités d'apprentissage non négligeables, mais les robots actuels, contrairement aux animaux, ne circulent pas en permanence dans l'environnement. Il s'ensuit que leurs capacités d'apprentissage sont fortement limitées par ce que leur programmeur veut bien leur communiquer. Les animaux, en revanche, interagissent en permanence grâce à leurs organes sensoriels, avec leur environnement, dans une sorte de dialectique ininterrompue avec leur milieu de vie. D'où, chez les animaux, et contrairement aux robots actuels, une «éducation ouverte», une éducation permanente par les contacts avec leur environnement, une éducation dans laquelle les aspects émotionnels, et par suite le rapport affectif à l'autre, jouent un rôle essentiel. Ce sont ces rapports affectifs à l'autre qui peuvent être considérés comme les racines des (proto) morales, qui ont pu être décrites chez les animaux.

Conscience et dialectique permanente avec leur environnement, toutes deux liées à l'émotion, voici deux différences essentielles du comportement des robots par rapport à celui des animaux.

6. LA PLACE DE L'ÊTRE HUMAIN

Toutes ces considérations nous amènent à mieux cerner la place de l'être humain. Nous venons de voir les limites des robots actuels par rapport aux animaux. Or l'homme est un animal et, en tant que tel, bénéficie des deux points essentiels que nous venons de mentionner : il dialogue en permanence avec son environnement et il possède des processus émotionnels conscients, qui sont une part essentielle de son être. Bien entendu, on peut imaginer, selon une approche voisine de la science-fiction, la création de robots futurs d'un nouveau genre, qui

possèderaient ces mêmes caractéristiques que les animaux. A notre avis, si cela devait se faire, il faudrait utiliser les mêmes procédés de miniaturisation carbonée que ceux que la nature a sélectionnés par l'évolution des espèces. En d'autres termes, le projet futuriste de robots qui se comporteraient comme les animaux, notamment pour les deux points essentiels qui viennent d'être mentionnés, n'est pas dissociable de la fabrication éventuelle par l'homme d'animaux artificiels. Nous en sommes, bien sûr, encore très loin.

Si les robots actuels parviennent à maîtriser de nombreux aspects de l'activité intellectuelle humaine, dans ses dimensions abstraites, voire cognitives, ils butent notamment sur le caractère émotionnel ou affectif de l'intelligence humaine. Par rapport aux robots actuels, l'homme se distingue clairement par son caractère animal. Il est amusant de constater qu'après avoir voulu si souvent se couper de « la bête » et de l'animalité, l'être humain, pour éviter la robotisation, devra justement faire appel à ses traits animaux. Ce sont justement les aptitudes émotionnelles héritées de l'animalité qui permettent à la conscience humaine d'accéder à un sens de l'empathie et de l'altérité, qui sont les conditions d'un vrai sens moral, par rapport à la morale du « comme si » des robots actuels. Ses ancêtres animaux ont donné à l'homme le seul accès à une morale véritable. Dans le domaine culturel, c'est donc par les aspects bénéfiques de son animalité que l'homme se distingue, de nos jours, des robots.

BIBLIOGRAPHIE

- Georges Chapouthier, (2006). *Biologie de la mémoire*, Paris, Odile Jacob.
- Chapouthier, Georges, (2009). *Kant et le chimpanzé. Essai sur l'être humain, la morale et l'art*, Paris, Éditions Belin.
- Chapouthier, Georges, (2012). « Les chiens comprennent-ils la parole? », *Cerveau et Psycho*, Mai-Juin, 51, p. 92-93.
- Chapouthier, Georges, Kaplan, Frédéric, (2011). *L'homme, l'animal et la machine, perpétuelles redéfinitions*, Paris, CNRS Éditions.
- Chapouthier, Georges, Policar, Alain, (2015). « La néoténie humaine, une idée à relancer », *Pour la Science*, 452, p. 14-15.
- De Waal, Frans, (1989). *Peace making among primates*, Harvard University Press, USA.
- Fodor Jerry A, (1975). *The language of thought*, New York, Crowell.
- Kreutzer, Michel, Aebischer, Verena, (2015). « The Riddle of Attractiveness: Looking for an 'Aesthetic Sense' Within the Hedonic Mind of the Beholders », in: Thierry Hoquet (ed.) *Current perspective in sexual selection*, Springer Verlag, Germany, p. 263-285

- Lestel, Dominique, (2001). *Les origines animales de la culture*, Paris, Flammarion.
- Oudeyer Pierre-Yves, (2015). «Quand les machines créent leur propre complexité», in: Jean Audouze, Georges Chapouthier, Denis Laming, Pierre-Yves Oudeyer (dir.), *Mondes mosaïques*, Paris, CNRS Editions, p. 206-207.
- Panksepp, Jaak, Burgdorf, Jeff, (2003). «'Laughing' rats and the evolutionary antecedents of human joy?», *Physiology and Behavior*, 79, p. 533-547.
- Vauclair Jacques, (1995). *Intelligence de l'animal*, Paris, Seuil.
- Watanabe, Shigeru, (2013). «Animal aesthetics from the perspective of comparative cognition» in: Shigeru Watanabe, Stan A Kuczaj (Eds.), *Emotions of Animals and Humans, Comparative Perspectives*, Springer, Japan, 2013, Chapter 7, p. 129-162.

CHAPITRE 20

La robotisation des environnements : de l'émancipation des corps à la domination de la nature

Louis-Étienne PIGEON

INTRODUCTION

L'être humain a pour particularité de transformer son environnement de manière parfois profonde. Au sein de la biosphère, il est certainement l'espèce la plus influente en rapport à sa structure écosystémique, au point tel que l'on parle désormais de notre ère comme de l'ère anthropocène. Un concept analogue peut être employé pour décrire le monde que nous façonnons actuellement, soit le concept de technosphère. En effet, si le temps géologique nous appartient, cette identification ne peut se faire qu'à partir de l'activité technique de l'être humain, principal mode de transformation de l'environnement que l'être humain utilise pour construire ses milieux.

Or, la technosphère se déploie sous un mode d'expansion qui mobilise un ensemble important et complexe de techniques, de technologies, de systèmes et d'actions, parfois concertées, parfois aléatoires, mais qui altèrent l'identité du substrat naturel des espaces, des lieux et des environnements dans lesquels nous évoluons. À l'heure où nous nous questionnons sur les enjeux éthiques et philosophiques de l'apparition

dans nos vies de robots et de nouveaux êtres hybrides de nature et de technologie, nous devons tout autant mettre en question la mise en œuvre de ces dispositifs à des échelles plus grandes, car l'émergence de la technosphère implique aussi des formes de robotisation des environnements. L'enjeu n'est pas banal, car l'hybridation des fonctions écosystémiques, ou simplement des environnements naturels, avec nos dispositifs techniques de robotisation ne peut être pensée de la même manière que l'hybridation des corps eux-mêmes. Un environnement est un espace, il inclut des multitudes de corps, de processus naturels, d'êtres et de choses. Un environnement est un contexte dans lequel se déploient des actions, et en ce sens, il peut soit être ouvert sur l'émergence des possibles ou alors fermé et déterminant. Alors que le projet de convergence de la biologie humaine et de la technique s'appuie sur le fantasme d'une émancipation de l'être humain de sa condition naturelle limitée, la robotisation des milieux, quant à elle, reconduit forcément un projet de domination de la nature, dans lequel les corps biologiques sont *de facto* soumis à un processus de contrôle.

Afin de traiter de cette problématique, nous commencerons par situer le phénomène de la robotisation des environnements dans le contexte de l'émergence de la technosphère (1). Ensuite, nous préciserons ce que nous entendons par la robotisation des corps et des environnements (2). Finalement, nous poserons plus précisément la question de la problématique à la fois ontologique et éthique de ce processus pour en tirer quelques conclusions (3).

1. LA TECHNOSPHERE, CONTEXTE ET HORIZON DE LA ROBOTISATION

C'est au scientifique russe Vladimir Ivanovitch Vernadski (1863-1945) que nous devons le premier emploi du terme « technosphère ». Vernadski n'était cependant pas technologue, mais expert en minéralogie et en chimie si bien qu'il s'intéressait principalement à la structure physique de la planète que nous occupons. Non plus avait-il comme objectif de fonder une philosophie de la technique, car ses travaux se sont plus attardés sur ce qu'il nomma la « biosphère », soit la partie vivante de la planète. Le système de Vernadski comportait également d'autres concepts visant à décrire les différentes « couches » de la planète : la géosphère, l'atmosphère, la lithosphère et la noosphère. Mais sa proposition de nommer le domaine de l'artificialité (la technosphère)

et de lui offrir une place analogue aux autres reste à ce jour fort pertinente, sinon plus qu'à l'époque de sa formulation. En effet, ce que Vernadski nomma la « technosphère », l'ensemble de la nature transformée par l'être humain (Vernadski, 2002), est au centre de la problématique de la robotisation des environnements.

Il est peu habituel de parler de la robotisation des environnements, parce que l'image que nous avons des robots pointe vers des entités circonscrites dans l'espace : des robots humanoïdes, par exemple. Il existe pourtant des robots invisibles, comme ceux qui agissent dans le monde numérique et bien d'autres systèmes automatisés qui peuplent nos environnements quotidiens. La place qu'occupent tous ces robots dans le système planétaire relève pour autant d'un même phénomène par lequel la technosphère tend à englober les autres dimensions de la planète. À travers ce processus, la hiérarchie des couches constitutives de la planète s'altère et se transforme. Un exemple de cela est bien connu sous l'appellation des « changements climatiques ». La couche atmosphérique de la planète, en interaction jusqu'ici assez stable avec les autres couches constitutives, se transforme. En fait, l'atmosphère devient objet de la technosphère dans la mesure où nous savons que les perturbations climatiques auxquelles nous assistons sont largement dues aux conséquences de l'activité technique humaine. La terminologie de Vernadski permet ainsi de nommer des interactions, de décrire l'évolution du système terrestre. Ainsi peut-on constater également que la robotisation est une étape dans un processus d'artificialisation de la biosphère. Ce processus, quoique semblable à la transformation de l'atmosphère, comporte ses caractéristiques propres que nous détaillerons ultérieurement.

D'abord, penser cette problématique dans une logique environnementale pose deux difficultés sur le plan de la logique : d'une part concernant les relations entre les différentes couches de la planète, et d'autre part, en ce qui a trait à la nature même des environnements.

