

RAPPORT DE LA COMEST SUR L'ÉTHIQUE DE LA ROBOTIQUE

Dans le cadre de son programme de travail pour 2016-2017, la COMEST a décidé d'aborder la question de l'éthique de la robotique, en s'appuyant sur sa réflexion antérieure au sujet des problèmes éthiques soulevés par la robotique moderne, ainsi que de l'éthique des nanotechnologies et des technologies convergentes.

Lors de la 9^e session (ordinaire) de la COMEST en septembre 2015, la Commission a créé un Groupe de travail pour initier la réflexion sur cette question. Le Groupe de travail de la COMEST s'est réuni à Paris en mai 2016 pour définir le plan et le contenu d'un avant-projet de rapport, qui a été discuté pendant la 9^e session extraordinaire de la COMEST en septembre 2016. Au cours de cette session, le contenu de l'avant-projet de rapport a encore été développé et précisé, et le Groupe de travail a poursuivi son travail par courrier électronique. Le Groupe de travail de la COMEST s'est ensuite réuni à Québec en mars 2017 pour développer plus avant le texte. Une version révisée sous la forme d'un projet de rapport a été soumise à la COMEST et au CIB en juin 2017 pour commentaires. Le projet de rapport final a ensuite été révisé en tenant compte des commentaires reçus. Ce projet final du rapport a été examiné et révisé lors de la 10^e session (ordinaire) de la COMEST, et il a été adopté par la Commission le 14 septembre 2017.

Ce document ne prétend pas à l'exhaustivité et ne représente pas nécessairement les vues des États membres de l'UNESCO.

RAPPORT DE LA COMEST SUR L'ÉTHIQUE DE LA ROBOTIQUE

RÉSUMÉ ANALYTIQUE

I. INTRODUCTION

II. QU'EST-CE QU'UN ROBOT ?

- II.1. Problèmes de définition
- II.2. Histoire de la robotique : fiction, imaginaire et réel
- II.3. Autonomie, interactivité, communication et mobilité
- II.4. Nanorobotique
- II.5. Robots, intelligence artificielle et algorithmes

III. LES ROBOTS ET LA SOCIÉTÉ

- III.1. Utilisation de robots dans l'industrie
- III.2. Usages militaires et civils des systèmes robotiques mobiles
 - III.2.1. Systèmes robotiques militaires (drones)
 - III.2.2. Armes autonomes
 - III.2.3. Robots de surveillance et de police et utilisation de technologies militaires dans des contextes non militaires
 - III.2.4. Utilisation privée et illicite de robots
- III.3. Utilisation de robots dans les transports
- III.4. Utilisation de robots dans le domaine de la santé et du bien-être
 - III.4.1. Robots médicaux
 - III.4.2. Robots infirmiers
 - III.4.3. Robots de soins pour personnes âgées
 - III.4.4. Robots compagnons
- III.5. Utilisation de robots dans le domaine de l'éducation
- III.6. Robots ménagers
- III.7. Agriculture et environnement

IV. RÉGLEMENTATION ÉTHIQUE ET LÉGALE

V. ENJEUX ÉTHIQUES

- V.1. Techno-pessimisme, techno-optimisme et au-delà
- V.2. Robots et responsabilité
- V.3. Une « capacité d'agir » non humaine
 - V.3.1. Les robots en tant qu'agents autonomes

- V.3.2. Les robots en tant qu'agents moraux
- V.4. Statut moral des robots
- V.5. Dynamisme des valeurs

VI. RECOMMANDATIONS

- VI.1. Un cadre éthique fondé sur la technologie
- VI.2. Valeurs et principes éthiques pertinents
 - VI.2.1. Dignité humaine
 - VI.2.2. Valeur de l'autonomie
 - VI.2.3. Respect de la vie privée
 - VI.2.4. Principe d'innocuité
 - VI.2.5. Principe de responsabilité
 - VI.2.6. Valeur de la bienfaisance
 - VI.2.7. Valeur de la justice
- VI.3. Recommandations spécifiques de la COMEST sur l'éthique de la robotique
 - VI.3.1. Recommandation sur le développement de codes d'éthique pour la robotique et les roboticiens
 - VI.3.2. Recommandation sur la conception éthique des robots
 - VI.3.3. Recommandation sur l'expérimentation
 - VI.3.4. Recommandation sur la discussion publique
 - VI.3.5. Recommandation sur la reconversion et la réadaptation de la force de travail
 - VI.3.6. Recommandations relatives aux transports et aux véhicules autonomes
 - VI.3.7. Recommandations sur les Systemes robotiques militaires armés (« drones armés »)
 - VI.3.8. Recommandations sur les armes autonomes
 - VI.3.9. Recommandations concernant la surveillance et le travail de police
 - VI.3.10. Recommandation concernant l'utilisation privée et commerciale de drones
 - VI.3.11. Recommandation sur l'égalité de genre
 - VI.3.12. Recommandation sur l'évalutaion de l'impact sur l'environnement
 - VI.3.13. Recommandations sur l'internet des objets

BIBLIOGRAPHIE

RAPPORT DE LA COMEST SUR L'ÉTHIQUE DE LA ROBOTIQUE

RÉSUMÉ ANALYTIQUE

I. INTRODUCTION

Les robots peuvent aider l'humanité, et ils le font depuis la moitié du XX^e siècle. Initialement utilisés principalement à des fins industrielles ou militaires, ils apparaissent aujourd'hui dans d'autres domaines comme celui des transports, de la santé et de l'éducation, ainsi que dans le contexte ménager. La robotique contemporaine repose de plus en plus sur les technologies d'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. La caractéristique principale de ces « machines cognitives » est le fait que leurs décisions sont imprévisibles et que leurs actions dépendent de l'expérience et de certaines conditions stochastiques. La question de la responsabilité des actions exécutées par des robots cognitifs est donc cruciale.

La présence de robots cognitifs au sein de la société, qui s'accroît rapidement, suscite de nouveaux défis. En effet, ces robots ont un impact sur les comportements humains, induisent certains changements sociaux et culturels, et soulèvent des questions en matière de sécurité, de respect de la vie privée et de protection de la dignité humaine. Le présent rapport poursuit un but de sensibilisation et cherche à promouvoir à la fois un examen public et un dialogue inclusif sur les questions éthiques concernant les différentes formes d'utilisation de robots cognitifs autonomes dans la société.

II. QU'EST-CE QU'UN ROBOT ?

Les robots contemporains se caractérisent par quatre attributs essentiels :

- la *mobilité*, qui constitue un atout important dans les environnements humains tels que les hôpitaux et les administrations ;
- l'*interactivité*, rendue possible par des capteurs et des actionneurs recueillant l'information pertinente de l'environnement, qui permet aux robots d'agir sur cet environnement ;
- la *communication*, rendue possible par des interfaces informatiques ou des systèmes de reconnaissance vocale ou de synthèse vocale ; et
- l'*autonomie*, c'est-à-dire la capacité à « penser » par eux-mêmes et à prendre leurs propres décisions pour agir sur l'environnement sans contrôle extérieur direct.

La robotique contemporaine intègre généralement diverses formes d'intelligence artificielle (IA) permettant de reproduire la cognition et l'intelligence humaines au moyen de systèmes informatiques et aboutissant ainsi à des machines qui peuvent exécuter des tâches requérant une forme spécifique d'intelligence, par exemple la capacité à percevoir et représenter formellement les modifications de leur environnement et à adapter leur fonctionnement en conséquence. L'intelligence artificielle est déterminante pour l'autonomie des robots car elle leur permet d'exécuter des tâches complexes dans un environnement non structuré et qui change – par exemple conduire une voiture en s'adaptant aux conditions de la circulation – sans être télécommandés ou contrôlés par un opérateur humain.

Les robots exécutent leurs tâches au moyen d'algorithmes, c'est-à-dire de règles ou d'instructions en vue de la résolution d'un problème. On distingue généralement deux types d'algorithmes : les algorithmes déterministes, qui contrôlent le comportement prévisible de robots *déterministes*, et les algorithmes IA ou stochastiques, dotés de capacités d'apprentissage, qui sont au cœur du fonctionnement des robots *cognitifs*. Le comportement d'un robot déterministe – même s'il s'agit d'un robot très complexe, intelligent et autonome, c'est-à-dire requérant très peu ou aucune surveillance humaine – est essentiellement préprogrammé et déterminé. Les robots cognitifs reposant sur l'intelligence artificielle, en revanche, seront capables d'apprendre à partir de leurs expériences passées et de recalibrer

eux-mêmes leurs algorithmes ; par conséquent, leur comportement ne sera pas entièrement prévisible et ce point exigera certainement à l'avenir une sérieuse attention et réflexion éthique.

III. ROBOTS ET SOCIÉTÉ

Dans l'*industrie*, les robots remplacent progressivement le travail humain dans un grand nombre de professions du secteur des services, où la croissance de l'emploi s'est concentrée pendant les dernières décennies. En l'absence d'ajustements structurels et de mesures visant à compenser cette évolution, cela pourrait contribuer à l'augmentation du chômage et des inégalités au sein de la société. Les robots vont modifier en profondeur les conditions de travail et les emplois. Travailler côte à côte avec des robots exige de nouvelles compétences professionnelles, la mise en place de mesures de sécurité nouvelles et adéquates sur le lieu de travail, l'aménagement des heures de travail et la formation. La robotisation de l'industrie suscitera également de nouveaux défis économiques et politiques : risque-t-elle de créer un nouvel écart entre pays en développement et pays développés et, dans l'affirmative, comment faire face à cette situation ?

Les drones désignent toute une catégorie de *systèmes robotiques militaires*. Un drone peut être soit guidé à distance par des opérateurs humains (télépilotage), soit guidé de façon autonome par des moyens robotiques (pilotage automatique). En pratique aujourd'hui, les drones combinent ces deux types de contrôle. La distance entre le pilote et le champ d'action peut susciter chez le pilote un état d'esprit proche du jeu, isolant les actes de toute considération morale, mais la possibilité d'être présent à distance peut aussi contribuer à aiguïser le sens moral des soldats. La question de savoir si le droit international humanitaire (DIH), applicable aux situations de conflit armé et d'occupation, est suffisamment clair s'agissant des drones est source de préoccupations. Cette question est encore plus problématique dans le cas des systèmes d'armes autonomes en cours de développement.

Dans le domaine des *transports*, les véhicules autonomes (VA) sont sur le point de devenir une réalité. Ces véhicules pourraient réduire fortement le nombre d'accidents de circulation et améliorer l'efficacité des transports, tout en provoquant potentiellement des pertes d'emploi et en creusant le « fossé robotique ». La question essentielle sur le plan éthique concerne les processus décisionnels intégrés aux VA. Comment le véhicule devra-t-il être programmé à se comporter dans l'éventualité d'un accident impossible à éviter ? Devra-t-il chercher à réduire au minimum le nombre de victimes, même si cela signifie sacrifier ses occupants, ou bien devra-t-il protéger ces derniers à tout prix ? De telles questions devraient-elles être réglementées par la loi, des normes ou des codes de conduite ?

Dans le domaine de la *santé* et du *bien-être*, les robots sont de plus en plus utilisés en chirurgie, en permettant parfois une plus grande précision mais à un coût élevé, tout en modifiant la pratique chirurgicale. Les robots sont aussi utilisés dans les soins thérapeutiques et de réadaptation, les soins des personnes âgées (robots sociaux), le traitement des enfants autistes et, sous forme d'exosquelettes, le traitement des personnes victimes d'un traumatisme médullaire. Les questions éthiques qui se posent à cet égard concernent généralement le caractère adéquat de l'utilisation de ce type de technologies par rapport aux soins : des robots peuvent-ils dispenser des soins ? Quelles en sont les implications en termes de sûreté et de sécurité ? Comment influencent-ils notre attitude à l'égard de la santé et du handicap, et à l'égard des personnes malades, âgées et vulnérables ? Les robots compagnons dans le domaine de la sexualité constituent une sous-catégorie particulière : comment la possibilité d'avoir une relation sexuelle à distance par l'intermédiaire d'un robot, ou même d'avoir des rapports sexuels avec un robot, affectera-t-elle nos valeurs concernant l'amour et l'intimité ?

Les robots apparaissent aussi dans le domaine de l'*éducation*. Les robots éducatifs soutiennent les activités d'apprentissage seul et en collaboration. L'introduction de robots en classe aura sans doute des implications sur le mode d'apprentissage des enfants, sur le rôle

de l'enseignant en tant que modèle et sur le développement affectif des enfants. Dans le *contexte ménager*, les robots de service aident les êtres humains à exécuter des tâches quotidiennes comme passer l'aspirateur, ramasser les ordures, nettoyer les vitres, faire le repassage et préparer le repas, en affectant potentiellement la qualité de vie et la définition des rôles sexués.

Enfin, les robots sont également utilisés dans le domaine de l'*élevage*. Les robots d'élevage modifient les pratiques agricoles et les relations entre êtres humains et animaux : des réseaux de capteurs permettent de surveiller les conditions de vie des animaux, en rendant possibles des soins adaptés et en réduisant potentiellement l'interaction entre êtres humains et animaux. Dans l'*agriculture*, l'utilisation de drones se développe aux fins de l'« agriculture de précision », qui vise à optimiser la productivité et la qualité des aliments en recueillant et en analysant des données en vue d'une gestion scientifique (par exemple, optimisation de l'utilisation d'engrais, irrigation goutte à goutte).

Les avantages potentiels des robots dans ce domaine doivent être mis en balance avec l'impact environnemental de l'ensemble du cycle de production des robots. Il est probable, en outre, que la robotique ajoute aux préoccupations suscitées par le volume croissant de déchets électroniques, en particulier dans les pays en développement.

IV. RÉGLEMENTATION ÉTHIQUE ET JURIDIQUE

La robotique et les robots soulèvent des enjeux juridiques et éthiques nouveaux et sans précédent. Étant donné le caractère complexe de la conception, de la construction et de la programmation des robots, l'aspect le plus souvent mis en avant d'un point de vue éthique est celui de la « traçabilité », c'est-à-dire la possibilité de déterminer les causes de toutes les actions (ou omissions) antérieures d'un robot. Cependant, dans le cas des robots dotés d'un haut niveau d'autonomie et de capacités de décision et d'apprentissage, l'exigence de traçabilité est problématique : en effet, ces robots ne sont pas simplement programmés pour exécuter des tâches spécifiques mais pour apprendre et continuer à se développer en interaction avec leur environnement, ce qui exige de modifier les notions éthique et juridique actuelles de traçabilité.

Le *Projet de Rapport contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publié en 2016 par la Commission des affaires juridiques du Parlement européen, est l'initiative la plus concrète à ce jour en vue de la réglementation légale et éthique de la robotique et de l'industrie des robots. Préoccupé par l'impact possible de la robotique sur la sécurité, la vie privée, l'intégrité, la dignité et l'autonomie des êtres humains, ce document aborde un certain nombre de questions éthiques et juridiques touchant à la robotique et à l'utilisation des robots.

En 2017, le Rathenau Instituut a publié un rapport sur « Les droits de l'homme à l'ère des robots », commandité par l'Assemblée parlementaire du Conseil de l'Europe (APCE). Ce rapport examine l'impact potentiellement négatif de la robotique dans un certain nombre de domaines relatifs aux droits de l'homme : le respect de la vie privée, la dignité humaine, la propriété, la sécurité et la responsabilité, la liberté d'expression, l'interdiction de la discrimination, l'accès à la justice et l'accès à un procès équitable. Il recommande l'introduction de deux nouveaux droits fondamentaux : le droit de ne pas faire l'objet de mesures, d'analyses ou de contrôles (en relation avec le mésusage possible de l'IA, de la collecte de données et de l'internet des objets) et le droit à un contact humain significatif (en relation avec les mésusages possibles, intentionnels ou non, des robots de soins).

Selon Leenes *et al.* (2017), la réglementation du secteur de la robotique est confrontée à quatre difficultés : (1) rester en phase avec l'évolution rapide des technologies ; (2) trouver un équilibre entre stimuler – ou, tout au moins, ne pas freiner – l'innovation et protéger les valeurs et droits humains fondamentaux ; (3) savoir quand affirmer les normes sociales existantes ou les pousser à évoluer dans de nouvelles directions ; et (4) mettre en balance l'efficacité et la légitimité de la réglementation des technologies.

V. ENJEUX ÉTHIQUES

Les machines robotiques brouillent les limites entre sujets humains et objets technologiques. Elles ont donc non seulement des implications sociétales qui demandent à être évaluées d'un point de vue éthique, mais elles obligent à s'interroger sur certaines notions essentielles de l'éthique, par exemple celles de « capacité d'agir » et de « responsabilité », et sur les cadres de valeur existants.

Étant donné l'autonomie accrue des robots, la question se pose de savoir qui exactement devrait porter la *responsabilité* éthique et/ou légale du comportement des robots. Il semble exister une responsabilité « commune » ou « partagée » entre le concepteur, l'ingénieur, le programmeur, le fabricant, l'investisseur, le vendeur et l'utilisateur du robot. Aucun de ces acteurs, en effet, ne peut être désigné comme la source ultime d'un acte. Cette approche, cependant, a pour effet de diluer tout à fait la notion de responsabilité : si tous ces acteurs ont une part de la responsabilité totale, aucun d'eux n'est entièrement responsable. Surmonter l'effet potentiellement paralysant de la difficulté à endosser ou assigner la responsabilité représente donc un enjeu majeur pour l'éthique de la robotique. Pour résoudre ce problème, on pourrait envisager de développer des techniques permettant d'anticiper dans la mesure du possible les impacts du développement robotique (Waelbers and Swierstra, 2014 ; Verbeek, 2013). Une autre solution consisterait à réfléchir soigneusement à l'apparition inévitable d'effets imprévus, en envisageant l'introduction dans la société de technologies robotiques comme une « expérimentation sociale » qui demande à être conduite avec le plus grand soin (Van de Poel, 2013).

À cause de leur autonomie, les robots remettent en cause la notion de *capacité d'agir*. Bien qu'il existe des différences très nettes entre la capacité d'agir des robots et celle des êtres humains, les robots « font » des choses qui sont le résultat de leurs processus de décision et interactions propres, et pas seulement des instructions reçues de leurs concepteurs, ce qui a des incidences sur la manière dont on définit un agent moral. La question principale, par conséquent, est de déterminer comment les robots modifient les pratiques humaines, et de quelle façon la qualité des relations homme-robot pourrait orienter la conception, la mise en œuvre et l'utilisation des robots. La discipline émergente de l'« éthique des machines », qui vise à équiper les machines de principes ou de procédures éthiques leur permettant de résoudre des problèmes éthiques et de fonctionner d'une manière responsable sur le plan éthique, est aussi une manière d'aborder la question de leur capacité morale. Une question intrigante au sujet des technologies robotiques est celle de leur *statut moral*. Les robots sont-ils appelés à devenir des entités moralement utiles, et pas seulement utiles en un sens instrumental comme de simples appareils fabriqués pour exécuter des tâches spécifiques ? Dans ce cas, mériteront-ils un respect moral et une protection contre les dommages et seront-ils dotés non seulement d'obligations et de devoirs mais aussi de droits moraux ?

Enfin, les technologies robotiques perturbent aussi quelque peu les cadres moraux : elles n'ont pas seulement des effets sociétaux qui peuvent être évalués d'un point de vue éthique, elles remettent aussi en cause les cadres mêmes avec lesquels on cherche à les évaluer. Les robots de soins, par exemple, modifient sans doute ce que les êtres humains valorisent le plus dans les soins, tandis que les robots éducatifs changent nos critères d'une bonne éducation et les robots sexuels pourraient avoir des incidences sur notre idée de l'amour et des relations intimes. Pour prendre en compte de manière responsable de tels effets normatifs, il est nécessaire de mettre soigneusement en balance anticipation et expérimentation, en suivant de près, à petite échelle et dans un contexte expérimental, les retombées des technologies robotiques sur les systèmes de valeurs, afin de pouvoir les prendre en compte dans les pratiques de conception des robots, dans les discussions publiques et dans l'élaboration des politiques.

VI. RECOMMANDATIONS

La distinction entre robots déterministes et robots cognitifs est importante aux fins de l'élaboration de recommandations concernant l'éthique de la robotique. En effet, le comportement d'un robot déterministe est dicté par le programme qui contrôle ses actes. La responsabilité de ces actes est donc claire et peut faire l'objet, pour l'essentiel, d'une réglementation légale. En revanche, les décisions et les actes d'un robot cognitif peuvent donner lieu uniquement à des estimations statistiques et sont donc imprévisibles. Par conséquent, la responsabilité des actes de ce type de robot n'est pas claire et son comportement dans un environnement autre que celui dans lequel il a été formé pendant sa phase d'apprentissage – c'est-à-dire un environnement « aléatoire » – peut être catastrophique. Assigner la responsabilité des actes d'une machine en partie *stochastique* ne va pas de soi.

C'est pourquoi la COMEST propose d'élaborer des recommandations prenant en compte la distinction susmentionnée. À un premier niveau, celui des machines déterministes dont on peut assigner la responsabilité du comportement, les recommandations de la Commission porteront essentiellement sur les outils juridiques nécessaires pour réglementer leur utilisation. À un second niveau, celui des machines cognitives dont le comportement n'est pas prévisible à 100 % et est donc en grande partie stochastique, les recommandations envisageront, outre l'adoption d'outils juridiques, l'élaboration de codes de pratique et de directives éthiques à l'intention des producteurs et des utilisateurs. Enfin, s'agissant des machines stochastiques utilisées dans un contexte où le risque de dommages existe (voiture ou arme autonome, par exemple), il conviendra d'examiner le degré d'autonomie qui peut raisonnablement être accordé à ces machines et les modalités de maintien d'un niveau significatif de contrôle humain.

Le tableau ci-dessous présente le modèle proposé sous une forme schématique. La structure envisagée est simple mais sa mise en œuvre sous forme de délimitation des responsabilités et de réglementation de l'utilisation des robots sera complexe et difficile, à la fois pour les scientifiques, les ingénieurs, les responsables de l'élaboration des politiques et les spécialistes des questions éthiques.

Décision du robot	Implication humaine	Technologie	Responsabilité	Réglementation
Effectuée à partir d'une série limitée d'options, sur la base de critères stricts prédéfinis	Critères mis en œuvre dans un cadre légal	Machine uniquement : algorithmes/robots déterministes	Fabricant du robot	Légale (normes, législation nationale ou internationale)
Effectuée à partir d'une gamme d'options, avec une certaine flexibilité, sur la base d'une politique prédéfinie	Décision déléguée au robot	Machine uniquement : algorithmes d'intelligence artificielle, robots cognitifs	Concepteur, fabricant, vendeur et utilisateur du robot	Codes de pratique des ingénieurs et des utilisateurs ; principe de précaution
Effectuée via l' interaction homme-robot	Contrôle humain des décisions du robot	Capacité des êtres humains à prendre contrôle du robot dans le cas où ses actes pourraient entraîner des dommages graves ou la mort d'un individu	Êtres humains	Éthique

Étant donné la diversité et la complexité des robots, établir un cadre de valeurs et de principes éthiques pourra être utile en vue de la définition de la réglementation à tous les niveaux – conception, fabrication et utilisation – et d'une manière cohérente, sous une forme allant de codes de conduite des ingénieurs à des textes de loi nationaux et à des conventions internationales. Le principe de la responsabilité humaine est le fil conducteur reliant les différentes valeurs examinées dans ce rapport. Les valeurs et principes éthiques pertinents incluent : (i) la dignité humaine ; (ii) l'autonomie ; (iii) le respect de la vie privée; (iv) l'innocuité ; (v) la responsabilité ; (vi) la bienfaisance ; et (v) la justice.

Les recommandations spécifiques de la COMEST sur l'éthique de la robotique sont les suivantes :

- i. Il est recommandé de poursuivre, tant au niveau national qu'international et selon des modalités pluridisciplinaires, le développement, la mise en œuvre, la révision et l'actualisation de codes d'éthique des roboticiens, en prenant en compte les avancées futures possibles de la robotique et leur impact sur la vie humaine et l'environnement (énergie, déchets informatiques, empreinte écologique). Il est également recommandé que les disciplines et les professions qui contribuent de façon importante à la robotique ou peuvent y recourir – notamment l'ingénierie électronique, l'intelligence artificielle, la médecine, les sciences animales, la psychologie et les sciences physiques – révisent, de préférence de manière coordonnée, leurs codes d'éthique propres, en cherchant à anticiper les défis qui pourront résulter de leurs liens avec la robotique et l'industrie des robots. De plus, il est recommandé que l'éthique – incluant les codes d'éthique, les codes de conduite, et autres documents pertinents – devienne une partie de l'étude des programmes pour tous les professionnels impliqués dans la conception et fabrication de robots.

- ii. Il est recommandé de prendre en compte les considérations éthiques dans la conception des technologies robotiques, en s'appuyant sur une approche telle que la « conception éthique ».
- iii. Pour assurer l'utilisation responsable des robots dans la société, il est recommandé d'introduire les nouvelles technologies robotiques à petite échelle, de manière minutieuse et transparente, dans des situations contrôlées, afin d'étudier ouvertement leur impact sur les pratiques et activités humaines, ainsi que sur les valeurs et cadres d'interprétation. Les résultats de telles expérimentations pourront ensuite être utilisés pour adapter la conception des robots, informer l'élaboration des politiques et de la réglementation, et fournir aux utilisateurs un point de vue critique sur l'emploi de ces technologies.
- iv. Il est recommandé d'organiser des discussions publiques au sujet des implications des nouvelles technologies robotiques eu égard à divers aspects de la vie sociale et de la vie quotidienne, y compris l'impact de l'ensemble du cycle de production des robots, afin d'aider les individus à développer une attitude réflexive à ce sujet et de mieux sensibiliser les concepteurs et les décideurs.
- v. Il est recommandé que les États, les organisations professionnelles et les établissements d'enseignement réfléchissent à l'impact de la robotique en termes de réduction des opportunités d'emploi et de création de nouveaux emplois, en prêtant attention aux secteurs de la société plus susceptibles d'être vulnérables à ces changements, et qu'ils prennent des mesures adéquates en vue de la reconversion et de la réadaptation de la force de travail pour permettre de réaliser leurs avantages potentiels.
- vi. En ce qui concerne les véhicules autonomes capables de fonctionner et de prendre des décisions sur la base de l'apprentissage automatique et d'algorithmes cognitifs, il est recommandé de recenser et de définir les situations dans lesquelles la responsabilité des résultats de l'action d'un véhicule autonome incombe uniquement à un être humain (« le conducteur »).
- vii. Les questions éthiques soulevées par les drones armés ne concernent pas seulement le droit international humanitaire. L'utilisation de drones armés contre des acteurs non étatiques soupçonnés dans une situation d'insurrection soulève également des questions éthiques et juridiques. La COMEST considère qu'outre les problèmes juridiques que cela pose, l'assassinat d'un être humain par un robot armé, que ce soit dans une situation de conflit armé déclaré ou dans le cadre d'opérations de contre-insurrection, est moralement inacceptable. Il est recommandé que les États réexaminent cette pratique.
- viii. En ce qui concerne les armes autonomes, il est vigoureusement recommandé, pour des raisons juridiques, éthiques et militaires-opérationnelles, de maintenir le contrôle humain sur les systèmes d'armement et l'utilisation de la force. Compte tenu de la rapidité potentielle du développement des armes autonomes, il est urgent, comme le préconise le CICR, « de déterminer le type et le degré de contrôle humain sur l'utilisation de systèmes d'armement qui sont jugés nécessaires pour satisfaire aux obligations légales et répondre à certains critères éthiques et sociétaux ». (ICRC, 2016).
- ix. Il est recommandé que les États établissent des politiques sur l'utilisation de drones dans les activités de surveillance et de police et dans les contextes non militaires. Les modalités d'utilisation de drones par la police devraient être décidées par les représentants du public et non par les forces de police. Les politiques en la matière devraient être claires, écrites et accessibles au public. Elles devraient veiller, au minimum, à ce que le déploiement d'un drone soit autorisé uniquement sur la base d'une commission rogatoire, dans une situation d'urgence ou lorsqu'il existe des raisons spécifiques et clairement définies de penser qu'un drone permettra de recueillir des éléments de preuve relatifs à une infraction pénale particulière. La conservation

des images ne devrait être autorisée que s'il existe des soupçons motivés qu'elles contiennent des éléments de preuve au sujet d'un délit ou sont pertinentes aux fins d'une enquête ou d'un procès en cours. L'utilisation de drones devrait faire l'objet d'un contrôle ouvert et d'une surveillance adéquate afin de prévenir toute utilisation abusive. Les drones utilisés par la police ne devraient pas être équipés d'armes létales ou non létales. L'utilisation d'armes autonomes par la police ou les services de sécurité ne devrait pas être autorisée.

- x. Il est recommandé que l'utilisation privée de drones soit soumise à l'octroi d'une licence et que leur aire d'utilisation fasse l'objet d'un contrôle strict pour des raisons juridiques, de sécurité et de respect de la vie privée. Équiper des drones privés d'armes létales ou non létales devrait être interdit.
- xi. Il est recommandé qu'une attention particulière soit prêtée aux questions de genre et aux stéréotypes concernant tous types de robots décrits dans ce rapport, et en particulier aux robots de jeu, aux compagnons sexuels, et autre remplacements d'emplois.
- xii. De la même façon que pour d'autres technologies avancées, il est recommandé que l'impact environnemental soit considéré comme faisant partie d'une analyse de cycle de vie, pour permettre une évaluation plus holistique pour savoir si un usage spécifique de robotique fournit plus de bien que de mal à la société. Il est aussi recommandé que lors de la construction de robots (nano, micro ou macro), un effort soit fait pour utiliser des matériels biodégradables et des technologies respectueuses de l'environnement, et pour améliorer le recyclage de matériels.
- xiii. Il est recommandé que la COMEST examine les enjeux éthiques associés à l'internet des objets (IdO) dans le prolongement de son travail sur ce rapport.

RAPPORT DE LA COMEST SUR L'ÉTHIQUE DE LA ROBOTIQUE

I. INTRODUCTION

1. Les robots peuvent aider l'humanité, et ils le font depuis la moitié du XX^e siècle. Pendant les premières cinquante années de leur existence, les robots ont été utilisés principalement pour exécuter des tâches industrielles : intégrés aux usines, ils permettaient de libérer les êtres humains de certaines tâches répétitives. Les robots ont aussi été utilisés traditionnellement dans des applications militaires. Depuis plusieurs années, les robots apparaissent dans d'autres domaines comme celui des transports, des soins de santé, de l'éducation et des travaux ménagers (robots de service), où leur interaction avec la société est plus visible.

2. Qu'il s'agisse des robots autonomes, des drones, des robots ménagers, des robots humanoïdes ou des robots industriels, médicaux et militaires, la robotique moderne repose de plus en plus sur l'intelligence artificielle (IA). Cette technologie, également appelée « informatique cognitive » rend possible l'interaction linguistique et sensorielle de type humain, ainsi que la réalisation de tâches d'analyse et de résolution de problèmes, et même la créativité. Les machines reposant sur l'intelligence artificielle disposent de capacités d'apprentissage comparables à celles des êtres humains et peuvent devenir des apprenants autonomes. Ces capacités sont très fortement améliorées par l'emploi d'algorithmes d'apprentissage profond et de l'informatique en nuage.

3. Avec les technologies avancées d'apprentissage automatique, les robots deviennent des robots cognitifs capables d'apprendre à partir de l'expérience, avec des enseignants humains, ou même par leurs propres moyens, en acquérant ainsi potentiellement la capacité d'interagir avec leur environnement. La caractéristique principale des machines cognitives est le fait que leurs décisions sont imprévisibles et que leurs actions dépendent de leur expérience et de certaines conditions stochastiques. La question de la responsabilité des actions exécutées par des robots cognitifs est donc cruciale. La présence de robots cognitifs au domicile, sur le lieu de travail et, plus généralement, dans l'ensemble de la société constitue un défi croissant car elle affecte les comportements humains et induit de profonds changements sociaux et culturels (par exemple, dans les relations familiales et de voisinage, au travail et en ce qui concerne le rôle de l'État). Elle soulève aussi certaines questions du point de vue de la sécurité, du respect de la vie privée et de la dignité humaine. Les technologies émergentes comme l'internet des objets (IdO) ne font qu'amplifier ces nouvelles questions éthiques.

4. Les robots soulèvent régulièrement des problèmes éthiques, surtout lorsqu'il s'agit de savoir s'ils vont remplacer les êtres humains, ou prendre des rôles normalement réservés aux humains. Jusqu'où les robots doivent-ils remplacer le travail humain ? Est-il moralement acceptable de donner aux robots un rôle dans la guerre ? Voulons-nous que les robots fournissent des soins pour les personnes âgées, ou pour les enfants autistes ? Quelles sont les dimensions éthiques lorsqu'on donne le rôle de 'compagnon' ou même de 'partenaire érotique' à un robot ? Y a-t-il une voiture automatique ? Ces questions se basent sur une préoccupation fondamentale avec la dignité humaine et les possibles menaces que la robotique lui pose. Ce rapport vise à répondre à ces et autres questions en explorant de quelles manières les technologies robotiques pourront jouer un rôle dans la société dans un futur proche. En effet, souvent, les robots ne remplacent les humains, mais plutôt ont une influence profonde sur les pratiques humaines : elles ne prennent pas le contrôle de la médecine, l'éducation, la police, le travail et les conflits armés, mais plutôt elles changent ces pratiques. En examinant quelles valeurs sont en jeu dans ces développements, ce rapport vise à identifier les dimensions éthiques des technologies robotiques en relation étroite avec et les développements technologiques eux-mêmes, mais aussi les questions éthiques qui surgissent dans les discussions éthiques et sociétales.

5. Le présent rapport poursuit un but de sensibilisation et cherche à promouvoir à la fois un examen public et un dialogue inclusif sur les questions éthiques concernant les différentes formes d'utilisation de robots autonomes dans la société. Au fur et à mesure des progrès techniques, ces questions deviennent plus complexes et il est nécessaire d'examiner de près la relation entre les robots et la société, telle qu'elle existe aujourd'hui et telle qu'elle pourrait évoluer à l'avenir.

6. L'étude détaillée de ces questions exige une large collaboration interdisciplinaire. Les aspects technologiques de la robotique requièrent des compétences dans un large éventail de domaines scientifiques : les ingénieurs mécaniciens et automaticiens s'occupent de la mobilité des robots ; les physiciens et les ingénieurs électriciens prennent soin de leurs capacités sensorielles et de communication ; les informaticiens et les spécialistes du traitement de signaux conçoivent les aspects cognitifs des robots ; et les ingénieurs de systèmes sont responsables de l'intégration de ces diverses composantes. En outre, les aspects techniques de la robotique entretiennent des liens étroits avec la fonctionnalité des environnements anthropocentriques et la participation d'experts de différents domaines des sciences humaines et des sciences sociales est donc essentielle. Le groupe de travail à l'origine de ce rapport est une équipe interdisciplinaire regroupant des compétences dans les domaines éthique, politique, social et technologique.

7. La COMEST a abordé pour la première fois les questions éthiques relatives à la robotique moderne lors de sa 7^e session extraordinaire en juillet 2012. Cette première discussion a été suivie par un atelier organisé par British Pugwash de deux jours sur l'éthique de la robotique moderne dans la surveillance, le travail de police et les conflits armés, qui a eu lieu à l'université de Birmingham (Royaume-Uni) du 20 au 22 mars 2013. Un bref rapport sur cet atelier, qui nous a fourni certains des éléments utilisés dans ce rapport, peut être consulté sur le site web de la COMEST (UNESCO, 2013). Lors de la 8^e session ordinaire de la COMEST (Bratislava, 20-30 mai 2013) a été organisée conjointement avec British Pugwash une Conférence sur les défis éthiques émergents en science et en technologie, dont l'une des sessions était consacrée aux robots autonomes. Enfin, lors de sa 9^e session ordinaire en septembre 2015, la COMEST a décidé de travailler spécifiquement sur l'éthique de la robotique en créant un groupe de travail pour poursuivre la réflexion à ce sujet. Le groupe de travail s'est réuni à Paris du 18 au 20 mai 2016 et a participé à un atelier interdisciplinaire international intitulé « Les machines morales : développements et relations. Nanotechnologies et hybridité », organisé conjointement par la COMEST et l'université Paris-VIII. Il s'est réuni de nouveau à l'Université Laval à Québec du 27 au 31 mars 2017 et a participé à un atelier international sur le thème « Les robots et la société : Quelles transformations ? Quelles réglementations ? ». Trois membres de l'équipe ont aussi participé au Sommet mondial « IA for Good », organisé par l'Union internationale des télécommunications (UIT) à Genève du 7 au 9 juin 2017, au cours duquel ils ont pu aborder un certain nombre de questions pertinentes avec des spécialistes de l'intelligence artificielle.

II. QU'EST-CE QU'UN ROBOT ?

II.1. Problèmes de définition

8. Définir ce qu'est un « robot » est une tâche difficile et peut-être infinie, étant donné les développements rapides de la robotique. Le terme « robot » (qui a remplacé le terme d'« automate » utilisé auparavant) est d'origine tchèque et a été utilisé pour la première fois par Karel Čapek dans une pièce de science-fiction de 1920 appelée *R.U.R. (Rossumovi Univerzální Roboti* [Les robots universels de Rossum]). Il dérive du mot *robota*, qui signifie « travail » ou « labeur » en tchèque. Cependant, l'étymologie du mot « robot » n'aide guère à définir ce qu'est un robot. Dire qu'un robot est un objet créé pour exécuter certaines tâches de travail n'est pas suffisamment précis puisqu'il existe de nombreux objets correspondant à cette description qui ne peuvent être considérés comme des robots (par exemple, les ordinateurs individuels ou les voitures).

9. L'idée la plus répandue de ce qu'est un robot – grâce aux films, aux séries télévisées et à la littérature de science-fiction – est celle d'une machine qui voit, pense et se comporte comme un être humain. Cette conception anthropomorphique des robots (« androïdes », s'ils ont été conçus pour ressembler à un homme et « gynoïdes » s'ils cherchent à ressembler à une femme) est seulement en partie exacte et ne correspond pas nécessairement aux définitions du mot « robot » que l'on trouve dans la littérature spécialisée. Dans ces définitions, en effet, il n'est pas nécessaire qu'un robot ait une forme humanoïde. Voici deux exemples de ces définitions :

- a. « (1) Machine équipée de capteurs ou d'instruments de détection de signaux d'entrée ou d'états environnementaux mais comprenant aussi des mécanismes de réaction ou d'orientation et pouvant effectuer des tâches de détection, de calcul ou autres, et des programmes enregistrés déterminant la consécution des actions ; par exemple, une machine fonctionnant par elle-même ; (2) une unité mécanique pouvant être programmée pour exécuter certaines tâches de manipulation ou de locomotion sous contrôle automatique » (Rosenberg, 1986, p.161).
- b. « Machine intelligente pouvant exécuter des tâches mécaniques courantes, répétitives ou dangereuses, ou d'autres opérations directement sur commande d'un être humain ou de façon autonome, en utilisant un ordinateur à logiciel intégré (contenant des commandes et des instructions préalablement enregistrées) ou reposant sur un niveau avancé d'intelligence machinique (artificielle) (qui permet de baser les décisions et les actions sur les données recueillies par le robot sur son environnement actuel) » (Angelo, 2007, p.309).

10. Dans son *Encyclopedia of Robotics*, Gibilisco (2003) distingue cinq générations de robots sur la base de leurs capacités respectives. Les robots de la première génération (avant 1980) étaient mécaniques, stationnaires, précis, rapides, physiquement robustes et basés sur des servomécanismes mais n'utilisaient ni capteurs externes, ni intelligence artificielle ; les robots de la deuxième génération (de 1980 à 1990), contrôlés par microprocesseurs, étaient programmables et incluaient des systèmes de vision, ainsi que des capteurs tactiles, de position et de pression ; les robots de la troisième génération (à partir du milieu des années 90) sont devenus mobiles et autonomes, capables de reconnaître et de synthétiser la parole et ils intègrent des systèmes de navigation ou de télécommande et l'intelligence artificielle. Il soutient en outre que les quatrième et cinquième générations sont les robots pensants de l'avenir, capables par exemple d'acquiescer et de reproduire certaines caractéristiques humaines comme le sens de l'humour.

11. Joseph Engelberger, l'un des pionniers de la robotique industrielle, a un jour déclaré : « Je ne peux définir ce qu'est un robot mais, quand j'en vois un, je sais le reconnaître » (cité par Tzafestas, 2016b, p.4). Comme on le verra dans la suite de ce rapport, une telle confiance dans la possibilité de distinguer intuitivement les robots des non-robots n'est peut-être pas justifiée car les robots sont déjà extrêmement divers par la forme, la taille, les matériaux et les fonctions. La robotique, en effet, ne relève plus seulement de l'ingénierie mécanique et électrique/électroniques (« robots »), mais aussi des nanosciences (« nanobots »), de la biologie (« biorobots » ou « cyborgs »), de la botanique (« plantoïdes ») et d'autres disciplines.