Premièrement, les différentes couches de la planète nommées précédemment se côtoient, se superposent et interagissent. La géosphère côtoie l'atmosphère et ces deux entités s'influencent mutuellement. Elles interagissent. De même, la technosphère, qui relève de l'activité humaine, est en interaction avec l'ensemble des couches composantes de la Terre. Nous ne pouvons donc pas penser le processus d'artificialisation environnemental dans une perspective linéaire, comme si le monde humain avançait, un pas à la fois, dans une zone qui lui aurait

été étrangère. Une telle approche impliquerait d'admettre une séparation radicale entre les domaines de la nature et de la culture, une thèse par ailleurs mise à mal par la littérature depuis un demi-siècle. Il s'agit plutôt ici d'une dynamique d'émergence. La technosphère, parce qu'elle est essentiellement une construction humaine, prend source dans la biosphère, dans laquelle elle est au départ *englobée*. L'être humain est un animal particulier, certes, mais tout aussi biologique que les autres. Conséquemment, son activité technique de transformation des milieux, de construction d'artefacts, etc., procède, *stricto sensu*, d'une naturalité antérieure¹. Lorsque nous observons la robotisation des environnements, nous assistons à une modification de cette hiérarchie des englobements, par laquelle les rôles d'englobé et d'englobant changent. L'environnement naturel a toujours été le contexte, le lieu de la technique. Se pourrait-il alors que par un phénomène de renversement le lieu technique, artificialisé, devienne le contexte de la naturalité? Peut-être une telle conclusion serait-elle hâtive à l'heure actuelle, mais étant donné la puissance des idées technopositivistes et de ses moyens techniques de réalisation, un tel scénario doit être, au moins, considéré.

Sur le plan théorique, le jeu des englobements ici mobilisé se réfère à la philosophie de Kitaro Nishida, qui, dans ses travaux sur la logique du lieu (*basho*)², postulait que rien n'existe en soi quelque part, mais que tout contenu est englobé dans un milieu: «Il n'y a donc pas de milieu séparé du contenu, et pas de contenu séparé du milieu.» (Tremblay, 2015, p. 171) Cette approche permet, entre autres, de penser la problématique de la robotisation des environnements au-delà de la métaphore d'une conquête linéaire issue de l'idée selon laquelle notre monde serait structuré par une dichotomie de la nature et de la culture. Mais il existe surtout entre le milieu et son contenu,

1. Il est désormais bien admis que l'une des caractéristiques les plus importantes de l'être humain sur le plan des stratégies évolutives est sa capacité à transformer son environnement et non seulement de fabriquer des outils. Dans tous les cas, il semblerait que la capacité à développer des techniques soit centrale à la constitution de l'*homo faber*. Pour de plus amples détails sur ce sujet, le lecteur se référera à l'ouvrage *L'homme artificier*, de Dominique Bourg (1996).

2. Le philosophe japonais Nishida Kitarô (1870-1945) a développé sa logique du lieu (*basho*) au fil de plusieurs ouvrages. Les plus importants ont récemment été rassemblés, traduits en français par Jacynthe Tremblay et publiés dans le recueil *De ce qui agit à ce qui voit* (2015). Le lecteur pourra se référer à cet ouvrage pour plus de détails concernant cette approche.

entre l'englobant et l'englobé, une relation essentielle et constitutive, la relation de «détermination réciproque». Par celle-ci tout contenu est déterminé par son milieu, mais par le fait même tout milieu est également déterminé par son contenu en tant que milieu (p. 171). Cette détermination peut se décrire comme: «une action, consciente ou non selon le cas, exercée sur son vis-à-vis dans la relation visant à le transformer (peu importe qu'il s'agisse d'une transformation en bien ou en mal)» (p. 172). En ce sens, la hiérarchie des englobements est aussi une hiérarchie des déterminations, car si le milieu est un élément englobant, il est aussi l'élément déterminant face à son contenu englobé. En d'autres mots, le milieu détermine son contenu, et sachant que les environnements sont des contextes, des espaces dans lesquels nous agissons, nous pouvons dès lors nous demander jusqu'à quel point nous souhaitons être déterminés par des milieux où la robotisation devient le moteur des actions.

Selon cette logique, le processus d'émergence de la technosphère se comprend comme une action de transformation de la biosphère par l'activité technique, action dans laquelle la robotisation prend désormais une place de choix étant donné la multiplication des systèmes automatisés et robotisés qui deviennent des éléments quotidiens de nos vies. Il va sans dire que la métaphore de la conquête linéaire ne correspond pas au portrait relationnel ici présenté, car c'est bien de l'intérieur du milieu humain lui-même qu'émergent les environnements robotisés.

Cela pose la deuxième difficulté évoquée plus haut, à savoir que l'environnement dont nous parlons est un lieu dans lequel nous sommes englobés. Or, loin d'être détruit par nos actions de transformation technique, cet environnement est plutôt altéré par le jeu des déterminations réciproques. Nous aurions affaire à une forme de renversement de la relation de détermination réciproque entre la biosphère et la technosphère, par lequel ce qui était englobant (la biosphère) tend à devenir ce qui est englobé. À l'inverse, la technosphère, que l'on identifiait à l'englobée, tend à devenir l'englobant. Autrement dit, la technosphère devient le lieu, le milieu, alors qu'elle était le contenu.

Mais pour mieux comprendre la portée de cette explication, il nous faut porter un regard sur ce que nous définissons comme un «robot».

2. QU'ENTEND-ON PAR ROBOTISATION ?

La robotisation de notre monde consiste en la multiplication de dispositifs automatisés au sein de nos systèmes sociaux ou alors naturels. Il s'agit d'une transformation importante de nos contextes de vie. Ce processus de transformation mobilise plus précisément l'application de différentes techniques d'artificialisation des corps, mais aussi de ce que nous nommons l'intelligence, donc de certaines fonctions dites immatérielles. Qui plus est, la robotisation ne peut qu'être ce qu'elle est, à savoir, différente des autres formes d'artificialisation, que dans la mesure où elle engendre une automatisation artificielle des objets et des processus qu'elle a pour objet. Il ne s'agit donc pas que d'artificialiser des éléments inertes de l'environnement, comme remplacer des plantes naturelles par des plantes artificielles en plastique dans un centre commercial, mais bien d'engendrer aussi des processus caractérisés par une certaine autonomie d'action. Pour reprendre l'exemple du centre commercial, le contrôle automatisé de la température procède, lui, de l'automatisation, alors que la présence de plantes artificielles ne répond pas à la caractéristique énoncée. Ainsi, lorsque nous parlons d'une « robotisation des environnements », nous voulons parler d'un ensemble de dispositifs techniques qui se côtoient, qui s'entremêlent, se complètent et interagissent entre eux, comme le font les différents êtres naturels et les fonctions qui constituent un écosystème.

En effet, si nous devons identifier ce qui définit un robot, essentiellement, nous rencontrons la notion « d'automatisation³ ». Le robot est une machine, un programme ou une installation artificielle qui opère de manière autonome. Un robot est avant tout un automate qui répond à certaines caractéristiques, dont celle d'avoir été programmé. Les développements récents de l'informatique ont propulsé nos moyens en termes de création robotique : « Alors que dans la plupart des machines l'utilisateur commande la machine pas à pas, celle-ci exécutant les instructions qui lui sont données au fur et à mesure qu'elle les reçoit, la programmation suppose un stockage des instructions puis le déroulement indépendamment de l'utilisateur » (Chazal, 1996, p. 27 [voir sous Von Neumann, 1996]).

L'autonomie face à l'utilisateur est ici la clé de voûte de la définition : un robot opère, agit et déploie ses actions dans le monde indépendamment

3. Pour approfondir la notion scientifique de l'automatisation à sa source, le lecteur pourra se référer au traité *Théorie générale et logique des automates* de John Von Neumann (1996), dont la première parution remonte à 1951.

des utilisateurs (dans la mesure où il a été mis en marche, bien sûr). Cette autonomie peut certes être banale, comme l'est l'automatisation d'une machine industrielle qui pose des rivets sur une carrosserie d'automobile sur une chaîne de montage, ou encore, comme l'est la mise en marche automatique d'une cafetière programmable; mais elle peut également atteindre des niveaux de complexité étonnants, comme le serait l'automatisation de la gestion du trafic routier dans une grande métropole moderne. Ce dernier exemple impliquerait une programmation des feux routiers, un dispositif de reconnaissance et de surveillance et la mise en fonction d'algorithmes décisionnels. Quand nous imaginons des robots, nous pensons plus aisément à la chaîne de montage, mais peu souvent à la cafetière, qui est trop banale; et encore moins avons-nous à l'esprit la gestion du trafic routier, qui pourtant demande des fonctions automatisées hautement plus perfectionnées et qui peut devenir extrêmement déterminante dans notre vie quotidienne.

Dans tous les cas, le processus de robotisation auquel nous assistons laisse entrevoir deux pratiques qui se distinguent : la robotisation des corps et la robotisation des milieux, dont les finalités individuelles ne sont pas analogues, mais qui se recoupent néanmoins car ayant pour conséquence commune d'opérer une fusion (contextuelle ou conceptuelle) entre le vivant et la machine.

3. LA ROBOTISATION DES CORPS

Le processus de robotisation de la nature implique nécessairement de faire entrer dans le monde du biologique des éléments artificiels et, du moins, partiellement autonomes. L'un des lieux importants de cette hybridation est le corps humain lui-même. En effet, les avancées scientifiques des dernières décennies nous ont démontré que cette fusion n'est plus du domaine strict de la science-fiction. Pour des motifs de recherche scientifique, par exemple, il est courant d'associer le corps à la machine, d'en faire un relai informationnel pour mieux le comprendre, ou alors pour mieux comprendre la machine? La même stratégie est aussi employée en médecine, discipline par laquelle s'est développée toute une culture de l'hybridation à travers la production de prothèses, internes ou externes au corps, qui, une fois mises en marche, laissent émerger la figure du cyborg, un être mi-humain,

mi-machine, dont l'existence relèverait plus de la cybernétique que de la biologie⁴.