II.2. Histoire de la robotique : fiction, imaginaire et réel

12. La création d'êtres intelligents artificiels est l'une des préoccupations les plus anciennes et les plus durables de l'imagination et de l'esprit humains. Des êtres de ce type apparaissent dans de nombreux récits humains : dans les mythologies, les religions, les littératures, et tout particulièrement les films et les séries télévisées de science-fiction. Il est impossible de donner une vue d'ensemble de tous ces récits, ou même seulement des plus importants d'entre eux, mais l'on peut citer quelques exemples paradigmatiques ou qui apparaissent fréquemment dans la littérature sur la robotique et la roboéthique.

13. Dans la mythologie grecque, Héphaïstos, fils de Zeus et d'Héra, façonne à partir de l'or deux servantes pour l'aider dans son atelier (Héphaïstos était infirme de naissance et ces deux « servantes d'or » peuvent être considérées comme des prototypes fictionnels de robots assistants personnels). Héphaïstos est aussi le créateur d'un géant de bronze artificiel, Talos, qui était chargé de contrôler et de défendre l'île de Crète (le parallèle avec les robots militaires et policiers est évident). Dans un autre mythe grec, Pygmalion, sculpteur de talent, crée une statue d'ivoire représentant une femme, Galatée, d'une telle beauté qu'il en tombe amoureux. Pygmalion traite Galatée comme si elle était une femme réelle, en lui offrant des cadeaux précieux, en l'habillant et en lui parlant. Comme elle n'est qu'une statue, Galatée ne peut bien entendu être considérée comme un robot au sens strict (ni même comme le prédécesseur d'un robot). Néanmoins, l'histoire de Pygmalion n'est pas sans intérêt du point de vue de la roboéthique car elle illustre le penchant humain à développer des liens affectifs étroits avec des objets anthropomorphiques inanimés.

14. La mythologie hébraïque contient un certain nombre de récits consacrés au Golem, une créature anthropomorphe qui, dans certaines versions, est créée à partir de terre ou d'argile, la vie lui étant insufflée au moyen de rituels religieux ou magiques. Dans certaines variantes du mythe, le Golem est un serviteur bienfaisant et fidèle, tandis que dans d'autres, il s'agit d'une créature dangereuse capable de se retourner contre son créateur. La mythologie inuite comprend aussi une légende similaire à propos d'un être artificiel appelé Tupilak, la seule différence étant que Tupilak est créé avec des parties d'animaux et d'êtres humains (souvent des enfants). Tupilak est aussi décrit comme potentiellement dangereux pour son créateur. *Les mille et une nuits*, le recueil classique en langue arabe de contes d'origine surtout populaire, contient plusieurs récits dans lesquels figurent des créatures artificielles telles que des robots humanoïdes, des animaux robotiques et divers automates.

15. Si l'on en vient aux représentations littéraires de robots ou de créatures robotiques, l'exemple classique est celui du roman de Mary Wollstonecraft Shelley, *Frankenstein ou le Prométhée moderne* (1818). C'est l'histoire de Victor Frankenstein, un savant qui crée à partir de fragments de cadavres humains un être vivant artificiel (qui, dans le roman, est appelé « la Créature » ou « le Monstre » mais qui est connu sous le nom de Frankenstein). Victor Frankenstein est horrifié par le résultat de son expérimentation, tout particulièrement par l'apparence de la Créature, et il l'abandonne ainsi que tous ses travaux. Cependant, la Créature se révolte contre son créateur et recourt à la menace et au chantage pour contraindre Frankenstein à lui créer une compagne féminine, en invoquant son droit au bonheur. Le roman s'achève par la mort violente, dans des conditions tragiques, de presque tous les personnages principaux.

16. Comme indiqué plus haut (voir paragraphe 9), la pièce de science-fiction de Karel Čapek, *R.U.R. (Rossumovi Univerzální Roboti* [Les robots universels de Rossum]), publiée en 1920, occupe une place à part non seulement dans l'histoire de la littérature sur les robots mais aussi dans la robotique. En effet, le terme de « robot » a été utilisé pour la première fois dans cette pièce. Čapek a attribué la paternité de ce mot à son frère Josef Čapek. La pièce représente une usine fondée par le vieux Rossum et son neveu, le jeune Rossum, qui fabrique à partir de la matière organique des robots intelligents et souvent impossibles à distinguer des êtres humains pour servir de force de travail à bon marché et être vendus dans le monde entier. La révolte des robots contre leurs créateurs est l'événement central de la pièce, qui s'achève par l'extinction quasi-complète de l'humanité et l'établissement d'un nouveau gouvernement mondial de robots (cependant, la nouvelle race robotique conserve certaines caractéristiques humaines comme l'aptitude à l'amour et au sacrifice de soi).

17. L'œuvre littéraire d'Isaac Asimov présente aussi un intérêt particulier pour la robotique car c'est dans l'une de ses nouvelles de science-fiction « Menteur ! » (publiée pour la première fois en 1941, reimprimé dans Asimov, 1950) que le mot « robotique » apparaît pour la première fois. De façon plus célèbre encore, c'est dans la nouvelle « Cercle vicieux » (1942) qu'Asimov a introduit les Trois Lois de la robotique :

- a. un robot ne peut porter atteinte à un être humain, ni, en restant passif, permettre qu'un être humain soit exposé au danger ;
- b. un robot doit obéir aux ordres qui lui sont donnés par un être humain, sauf si de tels ordres entrent en conflit avec la première loi ;
- c. un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la première ou la deuxième loi.

18. Plus tard, dans son livre *Les robots et l'Empire* (1985), Asimov a introduit la Loi Zéro de la robotique : « Un robot ne peut porter atteinte à l'humanité, ni, en restant passif, permettre que l'humanité soit exposée au danger ». Les œuvres de science-fiction d'Asimov et ses ouvrages de vulgarisation scientifique ont inspiré de nombreux roboticiens dans la vie réelle, ainsi que des spécialistes de l'intelligence artificielle comme Joseph Engelberger et Marvin Minsky.

19. Il existe, semble-t-il, des différences culturelles dans l'attitude à l'égard des êtres artificiels tels que les robots. Alors que, dans la culture occidentale, ces êtres sont généralement décrits comme des êtres malfaisants et un danger potentiel pour les humains, il n'en va pas toujours ainsi dans les cultures non occidentales. Veruggio et Operto (2008) indiquent que, dans la culture japonaise, « les machines (et, d'une manière générale, les produits de l'activité humaine) sont toujours bienfaisants et favorables à l'humanité » (p.6). Selon Bar-Cohen et Hanson (2009), les Japonais sont mieux préparés à accepter les robots humanoïdes à cause de la forte influence du bouddhisme et du shintoïsme dans leur culture. Ces auteurs expliquent qu'alors que les religions occidentales (christianisme et judaïsme) considèrent la création de tels êtres comme une manière d'interférer avec le rôle du Créateur, le bouddhisme et le shintoïsme n'éprouvent aucune hantise à cet égard à cause de leur fond animiste qui les amène à considérer tous les êtres vivants et non vivants comme dotés d'une âme ou d'un esprit.

20. Les robots, en particulier les robots humanoïdes, doivent une grande part de leur popularité aux films et aux séries télévisées de science-fiction. L'industrie du cinéma s'est très tôt intéressée aux robots. Parmi les premiers films présentant des robots, on peut citer *The Clever Dummy* (de Ferris Hartman, 1917), *L'uomo meccanico* (d'André Deed, 1921, film préservé seulement en partie) et *Metropolis* (de Fritz Lang, 1927). Les cinéastes ont toujours manifesté un vif intérêt pour la description des conséquences potentiellement catastrophiques de l'utilisation de robots et de leur révolte contre les êtres humains. Dans *2001, l'Odyssée de l'espace* (de Stanley Kubrick, 1968), par exemple, un ordinateur contrôlant toutes les fonctions d'un vaisseau spatial en vol vers Jupiter cherche à détruire son équipage humain. Dans *Le cerveau d'acier* (de Joseph Sargent, 1970), deux superordinateurs conçus à l'origine comme systèmes de défense des États-Unis et de l'Union soviétique pour le contrôle des armes nucléaires, s'unissent et prennent le pouvoir sur l'ensemble de la race humaine. Dans *Génération Proteus* (de Donald Cammel, 1977), un ordinateur qui gère le ménage d'un scientifique emprisonne et insémine artificiellement sa femme (elle donne ensuite naissance à un clone d'elle-même). *Terminator* (de James Cameron, 1984) décrit un monde futur devenu une dictature gouvernée au moyen de l'intelligence artificielle et d'une armée de robots.

21. Les films et séries télévisées de science-fiction ne décrivent pas tous, évidemment, les robots comme une menace ultime pour l'humanité. Nombre d'entre eux présentent les robots sous un jour neutre ou même les humanisent en leur attribuant toutes les vertus et les vices des êtres humains. Parmi les exemples les plus célèbres, on peut citer le duo de robots R2-D2 et C3PO dans *La guerre des étoiles* (de George Lucas, 1977), le répliquant Roy Batty créé par des moyens génétiques dans *Blade Runner* (de Ridley Scott, 1982), qui sauve la vie de son poursuivant et ennemi (le « blade runner »), l'androïde Data qui souhaite acquérir des caractéristiques humaines, en particulier les sentiments d'un être humain, dans la série *Star Trek* et l'enfant robot David dans *A.I. Intelligence artificielle* (de Steven Spielberg, 2001), qui souffre d'avoir été abandonné par sa « mère » biologique.

22. Dans la réalité, le développement des robots a été nettement plus lent et plus modeste que ce qu'ont imaginé les fictions littéraires et cinématographiques. Les prédécesseurs historiques des robots, les « automates », étaient des artefacts mécaniques ressemblant à des êtres humains ou à des animaux réels et capables d'exécuter des mouvements assez simples comme écrire ou jouer d'un instrument de musique. Selon Angelo (2007), en Chine, les artisans de l'antiquité ont fabriqué divers automates dont un orchestre mécanique. Au XV^e siècle, Leonardo da Vinci a réalisé les plans d'un chevalier médiéval automate ; au XVI^e siècle, Guinallo Toriano a fabriqué une joueuse de mandoline mécanique et, au XVIII^e siècle, des automates d'une très grande sophistication ont été fabriqués par Jacques Vaucanson (Le flûteur automate, Le joueur de tambourin et Le canard digérateur) et Pierre Jaquet-Droz (L'écrivain, La musicienne et Le dessinateur). Ces automates étaient généralement conçus dans un but de divertissement. Les scientifiques et ingénieurs arabes avaient des intérêts similaires. Le plus célèbre d'entre eux est sans doute Ismaël al-Djazari, un savant et ingénieur du XIII^e siècle qui nous a laissé un livre comprenant des croquis et des instructions indiquant comment fabriquer divers automates. Al-Djazari fabriqua lui-même nombre de ces automates, par exemple une servante offrant à boire, un groupe de musiciens et des paons mécaniques reposant pour la plupart sur un système hydraulique.

23. La plupart des historiens (voir, par exemple, Stone, 2005 ; Angelo, 2007) considèrent qu'« Unimate » et « Shakey » sont les premiers robots au sens moderne du terme. « Unimate » est considéré comme le premier robot industriel. Il a été conçu en 1954 par Joseph Engelberger et George Devol pour la chaîne de montage de General Motors à Trenton dans le New Jersey. « Unimate » exécutait des tâches dangereuses pour les travailleurs humains, par exemple le transport des pièces coulées sous pression et leur soudage à l'armature des véhicules. Le premier « Unimate », contrairement à ses successeurs plus complexes, n'était composé que d'un seul bras robotique et d'un programme sous forme de tambour magnétique. « Shakey » est le robot à roues conçu entre 1966 et 1972 par Charles Rosen et ses associés du Centre pour l'intelligence artificielle en Californie. Il s'agissait du premier robot généraliste basé sur l'intelligence artificielle et capable de raisonner et d'optimiser ses instructions. Commandé à partir d'une console informatique, il pouvait exécuter des tâches comme se déplacer d'une pièce à une autre, ouvrir et fermer une porte, déplacer des objets ou allumer et éteindre la lumière.

II.3. Autonomie, interactivité, communication et mobilité

24. Certaines caractéristiques générales des robots contemporains méritent d'être soulignées, non seulement parce qu'elles sont essentielles pour comprendre ce que sont ces robots mais aussi parce qu'elles soulèvent, individuellement et collectivement, des préoccupations éthiques particulières, même lorsque la technologie fonctionne parfaitement. Cependant, des pannes peuvent se produire et l'impossibilité de garantir une fiabilité absolue est également source de préoccupations. Ces caractéristiques sont la mobilité, l'interactivité, la communication et l'autonomie.

25. Il n'est pas nécessaire qu'un robot soit mobile ; il peut être stationnaire comme cela est le cas de la plupart des robots industriels. Néanmoins, la mobilité est un attribut essentiel de nombreux types de robots parce qu'elle leur permet d'exécuter des tâches à la place des êtres humains dans des environnements typiquement humains (hôpital, bureau ou cuisine, par exemple). Cette mobilité peut être obtenue par divers moyens techniques. Il existe aujourd'hui des robots qui peuvent marcher (robots bipèdes et à plusieurs jambes), ramper, rouler, avancer sur des roues, voler et nager. Alors que les dommages pouvant éventuellement être causés par un robot stationnaire se limitent aux personnes travaillant ou vivant à proximité, les robots mobiles (en particulier s'ils ont un fort degré d'autonomie et des capacités avancées d'interaction avec leur environnement) peuvent poser des risques plus graves. Pour illustrer ce point, il suffit de citer un exemple tiré de la vie réelle : en juin 2016, « Promobot » (un robot publicitaire capable d'interaction avec des êtres humains dans un lieu public) s'est échappé d'un laboratoire de la ville de Perm en Russie et a été retrouvé au milieu d'une rue très fréquentée.

26. La capacité d'interaction avec l'environnement est un autre attribut important des robots. On affirme souvent que cette capacité est ce qui distingue les robots des ordinateurs. Les robots sont équipés de capteurs et d'actionneurs. Les capteurs permettent à un robot de recueillir des informations pertinentes à partir de son environnement (par exemple, reconnaître et distinguer différents objets ou personnes et mesurer leur distance). Les capteurs incluent toute une gamme d'appareils (cameras, sonars, lasers, microphones, etc.) permettant à un robot de voir, d'entendre, de toucher et de déterminer sa position par rapport aux objets environnants et à certaines limites à ne pas dépasser. Les actionneurs sont des mécanismes ou appareils qui permettent à un robot d'agir sur l'environnement (bras robotiques ou préhenseurs, par exemple) ; ils peuvent être de type mécanique, hydraulique, magnétique, pneumatique ou électrique. La capacité d'interaction des robots (au moyen de capteurs et d'actionneurs) pose un problème particulier d'un point de vue éthique. D'un côté, en effet, les robots sont capables d'agir activement dans leur environnement et donc d'exécuter des tâches qui pourraient être dangereuses pour des êtres humains. La force d'un robot ménager ou d'un robot de soins, par exemple, n'est pas sans incidences d'un point de vue éthique car, plus puissant est ce robot, plus graves sont les dommages qu'il peut infliger à un être humain en cas de mauvais fonctionnement. D'un autre côté, cependant, les robots peuvent recueillir au moyen de leurs capteurs des données pouvant être utilisées, intentionnellement ou non, pour nuire à des êtres humains ou porter atteinte à leur vie privée (par exemple, dans le cadre d'activités criminelles, de l'espionnage industriel ou du journalisme jaune).

27. Contrairement aux premiers robots qui avaient besoin d'une interface informatique pour communiquer, de nombreux robots sont aujourd'hui équipés de systèmes sophistiqués de reconnaissance de la parole et de synthèse vocale, qui leur permettent de communiquer avec des êtres humains dans une langue naturelle. La nécessité de mettre au point des robots capables de recevoir de manière fiable des instructions ou même d'être programmés au moyen d'une langue naturelle (sur le modèle de la programmation neurolinguistique) devient d'autant plus pressante que leur présence s'accroît dans de très nombreux domaines de la vie humaine (par exemple, les robots de soins aux personnes âgées et les robots infirmiers doivent être capables de comprendre les commandes de l'utilisateur aussi précisément que possible et aussi d'expliquer leurs propres actions). Toutefois, selon Wise (2005), il est difficile de programmer des ordinateurs en se servant d'une langue naturelle parce que les langues humaines mettent en jeu à la fois de multiples significations, des aspects contextuels complexes et des présupposés implicites, alors que les ordinateurs ont besoin d'instructions tout à fait explicites. Gibilisco (2003) indique néanmoins que, bien que les systèmes de reconnaissance vocale et de synthèse vocale des robots soient encore loin d'être parfaits, ce domaine de la robotique se développe rapidement. Les robots humanoïdes Wakamaru (conçu par Mitsubishi au Japon) et Nao (développé par Aldebaran Robotics en France), qui peuvent communiquer avec des êtres humains par le geste et la parole, sont deux exemples des progrès obtenus en ce domaine (Bekey, 2012).

28. Selon Bekey (2012), « l'idée généralement admise de ce qu'est un robot implique de façon essentielle un certain degré d'autonomie ou l'aptitude à « réfléchir » seul et à prendre ses propres décisions pour agir sur l'environnement » (p.18). L'autonomie est définie comme « l'aptitude à fonctionner dans un environnement réel sans aucune forme de contrôle extérieur une fois la machine activée et, au moins dans certains domaines de fonctionnement, pendant une durée prolongée » (Bekey, 2012, p.18). Les premières générations de robots étaient des « automates » plus ou moins complexes, c'est-à-dire des machines capables d'exécuter des tâches répétitives assez simples, comme les robots de chaîne de montage par exemple. Les robots plus récents, par contre, sont plus autonomes et peuvent exécuter seuls des tâches plus compliquées, sans dépendre de commandes ou d'un contrôle humains. L'« autonomie » peut en fait être considérée comme l'attribut essentiel d'un robot parce que c'est elle qui le distingue d'un simple outil ou appareil commandé par un être humain.

29. Au vu de l'autonomie accrue des robots, la question qui se pose est de savoir quel degré de proximité les performances d'un robot autonome doivent atteindre par rapport à

celles d'un être humain exécutant une tâche identique avant qu'on puisse le « lâcher » dans son environnement de travail. Les performances humaines étant elles-mêmes variables, on soutient parfois que les systèmes robotiques autonomes doivent être soumis à des normes de performance *plus élevées* que celles que l'on peut attendre d'un être humain avant de leur donner les clés d'une voiture, par exemple.

30. Il est intéressant de noter à cet égard que certaines données expérimentales indiquent que les équipes homme-machine obtiennent de meilleurs résultats qu'un être humain ou une machine travaillant seuls. Outre ce que cela implique en termes d'efficacité, on peut voir là comme un possible « filet de sécurité » contre le mauvais fonctionnement d'un partenaire robotique. De puissants arguments, par conséquent, existent en faveur du maintien d'un être humain dans le circuit de ce type de systèmes.

II.4. Nanorobotique

31. On s'accorde en général à définir la nanorobotique comme une technologie dont le but est de créer des machines ou des robots dont les composants ont une dimension égale à quelques nanomètres (10^{-9} mètres). Ces appareils sont appelés nanobots, nanoïdes, nanites, nanomachines ou nanomites.

32. Ils sont généralement regroupés en deux catégories – organique et inorganique – en fonction du type de matériau utilisé dans leur conception. La fabrication des nanorobots inorganiques repose sur une application spécifique de la nanoélectronique et de la technologie des semi-conducteurs permettant d'obtenir des composants de taille nanoscopique d'un haut niveau de complexité (Smith *et al.*, 2012 ; Weir *et al.*, 2005). Les nanorobots organiques reposent, quant à eux, sur des machines biomoléculaires qui s'inspirent directement des « processus naturels à l'échelle nanométrique » (Ummat *et al.*, 2004). Les efforts de recherche se concentrent actuellement sur le développement de systèmes robotiques miniatures ou micro-nanométriques conçus à partir de processus biologiques, intégrés à des entités biologiques et pouvant être utilisés dans des applications biologiques ou biomédicales. La plupart des machines moléculaires naturelles sont constituées d'un assemblage de protéines soigneusement sélectionnées et sont dotées de fonctions biologiques préprogrammées pouvant être activées en réponse à des stimuli physicochimiques spécifiques *in vitro* ou même dans un milieu artificiel (Mavroidis *et al.*, 2004).

33. Les composants qui jouent le rôle de moteur des nanorobots – par exemple, la kinésine, l'ARN polymérase, la myosine et la synthèse de l'ATP (adénosine triphosphate), qui fonctionnent à l'échelle nanométrique comme des activateurs biologiques – ont retenu fortement l'attention car ils ont un niveau d'efficacité extrêmement élevé, sont capables de se reproduire et n'ont pas besoin d'être inventés puisqu'ils existent et fonctionnent déjà dans la nature, et peuvent donc être prélevés et adaptés à nos besoins (Ummat *et al.*, 2004).