Mais au-delà du domaine thérapeutique, certains entrevoient même l'avènement d'une posthumanité, envisageable grâce aux possibilités amélioratives que fournissent la robotique et l'hybridation du biologique et de l'artificiel. Les promoteurs de cette voie jouissent d'ailleurs d'une certaine popularité, car leur position reconduit un fantasme bien ancien et ancré dans l'idée d'une libération du corps humain de ses contraintes physiques: « Ils annoncent la venue d'esprits (*minds*) sans entravent, libérés des corps, affranchis des passions et accédant à l'immortalité » (Lecourt, 2007, p. 222). En somme, la coévolution du vivant et de la technique demande que nous nous unissions aux robots que nous créons, peu importe leur forme. Que cette hybridation soit simpliste, comme l'est la greffe de dispositifs mécaniques au corps humain, ou complexe, comme l'alliage de la pensée humaine à des systèmes informatiques en voie d'automatisation, ils seront alors perçus comme des jalons dans un processus de coévolution entrevu avec enthousiasme. La posthumanité représente, en effet, la maîtrise par l'humain de sa propre destinée naturelle. Elle est le chemin le plus légitime qui nous permettra de nous libérer de la douleur et du vieillissement, de laisser s'exprimer la créativité humaine la plus exacerbée et peut-être même de repousser les limites de la mort. Il s'agit, à juste titre, d'un fantasme d'émancipation⁵.

Le posthumanisme s'inscrit parfaitement dans la perspective d'une émergence de la technosphère, car sa finalité est proprement évolutive ou transformative, bien qu'elle n'ait directement pour objet que

4. On ne peut pas, en effet, comprendre les enjeux épistémologiques et éthiques de l'hybridation humain-machine sans accepter l'importance du paradigme cybernétique dans la compréhension du monde. Hérité des sciences du milieu du XX^e, le paradigme cybernétique permet la réduction des activités biologiques et technologiques à un dénominateur commun: l'information. Pour citer l'excellent essai de Céline Lafontaine: « La cybernétique subordonne en fait la vie au principe informationnel. Suivant cette logique, l'être humain ne possède qu'une valeur différentielle reliée à sa capacité de traiter de l'information complexe. Lui déniait tout statut ontologique particulier, le père de la cybernétique (Norbert Wiener) lui accorde toutefois une valeur prédominante dans la hiérarchie du vivant » (Lafontaine, 2004, p. 42-43).

5. Si nous employons ici le terme « fantasme », Dominic Lecourt n'hésite pas, de son côté, à employer un vocabulaire qui réfère au mythe et à la prophétie. En effet, il qualifie les enthousiastes du posthumanisme de « technoprophètes » (Lecourt, 2007). Un tel langage ne constitue pas un débordement dans la mesure où le rêve posthumaniste peut impliquer de viser l'immortalité de l'esprit par la substitution des corps artificiels (robotiques ou non) aux corps biologiques.

des parties très limitées du système terrestre : les corps humains. Il affirme l'idée selon laquelle l'être humain peut prendre en charge le processus de son évolution et mobilise directement un ensemble de technologies relevant ou associé à la robotique. Or, dans une logique de milieu, si le contenu englobé détermine par réciprocité la forme de son englobant, il va sans dire que la transformation des corps humains participe de la transformation du milieu lui-même, du moins dans une certaine mesure. Ce qui en ressort est que cette perspective est présentée comme émancipatrice des conditions biologiques limitées de l'être humain et qu'à travers la lunette du posthumanisme, la nature, ou l'évolution naturelle, est une frontière que les technologies telles que les biotechnologies, mais aussi la robotique et les procédés de programmation qui lui sont attribuables, nous aideront à dépasser.

4. LA ROBOTISATION DES ENVIRONNEMENTS

La robotisation des environnements naturels, ou des milieux humains, procède d'une logique d'automatisation et d'hybridation technologique semblable à celle appliquée aux corps au sens où elle mobilise les mêmes connaissances scientifiques, mais dont les effets n'émergent pas sous la même forme. Or, bien qu'il soit peu commun de parler de milieux robotisés, il existe néanmoins plusieurs exemples concrets qui illustrent la réalité de ce processus.

Nous connaissons tous la base des principes techniques de l'industrialisation, par lesquels des lieux de production peuvent être automatisés. Dans une installation industrielle, la robotisation peut s'appliquer à une machine en particulier, mais elle peut aussi s'étendre à toute une chaîne de montage. De plus, avec l'apparition des robots dotés d'une autonomie d'action plus importante, certaines tâches, comme le déplacement des produits dans un entrepôt, ou même prochainement la livraison, peuvent être automatisés. Dans la littérature de la philosophie de l'environnement, ce processus est décrit comme le résultat de l'émergence d'une véritable « culture techno-industrielle » (Naess, 2013, p. 51).

L'usine devient donc un milieu robotisé, car les activités qui s'y déroulent dépendent largement de la mise en œuvre de technologies robotiques. Cette même pratique touche désormais des milieux où la présence d'êtres vivants est essentielle et où, alors, la robotisation de certaines activités mettra directement en contact les êtres vivants avec

les processus et les fonctions automatisés. À titre d'exemple, les fermes laitières robotisées apparaissent comme de véritables laboratoires où l'on expérimente la possibilité de marier l'existence biologique des vaches à un environnement de plus en plus automatisé. En ce qui a trait à la production animale ou à ses dérivés (le lait, mais aussi les œufs, etc.), certains n'hésitent pas à décrire le processus de robotisation comme un phénomène qui tend à transformer les êtres vivants eux-mêmes, au sens où ils feraient désormais partie d'une logique industrielle et que cette incorporation au système de production aurait une influence sur leur développement et sur leur existence (Larrère et Larrère, 2015, p. 220).

Un élément important doit néanmoins être retenu : en aucun cas les êtres vivants ne sont complètement exclus de cette transformation, mais leur rôle change. Les animaux sont intégrés dans la fonction automatisée de production, et de son côté la tâche du fermier devient moins manuelle. Si les animaux deviennent des appareils de production, le fermier se transforme quant à lui en gestionnaire de système (ce qui est souhaité pour certains, et non pour d'autres); et la vache laitière devient, dans ce cas, productrice volontaire par habitude ou conditionnement (Holloway, Bear et Wilkinson, 2014).

Avec un peu de recul, il est possible de voir dans ces dispositifs environnementaux, dans les nouvelles installations de production agroalimentaires, des bancs d'essai. En effet, il est dans la pratique courante en recherche scientifique, médicale ou autre, de tester l'application de nos inventions sur les corps des animaux. Les différents protocoles de la recherche pharmaceutique en sont l'exemple le plus connu. Peut-on alors poser l'hypothèse selon laquelle la transformation des environnements propres à la production agroalimentaire serait une anticipation de ce que nous souhaitons réaliser sur nos propres environnements⁶? Nous oublions souvent que le mode de vie moderne des pays développés amène les êtres humains eux-mêmes à circuler et à agir au sein d'environnements de plus en plus robotisés. Non pas que ces environnements seraient eux-mêmes des robots, mais la présence de systèmes autonomes, parfois intégrés les uns aux autres, devient structurante, car constitutive de l'horizon spatial, symbolique

6. Les lieux de travail peuvent être ici cités en exemple dans des cas où, par exemple, la multiplication des outils de surveillance permet un contrôle de plus en plus serré des comportements humains. On pense directement aux caméras, mais aussi doit-on inclure dans ces dispositifs les outils de télécommunication qui permettent souvent de monitorer la présence des travailleurs sur les lieux de production.

et fonctionnel. En d'autres mots, la robotique investit les milieux et les teinte de ses caractéristiques. En effet, les techniques de robotisation et d'automatisation sont de plus en plus présentes dans notre quotidien. De la machine à café au téléphone intelligent ; des feux de circulation aux systèmes de contrôle de la température, la multiplication des techniques change la qualité du tissu du monde que nous habitons, de nos milieux⁷.

Cela amène à dire qu'il y a donc deux types d'environnements robotisés. Le premier est un environnement clos, circonscrit, où une activité, souvent productive, est optimisée par l'automatisation des actions techniques concertées. C'est un lieu habité de machines qui « travaillent » sous les ordres d'autres machines, ou alors ce sont des environnements numériques où des programmes (*bots*) agissent sur des données. Il s'agit du modèle de l'usine, dans lequel l'existence biologique est obsolète, du moins conceptuellement. Le deuxième type se rapporte plus à un modèle touchant les milieux ambiants. Il s'agit des lieux que nous habitons, où nous travaillons et où nous circulons. Dans ces espaces, la robotisation du milieu ne se réalise pas nécessairement par une gestion intégrée, ou par la mise en fonction d'une « méga machine », mais bien par l'addition successive de différents dispositifs automatisés dans l'espace. C'est l'horizon naturel de la culture techno-industrielle. Leur intégration en un système cohérent peut être verticale (par des techniques de gestion, automatisées ou semi-automatisées) ou alors horizontale (par des mises en réseau spontanées), ou les deux. Il en résulte une transformation du milieu, car les qualités relationnelles de celui-ci s'en trouvent transformées. Si nous sommes généralement frileux à l'idée de vivre dans un environnement robotisé intégré verticalement, qui rappelle des romans dystopiques tels que *1984* ou *Le Meilleur des mondes*, nous sommes beaucoup moins conscients qu'une forme d'intégration horizontale est néanmoins déjà à l'œuvre. Celle-ci se réalise par la mise en service d'une multitude de fonctions et d'appareils automatisés, librement intégrés dans nos vies

7. Cette constatation pointe vers un débat fort pertinent en philosophie de la technique, à savoir que la technique moderne ne serait pas en elle-même si différente de la *tekne* ancienne, comme l'affirmait Heidegger (1993), mais que ce serait plutôt la généralisation de l'usage de la technique qui serait le véritable moteur de changement. À ce sujet, voir Ellul (1990). Bien que notre but ne soit pas ici de proposer une position définitive, nous pouvons toutefois affirmer que, peu importe si la technique moderne est intrinsèquement différente de la *tekne* ou non, la problématique de la robotisation des environnements concerne directement la multiplication de l'usage des dispositifs automatisés.

par le biais du système économique libéral, qui en permet la vente, l'achat et l'utilisation sur la base de l'utilité⁸.

5. PROBLÉMATIQUE ONTOLOGIQUE ET ÉTHIQUE DES ENVIRONNEMENTS ROBOTISÉS

À partir de ces considérations, nous pouvons dire que le processus de robotisation des milieux, propre à l'émergence de la technosphère comme environnement global, doit se juger à l'aune de deux problématiques, l'une ontologique et l'autre éthique (ou normative). Dans les deux cas, nous pourrions être à même de constater que ce processus s'inscrit dans une logique de domination qui, sans nécessairement contraindre la corporalité du vivant en elle-même, comme elle le ferait en le dénaturant, agit néanmoins sur son contexte d'existence à travers son milieu.

Premièrement, sur le plan ontologique, en opposition aux êtres vivants dotés d'un corps défini, les environnements ne se déploient pas ou n'évoluent pas exactement en accord avec des lois biologiques. Dès lors, tout projet de robotisation des milieux s'appuyant sur l'analogie entre le vivant et la machine est en soi un faux départ. Les environnements et les milieux sont, par nature, des systèmes de coévolution où s'entrelacent une multiplicité de niveaux, de fonctions et d'éléments, que ceux-ci soient biotiques ou abiotiques. Or, l'automatisation de ces systèmes, ou alors parfois de certaines de leurs fonctions plus précises, demande une connaissance de leur fonctionnement. Mais si les sciences écologiques ou même sociales nous ont bien démontré une chose au courant du dernier siècle, c'est bien que notre compréhension de ces systèmes est encore parcellaire et que d'agir avec force en intervenant dans leur logique interne comporte des risques importants. Nous retrouvons ici le thème bien connu de l'incertitude, à savoir que nos actions de transformation technique, dont fait partie la robotisation des environnements, pourraient induire des chaînes causales dont nous ignorons les conséquences (Jonas, 1990, p. 33).

Les environnements et les milieux ne sont pas exactement des « choses », mais plutôt des contextes dans lesquels les choses apparaissent et existent. Ils sont des lieux déterminants de l'existence et non

8. Pour un approfondissement de cette problématique, le lecteur pourra se référer à Feenberg (2014).

pas uniquement des existants en soi. Dans une logique de milieux, nous dirions qu'ils sont des *englobants*. Leur déploiement tient ainsi à la mise en relation des choses les constituant. À ce titre, nous pouvons dire que les corps vivants, humains et non humains, ne peuvent exister qu'en relation à leur contexte, car les contextes sont les horizons, tant physiques qu'existentiels, dans lesquels les corps se déploient. Or, par le jeu des déterminations réciproques, les corps sont soumis à la détermination du milieu dans lequel ils existent.

Cette dernière définition mène à porter un regard éthique sur la tendance actuelle à robotiser les milieux humains, des milieux dont la liberté d'action et l'autonomie propre à la conscience sont des éléments constitutifs particuliers.

L'automatisation des milieux demeure une entreprise téléologiquement déterminée, que celle-ci se réalise sur le modèle de l'usine intégrée, ou sur le modèle de la multiplication des systèmes robotisés. Elle s'opérationnalise à partir des critères techniques de l'efficacité et de l'optimisation, en conformité d'une culture technicienne déjà bien implantée dans notre monde et qui tend à investir l'ensemble de nos domaines d'activités. La robotisation est une étape du développement de la technique qui « n'est plus limitée en rien » (Ellul, 1990, p. 73). Ainsi, la substitution d'une fonction écosystémique par un dispositif technique robotisé, comme peut l'être l'épuration des eaux par exemple, s'appuie sur l'optimisation d'une fonction d'abord identifiée. De même, la mise en œuvre d'algorithmes automatisés pour gérer la distribution des nouvelles sur les plateformes des médias sociaux répond à une demande d'optimisation relevant de la sélection de contenu. On peut donc constater que la robotisation, lorsque considérée dans une logique de milieu ou d'environnement, valorise toujours des fonctions ciblées au sein de ces systèmes : elle opère une sélection et hiérarchise, du fait même, les interactions propres au contexte. Il en découle des ouvertures et des fermetures possibles en ce qui a trait au contact des corps avec le monde, à leurs possibilités d'agir sur celui-ci par leur présence dans la nature, laquelle est déterminée autrement que par la volonté humaine technique. Autrement dit, la robotisation des environnements et des milieux est tout sauf une entreprise neutre sur le plan des finalités.

Agir en robotisant le contexte environnemental des êtres vivants, ou ici des milieux humains, équivaut à introduire des formes d'automatisation dans la toile complexe des univers relationnels. Cette démarche

ouvre des possibles, car elle peut effectivement optimiser certaines fonctions. Mais, puisque ces actions demeureront téléologiquement déterminées, c'est-à-dire attachées à des projets dont les finalités sont préétablies, n'y a-t-il pas raison de se questionner sur la mise en œuvre d'une nouvelle forme de pouvoir ?

CONCLUSION

À une époque où les avancées technologiques sont perçues avec un enthousiasme débordant, parfois aveugle, il peut être difficile de maintenir un discours critique. Si certaines prises de position peuvent facilement être déclassées et taxées de pessimisme, à tort ou à raison, il est tout aussi facile de sombrer dans un catastrophisme aussi contre-productif que réducteur. En effet, la robotique n'est pas une menace en soi, au sens où, comme toute connaissance technique et scientifique, elle peut être utilisée à bon escient. Il semble toutefois essentiel de mettre en question la portée de son application et les conséquences de sa mise en œuvre à des échelles qui échappent à la conception générale. La recherche ici effectuée dans la littérature sur la robotisation des environnements a laissé apparaître bon nombre d'articles spécialisés sur l'automatisation des fermes, mais peu de textes sur les environnements des humains eux-mêmes ! Ce simple fait amène à se questionner sur le degré de réflexivité de notre société quant à ses propres outils de transformation des milieux.

Un constant semble assez clair en ce sens. Le milieu qui émerge de nos activités technoscientifiques est de plus en plus qualifiée par les processus automatisés et les machines qui les rendent opérationnels. La construction de la technosphère ne relève pas d'un « plan » préétabli, mais d'un foisonnement culturel propre à une société désormais qualifiée de « technoscientifique ». La question devient donc : savons-nous réellement ce que nous produisons ? Ou alors, ne sommes-nous pas simplement transportés par les rêves de la technoscience, une pratique associable à « un processus historique qui transforme la nature et la société dans son ensemble en une vaste scène expérimentale » (Bensaude-Vincent, 2009, p. 15) ?

BIBLIOGRAPHIE

Bensaude-Vincent (2009). *Les vertiges de la technoscience*. Paris, Éditions La Découverte.

- Bourg, Dominique (1996). *L'homme artificie*. Paris, Gallimard.
- Briggle, Adam (2009). « *Technology* » dans John Baird Callicott et Robert Frodeman (dir.), *Encyclopedia of environmental ethics and philosophy*. Farmington Hills, Michigan, Gale.
- Callicott, J. B. (1999). *Beyond the Land Ethic. More Essays in Environmental Philosophy*. State University of New York Press.
- Devaux, Pierre (1960). *Automates, automatisme, automation*, collection Que sais-je? Paris, Presses universitaires de France.
- Diamond, Jared (2006). *Effondrement*. Paris, Gallimard.
- Ellul, Jacques (1990). *La technique ou l'enjeu du siècle*. Paris, Economica.
- Feenberg, Andrew (2014). *Pour une théorie critique de la technique*. Montréal, Lux.
- Heidegger, Martin (1993). *Essais et conférences*. Paris, Gallimard.
- Holloway, Lewis; Bear, Christopher, Wilkinson, Katy (2014). « Robotic milking technologies and renegotiating situated ethical relationships on UK dairy farms », *Agric. Hum. Values*, 31, p. 185-199.
- Jonas, Hans (1990). *Le principe responsabilité*. Paris, Les Éditions du Cerf.
- Lafontaine, Céline (2004). *L'empire cybernétique: des machines à penser à la pensée machine*. Paris, Éditions du Seuil.
- Lecourt, Dominique (2007). « La technique, la vie et la nature humaine » dans Marie-Hélène Parizeau et George Chapoutier (dir.) *L'être humain, l'animal et la technique*. Québec, Presses de l'Université Laval.
- Larrère, Catherine et Larrère, Raphaël (2015). *Penser et agir avec la nature*. Paris, Éditions La Découverte.
- Mihelcic, James R.; Zimmerman, Julie Beth (2014). *Ecological Engineering*. Hoboken, New Jersey, Wiley.
- Naess, Arne (2013). *Écologie, communauté et style de vie*. Paris, Éditions Dehors.
- Nishida Kitarô (2015). *De ce qui agit à ce qui voit*. Montréal, Presses de l'Université de Montréal.
- Tremblay, Jacynthe (2015). « Le milieu comme lieu interrelationnel chez Nishida » dans Marie-Hélène Parizeau et Jacynthe Tremblay (dir.), *Milieus modernes et reflets japonais*. Québec, Presses de l'Université Laval.
- Vernadski, Vladimir Ivanovitch (2002). *La biosphère*. Paris, Éditions du Seuil.
- Von Neumann, John (1996). *Théorie générale et logique des automates*. Précédé de *La pensée et les machines: le mécanisme algorythmique de John von Neumann* (Gérard Chazal). Seyssel, Champ Vallon.
- Watsuji, Tetsurô (2011). *Fûdo, le milieu humain*. Paris, CNRS Éditions.

CHAPITRE 21

Quelle place les robots peuvent-ils avoir dans la transition écologique ?

Catherine LARRÈRE

Ces derniers temps, on parle volontiers de « la révolution robotique¹ », alors qu'il est question, beaucoup plus modestement, de la « transition écologique² ». C'est que l'imaginaire technologique se déploie dans la démesure du « *hype* », promet un avenir radieux et tient pour négligeables d'éventuelles contraintes ou difficultés écologiques, auxquelles, de toute façon, on trouvera des solutions techniques³. De son côté, la conscience écologique, hantée par la catastrophe imminente, se détourne de ces promesses creuses, ou dangereuses. Elle n'est guère prête à faire de place à une robotique qui est à l'opposé de son imaginaire. Il n'est donc pas étonnant que, si la robotique et la transition écologique retiennent actuellement beaucoup l'attention, ce soit de façon séparée. On ne trouve guère de références écologiques dans les dossiers de presse ou dans les livres consacrés aux robots, pas plus que ceux qui traitent de la transition écologique n'y incluent volontiers la question des robots.

-
1. Par exemple « La révolution robotique est en marche », *Contrepoints*, 25 février 2013.
 2. Voir, notamment, Bourg, Kaufmann et Méda (2016), *L'Âge de la transition. En route pour la reconversion écologique*.
 3. Audétat M., Kaufman (2015), *Sciences et technologies émergentes, pourquoi tant de promesses?*

On ne peut cependant pas s'en tenir là. Certes, il y a de bonnes raisons de ne pas partager la croyance selon laquelle des moyens techniques suffiraient à maintenir notre mode de vie actuel, et notamment notre consommation énergétique⁴, cependant, il n'en reste pas moins que « l'importance cruciale de la technique et des infrastructures dans les transitions écologiques est comme une évidence » (Joly, 2016, p. 117). Quelle place les robots peuvent-ils avoir dans la transition écologique ? Si l'on envisage l'automatisation robotique comme une réalité en marche, quels rapports peut-elle entretenir avec la transition écologique également en cours ? Et s'il en est ainsi, à quelles conditions les robots peuvent-ils être intégrés dans la société écologique ? On en vient alors à se demander quel type de conduite il faudrait adopter à leur égard.