34. Parmi les capacités spécifiques des nanorobots organiques, on peut citer : leur durabilité, qui est liée à leur petite taille ; leur activation rapide par rapport à leurs homologues de plus grande taille ; leur aptitude à extraire de l'énergie de la chaleur ambiante ou de produire leur propre énergie à partir de substances décomposables ; leur intelligence collective, leur comportement coopératif, leur auto-assemblage et leur facilité à se reproduire ; leur aptitude à traiter et à programmer des nano-informations ; ainsi que leur architecture d'interfaces depuis le niveau nano jusqu'au niveau macro (Weir *et al.*, 2005).

35. Les applications potentielles des nanorobots incluent, entre autres, la détection et l'élimination de substances chimiques toxiques dans l'environnement. Leur capacité à se reconfigurer les rend également utiles dans le domaine des technologies spatiales et militaires où elles pourront servir à effectuer des tâches auparavant impossibles à réaliser à l'avance. Dans le domaine médical, outre l'acheminement de médicaments, les micro-robots médicaux (ou « nanomédibots ») permettront aussi de : capter de manière chimique ou physique et analyser certains aspects d'un organe humain au niveau nanoscopique ; surveiller des fonctions corporelles ; réparer des tissus endommagés ; déconstruire des matériaux ou

cellules pathologiques ou anormaux (par exemple, cancer ou plaque) ; et améliorer la santé et les performances des êtres humains.

36. Du fait de leur petite taille, qui les rend invisibles à l'œil humain, les nanorobots exigent un examen sérieux du niveau de sécurité et de sauvegarde de la vie privée requis pour protéger les êtres humains et les écosystèmes. L'impact des nanorobots sur l'économie, leur coût de fabrication, les contraintes de conception et les technologies de pointe correspondantes sont aussi des éléments préoccupants.

II.5. Robots, intelligence artificielle et algorithmes

37. Comme il n'est déjà pas facile de définir l'« intelligence », définir l'« intelligence artificielle » est encore plus difficile. L'intelligence artificielle (IA) désigne généralement deux choses qui sont liées entre elles.

38. L'IA peut, en effet, désigner un domaine de recherche ou « une approche transdisciplinaire de la compréhension, de la modélisation et de la réplication de l'intelligence et des processus cognitifs faisant appel à divers principes et appareils informatiques, mathématiques, logiques, mécaniques ou même biologiques » (Frankish and Ramsey, 2014, p.7). Comme le disait Marvin Minsky, l'un des fondateurs de l'IA, c'est « la science de faire faire à des machines des choses qui nécessiteraient de l'intelligence si elles étaient faites par des hommes » (cité par Copeland, 1993, p.1). L'IA est un exemple majeur d'un domaine de recherche interdisciplinaire parce qu'elle combine de nombreuses disciplines différentes comme l'informatique, la psychologie, les sciences cognitives, la logique, les mathématiques et la philosophie.

39. Les progrès de la robotique en matière de mobilité, de communication, de capacité d'interaction et d'autonomie ont été rendus possibles grâce au développement de l'IA. Selon Warwick (2012), « un robot intelligent n'est que l'incarnation d'une entité artificiellement intelligente : il est ce qui donne un corps à l'IA » (p.116). Toutefois, bien que tous les robots aient un « corps » (ce qui les distingue des ordinateurs), ils ne sont pas tous « intelligents ». La plupart des robots industriels ou des « manipulateurs industriels » ne peuvent être considérés comme intelligents puisque leur comportement est préprogrammé et adapté pour exécuter automatiquement un nombre limité de tâches extrêmement spécifiques et répétitives (par exemple, des tâches de soudure ou de peinture) dans un environnement immuable, c'est-à-dire un « monde structuré ». En cas de modification de son environnement, un robot industriel est incapable de s'adapter aux changements sans l'intervention d'un opérateur humain (la plupart des robots industriels, en outre, ne sont pas équipés de capteurs qui leur permettraient de percevoir ces changements). Il convient de noter, cependant, que les robots industriels sont constamment améliorés afin de pouvoir s'adapter le mieux possible à toute situation imprévue sur le lieu de travail (et qu'ils deviennent donc d'un meilleur rapport coût-efficacité).

40. Les robots ayant une certaine forme d'« intelligence artificielle » sont généralement capables de percevoir et de représenter formellement les changements de leur environnement et d'adapter leur fonctionnement en conséquence. L'intelligence artificielle est déterminante pour l'autonomie des robots car elle leur permet d'exécuter des tâches complexes dans un environnement qui change (non structuré) – par exemple conduire une voiture en s'adaptant aux conditions de la circulation – sans être télécommandés ou contrôlés par un opérateur humain. Selon Murphy (2000), « alors que l'origine des manipulateurs industriels et de l'ingénierie robotique remonte dans une certaine mesure à la course aux armements nucléaires, on peut dire que le développement de l'IA, en tant que nouvelle approche, a commencé avec la course à l'espace » (p.26). En effet, si les bras robotiques, les préhenseurs et les véhicules automatiques ou télécommandés étaient suffisants pour protéger les travailleurs humains des matériaux radioactifs, les robots spatiaux (sondes spatiales ou astromobiles, par exemple) avaient besoin d'une certaine forme d'« intelligence » pour pouvoir fonctionner de façon autonome dans des situations inattendues et des environnements

nouveaux (notamment en cas de rupture de la communication radio ou lors de l'exploration de régions encore inconnues de la lune ou d'une planète).

41. Comme l'ont montré Bekey *et al.* (2008), les robots spatiaux opèrent fréquemment dans des situations imprévisibles ; c'est pourquoi, lors de leur conception, les roboticiens spatiaux doivent prendre en compte les quatre aspects suivants : (1) la mobilité (le robot doit être capable de changer de localisation sans mettre en danger les astronautes humains, le matériel ou lui-même) ; (2) la manipulation (le robot doit pouvoir utiliser ses bras et ses outils avec suffisamment de précision et une force suffisante mais non dangereuse) ; (3) le laps de temps (le robot doit être capable de recevoir des ordres d'opérateurs humains distants et de fonctionner de manière autonome s'il ne peut recevoir ces ordres) ; et (4) les environnements extrêmes (le robot doit pouvoir supporter, par exemple, le rayonnement ionisant, l'ultravide et des températures extrêmes). L'exploration humaine au-delà de la lune continuant à poser des risques de sécurité majeurs, il est probable que les robots spatiaux (sous la forme de sondes, d'astromobiles, d'atterrisseurs équipés de bras robotiques et de manipulateurs autonomes, ou même d'« astronautes robotiques » humanoïdes) gagneront en sophistication et joueront un rôle de plus en plus important dans les futures explorations spatiales.

42. L'article d'Alan Turing « Computing machinery and intelligence », publié en 1950 dans la revue philosophique *Mind*, est la première formulation théorique de la possibilité de l'intelligence artificielle. Dans cet article, Turing présentait ce que l'on appelle maintenant le « test de Turing » selon lequel l'« intelligence » peut être attribuée à une machine ou à un programme informatique si un être humain placé dans un contexte expérimental particulier (conversation en langue naturelle via un terminal informatique) ne peut distinguer les réponses de la machine ou du programme des réponses données par un être humain. À ce jour, malgré de nombreuses tentatives et de nombreux progrès, aucune machine ou programme informatique capable de gagner à ce « jeu d'imitation » n'a pas encore été créé (Franklin, 2014). On considère généralement que le domaine de recherche de l'IA (ainsi que le nom même d'« intelligence artificielle ») ont été définis lors de ce qu'on appelle la « conférence de Dartmouth » organisée en 1956 au Dartmouth College par les pionniers en ce domaine, notamment John McCarthy, Marvin Minsky et d'autres.

43. L'« intelligence artificielle » peut aussi désigner le produit final de la recherche en IA : une machine ou un artefact qui incorpore une certaine forme d'« intelligence », c'est-à-dire qui est capable de « réfléchir » ou de résoudre des problèmes d'une manière semblable à la pensée humaine. On pourrait donc penser que la robotique est simplement une application des connaissances acquises dans le cadre général de la recherche en IA mais la relation entre IA et robotique est en fait plus complexe. Selon Murphy (2000), « la robotique a joué un rôle décisif dans les progrès de l'IA » (p.36) et contribué au développement de tous les grands domaines de recherche qu'englobe l'IA :

- a. représentation des connaissances ;
- b. compréhension du langage naturel ;
- c. apprentissage ;
- d. planification et résolution de problème ;
- e. inférence,
- f. recherche ;
- g. vision.

Autrement dit, ce qui semble s'être produit est une « coévolution » de la robotique et de l'IA.

44. Pour illustrer les progrès énormes de l'IA en robotique, il suffit de comparer deux générations de robots faisant appel à l'intelligence artificielle (Husbands, 2014) :

- a. Le robot « Shakey », conçu entre 1966 et 1972 au Stanford Research Institute, qui a été l'un des premiers robots à intelligence artificielle : il s'agissait d'un robot mobile, capable de percevoir et de représenter dans une certaine mesure son

environnement et aussi d'exécuter des tâches simples telles que planifier, effectuer des calculs d'itinéraire et re-disposer des objets simples. Cependant, ce robot devait être contrôlé par un ordinateur occupant une pièce entière et ne pouvait fonctionner en temps réel (il lui fallait parfois des heures pour trouver le chemin conduisant à l'autre extrémité d'une pièce).

- b. Le robot « Kismet » conçu au début du XXI^e siècle par le Laboratoire d'intelligence artificielle du MIT dans un but d'interaction sociale : ce robot pouvait remuer les yeux, changer d'expression faciale en fonction de son « humeur », communiquer par la parole avec des êtres humains et réagir à l'humeur et aux émotions de son interlocuteur. Ce premier prototype a ensuite conduit au développement de robots encore plus sophistiqués dans le domaine de l'interaction sociale.

45. Ces deux exemples montrent clairement le changement de paradigme qui est intervenu en robotique pendant les dernières décennies. Alors que les premiers robots à intelligence artificielle étaient programmés sur la base de ce qu'on appelle le « paradigme hiérarchique », les plus récents s'appuient sur le « paradigme réactif » ou le « paradigme mixte délibération/réaction ». Brièvement, le « paradigme hiérarchique » signifie que le robot exécute chaque opération de haut en bas : il commence par « capter » des informations, puis prépare ou « planifie » et finit par « agir ». Le « paradigme hiérarchique », qui était essentiellement une tentative d'imiter la pensée humaine, s'est heurté à plusieurs difficultés (en particulier ce qu'on appelle le « problème du cadre », c'est-à-dire la difficulté à reconstruire les éléments d'information manquants et à distinguer information pertinente et information non pertinente). Cela a conduit à la définition d'un nouveau paradigme, le « paradigme réactif », qui s'est inspiré notamment de la biologie et de la psychologie cognitive. L'innovation fondamentale a consisté à réduire l'élément de « préparation » ou de « planification », qui demandait trop de temps et de puissance de calcul, en liant directement certaines « actions » du robot à des données « sensorielles » particulières (en imitant le comportement des insectes). Toutefois, comme la « planification » ne pouvait être entièrement évitée, surtout dans le cas des robots généralistes, le « paradigme mixte délibération/réaction » a ensuite été mis au point en combinant les éléments les plus pertinents des deux paradigmes précédents (Murphy, 2000).

46. Quel que soit le paradigme dont ils relèvent et indépendamment de leurs différences de forme, de taille et de mouvement, les robots exécutent leurs tâches sur la base d'algorithmes. Le terme d'algorithme est surtout utilisé en informatique. Il est généralement défini comme « un ensemble de règles ou d'instructions bien définies à appliquer pour résoudre un problème, par exemple exécuter un calcul, en un nombre d'étapes fini » (Butterfield et al., 2016). Les algorithmes déterministes contrôlent le comportement prévisible des **robots déterministes**, tandis que les algorithmes IA disposant de capacités d'apprentissage sont au cœur du fonctionnement des **robots cognitifs**.

47. Il est important de noter que le comportement d'un robot déterministe – même s'il s'agit d'un robot très complexe, intelligent et autonome, c'est-à-dire requérant très peu ou aucune surveillance humaine – est essentiellement préprogrammé et déterminé (ce qui a des implications quant au statut moral des robots, comme on le verra plus loin dans ce rapport). D'un point de vue éthique, le fait qu'un robot soit déterministe est pertinent (et souhaitable) lorsque se posent des problèmes de traçabilité, c'est-à-dire lorsqu'il est nécessaire de reconstruire précisément les décisions et actions antérieures d'un robot pour résoudre un différend éthique ou juridique. Les robots cognitifs reposant sur l'intelligence artificielle, en revanche, seront capables d'apprendre à partir de leurs expériences passées et de recalibrer eux-mêmes leurs algorithmes ; par conséquent, leur comportement ne sera pas entièrement prévisible et ce point exigera certainement à l'avenir une sérieuse attention et réflexion éthique.

III. LES ROBOTS ET LA SOCIÉTÉ

III.1. Utilisation de robots dans l'industrie

48. La robotisation de l'industrie est aujourd'hui un fait. Pendant les dernières décennies, des robots industriels ont peu à peu remplacé les opérateurs humains pour les tâches répétitives dans le secteur manufacturier. Les robots les plus avancés (robotique mobile) équipés de capteurs et de manipulateurs plus performants et fonctionnant sur la base d'algorithmes IA plus puissants (technologies convergentes) commencent aussi de plus en plus à exécuter des tâches manuelles de type non répétitif. Grâce aux progrès des capteurs et de la mobilité, les robots sont capables de produire des marchandises d'une qualité et d'une fiabilité plus grandes que le travail humain dans de grands secteurs industriels comme l'industrie alimentaire, le secteur énergétique et la logistique. Plus le coût des robots diminue et plus leurs capacités technologiques s'accroissent, plus l'on peut s'attendre à ce qu'ils remplacent progressivement le travail humain dans une large gamme de fonctions peu rémunérées du secteur des services, où se concentre la croissance de l'emploi depuis plusieurs décennies. Par conséquent, nombre d'emplois manuels peu rémunérés qui sont restés jusqu'ici à l'écart de l'informatisation risquent de disparaître à l'avenir (Frey and Osborne, 2013). Cette évolution, qui devrait conduire à un accroissement de la productivité sur la base des technologies nouvelles, ne constitue pas nécessairement un progrès pour les travailleurs. En l'absence d'ajustements structurels et de mesures compensatoires, elle risque d'entraîner à la fois une augmentation du chômage dans certains secteurs de la force de travail et une inégalité sociale accrue si les profits résultant de l'amélioration de la productivité grâce aux nouvelles technologies bénéficient principalement aux propriétaires de ces technologies.

49. Certains auteurs ont décrit cette évolution comme le passage à une nouvelle période de l'histoire qui se caractériserait par « la fin du travail ». Toutefois, les interprétations à ce sujet divergent très fortement : pour certains, en particulier les scientifiques, les ingénieurs et les employeurs, un monde sans travail humain signifierait le commencement d'une ère historique nouvelle où les êtres humains seraient enfin libérés d'une vie de labeur ardu et de tâches répétitives dépourvues de sens ; pour d'autres, une société sans travailleurs évoque plutôt l'idée d'un avenir sombre caractérisé par le chômage de masse et la pauvreté mondiale, et ponctué de plus en plus fréquemment par des soulèvements et des troubles sociaux (Rifkin, 1995).

50. La deuxième question éthique importante que soulève la robotisation de l'économie concerne le financement et la promotion de la recherche et du développement en robotique. Des institutions comme la Foundation for Responsible Robotics (FRR) et son fondateur Noel Sharkey, pionnier de l'intelligence artificielle, attirent l'attention sur le fait que, malgré les investissements importants des pouvoirs publics et du secteur privé en faveur de la R-D robotique, les effets sociaux de la robotique ne sont aucunement pris en compte et ses implications éthiques sont assez peu mentionnées dans les documents de fond sur les politiques.

51. Dans ce contexte, il importe de souligner une nouvelle fois la responsabilité éthique des concepteurs de robots industriels et des ingénieurs en robotique. Divers codes éthiques ont été proposés pour définir la responsabilité des ingénieurs en robotique. Le code d'éthique des ingénieurs en robotique du Worcester Polytechnic Institute indique, par exemple, qu'un ingénieur doit :

prendre en considération, autant que faire se peut, les utilisations non éthiques possibles de ses créations et limiter les possibilités de telles utilisations. Un ingénieur en robotique soucieux d'éthique ne peut prévenir tous les dangers potentiels et les utilisations non souhaitables de ses créations mais il devrait faire tout ce qui est en son pouvoir pour les réduire au minimum. Cela peut vouloir dire ajouter certains éléments de sécurité, avertir autrui d'un danger ou refuser entièrement les projets dangereux. (Ingram *et al.*, 2010, p.103)

Un ingénieur en robotique doit aussi examiner les conséquences de l'interaction de ses créations avec leur environnement. Toute préoccupation relative aux dangers potentiels ou aux comportements non éthiques d'un robot doit être divulguée, indépendamment du fait de savoir si un ingénieur en robotique en est directement responsable. Et, dans le cas où une utilisation non éthique d'un robot apparaît seulement après sa mise en circulation, un ingénieur en robotique doit faire tout ce qui est possible pour résoudre le problème.

52. Les progrès de la robotisation modifient en profondeur les conditions de travail et les emplois. Travailler côte à côte avec des robots exige de nouvelles compétences professionnelles et la mise en place de mesures de sécurité nouvelles et adéquates sur le lieu de travail. La question de la durée du travail des êtres humains travaillant côte à côte avec des machines de plus en plus autonomes se pose également, comme cela s'est produit lors des autres révolutions technologiques dans l'histoire du travail. Enfin, il est nécessaire de créer et d'adapter des systèmes d'éducation et de formation continue des travailleurs, afin d'éviter que la dernière révolution technologique ne finisse par produire un très grand nombre de travailleurs inemployables à cause de leur manque de familiarité avec les technologies les plus récentes.

53. Des études récentes (Murashov *et al.*, 2015) sur la sécurité dans le travail avec les robots recensent trois catégories de robots présentes sur les lieux de travail :

- a. les robots industriels ;
- b. les robots de service professionnels et personnels ;
- c. les robots collaborateurs.

54. La plupart des *robots industriels* ne sont pas conscients de leur environnement et peuvent par conséquent être dangereux pour les travailleurs. La principale méthode de sécurité adoptée à l'égard des robots industriels consiste à maintenir une distance suffisante entre les travailleurs humains et les robots en activité en créant des « espaces réservés ».

55. Un certain nombre de documents directeurs et de normes internationales portent sur la sécurité au travail autour des robots industriels. Cependant, la poursuite des incidents aboutissant à des blessures ou au décès de travailleurs montre que des mesures supplémentaires, notamment le doublement des mesures de sécurité ainsi que l'amélioration de la formation et de la culture générale en matière de sécurité, sont nécessaires.

56. Les *robots de service* sont utilisés principalement en dehors des sites industriels, dans des environnements non structurés et extrêmement imprévisibles, soit en l'absence d'êtres humains, par exemple sur le site de catastrophes, soit avec des êtres humains, par exemple dans les hôpitaux et les maisons de repos. La proximité physique est beaucoup plus fréquente entre robots de service professionnels et travailleurs humains qu'entre robots industriels et travailleurs car ils partagent souvent le même lieu de travail. Il n'est donc pas possible dans ce cas d'isoler les travailleurs des robots comme méthode principale de sécurité. En outre, l'environnement plus complexe dans lequel fonctionnent les robots de service exige de leur accorder un degré d'autonomie et de mobilité beaucoup plus grand. Cette autonomie et cette mobilité peuvent créer des situations dangereuses pour les travailleurs. Malgré les multiples inquiétudes en matière de sécurité que suscite l'utilisation de robots de service dans le même espace de travail que des travailleurs humains, aucune norme n'a encore été développée au niveau international sur la sécurité des travailleurs humains dans les situations d'entretien et de fonctionnement des robots de service professionnels et personnels.

57. Les *robots collaborateurs* sont des robots conçus pour l'interaction directe avec un être humain ; ils comprennent à la fois des robots industriels et des robots de service professionnels ou personnels. Les robots collaborateurs combinent la dextérité, la flexibilité et la capacité à résoudre des problèmes des travailleurs humains avec la force, l'endurance et la précision des robots mécaniques. La robotique managériale est un nouveau domaine des technologies robotiques collaboratives. Au lieu d'être relégués à exécuter des tâches ordinaires, répétitives et de précision, les robots, grâce à leur mémoire parfaite, à leur

connectivité internet et à leurs ordinateurs hautement performants pour l'analyse des données, pourraient remplir avec succès certaines fonctions de direction ou de gestion. Comme les robots collaborateurs travaillent côte à côte avec des travailleurs humains, l'isolement n'est pas une option envisageable dans un but de sécurité et d'autres méthodes de sécurité (capteurs de proximité, matériaux appropriés, outils logiciels) devront donc être conçues et mises en place.