L'examen des questions ainsi soulevées nous mènera du revenu universel, envisagé comme la médiation entre l'automatisation et la transition écologique, à la question, rarement posée, de l'empreinte écologique des robots, pour nous conduire, finalement, à traiter de problèmes éthiques. Preuve, s'il en est besoin, que des questions à première vue techniques, comme celle des robots, ne sont jamais purement telles, mais engagent analyse sociale et réflexion morale.

1. QUEL IMPACT LES ROBOTS ET L'AUTOMATISATION CROISSANTE ONT-ILS SUR L'ENGAGEMENT DE LA TRANSITION ÉCOLOGIQUE ?

Pour ceux qui s'intéressent aux conséquences sociales et à l'insertion économique des technologies, l'importance des robots tient à leurs capacités à contribuer aux transformations du capitalisme. Le capitalisme industriel a fait place au capitalisme cognitif, ou à une économie de la connaissance, apte à faire des connaissances, comme plus largement du vivant, une production rentable, insérée dans les circuits économiques. Cela a été rendu possible par l'innovation technologique du numérique, qui, lentement murie depuis les années 1930, s'est imposée, dès les années 1980, avec les ordinateurs personnels et Internet. L'intelligence artificielle et les robots peuvent être ainsi vus comme participant à la fois d'« une deuxième vague de déploiement

4. C'est la vérité « désagréable » par laquelle Philippe Bihoux commence sa présentation (2016, p; 158).

du capitalisme cognitif», et d'« un deuxième âge des machines », celui de la mécanisation de l'esprit (Moulier-Boutang, 2015, p. 21). Cela se marque par une automatisation des activités intellectuelles (numériques, informationnelles) de plus en plus complexes, au point que des activités qui ne semblaient pas du tout devoir être automatisées, comme les tâches sociales et affectives, le sont à leur tour. On assiste ainsi, avec la production de robots destinés à s'occuper des autistes ou des vieillards, à la robotisation du *care* (Dumouchel, Damiano, 2016). Une fois que le domaine de l'affectif et du relationnel est en cours de robotisation, on ne peut même plus dire qu'il reste encore les emplois impliquant la créativité, la motivation, les relations interpersonnelles, le soin des autres. Aucune tâche ne semble pouvoir être sauvée de la robotisation.

On s'attend en même temps, du fait de cette automatisation, à une explosion spectaculaire des chiffres du chômage. Certes, celui-ci n'a pas attendu les robots et ne dépend pas seulement d'eux. Dans les années 1980, Michel Rocard annonçait qu'il n'y aurait pas de retour au plein emploi et, depuis, les invocations rituelles du retour d'une croissance qui apporterait nécessairement des emplois n'ont guère été suivies d'effets. Mais, venant s'ajouter à ce mouvement de fond, beaucoup d'études sur l'automatisation prévoient une diminution très importante du nombre des emplois, et une extension des domaines atteints par cette diminution. Ce ne sont plus les ouvriers et les employés d'exécution qui sont menacés par ces nouveaux automates, mais les classes moyennes : ingénieurs, programmeurs, marketeurs, médecins, avocats, avoués, administrateurs, agences de voyage, experts, enseignants, etc. Quand le ministre français de l'éducation nationale annonce, début 2018, qu'enseigner est une science, qui relève de la neurobiologie et des sciences cognitives, ne peut-on, à l'heure de la mécanisation de l'esprit, entendre cette annonce comme celle de la possibilité que les enseignants, qui font bien mal leur travail, soient remplacés par des robots beaucoup plus efficaces ?

Certes, ces évaluations et les conséquences que l'on peut en attendre sont discutées. On voit mettre en cause aussi bien les chiffres estimés des suppressions d'emplois à venir que la durée du chômage. Certains font valoir que les emplois éliminés par l'automatisation seraient systématiquement compensés à terme par ceux créés par l'innovation technologique – selon la théorie de la « destruction créatrice » de Joseph Schumpeter. Mais, pour d'autres, on assiste, avec ces modifications du capitalisme et les développements qui les accompagnent, à une

transformation de fond qui met en cause le salariat comme institution de base de nos sociétés.

Ce sont ces derniers analystes qui voient, dans le revenu universel, une réponse à la perte d'emploi consécutive à la robotisation. « Instaurons un revenu de base attribué à tout membre de la société qui remplacera la partie de base du salaire », propose Yann Moulier Boutang, à la fin de son article consacré à l'automatisation intellectuelle (Moulier-Boutang, 2015, p. 26). Floran Augagneur en précise la définition : « tous les membres de la communauté se verraient verser un revenu identique par la puissance publique, sans aucune condition, cumulable avec tous les autres revenus, y compris ceux du travail » (Augagneur, 2017, p. 50).

Mais quels rapports cette proposition entretient-elle avec la société écologique ? D'une part, les projets de revenu universel sont divers, et antérieurs à l'importance prise par la robotisation (Swaton, p. 13-62). D'autre part, le lien entre revenu universel et orientation écologique ne va nullement de soi. Dès les années 1970, André Gorz s'est intéressé aussi bien aux effets des limites écologiques (révélées par le rapport Meadows en 1972) sur l'économie, le travail et, plus généralement, le monde vécu, qu'aux possibilités sociales des nouvelles technologies, notamment numériques et informatiques⁵. Or, il a changé d'avis sur le revenu universel. Après l'avoir critiqué comme un mode de redistribution des revenus qui se ferait à l'avantage des riches et y avoir vu la menace de casser le lien entre revenu et travail (Gorz, 1985), il en est venu, en 1997, dans *Misères du présent, richesse du possible*, à l'accueillir comme une façon positive de sortir du salariat. Mais cela en fait-il vraiment un projet écologique ? Qu'est-ce qui a amené, notamment, la Fondation Nicolas Hulot pour la Nature et l'Homme, qui est une fondation écologique, à s'y intéresser de très près comme à une des composantes de la société écologique en préparation ? Et quelle place y a l'automatisation des tâches, y compris intellectuelles ou affectives ?

On n'appréhende pas de la même façon la robotisation dans ses effets individuels ou collectifs. Individuellement, les robots sont censés décharger les personnes d'un certain nombre de tâches pénibles, ce qui leur donne un rôle libérateur, à la façon dont Aristote jugeait que, quand les navettes iraient toutes seules, on n'aurait plus besoin d'esclaves (Aristote, *Politiques*, I, 4, 1253, b 35). Mais, collectivement,

5. Voir Willy Gianinazzi 2016.

la robotisation accroît la dépendance sur le plan social, en plaçant un nombre croissant de gens hors de la sphère du travail salarié. Or, la reconnaissance sociale, la dignité sont, depuis au moins le XIX^e siècle, attachées à la capacité à occuper un emploi, c'est-à-dire une activité salariée, validée par le marché ou par l'État, toutes les autres activités étant considérées comme improductives, comme du loisir, ou de l'oisiveté, qui se font aux dépens de la société.

Lorsque les emplois salariés ne sont plus accessibles à tous, il faut se donner les moyens d'attacher également une reconnaissance ou une validation sociale à des activités jusque-là exclues de l'estime sociale. C'est ce que devrait permettre l'existence d'un droit universel à un revenu de base. Celui-ci ne serait pas un revenu correctif sanctionnant un besoin (comme c'est le cas avec le revenu de solidarité active ou l'allocation chômage), mais représenterait un droit appartenant à tous, quels qu'ils soient. Il serait alors possible, comme l'explique Julien Dourgnon, de conjuguer les avantages individuels et la dimension collective. Le revenu universel permettrait d'apporter un surcroît de liberté à l'individu dans l'organisation de son temps. Le socle de revenu dont il disposerait inconditionnellement serait susceptible de modifier ses arbitrages : l'individu qui en disposerait aurait un peu de recul et pourrait échapper à la pression patronale (Dourgnon, 2017, p. 63). En même temps, comme le montrait André Gorz, le revenu ainsi attribué aurait une signification sociale, car il serait

en mesure d'opérer un changement radical par inversion du sens. La figure centrale du précaire qui est la nôtre, c'est elle qu'il s'agit de civiliser et de reconnaître au double sens du mot pour que, de condition subie, elle puisse devenir un mode de vie choisi, désirable, socialement maîtrisé et valorisé, source de cultures, de libertés et de socialités nouvelles, pour qu'elle puisse devenir le droit pour tous de choisir les discontinuités de son travail sans subir de discontinuité de revenu. (Gorz, 1997, p. 145)

Il s'agit donc de faire redevenir socialement actif le temps « libéré » par la robotisation et l'automatisation, en faisant, de ce qui, dans l'évaluation sociale actuellement dominante, n'est que du temps « vide », un travail socialement reconnu et individuellement enrichissant, « autonome » au sens de Gorz. Alors que le partage actuel entre actifs et inactifs freine le développement des activités menées hors du cadre salarial, le revenu universel en autoriserait le développement au service d'une société écologique. Cela permettrait d'encourager

les activités bénévoles, solidaires ou collaboratives, le *care*, comme de développer les activités hors marché ou démarchandisées (Perret, 2015). Avec le revenu universel peuvent se développer des activités péricapitalistes autour desquels s'organisent d'autres façons de vivre. C'est ce dont la transition écologique a besoin, directement et indirectement : une société plus sobre, quand la consommation n'est pas le seul objectif de l'amélioration des conditions de vie, quand il s'agit non pas de travailler plus pour gagner plus – quitte à se demander ce que l'on va faire de l'argent –, mais où l'objectif est de travailler moins (ou autrement) pour vivre mieux.

La société écologique peut trouver, dans la robotisation et l'automatisation en cours, les possibilités pour développer le type d'activités dont elle a besoin : des activités bénévoles et échappant à la marchandisation. Mais, si le revenu universel peut ainsi servir de médiation entre la robotisation et la transition écologique, les deux mondes, à la fois techniques, économiques et sociaux qu'il met ainsi en contact sont-ils vraiment compatibles ? La société écologique, explique-t-on, donne toute sa place au *care*, alors que la robotisation mécanise les activités fonctionnellement équivalentes au *care*, les rendant par là marchandes. Une société écologique se veut plus sobre, comptant plus sur les relations sociales que sur la course à la consommation pour s'affirmer et trouver son identité. N'est-ce pas au contraire la course à la consommation distinctive qu'engage une économie de la connaissance fondée sur les robots et l'intelligence artificielle ? Le rapport qui s'établit, par le biais du revenu universel, entre robotisation ou automatisation et société écologique, n'est qu'un rapport indirect. Cela ne répond pas de façon satisfaisante à la question de la place des robots dans la transition écologique, cela montre simplement que les conséquences n'en sont pas inévitablement négatives, si l'on sait s'y prendre.