58. Un autre série d'enjeux éthiques peut être décrite en utilisant la notion de « fossé robotique », qui est calquée sur celle de fossé numérique. Si le fossé numérique désigne la manière dont les technologies numériques – et l'accès aux technologies – redéfinissent la structure de pouvoir au sein des sociétés contemporaines, la notion de fossé robotique permet d'aborder des questions comme : la robotique, qui se caractérise par une convergence des technologies, des progrès rapides et une réduction progressive des coûts, peut-elle être intégrée à nos sociétés sous le modèle de marché actuel ? Sera-t-elle cause d'un nouveau fossé technologique, social et politique ? Va-t-elle transformer les relations de pouvoir entre individus, groupes, communautés et États, comme le ferait une technologie militaire cruciale ? Va-t-elle modifier les échanges sociaux quotidiens en devenant un élément de la vie quotidienne, tout comme elle a transformé le secteur de l'industrie où la robotique industrielle est aujourd'hui arrivée à maturité et pleinement établie ? (Peláez, 2014). De plus, comme pour le fossé numérique, la question se pose de l'écart mondial entre pays en développement et pays développés dans l'accès aux technologies robotiques, question qui demande à être examinée dans une perspective de justice mondiale.

59. L'impact de la robotisation sur l'économie mondialisée risque d'être fortement asymétrique. Les robots devenant moins chers et plus efficaces, l'intérêt économique d'employer des travailleurs peu qualifiés commence à s'estomper. Les travailleurs peu rémunérés pourront difficilement faire concurrence à la productivité des robots. Alors qu'auparavant, les pays plus pauvres pouvaient faire valoir leurs avantages en termes de coûts pour attirer les industriels, ces avantages risquent de disparaître à l'ère de la robotique. Le coût d'utilisation d'un robot est le même en Chine, aux États-Unis et à Madagascar. Certaines grandes entreprises occidentales rapatrient leur production dans des usines en grande partie automatisées et situées à plus grande proximité de leurs clients, en remédiant ainsi aux risques, aux coûts et aux difficultés associées à une longue chaîne d'approvisionnement (Balding, 2016).

60. Comment les pays en développement pourront-ils répondre à cet élargissement possible du fossé numérique et robotique et au danger potentiel qu'il représente pour leurs stratégies de développement ? Comme indiqué dans l'Agenda 2030 des Nations Unies pour le développement durable et dans le Plan d'action d'Addis-Abeba, ces pays doivent s'efforcer de retourner à leur avantage l'impact potentiellement libérateur des données ouvertes et des méga-données (UN, 2015a; UN, 2015b). Les données sont le moteur de la révolution robotique et elles doivent donc aussi être utilisées pour défendre et protéger des acteurs qui, autrement, risqueraient d'être laissés à la traîne par ces nouvelles technologies (OECD, 2016).

III.2. Usages militaires et civils des systèmes robotiques mobiles

61. La COMEST voudrait d'emblée réaffirmer son engagement envers la mission et le mandat de l'UNESCO de promouvoir et construire la paix. Puisque l'objectif de ce rapport est de fournir une analyse critique et éthique de l'usage des technologies robotiques, il est important que la COMEST aborde les discussions en cours sur l'usage militaire des robots. Ceci en aucun cas affaiblit son engagement pour la paix.

62. Le droit international humanitaire (DIH) couvre deux domaines : (a) la protection des personnes qui ne participent pas, ou plus, aux combats ; et (b) les restrictions aux moyens de guerre, principalement les armes, et aux méthodes de guerre, comme certaines tactiques militaires (ICRC, 2014). L'utilisation accrue de la robotique dans les conflits armés a des incidences fondamentales sur ces deux volets du DIH. Elle soulève des problèmes juridiques

et éthiques réels auxquels il est nécessaire de réfléchir aux fins de la protection des droits de l'homme et de la dignité humaine.

III.2.1. Systèmes robotiques militaires (drones)

63. Selon Chamayou (2015), l'armée américaine définit un drone comme « un véhicule terrestre, marin ou aérien qui est contrôlé à distance ou de manière automatique » (p.11). La famille des drones ne comprend pas uniquement des objets volants. Elle comprend autant de variétés de drones qu'il existe de catégories d'armes : drones terrestres, drones marins, drones sous-marins, drones souterrains même. N'importe quel type de véhicule ou de machine guidée peut en principe être « dronisé ». Un drone peut être soit guidé à distance par des opérateurs humains (télépilotage), soit guidé de façon autonome par des moyens robotiques (pilotage automatique). En pratique aujourd'hui, les drones combinent ces deux types de contrôle. « Drone » est le terme utilisé dans le langage courant. Les militaires parlent plutôt de « véhicules aériens sans pilote » (UAV) ou d'« avions de combat sans pilote » (UCAV), selon que l'appareil est ou non équipé d'armes.

64. Les opérateurs ("pilotes") de drones – aériens, terrestres ou sous-marins - se trouvent en fait à une grande distance de la machine armée. Ils sont assis à un poste de contrôle à des milliers de kilomètres de l'action et se servent d'une console de jeu pour guider l'appareil en vol et attaquer des cibles au sol. Comme les autres systèmes d'armement nouveaux, les drones ouvrent de nouvelles possibilités sur le plan militaire mais créent aussi de nouveaux risques sur le plan moral. On a observé, par exemple, que la distance entre le pilote et le champ d'action peut susciter chez le pilote un état d'esprit proche du jeu qui, en isolant les actes de toute considération morale, risque de conduire à l'indifférence dans le choix et l'attaque des cibles. C'est ainsi qu'un ancien opérateur à distance a pu déclarer : « Est-ce que ça vous est arrivé de marcher sur des fourmis et de ne plus jamais y penser ? C'est comme ça que vous êtes conduits à vous représenter les cibles : comme des tâches noires sur un écran » (Pilkington, 2015).

65. Le droit international humanitaire (DIH), appelé aussi droit de la guerre ou droit des conflits armés, est le cadre légal applicable aux situations de conflit armé et d'occupation. En tant qu'ensemble de normes et de principes, il cherche, pour des raisons humanitaires, à limiter les effets des conflits armés. Il comprend les Conventions de Genève et de La Haye, ainsi que les traités ultérieurs, la jurisprudence et le droit international coutumier. Les violations graves du DIH sont considérées comme des crimes de guerre.

66. Le DIH s'applique indubitablement aux nouveaux armements et à l'utilisation des nouveaux développements technologiques à des fins militaires. Ce point, fondamental pour la mise en œuvre effective du DIH, est reconnu à l'article 36 du Protocole additionnel aux Conventions de Genève qui stipule que les États ont l'obligation de déterminer si l'emploi de nouveaux moyens ou d'une nouvelle méthode de guerre serait interdit, dans certaines circonstances ou en toutes circonstances, par le DIH. Bien que le développement de systèmes d'armement télépilotés ou automatisés puisse présenter des avantages pour les forces armées, et même pour les civils dans certaines situations de conflit, des inquiétudes subsistent sur le point de savoir si les règles actuelles du DIH sont suffisamment claires eu égard aux caractéristiques spécifiques des technologies et à l'impact humanitaire prévisible de leur utilisation. Il est donc important de rappeler aux États la nécessité cruciale d'évaluer l'impact humanitaire potentiel des nouvelles technologies et leurs implications au regard du DIH, afin d'assurer qu'elles puissent être et soient effectivement utilisées d'une manière garantissant le respect du DIH. Il est aussi nécessaire que des discussions éclairées aient lieu au sujet de ces armes, analysant les questions juridiques, éthiques et sociétales en jeu, bien avant leur développement effectif.

67. Eu égard aux développements scientifiques qui ont rendu les systèmes robotiques possibles, il vaut sans doute la peine de noter que les Conventions de Genève datent de 1949 (ICRC, 1949a ; ICRC, 1949b ; ICRC, 1949c ; ICRC, 1949d), époque où la grande invention de l'année était le disque 45 tours. De même, les Protocoles additionnels aux Conventions ont

été approuvés en 1977(ICRC, 1977a ; ICRC, 1977b), alors que l'ordinateur personnel en était encore aux premières étapes de son développement et que les capacités existantes en matière de technologies de l'information et de la communication (TIC) étaient un million ou plus de fois inférieures à ce qu'elles sont aujourd'hui. Il est donc tout à fait approprié d'examiner quelles sont les conséquences de l'utilisation de systèmes robotiques dans les conflits armés à l'égard de l'application des principes les plus importants du DIH, à savoir les principes de *distinction*, de *proportionnalité* et de *responsabilité*.

68. *Distinguer* un combattant d'un civil dans une situation de télépilotage est par nature difficile – il ne s'agit pas d'un simple problème de classification visuelle. Les données dont dispose le pilote pour prendre une décision sont très différentes de celles auxquelles a accès un commandant sur le terrain. La substitution de bases de données à l'information concrète modifie en profondeur la conscience de la situation. La complexité très fortement accrue des séries de données sur lesquelles reposent les décisions de tir augmente les risques d'erreur dans l'identification des cibles. La qualité technique des images que reçoit le pilote, ainsi que la qualité des renseignements existants, sont toutes deux déterminantes pour informer la décision de faire feu. Télépiloter un drone armé a été comparé à « piloter un avion en regardant à travers un brin de paille », ce qui montre bien les limites de l'information spatio-contextuelle dont dispose le pilote lorsqu'il décide d'attaquer une cible. Ces limites peuvent favoriser la tendance à une représentation stéréotypée des cibles potentielles, augmentant ainsi le risque d'erreur d'identification des cibles.

69. Le fait qu'une arme permette d'anéantir avec précision tout individu que l'on souhaite éliminer ne signifie pas pour autant que l'on est mieux capable de déterminer qui constitue ou non une cible légitime. Comme l'explique Chamayou (2015), la précision du tir est sans relation avec la pertinence du choix de la cible.

70. Le partage de la prise de décision entre les membres de l'équipage d'un appareil télépiloté et les analystes qui observent l'action favorise la « pensée de groupe », l'équipage recherchant activement les éléments qui peuvent justifier un tir. Les systèmes de différenciation sont basés sur ce que l'on recherche et tenter de faire cela à distance aura tendance à simplifier les caractéristiques utilisées pour identifier une cible légitime. Une des conséquences en est l'attention privilégiée accordée aux hommes en âge de porter des armes, ce qui mène à l'emploi de stéréotypes grossiers dans la prise de décision. Un civil directement impliqué dans les hostilités peut constituer une cible légitime aux termes du DIH mais un télépilote est-il à même de faire une telle distinction de manière fiable ?

71. Le délai incompressible de transmission des signaux entre l'opérateur et le véhicule, qui varie entre 1,5 et 4 secondes, fait que la situation sur le terrain peut changer de manière significative entre la décision de faire feu et le moment où l'on « tire sur la gâchette ». Dans ce bref intervalle, la situation sur le terrain peut évoluer, avec le risque possible de tuer des non-combattants. Même si ce délai temporel diminue grâce aux progrès technologiques, les contraintes physiques intrinsèques ne permettront pas de le réduire à zéro. Sur le terrain, un combattant peut aussi décider de se rendre mais on voit mal comment il pourrait se rendre à un véhicule télépiloté.

72. Évaluer la *proportionnalité* à distance est encore plus problématique : la décision d'appliquer ou non une force létale ou cinétique doit en effet être prise dans un contexte particulier. Le pilote à distance dispose-t-il de données contextuelles suffisantes pour mettre en balance l'avantage militaire attendu et le risque de pertes de vies civiles ? Une frappe va-t-elle contribuer à la réalisation de l'objectif militaire global ou au contraire l'empêcher en créant des difficultés majeures pour la population locale ? Cette décision subjective et qualitative doit être prise par le pilote mais, en raison de sa distance, il est douteux qu'il puisse parvenir à une décision équilibrée.

73. L'examen de la proportionnalité par un pilote situé à distance est d'autant plus problématique qu'elle n'exige pas seulement de considérer les victimes civiles potentielles. D'autres types de dommages collatéraux doivent être pris en compte mais il est difficile

d'imaginer que cela puisse être effectué à distance. En dépit des affirmations selon lesquelles la précision des frappes effectuées au moyen de drones permet de minimiser les dommages collatéraux sur les biens, on voit mal comment l'endommagement potentiel des capacités d'approvisionnement et les dommages psychologiques pourraient être évalués à distance.

74. Les études consacrées aux effets des frappes à l'aide de drones mettent en évidence certains aspects moins tangibles de la proportionnalité. Nombre des personnes interrogées après de telles frappes sont dans un état de stress post-traumatique et souffrent d'autres troubles psychologiques graves tels que la dépression, les crises de panique, les réactions hystériques-somatiques, les réactions de peur excessives et les réactions de deuil anormales. L'impact sur les enfants est particulièrement inquiétant : troubles de l'attachement, peur grave du bruit, manque de concentration et perte d'intérêt pour les activités agréables. Une autre conséquence est l'abandon de la scolarité ou le manque d'assiduité à l'école.

75. D'un point de vue juridique, ces « dommages collatéraux » peuvent être considérés comme une forme de « châtement collectif » et comme allant à l'encontre de l'interdiction des représailles. L'impact psychologique des drones militaires sur les populations civiles est-il suffisamment reconnu ? Est-il reconnu dans le DIH ? Et qu'en est-il de leur impact social et économique et des conséquences à long terme de la perturbation sociétale qui en résulte ?

76. En ce qui concerne la *responsabilité*, lorsque des véhicules aériens sans pilote sont utilisés par des forces militaires régulières aux mêmes fins que celles pour lesquelles les avions classiques sont généralement utilisés, la situation au regard du DIH semble assez peu problématique : les troupes devront avoir reçu une formation au DIH et faire partie d'une chaîne de commandement militaire disciplinée.

77. Comme dans les autres formes de conflit armé, le DIH offre une protection juridique aux civils à l'égard des guerres menées à distance. Cependant, en l'absence de *transparence* sur l'utilisation des véhicules robotiques armés, le contexte précis où ils sont utilisés demeurera incertain, ce qui rendra difficile l'évaluation après-coup de la légalité et de l'impact des frappes effectuées au moyen des télécommandes.

78. La transparence brille aussi le plus souvent par son absence dans l'évaluation des armements robotiques conformément à l'article 36 du Protocole additionnel aux Conventions de Genève (ICRC, 1977a). Certains éléments indiquent que de telles évaluations ont lieu mais leurs résultats ne sont pas rendus publics. Bien que l'article 36 exige de déterminer si une arme *peut être utilisée* d'une manière conforme au DIH, aucune transparence n'existe sur la manière dont elle *sera effectivement utilisée* d'une manière conforme au DIH.

79. La question juridique du consentement au déploiement de véhicules robotiques armés ne semble pas différente de celle qui se pose lors de l'utilisation d'autres moyens militaires. La légalité du recours à la force de la part d'un État sur le territoire d'un autre État si ce dernier n'est pas lui-même en mesure de combattre une menace imminente est disputée. Même si une telle intervention est jugée légale, la menace doit constituer une menace directe pour l'État qui recourt à la force. Néanmoins, certaines questions se posent quant à la nature du consentement. Par exemple, doit-il être explicite pour être légal ? Qui donne le consentement ? Peut-il être retiré ? Comment déterminer si le consentement est réel et non contraint ? Ces questions sont valides indépendamment du fait de savoir si l'intervention inclut l'usage de véhicules télécommandés ou non, mais l'absence générale de transparence au sujet de leur déploiement risque de compliquer la question du consentement.

80. Les armes robotiques ont aussi des implications stratégiques qui sont de nature à changer la donne d'un point de vue militaire. Leur utilisation abaisse gravement le seuil de déclenchement d'un conflit armé en minimisant le rôle – ou en supprimant complètement la nécessité – de forces militaires humaines dans les zones de conflit, en assouplissant par conséquent les critères du *jus ad bellum* et en déplaçant l'évaluation de la nécessité militaire. Nous sommes donc confrontés au danger réel d'un changement stratégique majeur, avec la possibilité de conduire une « guerre sans risque » ininterrompue et de longue durée. On peut même commencer à envisager un monde où la guerre de basse intensité au moyen

d'assassinats ciblés serait la règle : un monde tout à fait nouveau où la possibilité constante d'assassinats ciblés prendrait la place du développement durable de formes de coexistence stables. Une fois entrés dans un tel monde, on voit mal comment il serait possible d'en sortir.

81. L'utilisation d'armes téléguidées soulève d'autres questions ayant des implications éthiques. On peut y inclure la modification des notions d'auto-défense et de menace : est-il acceptable d'un point de vue éthique de tuer lorsque l'on n'est pas soi-même menacé ? L'impact psychologique sur les personnes qui vivent sous la menace des frappes a été évoqué plus haut mais il faut également mentionner les effets psychologiques sur les opérateurs de robots armés, qui pourront voir les conséquences immédiates de leurs actions, ainsi que les retombées sur leur famille et la société à laquelle ils appartiennent. La possibilité que la guerre à distance remette en cause les valeurs militaires traditionnelles d'honneur et de courage est déjà devenue une réalité dans les armées.

82. Enfin, ces systèmes posent des problèmes importants en termes de sécurité. Comme toutes les machines, ils sont sujets à défaillance (interne ou provoquée de l'extérieur) et ceci pourrait avoir des conséquences fatales ou provoquer d'autres types de dommages. Contrairement à d'autres systèmes, cependant, les systèmes télépilotés sont vulnérables au risque de détournement ou de piratage – qui est déjà une réalité dans les conflits actuels au Moyen-Orient – pouvant aboutir, par exemple, à une prise de contrôle par l'« ennemi » ou même à un retournement au détriment des intérêts de l'attaquant.

83. Compte tenu des aspects problématiques, décrits ci-dessus, du déploiement de drones armés, la situation juridique relative à leur utilisation dans un conflit armé n'est en principe pas différente de celle des aéronefs pilotés. Les processus de décision gouvernementaux concernant le déploiement de véhicules télécommandés armés ne devraient donc pas être différents de ceux qui sont activés en vue du déploiement d'aéronefs pilotés ou de troupes sur le terrain. Il convient de noter en outre que, dans une situation de conflit armé, un opérateur distant, bien que situé à plusieurs milliers de kilomètres, constitue une cible légitime pour la partie adverse. Cependant, la capacité de cibler à une grande distance a ouvert la voie à une pratique, celle des *assassinats ciblés*, au moyen de laquelle les États ont la possibilité de viser à distance des individus particuliers (décrits dans les cas récents comme « terroristes »), en dehors d'une situation de conflit armé déclaré. Les assassinats ciblés ne sont pas nouveaux mais les systèmes robotiques armés les ont rendus plus faciles à exécuter et plus fréquents et ils sont devenus une composante essentielle de la « lutte contre le terrorisme ».

84. Philip Alston, Rapporteur spécial de l'ONU sur les exécutions extrajudiciaires, sommaires ou arbitraires a défini les assassinats ciblés comme « le recours intentionnel, prémédité et délibéré à la force létale par un État ou son agent agissant sous couvert de la loi, ou par un groupe armé dans le cadre d'un conflit armé, contre un individu en particulier qui n'est pas sous le contrôle physique de l'auteur de la frappe » (UN, 2010, p.3). Dans une situation de conflit armé, une telle action peut être légale du point de vue du DIH (même si on peut en contester la validité d'un point de vue éthique). Cependant, on doit s'interroger sur la légalité au regard du DIH de l'utilisation de drones pour exécuter des assassinats ciblés en dehors d'une zone de conflit, qui devient de plus en plus fréquente.

85. Deux points méritent sans doute d'être pris en compte à cet égard. Tout d'abord, l'article 51 de la Charte de l'ONU autorise l'emploi de la force armée, individuelle ou collective, à des fins d'auto-défense (UN, 1945). Cela inclut le « droit naturel » de recourir à la force dans un but d'auto-défense contre une agression armée imminente, et la force utilisée doit être à la fois nécessaire et proportionnée à la menace. Cet argument a été mis en avant par les États pour justifier les assassinats ciblés en dehors des zones de conflit. Cependant, l'évaluation de l'imminence, de la nécessité et de la proportionnalité est entièrement de l'ordre du jugement et peut faire l'objet, par conséquent, de manipulations politiques – il est bien trop facile de s'appuyer sur une définition si étendue qu'elle peut englober de nombreuses menaces vagues. En outre, la sécurité nationale est trop facilement invoquée pour justifier la non-divulgence d'information sur les raisons d'une décision.