Les robots ne sont pas de simples dispositifs techniques, adaptables à loisir à toutes les formes sociales. Ils arrivent portés par des structures sociales et économiques (capitalisme cognitif, nouvelles formes de mécanisation) qu'ils contribuent à renforcer. Symétriquement, la transition écologique est un projet qui a une extension variable. On peut s'en tenir au minimum, qui est celui de la transition énergétique : l'objectif est de substituer des énergies renouvelables aux énergies fossiles. Il s'agit là, comme le dit Gael Giraud, du noyau ou du cœur : « la transition écologique, c'est le processus par lequel nos sociétés pourraient évoluer d'une organisation économique centrée essentiellement sur la consommation d'énergies fossiles et dont l'un des sous produits

est une émission massive de gaz à effet de serre vers une économie de moins en moins énergivore et polluante» (Giraud, 2016, p. 87). Mais on peut aussi montrer que ce minimum ne suffit pas et qu'il faut viser le développement maximum, celui de nouvelles formes d'organisation économique (détournant les investissements financiers du carbone), accompagnées par une modification des rapports sociaux, portant sur les modes et les formes de vie (une société sobre, plus conviviale). La transition écologique renvoie ainsi à un véritable projet de société qui donne son sens à l'écologie politique. Or ce projet ne vise pas seulement à une modification des rapports sociaux et économiques, il implique aussi que l'on adopte une vision systémique, qui situe la société dans son environnement physique (biogéochimique).

Suivant l'extension que l'on accorde à la transition écologique, il faut faire appel à des critères différents pour évaluer la place que peuvent ou non avoir les robots dans la transition écologique. À l'étude de leur insertion énergétique et physique, nous ferons donc succéder celle des modèles sociaux qui leur sont associés, avant d'en venir à «l'écologie» des robots, dans un monde social ainsi redéfini.

2. QUELLE PLACE POUR LES ROBOTS DANS LE PROJET ÉCOLOGIQUE?

Dans le discours de la promesse qui soutient aujourd'hui l'innovation technologique, les avantages environnementaux occupent une place importante. Qu'il s'agisse de biotechnologies (organismes génétiquement modifiés, biologie de synthèse), de nanotechnologies ou de robots, c'est toujours la même chose: tous peuvent être mis au service de l'environnement, soit qu'ils permettent de diminuer la consommation énergétique, soit qu'ils aident à réparer l'environnement dévasté. À la façon dont on vante les mérites des robots tueurs pour remplacer à la guerre les soldats humains, on explique comment, en cas d'accident nucléaire, on pourrait envoyer, à la place des pompiers sacrifiés de Tchernobyl, des robots décontaminateurs: à supposer qu'ils soient capables de le faire, ce qui ne va pas de soi, les robots pouvant se montrer particulièrement maladroits pour effectuer des gestes qui ne posent pas de problème particulier aux humains.

À côté de ces interventions ponctuelles, on affirme pouvoir attendre du développement de l'économie de la connaissance, et de celui de l'intelligence artificielle et du numérique qui l'accompagne, des bénéfices plus

globaux, ceux qu'apportera une économie que l'on dit « immatérielle », assurant le découplage tant recherché entre la consommation énergétique et la croissance économique : on pourrait produire plus – en valeur – en détruisant moins d'énergie. Un rapport de l'OCDE de 2009 promet ainsi que la bioéconomie assurera une croissance verte. Le modèle en repose sur le principe simple selon lequel les organismes vivants représentent une source d'énergie renouvelable dont on peut économiquement tirer profit (en intégrant le savoir sur les gènes, sur la biomasse renouvelable, et les savoirs biotechnologiques) pour échapper à l'entropie et à ses limites (OCDE, 2009).

Mais cette croyance persistante dans l'idée que, grâce au développement technique, les hommes peuvent modifier leur situation écologique sans avoir besoin de toucher à leurs modes de vie et sans modifier leurs rapports à la nature (mais au contraire en accentuant l'emprise qu'ils ont sur elle) est mise en cause par des critiques comme Philippe Bihouix. Pour celui-ci, l'opposition pertinente n'est pas celle de l'immatériel et du matériel, mais celle du *low tech* et du *high tech*. C'est un leurre de croire que le numérique est virtuel et que l'intelligence artificielle enserme l'ensemble du globe de réseaux immatériels. Cette intelligence a un coût en énergie : les ordinateurs sont gros consommateurs, et les *big data* tout autant. Tout cela doit d'abord être fabriqué et fait appel à beaucoup de métaux rares, qu'il est à peu près impossible de recycler, la fabrication entraînant une fragmentation qui ne peut pas être surmontée (Bihouix, 2014, p. 222-234).

Si l'on entend souvent vanter l'économie d'énergie que réaliseront les robots par rapport aux activités qu'ils remplaceront, il est plus rare de s'interroger sur la consommation énergétique des robots. « Elle est pourtant considérable », commente Sébastien Briot, chargé de recherches CNRS au LS2N⁶. Depuis la crise de 2008, l'industrie européenne s'est lancée dans une course à la robotisation pour rester compétitive. Mais malgré la COP21 et le changement climatique, personne ne se pose vraiment la question de l'impact des robots sur l'environnement. Or, la fabrication du stock mondial de robots industriels correspond à 11 millions de tonnes en équivalent CO2 rejetées dans l'atmosphère. Mais comment diminuer cette dépense ? En changeant de matériaux (on propose ainsi de fabriquer les robots

6. Laboratoire des sciences du numérique de Nantes, issu de la fusion de l'IRCCyN (Institut de recherche en communications et cybernétique de Nantes) et du LINA (Laboratoire d'informatique de Nantes atlantique) en janvier 2017.

avec du bois) ou en apprenant à moins compter sur les machines? La première solution est un raffinement du *high tech*, la deuxième passe au *low tech*.

Mais, à ce niveau, les robots sont envisagés individuellement, comme des êtres isolés sans que l'on mette en question leur insertion dans des ensembles sociaux et matériels. Or, le point de vue de la transition écologique est un point de vue intégré qui tient compte non seulement de la production, mais aussi de la reproduction. C'est ce qui fait la différence entre une énergie fossile, détruite dans sa consommation, et une énergie renouvelable, qui peut être reproduite. Envisagés de ce point de vue, les robots posent cependant problème. Ils ne se reproduisent pas d'eux-mêmes, et c'est en cela, entre autres, qu'ils ne sont pas des êtres vivants. Mais ils ont besoin d'être rechargés, et cela ne se fait pas sans difficulté (même métaphoriquement, ils ne sont guère autopoïétiques). Bien loin donc d'être indépendants, les robots sont étroitement dépendants du réseau fournisseur d'énergie dans lequel ils s'insèrent: ce milieu technique dont ils sont inséparables tend à couper le réseau social des robots de leur environnement physique. Cela accroît leur vulnérabilité. Les robots ne semblent pas des êtres écologiques prêts à partir à l'aventure dans la nature, sans protection arrière. Ils manquent singulièrement de résilience.

Fort consommateurs d'énergie, incapables de se reproduire, prisonniers d'un réseau technique dont ils dépendent pour exister et qui les fragilise, les robots ne paraissent pas tellement souhaités dans une société écologique sobre, résistante et résiliente. En examinant leurs finalités, et pas seulement leurs conséquences, peut-on mieux découvrir le milieu social susceptible de les intégrer?

On accorde beaucoup d'importance, de ce point de vue, à l'agriculture, qui est présentée comme le second marché de la robotique de service à l'horizon 2025 (Lenain, 2017). L'appel à la robotique est censé rendre compatibles un certain nombre de contraintes: réduire la pénibilité du travail (tâches mécaniques et répétitives), réduire l'impact environnemental des activités agricoles tout en maintenant (ou augmentant) la production: c'est ce que des robots coopérant entre eux devraient permettre de faire. Et sans doute des robots décontaminateurs (à la façon des robots tueurs ou pompiers nucléaires) pourraient continuer à épandre intensivement des pesticides ou des phytosanitaires dont l'emploi est extrêmement nocif pour les agriculteurs qui les projettent ou les déversent. Ces tâches sont si dangereuses que, pour

les accomplir, les agriculteurs se masquent et se déguisent en robots. N'est-il pas préférable que de vrais robots fassent le travail à leur place sans mettre leur santé en danger ? Mais faut-il se faire remplacer par des robots ou développer plutôt une agriculture qui n'ait pas besoin de tels pesticides ? Faut-il confier à des robots des tâches devenues tellement mécaniques qu'elles n'ont plus aucun intérêt humain, ou se tourner vers des formes d'agriculture (agroécologie ou permaculture) qui, plus attentives au sol, vu comme un milieu et non comme un support, fassent appel à l'initiative et à l'inventivité de l'agriculteur ?

Sans doute les robots peuvent-ils être utiles à l'agriculture, mais ce n'est pas n'importe quelle agriculture qui a besoin de robots. C'est une agriculture qui repose sur un contrôle très fin des différentes étapes de production, décomposables et mesurables. Cela demande une surveillance constante, une grande précision technique pour faire le traitement qu'il faut, avec la quantité qu'il faut, au moment où il faut : c'est ce que l'on appelle l'agriculture de précision. L'objectif est d'intégrer la durabilité à l'objectif productiviste. Ou, pour le dire autrement, il s'agit de polluer moins pour polluer plus longtemps. Cela ne peut se faire qu'avec une agriculture à grande échelle et à gros capitaux (en France, 3 316 exploitations laitières sur 514 800 sont équipées de robots d'affouragement ou de traite, ce ne sont pas les plus pauvres). Il y a d'autres agricultures possibles, plus conformes à un projet de société écologique, non seulement parce qu'elles sont moins consommatrices en énergie, mais surtout parce qu'elles combinent des formes sociales plus solidaires et un rapport à la nature moins dominateur.