86. Il serait donc souhaitable d'envisager comme bonne pratique de soumettre les décisions d'assassinats ciblés à un ensemble de procédures adéquates, transparentes et convenues à l'avance. Il existe déjà en fait des exemples de procédures de ce type qui ont été définies publiquement et sont soumises à un degré significatif de contrôle légal, de transparence et d'évaluation après une frappe, notamment aux États-Unis et en Israël ; le Royaume-Uni a, quant à lui, rendu publics les critères sur la base desquels il évalue l'imminence d'une menace. Certains acteurs, en particulier le gouvernement britannique, considèrent cependant qu'avec le développement de la technologie des drones et l'apparition du terrorisme international, une réévaluation du cadre légal est nécessaire. Les aspects strictement éthiques des assassinats ciblés continuent de faire l'objet d'intenses discussions.

III.2.2. Armes autonomes

87. Une arme autonome peut être définie comme une arme qui, une fois activée, sélectionne et attaque des cibles sans intervention humaine supplémentaire. Tant les véhicules armés télépilotes que les systèmes d'armement entièrement autonomes soulèvent certaines questions au regard des principes de distinction, de proportionnalité, de responsabilité et de transparence. Cependant, l'absence d'intervention humaine dans le cas des seconds aggrave ces questions et suscite des inquiétudes supplémentaires sur les plans éthique et juridique. Le point clé est la capacité de logiciels à prendre des décisions qui devraient normalement être prises par un être humain, ainsi que l'acceptabilité éthique du fait que des décisions de vie et de mort soient prises par des machines.

88. En ce qui concerne les capacités des logiciels, il est crucial de reconnaître que les machines – quel que soit la qualité de leur préparation ou leur niveau cognitif – restent des machines, et sont donc intrinsèquement incapables de mettre en œuvre les capacités – souvent nuancées – de l'esprit humain. Un rapport récent de la Royal Society (2017) note la difficulté de mettre au point des machines capables d'aborder un problème de manière contextuelle, c'est-à-dire en faisant usage du « sens commun ». Ce rapport précise : « Lorsque leurs compétences sont prises en défaut, les êtres humains se fient en dernier ressort au sens commun et décident souvent d'agir d'une manière qui, bien que non optimale, ne risque pas d'entraîner de dommages importants. Les systèmes actuels d'apprentissage machine ne définissent ni ne programment ce type de comportement et, par conséquent, lorsqu'ils échouent, cela a parfois des conséquences graves qui révèlent leur fragilité » (The Royal Society, 2017, p.30).

89. Le rapport de la Royal Society remarque en outre que « dans le monde réel existent de nombreuses contraintes que nous connaissons d'après les lois naturelles de la physique, par exemple, ou les lois mathématiques que formalise la logique » mais « intégrer ces contraintes aux méthodes d'apprentissage machine ne va pas sans difficultés » (p.30). De plus,

(...) comprendre les intentions d'un être humain exige un très haut niveau de complexité et une analyse sophistiquée de notre mode de fonctionnement. Les méthodes actuelles n'ont qu'une compréhension limitée des êtres humains, et uniquement dans des domaines particuliers. Cela constituera une source de difficultés dans les environnements collaboratifs, par exemple avec les robots assistants, ou même dans le domaine des voitures sans chauffeur. (The Royal Society, 2017, p.30)

Les difficultés seront évidemment encore plus grandes – compte tenu des conséquences catastrophiques si elles ne sont pas résolues – dans le domaine des armes autonomes.

90. Lors de l'examen des véhicules télépilotes, nous avons abordé le problème de la conformité aux critères de distinction et de proportionnalité, qui tiennent une place essentielle dans le DIH. Dans le cas des armes autonomes, la conformité à ces critères pose des difficultés encore plus graves.

91. En ce qui concerne le critère de *distinction*, les systèmes de différenciation dépendent nécessairement de l'objectif recherché. Si celui-ci est bien défini, un logiciel sophistiqué de

reconnaissance d'images pourra identifier correctement un certain type d'objet mais, en dehors du champ strictement défini, il ne pourra agir de manière fiable. Par exemple, lorsqu'un combattant acquiert de nouveaux droits – en tant que prisonnier de guerre ou soldat blessé –, comment un logiciel tiendra-t-il compte de ces nouveaux droits ? De même, il existe toute une série de directives du CICR (Comité International de la Croix-Rouge) sur les civils directement impliqués dans un conflit armé, directives qui requièrent une interprétation humaine (ICRC, 2007). Comment un logiciel pourra-t-il distinguer un système offensif d'un système défensif du côté adverse ? Même si les systèmes avancés de reconnaissance d'images réussissent à l'avenir à identifier clairement un homme armé, pourront-ils déterminer si cet homme est un combattant ou un policier civil ?

92. Le problème de la reddition d'un combattant à un robot télépiloté a été mentionné plus haut. Avec un système autonome, en l'absence de critères préétablis des modalités de reddition, la réponse à une tentative de reddition ne pourra, par définition, être préprogrammée dans la machine ou apprise par elle. Reconnaître une tentative de reddition exige fondamentalement un jugement humain.

93. Le critère de *proportionnalité* est encore plus difficile à satisfaire par un système autonome. Évaluer une nécessité militaire exige de prendre en compte de très nombreux éléments, dont beaucoup sont par définition impossibles à quantifier. Un système autonome peut-il le faire en l'absence de contrôle humain significatif ? Peut-il déterminer le niveau précis des dommages requis – puis infliger ces dommages – en réponse à une situation particulière ?

94. Décider d'une réponse proportionnée ne se limite pas à estimer le nombre de victimes collatérales éventuelles dans la zone visée par un projectile. Une machine programmée peut-elle décider du nombre de victimes collatérales acceptable ? Comment peut-elle estimer les pertes humaines pouvant résulter, par exemple, de la destruction des infrastructures et de la perte de la capacité à assurer l'alimentation des civils, ou intégrer à l'équation d'autres conséquences sociales et psychologiques ? L'évaluation de la proportionnalité est entièrement une affaire de jugement, qui doit normalement être le fait d'un « commandant militaire raisonnable », ce qui exige une appréhension humaine du contexte. La conclusion inévitable est donc que cette évaluation implique nécessairement un jugement humain et ne peut, d'un point de vue éthique, être laissée à une machine.

95. Pour tuer légalement, un commandant militaire sur le champ de bataille doit agir conformément aux règles du droit international humanitaire (DIH) et cela vaut également pour une machine autonome. On a beaucoup parlé parmi certains roboticiens de la possibilité de créer un « régulateur éthique » programmé pour suivre ces règles, et qui, par conséquent, contrairement à un combattant humain, ne violerait pas ces règles. Le problème est que l'on a le plus souvent affaire ici à des directives générales et non à des règles reposant sur une alternative simple du type oui/non. Pour créer un tel « régulateur », il faudrait que ces directives soient ramenées à des termes simplistes sur la base desquels programmer le robot. Masi il serait inapproprié – ou même dangereux – de tenter de les réduire ou de les simplifier.

96. Un autre élément à évoquer à ce propos concerne le pouvoir de tuer ou de déléguer le pouvoir de tuer. Le DIH exige spécifiquement une décision humaine. La plupart des codes juridiques et moraux partagent à ce sujet un présupposé implicite, à savoir que, lorsque la vie est en jeu, la décision doit être prise par des êtres humains. La Convention de La Haye requiert, par exemple, qu'un combattant soit « commandé par une personne » (Conférence internationale de la paix, 1899, Art. 1 de l'Annexe de la Convention). La clause de Martens, qui est depuis longtemps une règle impérative du DIH, exige spécifiquement l'application des « principes humanitaires » dans un conflit armé (Ticehurst, 1997). Seul un être humain peut outrepasser le droit d'une personne à la vie ; cette décision ne peut être déléguée à une machine. Même s'il était techniquement possible d'introduire les règles du DIH dans la programmation d'un robot indépendamment de modifications juridiques majeures, cela resterait inacceptable.

97. On peut considérer que la délégation du pouvoir du tuer à une machine, aussi complexe que soit sa programmation, revient à nier ce qui constitue le fond de l'être moral d'un individu et sa valeur intrinsèque, à savoir la dignité humaine. La dignité humaine est fragile et doit être respectée et protégée : elle constitue un élément éthique et juridique fondamental, notamment dans la Déclaration universelle des droits de l'homme de l'ONU.

98. Un aspect essentiel de la robotique, qui n'est pas sans relation avec nombre des points évoqués ci-dessus, tient au fait que les robots fonctionnent dans des environnements « structurés » et non « ouverts ». Les robots sont conçus pour remplir une ou plusieurs fonctions spécifiques à l'intérieur d'un environnement défini (ou « structuré ») ou d'une série d'environnements dans lesquels ils ont été préparés à fonctionner. Ils sont généralement performants dans ces environnements mais, lorsqu'on les extrait de ces situations prédéfinies pour les placer dans un environnement plus ouvert, leur comportement devient imprévisible. Comme le montre l'utilisation de certains logiciels sophistiqués de reconnaissance d'images, même des modifications mineures de l'environnement peuvent être cause de dysfonctionnements. On se souvient de l'accident mortel provoqué en 2016 par une voiture Tesla en pilotage automatique lorsque le véhicule s'est trouvé en présence d'une configuration imprévue incluant un camion et une remorque. Il n'est pas possible de programmer tous les environnements et toutes les probabilités dans un système robotique, non plus que de former une machine de façon à couvrir toutes les éventualités ; en dehors de sa zone de confort, son comportement peut devenir catastrophique. Les conséquences potentielles de comportements imprévisibles sont particulièrement inquiétantes dans le cas, par exemple, de deux types de véhicules autonomes qui seraient programmés avec des règles de fonctionnement différentes non connues de l'autre.

99. De nouveaux prototypes de systèmes autonomes sont de plus en plus fréquemment testés à des vitesses supersoniques et hypersoniques. Pour contrer de tels systèmes, des appareils à déclenchement autonome encore plus rapides seront nécessaires, ce qui conduira à son tour au développement de systèmes d'armement toujours plus rapides. La course de vitesse qui en résultera ne laissera guère aux êtres humains de contrôle sur le champ de bataille. Si le développement et la prolifération d'armes autonomes se poursuit, les systèmes de défense (super/hypersoniques) d'un État pourront entrer dans une relation interactive avec les armes autonomes de vitesse équivalente d'un autre État. La vitesse de ce processus imprévisible pourrait conduire par inadvertance au déclenchement d'un conflit armé avant même que des êtres humains aient eu la possibilité de réagir.

100. Il convient de noter dans ce contexte que même des systèmes assez simples défont parfois et qu'une intervention humaine est alors nécessaire pour remédier à leur dysfonction. En outre, il est possible de rendre un système complexe imprévisible par des moyens assez simples, par exemple en introduisant un point dans un circuit imprimé ou en insérant un virus informatique dans le corps de l'appareil au moment de la fabrication ou de la mise à jour d'un logiciel. Les conséquences sur le champ de bataille pourraient être catastrophiques, des événements imprévus faisant rapidement boule de neige et conduisant à une escalade des hostilités sans aucune intervention humaine. Les conséquences du détournement ou de la prise de contrôle de machines autonomes seraient encore plus dramatiques que dans le cas des véhicules télépilotés.

101. Si l'on se place à un autre niveau que celui des principes conventionnels de distinction et de proportionnalité de l'éthique militaire, la question fondamentale qui se pose est celle de l'aptitude des armes autonomes à éliminer le combat et à transformer la guerre d'un conflit simplement asymétrique en une relation unilatérale de mise à mort dans laquelle l'ennemi serait privé de la possibilité même de riposter. Dans une telle éventualité, ces armes échapperaient au cadre normatif conçu initialement pour s'appliquer aux conflits armés.

III.2.3. Robots de surveillance et de police et utilisation de technologies militaires dans des contextes non militaires

102. Il existe de nombreuses formes d'utilisation positive de la robotique dans la surveillance, par exemple la surveillance des travaux agricoles, la surveillance de la faune, la détection des enfants perdus, les missions de sauvetage après une catastrophe et l'évaluation des dommages environnementaux dans les milieux extrêmes. Ces utilisations ne posent guère de problèmes éthiques.

103. Cependant, les véhicules robotiques sont de plus en plus utilisés pour des fonctions de surveillance qui soulèvent des questions potentiellement graves dont les implications éthiques ne sont pas négligeables. Les autres moyens de surveillance – vidéosurveillance, satellites – ont une portée beaucoup plus limitée. En effet, les drones et d'autres véhicules robotiques ne voient pas seulement, ils entendent aussi. Ils peuvent explorer des espaces à trois dimensions (en allant là où les autres techniques de surveillance ne peuvent aller), sont flexibles, constants et fonctionnent 24 jours sur 24 et 7 jours sur 7. Ils permettent donc de combler des lacunes majeures des capacités de surveillance et sont actuellement utilisés par les forces de police européennes pour : le contrôle des frontières, de foules ou d'événements ; la collecte d'éléments de preuve ; le contrôle de la circulation ; l'observation et les perquisitions ; la surveillance des « fauteurs de trouble » ; la surveillance de bâtiments et la protection de VIP ; la détection, le contrôle et la surveillance de migrants sans papiers, de travailleurs et de manifestants. Il en résulte des implications majeures pour le respect de la vie privée et la protection des données car ces nouvelles méthodes de surveillance permettent non seulement de recueillir des données de façon potentiellement massive, mais aussi de mettre en relation et d'utiliser différents types d'information ; en outre, la nécessité de classer les données à caractère personnel conduit à stéréotyper les individus. Une surveillance omniprésente 24 heures sur 24 et 7 jours sur 7 peut aussi nuire à la participation politique dans les démocraties actuelles, tout en constituant une source massive de pouvoir supplémentaire pour les régimes répressifs et dictatoriaux.

104. La surveillance civile se caractérise de plus en plus par le transfert de technologies de sécurité du domaine militaire au domaine civil. Ce transfert n'est pas sans conséquences sur le maintien de la sécurité dans des domaines et des secteurs de la société envisagée traditionnellement dans une optique non militaire, et il a aussi des implications du point de vue des droits de l'homme. L'adaptation et la prolifération de technologies militaires peuvent conduire à une « militarisation » de la société et à l'imposition d'une logique de gestion du monde reposant sur une économie de la peur, celle d'« un ennemi omniprésent parmi nous et pouvant frapper à tout moment » (Crandall and Armitage, 2005, p.20). Le développement de ces dispositifs de surveillance rend possibles non seulement la poursuite de délits spécifiques et le suivi de soupçons concrets mais aussi la surveillance systématique et approfondie d'une population de manière quotidienne. Il y a là quelque chose de similaire au concept militaire C4ISR (commandement, contrôle, communication, renseignement, surveillance et reconnaissance), à savoir la mise en réseau de tous les systèmes de contrôle pour avoir une vue globale du théâtre de guerre. L'effet sur les futures générations d'une surveillance aussi visible et considérée comme normale a des implications fondamentales quant à notre idée de la société future.

105. L'utilisation de robots équipés d'armes dites « non létales » dans la répression des formes de contestation publiques et le contrôle des régions frontières soulève aussi certaines questions. On peut s'inquiéter que les prochaines générations de dirigeants qui hériteront de cette technologie s'en servent comme d'outils d'oppression. En cas d'utilisation de force meurtrière dans la répression des populations, les robots ne risquent pas de désobéir à un régime autoritaire comme pourraient être amenés à le faire, par exemple, des soldats ou des policiers.

106. Il convient de noter qu'un appareil robotique a déjà été utilisé dans un but meurtrier dans un contexte civil. Une force de police d'une ville des États-Unis s'est servie d'un robot

télécommandé de déminage équipé d'un engin explosif sur son bras manipulateur pour tuer un suspect à distance. Le développement des technologies de ciblage automatique fait naître la possibilité de violations des droits fondamentaux et des droits civils par la police et des forces de sécurité privées.

III.2.4. Utilisation privée et illicite de robots

107. Des véhicules téléguidés de petite taille sont déjà en vente sur le marché et le développement rapide de ce type de technologie soulève un certain nombre de questions juridiques, notamment en matière de respect de la vie privée, étant donné les risques d'utilisation par les médias ou par des criminels (en particulier dans le trafic de drogues), ainsi que dans un but de harcèlement ou de voyeurisme.

108. Déployés en l'absence d'une réglementation adéquate, les drones équipés de logiciels de reconnaissance faciale, de technologie infrarouge et de micros permettant d'enregistrer des conversations personnelles constitueraient une atteinte sans précédent au droit au respect de la vie privée. Des drones interconnectés permettraient la surveillance massive de personnes et de véhicules sur des aires très étendues. Des drones de très petite taille pourraient aussi être utilisés pour épier sans être vus à travers la fenêtre d'un domicile ou d'un lieu de culte, par exemple.

109. Le risque d'utilisation de drones, en particulier de drones de petite taille, par des acteurs non étatiques doit aussi être mentionné. Il existe déjà plusieurs cas documentés d'utilisation ou de tentative d'utilisation de drones armés par des groupes non étatiques. Certaines organisations terroristes et des individus déclarant être affiliés à ces organisations ont montré qu'ils étaient prêts à employer des drones pour acheminer des engins explosifs improvisés, ainsi que des armes chimiques ou biologiques. Les responsables de la sécurité ont déclaré qu'avec les progrès de la diffusion et des performances des drones en vente sur le marché, la probabilité que de tels événements se produisent à l'avenir augmente.

110. Le développement plus général de l'utilisation des drones pose aussi des questions de sécurité. La communauté des amateurs de drones pourrait être amenée à déguiser des drones commerciaux pour transporter une charge importante sur de grandes distances. Des cas d'incidents aériens avec des drones télépilotés par des amateurs ont déjà été rapportés. La facilité d'accès de plus en plus grande à la technologie des drones s'est généralement produite en l'absence de réglementation significative des modalités de déploiement de ces appareils.

III.3. Utilisation de robots dans les transports

111. Le secteur des transports est l'un des domaines où la robotique est largement utilisée et bien acceptée, et l'industrie automobile a joué un rôle clé dans le développement de l'automatisation. Les systèmes de transmission automatique, avec la suppression du changement de vitesse manuel, en sont l'un des premiers exemples. Avec l'évolution des capteurs et les progrès technologiques, un plus grand nombre de fonctions de conduite ont été déléguées à des robots, par exemple la régulation de la vitesse, le changement de voie et la mise en stationnement.

112. Le transport de masse, c'est-à-dire principalement les divers systèmes ferroviaires, est en train de devenir pleinement automatisé, le conducteur ne jouant plus qu'un rôle d'appoint, notamment dans les situations d'urgence. Les avions et les navires sont, eux aussi, équipés de systèmes de pilotage automatique qui assistent les pilotes en effectuant la plupart des tâches répétitives.

113. Un véritable saut en avant quant à l'utilisation des robots dans les transports va se produire avec la technologie émergente des véhicules autonomes (VA). Dans un avenir proche, cette technologie permettra non seulement d'assister le conducteur mais de le remplacer complètement. Les VA reposent sur un système de capteurs multimodaux de l'environnement incluant caméras, radar, GPS et détection laser et sur quelques-uns des systèmes d'intelligence artificielle les plus avancés pour le traitement des données et la prise

de décisions. En outre, les VA s'appuient sur des algorithmes cognitifs et peuvent donc devenir des apprenants autonomes.

114. L'utilisation à grande échelle de véhicules autonomes présente de nombreux avantages. D'après le cabinet de consultants McKinsey, elle devrait permettre de réduire de 90 % les accidents automobiles aux États-Unis, économiser jusqu'à 190 milliards USD par an en frais de dommages et dépenses de santé, et sauver des milliers de vies (Bertoncello and Wee, 2015). Parmi les autres avantages potentiels importants, on peut citer l'amélioration de l'efficacité des transports (réduction du temps de déplacement, baisse des coûts) et de la qualité de la vie (autonomie) des personnes dépendantes.

115. Malgré les progrès rapides de la technologie sur laquelle reposent les VA sous l'impulsion de plusieurs grands producteurs et le succès de plusieurs prototypes récents qui ont été présentés au public, la technologie ne sera pleinement opérationnelle qu'une fois résolus les problèmes techniques qui subsistent. Ces problèmes concernent principalement la fiabilité de la technologie (matériel et logiciels) et la cybersécurité.