3. L'ÉCOLOGIE DES ROBOTS

S'il apparaît ainsi que la transition écologique peut tirer indirectement avantage, grâce au revenu universel, de l'automatisation croissante et des pertes d'emplois qu'elle entraîne, mais s'il s'avère également que la société écologique n'a pas tout à attendre des robots et qu'elle ferait mieux plutôt de se méfier tant de leur empreinte écologique que des formes sociales qui les accompagnent, on peut, pour terminer, reformuler la question. Il ne s'agira plus alors de savoir dans quel type de société les robots peuvent trouver une place, mais quelle place une société écologique peut-elle offrir aux robots ? Sans s'en rendre complètement dépendants, il n'est pas non plus question d'éviter complètement d'y avoir recours, et, dans ce cas, quelle attitude faudrait-il adopter à leur égard ? Comme le dit Dominique Lestel, pour donner une place

aux robots, nous devons leur « offrir une écologie dans laquelle ils peuvent vivre » (Lestel, 2015, p. 327). Sans doute se demande-t-on quels comportements nous allons exiger des robots (et ce, d'autant plus qu'avec l'arrivée des robots affectifs, ils entrent dans le domaine des comportements considérés comme moraux), mais la question est aussi de savoir comment nous devons nous comporter avec les robots que nous avons installés parmi nous.

En éthique animale comme en éthique environnementale, la question s'est posée de l'extension aux non humains de la communauté morale, jusque-là réservée aux seuls humains, et l'on a recherché les qualités qui permettent de se voir attribuer une place dans la communauté. Pour les animaux, la qualité d'être sensible est le plus souvent considérée, depuis Bentham (1970 [1789], p. 283), comme celle qui marque le seuil de l'attention morale (Nussbaum, 2007, p. 361); pour le vivant en général, en éthique environnementale, il s'agit de la capacité à se maintenir en vie, ce que l'on a nommé la valeur intrinsèque (Jamieson, 2008). Dans tous les cas, c'est cette qualité partagée par tous qui détermine l'appartenance à la communauté morale ainsi élargie à de nouveaux arrivants. Il n'en reste pas moins que l'exploration de cette communauté, animale ou plus généralement naturelle, conduit à prendre ses distances par rapport aux hommes et aux qualités qui n'appartiennent qu'à eux. L'écologie morale des mondes animaux ou naturels est ainsi l'exploration, tant descriptive que normative, de milieux dont les humains ne font pas toujours partie, ou, en tout cas, ne sont pas les personnages centraux, les seules références possibles. Une telle écologie morale est une expérience de diversité et de dépaysement, par rapport à l'anthropocentrisme usuel. Et n'est-ce pas là ce que nous attendons de l'écologie? Qu'elle nous apprenne à sortir de nous-mêmes, à découvrir des aspects qui nous échappent quand nous nous absorbons dans notre propre inspection?

Or ce n'est nullement le cas lorsqu'il s'agit des robots. L'écologie de la robotique est celle de la ressemblance à l'humain. On n'en retient généralement que la partie qui concerne les « créatures artificielles », et, parmi celles-ci, la partie encore plus petite qui est constituée des robots à forme humaine, ou « humanoïdes » (Grimaud, 2015, p. 49). Pour qu'un robot retienne notre attention, il suffit de le doter de deux yeux, mais encore faut-il que la ressemblance ne soit pas trop parfaite. Sous le nom de « vallée de l'étrange », le roboticien Masahiro Mori a représenté le graphique de la conjecture de nos rapports avec

les robots : nous interagissons d'autant plus facilement avec eux qu'ils nous ressemblent, mais quand la ressemblance devient trop grande, nous prenons nos distances, inquiétés par cette proximité «étrange» (*uncanny*) (Mori, 1970, dans Dumouchel et Damiano, 2016, p. 35).

Nous reconnaissons si peu aux robots un être qui leur serait propre que la seule possibilité que nous leur laissons est celle de se faire passer pour ce qu'ils ne sont pas, mais que nous les poussons à être ou à prétendre être : nous-mêmes. Ce sont des imposteurs par destination. Cela a commencé avec Descartes quand celui-ci, dans le *Discours de la méthode*, évoque la possibilité qu'une machine (fabriquée par Dieu) ressemble si fidèlement à un animal que l'on serait incapable de faire la différence (Descartes, 1958 [1637], p. 164). Encore l'expérience de pensée de Descartes avait-elle pour but de distinguer entre le corps et l'esprit, l'animal et l'homme. L'homme échappait à la confusion possible entre l'automate et le vivant. Mais avec la mécanisation de l'esprit, ou l'artificialisation de l'intelligence, l'expérience a simplement été repoussée d'un cran, englobant l'homme. Turing ne se demande pas si les machines pensent, mais « dans quelles conditions une machine peut-elle nous tromper sur le fait qu'elle pense ? » et à nouveau, on est mis au défi de produire l'expérience où il ne sera pas possible de faire la différence, où nous ne saurons pas si nous avons affaire à une machine ou à une intelligence humaine (Turing, 1950).

La prégnance du paradigme illusionniste en robotique vient sans doute de la fonction que nous lui assignons : nous fabriquons des robots pour qu'ils accomplissent des tâches dont nous voulons nous décharger. Ils doivent donc nous ressembler le plus possible, mais, ce faisant, ils menacent de prendre notre place. Pour qu'ils soient efficaces dans les tâches mécaniques que nous leur confions, nous voulons en faire de purs automates, mais la contingence faisant partie du monde, il faut que ces automates aient une marge d'indétermination, une possibilité de décision, une forme de liberté, et les voilà autonomes, aptes à nous supplanter ou à nous faire du mal. Ce passage de l'automatisme à l'autonomie se donne à lire dans la remarque fameuse d'Aristote sur les navettes :

Si donc il était possible à chaque instrument parce qu'il en aurait reçu l'ordre ou par simple pressentiment de mener à bien son œuvre propre, comme on le dit des statues de Dédale ou des trépieds d'Héphaïstos qui, selon le poète, entraient d'eux-mêmes dans l'assemblée des dieux, si, de même, les navettes tissaient

d'elles-mêmes et les plectres jouaient tout seuls de la cithare, alors les ingénieurs n'auraient pas besoin d'exécutants ni les maîtres d'esclaves. (Aristote, I, 4, 1253 b 35, 1990, p. 97)

La logique du substitut est celle à la fois de la continuité et de la discontinuité : la séparation de l'animé (esclave) et de l'inanimé (navette) est annulée dans l'instrumentalisation (pour Aristote, l'esclave est un instrument animé), et la hiérarchie du maître et de l'esclave n'a plus lieu d'être lorsque l'esclave est libéré de sa tâche mécanique : on passe ainsi sans transition de l'objet à l'homme, et de l'esclave au maître. Une fois lancée, la continuité de la substitution peut absorber le degré le plus élevé de l'humain. Il n'y a pas de limites et la continuité peut se lire dans les deux sens : celle d'une promotion de l'activité mécanique qui peut remplacer l'homme, comme celle d'une dévalorisation des actions humaines qui peuvent être remplacées par des gestes mécaniques. Menacés d'être remplacés par des choses, les hommes se trouvent aussi dévalorisés de ce qu'ils ne valent pas mieux qu'une chose : ils risquent de perdre et leur emploi et leur estime d'eux-mêmes. Tant que la tâche instrumentale reste dépréciée, la libération de l'esclave est précaire. C'est ce que fait remarquer Simondon quand, dans le passage d'Aristote, il refuse de ne retenir que la promesse de libération de l'esclave, et pose la question des navettes : sont-elles les sacrifiées dans l'affaire ? Il n'est peut-être pas mieux de caractériser les robots comme des esclaves (c'est le sens du mot robot en tchèque) que de réduire l'esclave à un instrument.

Une société qui se contente de déléguer les tâches pénibles dont elle ne veut pas se charger à des instruments passifs baptisés, métaphoriquement ou non, esclaves, se prépare à subir ce qu'elle essaie de faire subir à d'autres. On peut lui appliquer la remarque de Rousseau au début du *Contrat social*, « Tel se croit le maître des autres qui ne laisse pas d'être plus esclave qu'eux » (1966, p. 351). La question n'est pas tant de transférer aux artefacts ce qui était imposé aux humains, que de redéfinir nos rapports à nos instruments, de requalifier l'écologie des robots, le monde social que nous partageons avec eux. Comme le remarque Simondon, « [l]oin d'être le surveillant d'une troupe d'esclaves, l'homme est l'organisateur permanent d'une société des objets techniques qui ont besoin de lui comme les musiciens ont besoin du chef d'orchestre » (Simondon, 1989, p. 11).

Une grande partie des inquiétudes, mais aussi des incompréhensions, que provoquent en nous les robots vient de ce que nous les enfermons dans l'écologie de la ressemblance. La complexification

de l'automatisme en autonomie est tout à fait normale. Comme le fait remarquer Emmanuel Grimaud, «il suffit de peu de choses pour qu'un effet d'autonomie ou d'intelligence se produise dans un artefact» (Grimaud, 2015, p. 49). Pour peu que l'on s'en tienne à une définition technique de l'autonomie, comme une possibilité de jeu à l'intérieur d'un ensemble de règles prédéfinies, ce qui permet qu'adviennent des actes qui n'ont pas été imposés de l'extérieur, il n'y a pas là de quoi s'inquiéter. Mais la fascination pour la ressemblance, la projection anthropomorphique qui caractérise le rapport aux robots, conduit à passer à une autre conception de l'autonomie qui est celle de la capacité à se donner des règles (Ganascia, 2017, p. 43-54). Ce qui fait non seulement de l'humanoïde un rival menaçant, un esclave susceptible de se transformer en maître, mais surtout conduit à voir dans l'autonomie, comme parfaite indépendance, l'étape ultime de la robotique, alors que toute étude de la robotique montre à quel point les robots sont dépendants de leur milieu associé, si bien que leur perfectionnement ne se trouve pas dans une indépendance croissante, mais dans la multiplication de leurs relations: si autonomie il y a, elle ne peut être que relationnelle, elle n'est pas celle qui place celui qui en jouit en position éminente de chef qui s'impose à des serveurs, mais celle qui multiplie des liens entre différents dispositifs de coopération.

Plutôt que de parler de robot (ce mot qui désigne un esclave mais révèle la menace d'un maître), deux Américains travaillant pour l'industrie automobile ont inventé, en 1996, le terme de «cobots», pour robotique collaborative (Kleinpeter 2015). Il s'agissait de concevoir des machines asservies ou pseudoautonomes susceptibles d'être utilisées au sein d'un environnement humain. Tant que l'on a comme unique référence l'autonomie, des dispositifs d'assistance qui requièrent la présence d'un conducteur ne seront pas considérés comme des robots: la seule alternative est de commander ou d'obéir. Mais si l'on s'intéresse à la coopération, c'est alors la diversité et la complémentarité des tâches qui sont importantes. L'importance d'un robot se mesurera, non à sa capacité à me remplacer, mais à celle de faire des choses que je ne peux pas forcément faire mais qui me seront utiles. Le robot n'apparaît plus comme un concurrent, dont les succès dévalorisent les tâches humaines qu'il remplace, mais comme un facteur de diversité enrichissante. Le narcissisme humain y gagne, et le robot y trouve une possibilité de se définir autrement que comme un imposteur.

L'écologie du robot, finalement, comme l'écologie des animaux ou de la nature, c'est celle de la bonne distance. Mais les risques de mauvaise distance sont, dans l'un et l'autre cas, opposés. Là où, pour la nature, le risque est de ne la placer que dans l'extériorité radicale du sauvage (la nature n'est vraiment la nature que quand nous nous en absentes, posent ceux qui assimilent le naturel et le sauvage), le risque, pour l'écologie du robot, est inverse. Nous sommes absorbés par la ressemblance. La robotique est une expérience tellement anthropocentrique – et anthropomorphique – que nous en sommes conduits à ne plus pouvoir nous distinguer des robots, pour le pire comme pour le meilleur. Mais pourquoi, se demande avec raison Emmanuel Grimaud, une relation épanouie avec une machine devrait-elle nécessairement passer par un moment de confusion ontologique avec l'humain? (Grimaud, 2015, p. 54). Ni animaux ni humains, des robots: une relation à étudier. Une chance à leur donner.

Quelle place pour les robots dans la transition écologique? Celle que nous arriverons à leur découvrir quand nous cesserons de les considérer comme d'inquiétants doubles de nous-mêmes, quand nous accepterons d'en étudier les différents aspects au lieu de les ranger *a priori* dans les catégories préétablies d'un imaginaire mythifiant.

BIBLIOGRAPHIE

- Aristote, (1990). *Les Politiques*, traduction Pellegrin P., Paris, GF Flammarion.
- Audétat M., Kaufman A. (2015). *Sciences et technologies émergentes, pourquoi tant de promesses?* Paris, Hermann.
- Augagneur, Florian (2017). « L'allocation universelle est-elle juste? », *Projet, Inégalités, un défi écologique?* n° 356, p. 50-58.
- Bentham, Jeremy (1970). *An Introduction to the Principles of Morals and Legislation* (1789), chapitre XVII, § 4 note b, Oxford, Oxford University Press.
- Bihoux, Philippe (2014). *L'Âge des Low Tech*, Paris, Éditions du Seuil.
- Bihoux Philippe (2016). Comment entrer dans le règne des Low-Tech », dans D. Bourg, A. Kaufmann et D. Méda (dir.), *L'Âge de la transition*, p. 158-163.
- Bourg, D., Kaufmann, A., et Méda, D. (dir.) (2016). *L'Âge de la transition. En route pour la reconversion écologique*, Paris, Les Petits matins/ Institut Veblen.
- Descartes, René (1637). *Discours de la méthode*, V^e partie, in Descartes, *Œuvres et Lettres*, Paris, Gallimard, bibliothèque de la Pléiade, 1958, p. 164.
- Dourgnon, Julien (2017). *Revenu universel. Pourquoi? Comment?* Paris, Les petits matins/Institut Veblen.
- Dumouchel, Paul, Damiano, Luisa (2016). *Vivre avec les robots, Essai sur l'empathie artificielle*, Paris, Éditions du Seuil.

- Ganascia, Jean-Gabriel (2017). *Le mythe de la singularité, Faut-il craindre l'intelligence artificielle?* Paris, Éditions du Seuil.
- Gianinazzi, Willy (2016). *André Gorz, Une vie*, Paris, La Découverte.
- Giraud, Gaël (2016). *Illusions financières, Des subprimes à la transition écologique*, 3^e édition revue et augmentée, Paris, Les éditions de l'atelier.
- Gorz, André (1985). « Allocation universelle : version de droite et version de gauche », *La Revue nouvelle*, n° 81, p. 419-428.
- Gorz, André (1997). *Misères du présent, richesse du possible*, Paris, Galilée.
- Jamieson, Dale (2008). *Ethics and the Environment, An Introduction*, Cambridge, Cambridge University Press.
- Kleinpeter, Edouard (2015). « Le Cobot, la coopération entre l'utilisateur et la machine », *Multitudes*, 58, p. 70-75.
- Grimaud, Emmanuel (2015). « Les robots oscillent entre vivants et inertes », *Multitudes*, 58.
- Joly, Pierre-Benoît (2016). « Verrouillages socio-techniques et transitions écologiques », in *L'Âge de la transition*, p. 117-125.
- Lenain, Roland (2017). « Quel avenir pour la robotique agricole ? » *The Conversation*, 15/3/2017.
- Lestel, Dominique (2015). Would you like to be alive with me? dans Sandra Laugier (dir.), *Nodi / Symposia Nuova robotica, nuovi viventi / New Robotics, New Living Beings*.
- Meadows, Donella, Meadows, Dennis, Randers, Jorgen, Behrens, William W. III (1972). *The Limits to Growth, A Report to the Club of Rome*, New York, Universe Books; *Halte à la croissance?*, trad. fr. Paris, Fayard.
- Mori Masahiro, (1970). « La vallée de l'étrange », *Energy* 7 (4), p. 33-35.
- Moulier-Boutang, Yann (2015). « L'automatisation intellectuelle, la mort de l'emploi et le revenu de pollinisation », in *Multitudes*, n° 58, p. 17-26.
- Nussbaum, Martha (2007). « Beyond « Compassion » and « Humanity » », *Frontiers of justice*, Cambridge, London, Harvard University Press.
- OCDE (2009). *La Bioéconomie à l'horizon 2030. Quel programme d'action?*
- Perret, Bernard (2015). *Au-delà du marché. Les nouvelles voies de la démarchandisation*, Paris, Les Petits Matins.
- Rousseau, Jean-Jacques (1762). *Du Contrat social*, in 1964, *Œuvres complètes*, édition publiée sous la direction de Bernard Gagnebin, et Marcel Raymond, tome III, *Œuvres politiques*, Paris, Gallimard.
- Simondon, Gilbert (1989). *Du mode d'existence des objets techniques*, Paris, Aubier.
- Swaton, Sophie (2018). *Pour un revenu de transition écologique*, Paris, PUF.
- Turing, Alan (1950). « Computing Machinery and Intelligence », *Mind*, 36, p. 433-460.

Déjà parus :

Marie-Hélène Parizeau et Soheil Kash (dir.), *De l'inégalité dans le dialogue des cultures: mondialisation, santé et environnement* (2005).

Jean-Paul Amann, Catherine Chiron, Olivier Dulac et Anne Fagot-Largeault (dir.), *Épilepsie, connaissance du cerveau et société* (2006).

Marie-Hélène Parizeau et Soheil Kash (dir.), *Néoracisme et dérives génétiques* (2006).

Marie-Hélène Parizeau et Georges Chapouthier (dir.), *L'être humain, l'animal et la technique* (2007).

Marie-Luce Delfosse, Marie-Hélène Parizeau, Jean-Paul Amann (dir.), *La recherche clinique avec les enfants : à la croisée de l'éthique et du droit. Belgique, France, Québec* (2009).

Bernard Reber, *La démocratie génétiquement modifiée. Sociologies éthiques de l'évaluation des technologies controversées* (2011).

Jean-Philippe Pierron et Marie-Hélène Parizeau (dir.), *Repenser la nature. Dialogue philosophique, Europe, Asie, Amériques* (2012).

Catherine Baudoin, *Plantes transgéniques: quelle évaluation éthique?* (2014).

Marie-Hélène Parizeau et Jacynthe Tremblay (dir.), *Milieus modernes et reflets japonais. Chemins philosophiques* (2015).

Marie-Hélène Parizeau et Soheil Kash (dir.), *À chacun son développement durable? De la diversité culturelle aux nanotechnologies* (2017).

PETITE COLLECTION « BIOÉTHIQUE CRITIQUE »

Margarita Boladeras, *Parlons bioéthique. Entretiens avec A. Fagot-Largeault, J-Y Goffi, G. Hottois, J-N Missa, M-H Parizeau* (2017).

Mobiles, interactifs, capables de communiquer, les robots peuvent-ils pour autant « penser » ou prendre des décisions à la place des humains ? Faut-il les considérer comme des agents moraux ayant une « autonomie » ou leur donner un statut juridique particulier ? Qui est alors responsable de leurs actions – le concepteur informaticien, le fabricant industriel, l'utilisateur consommateur ? Pour quelles finalités tous ces robots sont-ils conçus ?

Depuis peu, le public a découvert comment des robots pouvaient remplacer l'être humain dans un nombre croissant d'activités économiques, sociales et politiques. Les robots-drones sont utilisés dans les conflits armés ou encore dans des contextes non armés pour la surveillance ou l'assassinat ciblé. Des robots aux formes androïdes ont fait leur apparition dans le domaine de la santé et du bien-être. Dans les hôpitaux, des robots opèrent sous la direction du chirurgien, d'autres robots aux formes animales deviennent des « compagnons » pour les personnes âgées. Dans des écoles, des robots sont utilisés par des enseignants pour l'apprentissage des langues ou des matières scientifiques. Dans des maisons, des robots de service aspirent la poussière des tapis tandis que des jouets-robots s'occupent des enfants. Dans le domaine de l'agriculture, des robots traient les vaches et nettoient l'étable. Dans les usines, les robots accélèrent la productivité et l'efficacité de la production industrielle.

Tous ces robots qui remplacent les êtres humains dans des tâches devenues « robotisables » font-ils de notre société « une société robotisée » ? Tant par la réflexion théorique qu'à l'aide d'exemples précis, cet ouvrage multidisciplinaire examine comment les robots modifient la qualité de nos relations humaines, en quoi ils transforment certaines valeurs fondamentales comme la liberté et l'égalité, ou encore de quelle façon ils entraînent des changements sociaux et culturels, par exemple dans nos relations aux animaux ou à l'environnement.

Avec la participation de :

Jean-Michel BESNIER, Vanessa NUROCK, Tomislav BRACANOVIĆ,
Charles-Étienne DANIEL, Hoda NEHMÉ, Frédéric DUBOIS, Alexandre PITTI,
Guillaume PELLETIER, Chantale PILON, John FINNEY, Soheil KASH,
Josée-Anne GAGNON, Félix-Antoine PAGEAU, Monique LANOIX,
Raoul KAMGA, Margarita ROMERO, Marie-Hélène PARIZEAU,
Sophie CLOUTIER, Sandra BLOUIN, Georges Chapouthier,
Louis-Étienne PIGEON, Catherine LARRÈRE.



Aussi en version numérique

COLLECTION
bioéthique
critique

www.pulaval.com

ISBN 978-2-7637-3993-9

9 782763 739939

**Presses de
l'Université
Laval**

Éthique