116. Cependant, les problèmes les plus difficiles associés au développement des VA sont avant surtout d'ordre social. En effet, ce nouveau type de véhicules risque d'entraîner des pertes d'emplois, de creuser l'écart avec les pays en développement et de susciter de nouvelles questions d'ordre éthique, comme par exemple : s'il devient possible de programmer la prise de décision dans des machines sur la base de principes moraux, qu'est-ce qui doit prévaloir : l'intérêt individuel ou le bien public ? (Editor's Summary of Bonnefon et al., 2016)

117. La question de l'opposition entre intérêt individuel et bien public est particulièrement pertinente s'agissant des VA. Comment un véhicule doit-il être programmé à se comporter dans l'éventualité d'un accident impossible à éviter ? La priorité doit-elle être donnée à la réduction maximum du nombre de victimes, même si cela implique de sacrifier les occupants du véhicule, ou à la protection à tout prix des passagers ? Le choix entre ces deux extrémités devrait-il avoir lieu de manière aléatoire ? Bonnefon *et al.* (2016) posent également les questions suivantes :

Est-il acceptable qu'un véhicule autonome, pour éviter une moto, fasse un écart et entre dans un mur, la probabilité de survie en pareil cas étant plus grande pour le passager du véhicule que pour le conducteur de la moto ? La décision doit-elle être différente lorsque des enfants sont à bord du véhicule, puisqu'ils ont plus longtemps à vivre en tant que futurs adultes et moins de contrôle sur la décision qui les a conduits à être présents dans le véhicule ? Si un fabricant offre des versions différentes de son algorithme moral et qu'un acheteur en choisit une en connaissance de cause, l'acheteur doit-il être tenu pour responsable des conséquences néfastes des décisions de l'algorithme ? (p.1576)

118. De telles questions devraient-elles faire l'objet d'une réglementation à l'échelon national ou international ? Soumises à des normes ? À des codes de conduite des véhicules autonomes ? Laisserées au jeu des forces du marché ? L'industrie des VA ne pourra atteindre le stade de la commercialisation si elle n'est pas réglementée. Ce rapport fournira plusieurs recommandations pertinentes à cet égard.

III.4. Utilisation de robots dans le domaine de la santé et du bien-être

III.4.1. Robots médicaux

119. Des robots semi-autonomes sont utilisés en chirurgie (cancer de la prostate) depuis une dizaine d'années. Cependant, leur utilisation continue à être discutée pour diverses raisons. Les partisans de ces robots mettent en avant le confort qui en résulte pour le chirurgien (une plus grande précision, l'absence de tremblement naturel de la main) et certaines études font apparaître une réduction de la durée d'hospitalisation et aussi une perte de sang moindre pour le patient, minimisant ainsi le traumatisme causé par l'incision des tissus. Ces études, cependant, ne détectent pas une différence significative en termes d'efficacité entre la chirurgie habituelle et l'utilisation d'un robot chirurgical (Kappor, 2014).

Les chirurgiens expriment certaines réticences à l'égard de la chirurgie robotique : ils perdent le sens du toucher ; le robot ne leur donne aucune impression tactile, uniquement une image ; le travail du chirurgien devient ainsi entièrement d'ordre visuel (Tzafestas, 2016a). Un problème majeur, en outre, est le coût des robots chirurgicaux et leur entretien. Le coût d'une opération avec un robot est actuellement plus élevé que celui d'une opération avec un chirurgien « normal » (Gagnon, 2018). La multiplication des robots en chirurgie aura des incidences sur l'allocation des ressources dans les systèmes de santé publique, le nombre des chirurgiens et leur aptitude à pratiquer la chirurgie sans robots.

III.4.2. Robots infirmiers

120. Des robots sont également utilisés dans certaines approches thérapeutiques des enfants autistes. Le robot Nao est un exemple de robot utilisé de manière expérimentale pour soutenir les capacités d'adaptation des enfants autistes. Des robots de ce type ont également été utilisés pour améliorer l'apprentissage des enfants trisomiques.

121. L'autisme (ou troubles du spectre autistique) est une maladie sans aucune cause directe identifiée et pour laquelle il n'existe pas de remède. Il se caractérise par divers symptômes décrits par les parents comme un comportement stéréotypé, des mouvements répétitifs, une attention réduite, par exemple. Les enfants atteints d'autisme ont des difficultés à communiquer ; ils parlent sur un ton uniforme, répètent certains mots ou des phrases inappropriées. L'évaluation clinique d'un jeune enfant par une équipe pluridisciplinaire débouche normalement sur une intervention rapide et constante en vue d'améliorer ses capacités linguistiques, son aptitude à l'interaction et son adaptabilité (amélioration de la fonctionnalité sociale). Des robots sont utilisés sur une base expérimentale pour renforcer les capacités d'adaptation de ces enfants (Scassellati *et al.*, 2012). Ceux-ci sont souvent attirés par les robots à cause de la constance de leurs « affects » et du caractère prévisible et répétitif de leur comportement. Un robot n'est pas menaçant parce qu'il n'est pas imprévisible, contrairement aux comportements et aux sentiments des êtres humains.

122. Les études consacrées à l'utilisation de robots dans ce contexte ne sont pas très nombreuses mais aucune d'entre elles n'indique que la dyade robot-enfant autiste est supérieure d'un point de vue fonctionnel à la dyade parent-enfant autiste (Diel *et al.*, 2012 ; Simut *et al.*, 2016). L'utilisation de robots avec des enfants autistes n'est d'ailleurs pas considérée comme une forme de soins standardisée car les patients atteints d'autisme constituent un groupe fortement hétérogène. Certains cliniciens ont signalé les difficultés qui en résultent. Il est probable que les robots ne peuvent avoir un rôle bénéfique pour tous les enfants autistes en raison de la diversité des manifestations de cette maladie. Un enfant autiste entre-t-il effectivement dans une forme d'interaction ou de communication avec un robot ? Communiquer signifie éprouver des sentiments, réagir à ce que l'on éprouve, l'exprimer par des mots, transmettre des affects puis répéter à nouveau ce processus. Les affects jouent un rôle essentiel dans la communication et, dans la vie réelle, ils sont imprévisibles ; leurs paramètres ne peuvent être contrôlés. Améliorer les capacités d'adaptation des enfants autistes par le biais de l'apprentissage de la communication reste le meilleur moyen d'aider ces enfants mais l'interaction entre enfants autistes et robots ne doit pas être conçue comme la seule forme de thérapie possible pour ces enfants (Gagnon, 2018).

123. Les exosquelettes utilisés pour aider les personnes handicapées à accroître leur mobilité (par exemple les prothèses Ekso) s'inscrivent aussi dans l'évolution actuelle et le développement de la robotique médicale. Des robots sont utilisés sur une base expérimentale au stade de la réhabilitation avec les patients victimes d'un traumatisme médullaire ou ayant survécu à un accident vasculaire cérébral. Certains modèles ont également été conçus pour les personnes ayant subi des blessures aux chevilles (Zhang *et al.*, 2013).

124. Cependant, ces appareils robotiques peuvent aussi être utilisés à des fins non thérapeutiques (par exemple, pour améliorer les performances sportives d'athlètes ou l'entraînement de soldats). Les exosquelettes, les implants neurologiques, les nanorobots et d'autres appareils robotiques similaires ouvrent la voie à une transformation du corps et de

l'esprit humains. La question qui se pose sur le plan éthique est de savoir jusqu'où doit aller le développement de l'hybridation et à quelles fins ? Les parties de l'être humain, ainsi que l'être humain dans sa totalité (corps et esprit), doivent-ils être conçus comme quelque chose de modulable qui peut et doit être amélioré, de la même façon qu'un robot, lorsque l'on utilise des appareils robotiques à des fins non thérapeutiques ? Comment justifier cette transformation : par des idées de perfectionnement, de soumission ou de développement de pouvoirs super-humains comme on le voit dans les films ?

III.4.3. Robots de soins pour personnes âgées

125. Un *robot social* peut être défini 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 caractérise tout particulièrement les appareils robotiques que l'on appelle « robots compagnons » (Oost and Reed, 2010).

126. L'utilisation de robots sociaux et de robots compagnons se répand très rapidement dans le domaine des soins de santé et, en particulier, des soins des personnes âgées. Au vu des prévisions démographiques et du vieillissement rapide de la population dans certains pays, et compte tenu également du statut social généralement peu élevé et sexospécifique du personnel de soins, les robots apparaissent comme des appareils fonctionnels capables de remédier à l'écart entre le besoin et l'offre de services de soins.

127. D'une manière générale, on peut classer les robots de soins en trois catégories : les robots qui fournissent une aide physique, les robots compagnons et les robots de surveillance de la santé et de la sécurité. Tous les systèmes robotiques, cependant, ne relèvent pas d'une seule de ces catégories. Certains sont plus directement impliqués dans la fourniture des soins, tandis que d'autres sont conçus comme des robots de compagnie et cherchent à avoir un impact positif de cette façon (Broadbent *et al.*, 2009).

128. Du point de vue médical, les patients âgés les mieux à même de bénéficier de l'interaction avec un robot sont principalement les patients atteints à différents degrés de la démence et les patients présentant différents niveaux de handicap physique.

129. Un robot capable de stimuler les activités cognitives d'un patient atteint d'une maladie neurodégénérative ou d'exécuter de manière quotidienne certaines tâches (nettoyage, cuisine, hygiène personnelle) difficiles à effectuer par une personne âgée peut être bénéfique (Vandemeulebroucke *et al.*, 2017 ; Mordoch *et al.*, 2013). Un objectif majeur est d'assurer la sécurité de ces personnes (risques de chute, défaillance cardiaque, troubles de la mobilité) à leur domicile, à l'hôpital et dans les établissements de soins.

130. Cela soulève la question du caractère approprié des moyens : faut-il utiliser des robots dans de telles situations ? Et à quelle fin ? L'automatisation (caméras, microphones, capteurs) doit-elle être l'option privilégiée ? Comment mettre en balance la sécurité et le contrôle comportemental (respect du traitement médical, cessation de la consommation d'alcool) résultant de l'utilisation de robots et l'autonomie des personnes âgées ? Et qu'en est-il du caractère intrusif de ces appareils et de l'atteinte à la vie privée qui en résulte ? Autrement dit, quel est le rôle du robot ? Par qui est-il conçu et à quelle fin ? Le but est-il d'améliorer la qualité de vie des personnes âgées ? Ces nouveaux robots peuvent-ils aider les personnes âgées à vivre de façon autonome chez elles et prolonger la période de vieillissement sur le lieu de vie en évitant le transfert dans un établissement de soins ? Le but est-il de réduire le travail du personnel de soins ? De décharger la société du soin des personnes âgées ? (Wu *et al.*, 2010)

131. Un autre type de robot, le robot compagnon, peut être utilisé dans un contexte médical pour réduire la solitude des personnes âgées et prévenir les comportements associés à la démence. Certaines études ont démontré les effets de robots – comme Paro le robot phoque et Aibo le chien robot – sur le degré d'interaction sociale, d'attention ou de coopération des patients âgés atteints ou non de démence. Avec Paro, par exemple, le développement d'une

relation affective est possible parce que ce robot imite fortement le comportement d'un animal réel : il respire, dégage une certaine chaleur physique, émet des sons et fait des mouvements qui suscitent une réaction affective, et réagit au toucher, au bruit et à la parole. Pour le moment, cependant, les résultats de l'utilisation de ces robots compagnons dans un contexte médical ou de soins ne sont pas entièrement concluants.

132. Les questions suivantes devraient être prises en compte à cet égard : est-il possible d'adapter les robots aux caractères de la personnalité et aux changements comportementaux d'une personne atteinte de démences ? Les travailleurs du secteur des soins aux personnes âgées sont généralement sous-rémunérés et les robots très coûteux : par conséquent, quelle charge en résultera pour les systèmes publics de santé ? Les robots compagnons provoquent-ils des sentiments faux, illusoire et sans réciprocité, en leurrant et en infantilisant les personnes âgées ? (Sharkey and Sherkey, 2012) Les systèmes robotiques sont-ils capables de fournir des soins adéquats ? Dans quelle mesure les soins nécessitent-ils des relations interpersonnelles, c'est-à-dire des relations entre êtres humains ? Les robots sont-ils capables de dûment respecter les personnes âgées ?

133. Les robots infirmiers peuvent-ils donc être utiles ? Seront-ils acceptés par les personnes âgées et leurs familles ? Cette acceptation signifie que le robot doit être intégré volontairement à la vie d'une personne âgée. Trois conditions essentielles sont requises à cet égard : la motivation à utiliser un robot, une facilité d'utilisation suffisante et l'absence d'inconfort physique, cognitif et affectif en présence d'un robot. D'autres éléments doivent aussi être pris en compte : le coût et l'entretien du robot, et ses effets sur le rôle des aidants naturels. L'acceptation sociale et culturelle des robots de soins varie aussi d'un pays à l'autre.

III.4.4. Robots compagnons

134. Comme indiqué dans la section précédente, les robots sont de plus en plus utilisés comme compagnons. Outre les soins des personnes âgées, il existe un autre domaine d'application courant : celui de la sexualité. Et, comme précédemment, des questions fondamentales se posent au sujet des implications potentielles et du caractère souhaitable de ce type d'utilisation.

135. Deux types de systèmes robotiques se développent aujourd'hui dans le domaine de la sexualité : les systèmes qui rendent possible des rapports sexuels à distance (« télé-dildonics ») et les systèmes qui fonctionnent comme un partenaire sexuel de substitution. Les appareils du premier type sont des systèmes robotiques interactifs permettant d'avoir des rapports sexuels à distance à l'aide d'un manchon interactif et d'un godemiché interactif (Liberati, 2017). Les questions éthiques que soulèvent ces appareils portent sur le risque de réduire les contacts intimes et les rapports sexuels à la seule pénétration, en introduisant un nouveau mélange de distance et de proximité dans les relations intimes.

136. Les robots sexuels se développent également sous la forme de poupées interactives d'apparence réaliste. La première maison de passe robotique, qui offre plusieurs versions d'un robot sexuel appelé Roxxy, a ouvert ses portes à Barcelone. Les défenseurs des robots sexuels mettent en avant leur aptitude à aider des personnes esseulées et frustrées, en prévenant ainsi la souffrance psychologique ou l'inconduite sexuelle. Les robots sexuels pourraient aussi entraîner une baisse de la prostitution. Leurs opposants déclarent qu'ils réduisent les rapports sexuels à la seule gratification individuelle, en potentiellement rendant leurs utilisateurs moins capables de nouer des relations familiales, d'amitié et amoureuses solides. Une autre évolution préoccupante sur le plan éthique est la potentielle incitation à la violence sexuelle ainsi qu'aux comportements pédophiles qui se manifestent par l'utilisation de robots sexuels ayant l'apparence d'enfants ou programmés pour ou programmés pour être abusés.

III.5. Utilisation de robots dans le domaine de l'éducation

137. Avec le développement rapide des technologies de communication, des outils multimédias de plus en plus nombreux sont utilisés dans le domaine de l'éducation, y compris

des robots éducatifs. La robotique éducative permet l'exploration, la conception, la modélisation, la programmation, la construction et la mise à l'épreuve de concepts d'apprentissage unitaires (mouvement, force, traction, par exemple) mais aussi de systèmes réalistes plus complexes exigeant la combinaison de concepts et de méthodologies relevant de plusieurs disciplines. Elle soutient les activités d'apprentissage individuelles et en collaboration, en prenant en compte des objectifs des programmes et des compétences différentes. Les objectifs d'apprentissage sont en général de deux types (Eguchi, 2012). Le premier consiste à utiliser les robots pour susciter chez les enfants le désir d'en savoir plus sur le monde de la technologie, en intégrant à l'enseignement des cours et des activités visant à les familiariser avec les robots. Le deuxième est de développer de nouveaux projets en se servant de robots comme outils d'apprentissage pour faire participer les enfants à des activités et leur impartir des notions difficiles à enseigner avec les méthodes classiques. Des outils robotiques éducatifs à faible coût, équipés de capteurs et de régulateurs simplifiés, peuvent maintenant être utilisés en classe. Les thèmes d'apprentissage de la robotique éducative sont généralement décrits comme incluant : l'intérêt pour la science et l'ingénierie (certains programmes de robotique éducative semblent un moyen prometteur d'améliorer le taux de rétention des élèves de sexe féminin qui sont sous-représentées dans les matières technologiques), le travail en équipe et la résolution de problèmes (Miller *et al.*, 2008).

138. Les robots servent à des fins éducatives ou remplissent des fonctions pédagogiques différentes en classe. Un robot humanoïde est, par exemple, utilisé sur une base expérimentale pour enseigner une deuxième langue dans le primaire. Ce robot provoque l'interaction en classe en répétant des mots et des phrases et contribue à renforcer la motivation des élèves (Chang *et al.*, 2010). Des robots sont aussi utilisés pour produire des récits et inciter les élèves du primaire à s'exprimer et à créer en collaboration avec eux (Sugimoto, 2011). Le plus souvent, l'utilisation d'un robot en classe a pour but d'aider les élèves à comprendre des processus scientifiques ou des concepts mathématiques.

139. Parmi les études récentes de l'utilisation de robots à des fins éducatives, seul un petit nombre présentent pour le moment des résultats empiriques (quantitatifs et qualitatifs) utiles pour comprendre leur impact concret dans ce domaine (Benetti, 2012 ; Toh *et al.*, 2016). L'une des principales questions qui se pose dans ce contexte concerne le rôle assigné au robot : doit-il être considéré comme un précepteur, un outil ou un pair ? Dans quelle mesure un robot peut-il remplacer un enseignant ? Certains enfants attribuent des caractéristiques cognitives, comportementales et affectives aux robots (Beran *et al.*, 2011). Par conséquent, la question de l'attachement des enfants aux robots et de la transformation de leur relation à l'enseignant est également un domaine de préoccupation. La vulnérabilité des enfants du primaire à une relation affective simulée a, par exemple, été soulignée à cet égard.

140. Du point de vue de la conception de l'éducation, l'utilisation de robots s'inscrit dans un modèle d'apprentissage constructiviste, qui implique une participation active des élèves à la construction et à l'acquisition des connaissances. Les activités de robotique éducative engagent les apprenants à travailler en collaboration lorsque le niveau de complexité exige d'eux la résolution de problèmes (Kamga *et al.*, 2016). Toutefois, malgré l'opportunité qu'offre la robotique éducative en termes de participation des apprenants, on peut se demander si les robots stimulent véritablement la motivation des enfants ou s'ils ne font qu'éveiller chez eux une curiosité et un intérêt temporaires. Dans quelle mesure le processus d'apprentissage avec un robot doit-il prendre une forme récréative pour induire une motivation chez les enfants ? De plus, la question de l'accès à la robotique conduit à s'interroger sur le risque de création de nouvelles inégalités et de marginalisation de certains enfants. Ces questions méritent d'être examinées.

141. 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. Dans une enquête réalisée en 2012 par le réseau TNS Opinion & Social (EC, 2012), 34 % des répondants européens se sont déclarés

favorables à l'interdiction de l'utilisation de robots dans le domaine de l'éducation, alors que 3 % seulement ont déclaré que l'introduction de robots dans l'éducation devrait constituer une priorité. Au Japon et en Corée du Sud, l'utilisation de robots dans l'éducation semble faire l'objet de plus nombreuses expérimentations et être mieux acceptée d'un point de vue culturel. Les traditions religieuses shintoïstes du Japon () ont favorisé le développement d'images positives des robots, qui sont perçus comme des objets harmonieux, jolis, sympathiques ou même curatifs. Les robots sont considérés comme des machines permettant le développement de liens affectifs aptes à favoriser des relations harmonieuses avec les êtres humains (Tzafestas, 2016a). Par conséquent, savoir s'il faut ou non introduire des robots dans le système éducatif dépend de certaines valeurs culturelles et d'un choix à la fois public, économique et politique.

III.6. Robots ménagers

142. Les robots ménagers – également appelés « robots de service » ou « robots domestiques » – ont pour but d'aider les êtres humains à exécuter certaines tâches considérées comme sales ou ennuyeuses comme passer l'aspirateur, ramasser les poubelles, nettoyer les vitres, arroser les plantes, nettoyer la piscine, repasser ou préparer à boire et à manger. Les appareils tels que les bacs à litière automatiques pour chats, les robots alarmes et de sécurité, les tondeuses à gazon robotiques, les robots de surveillance d'animaux domestiques, les berceaux robotiques et les assistants d'achat robotiques sont aussi souvent classés dans la catégorie des robots ménagers.

143. Bien qu'ils ne soient pas considérés uniquement comme des robots ménagers parce qu'ils apparaissent fréquemment dans d'autres contextes (par exemple, les crèches et les maisons de retraite), les jouets robotiques et les robots compagnons ou de divertissement peuvent raisonnablement être classés dans la catégorie des robots ménagers. La raison principale de cette classification est le fait que de nombreuses activités humaines comme le jeu, la compagnie, le divertissement ou divers passe-temps, que peuvent soutenir ou renforcer les robots, ont lieu généralement dans l'espace de la maison ou du foyer.

144. La vie domestique ou ménagère est l'un des domaines où l'utilisation de robots et d'appareils robotiques ne cesse de se développer. D'après les prévisions de la Fédération internationale de la robotique, « les ventes de toutes les catégories de robots servant à des tâches domestiques (passer l'aspirateur, tondre la pelouse, nettoyer les vitres ou autres) pourraient atteindre près de 31 millions d'unités pendant la période 2016-2019 » (IFR, 2016b, p.3). Ce chiffre est encore plus élevé si on inclut dans la catégorie des robots ménagers d'autres types de robots comme les jouets robotiques, les robots de divertissement et divers systèmes robotiques servant de passe-temps.

145. Il est hors de doute que les robots ménagers et les appareils robotiques ont amélioré et vont continuer à améliorer la vie et la condition humaines en exécutant des tâches que les êtres humains trouvent en général ennuyeuses, sales ou simplement fatigantes. Ils permettent à leurs utilisateurs d'économiser du temps et de l'énergie et de se consacrer à des activités plus intéressantes et plus épanouissantes. Ce type de robots devrait également contribuer à réduire les différences entre les genres qui, dans la plupart des sociétés et des cultures, restent très marquées et l'inégalité entre les femmes et les hommes eu égard aux tâches ménagères.

146. Comme beaucoup d'autres types de robots, cependant, les robots ménagers introduisent des risques et des dangers nouveaux et suscitent des enjeux juridiques et éthiques particuliers, qui devraient être pris au sérieux par leurs concepteurs, leurs fabricants et leurs utilisateurs. L'un des risques les plus évidents concerne les dommages dont ils peuvent être la cause non seulement pour leurs utilisateurs directs mais aussi pour d'autres personnes vivant sous le même toit. De tels dommages peuvent être la conséquence d'une force ou d'une vitesse inappropriée, d'une incompréhension du mode d'emploi ou d'une erreur dans l'exécution de tâches spécifiques (en particulier les tâches délicates comme jouer avec des enfants, prendre soin d'animaux domestiques ou préparer à boire ou à manger).

147. Un danger fréquemment mentionné à propos de nombreux types de robots ménagers est le fait qu'ils sont assez faciles à détourner à des fins non éthiques ou même criminelles. De nombreux robots ménagers étant équipés de capteurs tels que caméras et micros, il est facile de les transformer en outils pour épier un individu ou s'introduire dans sa vie privée et son intimité. En outre, ces robots, s'ils deviennent une présence permanente au foyer, auront accès à et pourront stocker une énorme quantité de données privées et confidentielles (par exemple, photos de l'intérieur d'une maison et de ses habitants, données sur leurs habitudes, mot de passe de leur système d'alarme, emplacement de leurs objets précieux). Le danger s'accroît encore lorsque les robots ménagers sont connectés à l'internet ou à un autre réseau insuffisamment protégé et facile à pénétrer. Dans ce type de situations, le piratage du robot peut aussi faciliter des activités criminelles comme le vol ou le chantage. Les concepteurs, fabricants et vendeurs de robots ménagers et des technologies connexes devraient réfléchir soigneusement aux mesures de précaution contre de tels abus.

148. Le mésusage ou le mauvais fonctionnement de robots ménagers de jeu ou de divertissement peut avoir des conséquences encore plus graves, en particulier si leurs utilisateurs sont des enfants ou des personnes âgées (deux groupes très vulnérables à la tromperie et à l'exploitation, intentionnelles ou non). Un degré particulièrement élevé de sensibilité aux risques et d'anticipation des dangers et mésusages possibles devrait être requis des concepteurs et des fabricants de jouets robotiques et de robots compagnons conçus pour être utilisés à l'intérieur d'une maison sans la supervision d'un personnel qualifié (comme cela est le cas dans les crèches, les établissements éducatifs, les hôpitaux et les maisons de retraite).

149. L'apparence extérieure des jouets robotiques et des robots compagnons est aussi un aspect sensible d'un point de vue éthique, qui exige une soigneuse réflexion. Pearson et Borenstein (2014), par exemple, ont identifié plusieurs domaines dans lesquels les considérations éthiques devraient être très sérieusement prises en compte lors de la conception de jouets robotiques pour enfants. Leur forme et leur apparence devraient, en particulier, refléter les préférences esthétiques collectives d'une culture, ne pas renforcer les stéréotypes de genre (comme cela est le cas avec les jouets robotiques d'apparence excessivement sexospécifique), tenir compte du contexte et être adaptées au niveau de développement des enfants. Il convient aussi de prêter une attention particulière au degré d'apparence humaine de ces jouets, l'impact positif ou négatif d'une telle apparence variant en général selon l'âge ou les traits de personnalité de leurs utilisateurs.

III.7. Agriculture et environnement

150. Les robots sont présents dans l'agriculture, par exemple dans le secteur laitier où des vaches en liberté dans une étable peuvent être traitées par un robot. Concrètement, ce robot commence par identifier une vache et à recueillir des données sur sa santé (température, niveau d'hormones, infections, etc.) et sur sa production laitière. Puis il localise ses mamelles et la traie pendant qu'elle reçoit un supplément alimentaire. Une fois que les vaches sont formées à aller vers le robot, elles peuvent être traitées trois fois par jour, ce qui constitue un gain de productivité pour les agriculteurs. Un autre robot s'occupe de nettoyer l'étable. La productivité de chaque vache est suivie par le robot qui adapte son alimentation en fonction du niveau de production et de l'âge de l'animal (la vie productive d'une vache à lait est d'environ quatre ans, après quoi on l'envoie normalement à l'abattoir). Il en résulte aussi pour l'agriculteur un gain de temps et une plus grande flexibilité. L'utilisation de robots de traite réduit le travail physique et le nombre de personnel nécessaires. Pour être productif, ce type d'organisation robotisée du travail requiert environ 60 à 70 vaches par robot. Le coût de ces robots est encore très élevé (environ 200.000 USD) et de nombreux agriculteurs doivent, pour les utiliser, adapter leur étable, ce qui entraîne de nouveaux investissements. L'installation de robots exige un entretien professionnel régulier. Elle nécessite aussi une longue période d'adaptation des vaches ; en définitive, celles qui refusent de s'adapter doivent être conduites directement au robot par l'éleveur. Certaines vaches peuvent aussi s'opposer au retrait du système à cause de l'autonomie qui en résulte pour elles. L'introduction de robots a modifié la

conception de l'animal, qui repose maintenant sur les nombreuses données transmises par le robot. L'agriculteur se fie au robot pour les décisions à prendre au sujet de chaque vache et s'appuie donc moins sur sa connaissance propre de l'animal. La relation avec l'animal s'en trouve également transformée. L'exposition des vaches au contact d'êtres humains est réduite au minimum et certaines d'entre elles perdent toute familiarité et relation de réciprocité avec les êtres humains (communication orale, échanges affectifs, étreintes affectueuses). La question du bien-être de l'animal, qui implique un certain degré d'individualisation, est en jeu. L'utilisation de robots a encore accentué le modèle de productivité reposant sur l'idée de l'animal-machine. Il n'existe pas encore d'analyse de l'interaction animal-robot (Driessen and Heutinck, 2015 ; Holloway *et al.*, 2014).

151. Les drones peuvent aussi contribuer à une agriculture durable, ainsi qu'à l'aquaculture. Cette application technologique est liée à l'agriculture de précision. Les drones obtiennent des données qui peuvent être analysées pour une utilisation plus efficace d'intrants chimiques (pesticides et engrais) or water (irrigation goutte à goutte). Ils permettent aussi la sélection de traits intéressants des plantes dans les champs (par exemple la tolérance à la sécheresse, la salinité ou les stress, la résistance aux parasites et maladies) afin d'utiliser les plantes sélectionnées dans les programmes d'amélioration des cultures pour faire face des défis du changement climatique. La contribution à la sécurité sanitaire et la production des cultures est haute, car ils renforcent les rendements agricoles. Le coût de la technologie est variable, mais elle devient plus accessible à temps, même pour les petits agriculteurs dans les pays en développement dépendant de leur conditions et spécifications. De même que dans le secteur laitier, la relation de l'agriculteur avec ses terres et ses cultures, et la connaissance qu'il en a, se trouvent transformées. Le modèle d'agriculture intensive de précision induit une modification de la relation à la terre qu'avait analysée il y a déjà longtemps Aldo Leopold (1949). L'utilisation de drones dans l'agriculture renforce encore le modèle de l'agriculture productive.

152. Il existe un certain nombre de domaines où les robots peuvent être utiles d'un point de vue environnemental, par exemple le recyclage, la surveillance et la restauration de l'environnement et le nettoyage après un accident nucléaire ou chimique (Lin, 2012). L'utilisation de robots dans la recherche océanique en eau profonde et l'exploration spatiale a apporté d'importantes contributions à la science environnementale. Toutefois, les avantages potentiels de l'utilisation de robots doivent être mis en balance avec l'impact environnemental de l'ensemble du cycle de production des robots. Celui-ci inclut l'extraction d'éléments terrestres rares et d'autres matières premières, l'énergie nécessaire pour produire et alimenter les machines et les déchets générés pendant la production des robots puis au terme de leur cycle de vie. Il est probable que la robotique va renforcer les préoccupations que suscitent le volume croissant de déchets électroniques et l'amenuisement des ressources en éléments terrestres provoqué par l'industrie informatique (Wildmer *et al.*, 2005 ; Alonso *et al.*, 2012). Outre leur impact sur la santé et sur l'environnement, les déchets électroniques ont d'importantes implications sociopolitiques, notamment à cause de leur exportation vers des pays en développement et de leur déversement à proximité des lieux d'habitation de populations vulnérables (Heacock *et al.*, 2016). Il ne semble guère exister d'analyses de l'impact environnemental ou de l'empreinte écologique d'un robot.

IV. RÉGLEMENTATION ÉTHIQUE ET LÉGALE

153. La robotique et les robots soulèvent des enjeux juridiques et éthiques nouveaux et sans précédent. Ces enjeux tiennent au caractère pluridisciplinaire de la robotique et, spécifiquement, des diverses sciences et technologies impliquées dans la conception, la fabrication, la programmation et l'application des robots. La création d'un robot nécessite l'apport de nombreux spécialistes appartenant à des disciplines très diverses comme l'ingénierie électrique/électronique, l'ingénierie mécanique, l'informatique et même la biomédecine. Selon l'utilisation prévue d'un robot, la catégorie à laquelle il appartient, son degré de complexité et le type d'environnement où il doit être utilisé, il peut être nécessaire de

faire appel au concours d'un éventail de spécialistes encore plus large (par exemple, spécialistes de l'intelligence artificielle, chercheurs en sciences cognitives, ingénieurs aéronautiques, psychologues, dessinateurs industriels, artistes).

154. La création d'un robot passe par un certain nombre d'étapes, notamment : la définition de sa fonction ou de ses tâches, la définition des modalités d'exécution de ces tâches, la mise au point des actionneurs et des effecteurs, le choix des matériaux à utiliser dans la construction du robot, la création et le test du prototype, et la conception et la programmation du logiciel. Le sous-domaine émergent de la « biorobotique », par exemple, recourt, outre aux technologies et matériaux habituels de la robotique, à des tissus organiques et à des organismes vivants (animaux de laboratoire), ce qui soulève un ensemble de questions légales et éthiques particulières.

155. Étant donné le caractère complexe de la conception, de la construction et de la programmation des robots, la question la plus souvent mise en avant d'un point de vue juridique et éthique est celle de leur « traçabilité ». Ce terme désigne la nécessité de pouvoir déterminer précisément d'un point de vue technique les causes de toutes les actions (ou omissions) antérieures d'un robot. La traçabilité est importante sur les plans éthique et juridique dans la mesure où « les circuits de décision d'un robot doivent pouvoir être reconstruits aux fins de la résolution des litiges et des différends » (Riek and Howard, 2014, p.6).

156. Cependant, l'exigence de traçabilité n'est pas nécessairement compatible avec le développement de robots dotés d'un degré élevé d'autonomie, de capacités de décision et d'aptitudes à l'apprentissage. Le développement de robots possédant de telles propriétés va néanmoins se poursuivre car ces propriétés les rendent mieux aptes à exécuter différentes tâches actuellement effectuées par des êtres humains. Il n'en reste pas moins que les processus d'apprentissage automatique que permettent les nouveaux algorithmes, et sur lesquels repose l'autonomie d'un robot et ses capacités de décision, réduisent la possibilité de déterminer précisément l'origine du comportement d'un robot. En effet, ce type de robot est programmé non seulement pour exécuter des tâches spécifiques mais pour apprendre et continuer à se développer en relation avec l'environnement. La capacité des robots à « apprendre » et à « se programmer eux-mêmes » nécessitera donc de modifier les notions juridique et éthique de traçabilité.

157. Comme l'observent à juste titre Matsuzaki et Lindemann (2016), la robotique doit poursuivre « deux tâches contradictoires : accroître l'autonomie des robots et, en même temps, assurer qu'ils sont sans risques » (p.502). Ce problème est inhérent à la robotique et il est difficile de le traduire dans des catégories et normes juridiques.

158. On trouvera ci-dessous un bref aperçu sélectif des réflexions et propositions les plus importantes émanant d'individus, d'institutions ou d'organisations en vue de la réglementation éthique et/ou légale de la robotique contemporaine et de l'industrie des robots.

159. Au niveau de la recherche universitaire individuelle sur les enjeux éthiques associés aux robots et à la robotique, une discussion animée se poursuit depuis près de deux décennies sous des étiquettes comme l'« éthique de la robotique », l'« éthique des robots » ou la « roboéthique ». Le terme fréquemment utilisé de « roboéthique », forgé initialement par Gianmarco Veruggio (2002), désigne ce qu'on pourrait appeler le « code d'éthique des roboticiens ». « L'objet de la roboéthique », selon Veruggio et Operto (2008), « n'est pas le robot et son éthique artificielle mais l'éthique humaine des concepteurs, fabricants et utilisateurs de robots » (p.1501). La roboéthique, en tant que composante de l'éthique appliquée s'occupant de la manière dont les êtres humains conçoivent, construisent et utilisent des robots, ne doit pas être confondue avec l'« éthique des machines », la discipline consacrée aux moyens d'intégrer des règles ou codes éthiques dans la programmation des robots, qui sera examinée plus loin dans ce rapport.

160. En dépit du nombre et de la grande diversité des discussions sur les implications éthiques actuelles de la robotique, peu de chercheurs ont développé des propositions

formulant précisément les principes fondamentaux qui devraient être ceux d'une future « éthique de la robotique ». Néanmoins, deux propositions méritent d'être mentionnées : celle d'Ingram *et al.* (2010) et celle de Riek et Howard (2014).

161. Dans leur « code d'éthique des ingénieurs en robotique », Ingram *et al.* (2010) déclarent qu'un ingénieur en robotique soucieux d'éthique doit garder présent à l'esprit le bien-être des communautés locales et nationales et de la communauté mondiale, ainsi que le bien-être des ingénieurs roboticiens, des clients, des utilisateurs ultimes et des employeurs. Le premier principe de ce projet de code d'éthique affirme :

Un ingénieur en robotique se doit, autant qu'il est possible, de prévoir d'éventuelles utilisations non éthiques de ses créations et de limiter les possibilités d'utilisation non éthique. Un ingénieur en robotique soucieux d'éthique ne peut prévenir tous les risques potentiels et les utilisations non souhaitables de ses créations mais il doit faire tout ce qui est en son pouvoir pour les réduire au minimum. Cela peut inclure : introduire des dispositifs de sécurité supplémentaires, signaler un danger à autrui ou refuser entièrement les projets dangereux. (p.103)

Ce projet de code d'éthique aborde aussi des questions éthiques telles que le respect du bien-être physique et des droits des individus, les règles à envisager pour prévenir la mésinformation et les conflits d'intérêts, l'obligation de critique constructive et le développement personnel.

162. Le « code d'éthique des professionnels de l'interaction homme-robot » de Riek et Howard (2014) cherche à définir des obligations spécifiques s'appliquant non seulement aux roboticiens mais aussi à toutes les personnes généralement impliquées dans la robotique et la fabrication ou la commercialisation de robots, comme les chefs de produit, les distributeurs, les travailleurs de l'industrie des robots et les agents publics concernés. Il propose la « directive primordiale » suivante :

Toutes les activités de recherche, de développement et de commercialisation concernant l'interaction homme-robot doivent être fondées sur le principe général du respect de la personne humaine, y compris le respect de l'autonomie humaine, le respect de l'intégrité corporelle et mentale des êtres humains, et la reconnaissance de tous les droits et protections normalement présupposés dans les relations entre êtres humains. (p.5)

Ce code mentionne aussi des principes à caractère plus spécifique tels que la nécessité de respecter les besoins affectifs humains, le droit au respect de la vie privée, la prévisibilité souhaitable du comportement des robots, l'existence de dispositifs d'interruption (coupe-circuit) et la nécessité de limiter les aspects humanoïdes de la morphologie des robots.

163. Comme le notent Ingram *et al.* (2010), les professions et disciplines particulières qui font partie de la robotique ou y contribuent de manière significative se sont généralement dotées de codes de conduite conséquents comme, par exemple, le code d'éthique de l'IEEE (Institute of Electrical and Electronics Engineers), le code d'éthique de l'ASME (American Society of Mechanical Engineers), le code d'éthique de l'ACM (Association for Computing Machinery) et le code d'éthique de l'ABET (Accreditation Board for Engineering and Technology). Néanmoins, les codes d'éthique spécifiquement conçus à l'intention des roboticiens en sont encore à leurs débuts. Un projet initial qui nous paraît positif est l'« Initiative de l'IEEE pour l'examen des considérations éthiques relatives à l'intelligence artificielle et aux systèmes autonomes » en cours de développement (IEEE Standards Association). Étant donné que la conception, la fabrication et l'utilisation de robots concernent aujourd'hui de nombreuses autres professions (infirmiers, chimistes, médecins, biochimistes, biologistes, par exemple), il serait tout à fait indiqué que d'autres organisations professionnelles suivent cet exemple, dans le but ultime de mettre au point un cadre éthique commun aux diverses disciplines.

164. Les projets de recherche en robotique menés dans des universités et des organisations de recherche doivent généralement être approuvés par une commission institutionnelle d'évaluation ou une commission d'éthique mais il n'existe pas de directives éthiques spécifiques sur les modalités de mise en œuvre de ces projets, en particulier ceux qui ont des incidences directes ou indirectes sur les êtres humains. N'oublions pas non plus que l'industrie de la robotique est devenue un secteur d'activité extrêmement lucratif et que l'expérience montre que les codes de conduite et les directives éthiques sont fréquemment perçus comme un obstacle aux activités de recherche et développement.

165. En dépit de l'absence de codes de conduite universellement reconnus pour les roboticiens, certaines initiatives institutionnelles et organisationnelles importantes en vue d'une réglementation éthique – ou pertinente d'un point de vue éthique – de la robotique existent et méritent d'être mentionnées.

166. Selon Veruggio et Operto (2008), le gouvernement de la République de Corée a créé en 2007 un groupe de travail chargé d'élaborer une charte de la roboéthique. Cette charte devait être présentée au Forum sur les politiques relatives à l'industrie des robots et au Conseil pour le développement industriel, après quoi des normes et lignes directrices spécifiques devaient être développées. Toujours selon Veruggio et Operto (2008), le ministère japonais de l'Économie, du Commerce et de l'Industrie a lancé en 2007 une initiative de travail sur un ensemble très détaillé de lignes directrices concernant le déploiement des robots. Dans ce travail, l'accent était mis avant tout sur le besoin de sécurité à toutes les étapes de la conception, de la fabrication, de la gestion, de la réparation, de la vente et de l'utilisation des robots. Il était recommandé, par exemple, que tous les robots « signalent à une base de données centrale tous les cas de blessures infligées à des personnes qu'ils sont censés aider ou protéger » (Veruggio et Operto, 2008, p. 1506). On ignore à ce jour si la charte coréenne ou les lignes directrices japonaises ont été finalisées.

167. En 2016, la British Standards Institution a publié un document sous le titre *Robots and robotic devices : Guide to the ethical design and application of robots and robotic systems* [Robots et appareils robotiques : Guide pour une conception et une application éthiques des robots et des systèmes robotiques]. Ce document, rédigé par un certain nombre de spécialistes de différents domaines, aborde des questions comme : les responsabilités relatives au comportement des robots, la vitesse et la force maximales des robots, les normes éthiquement acceptables de conception des robots industriels, médicaux et de soins personnels, le développement de liens affectifs avec des robots, la non-nécessité d'un anthropomorphisme des robots, l'apprentissage des robots, les implications sexistes ou racistes de la robotique, le respect de la vie privée, la confidentialité des données et la déshumanisation associée à l'utilisation de robots. Conçu à l'intention des concepteurs et opérateurs de robots, ce document de la BSI fournit « des indications utiles aux fins de l'identification des risques éthiques potentiels associés à la conception et à l'application des robots, ainsi que des directives sur les mesures de sécurité, de protection et d'information pertinentes » (BSI, 2016, p.1).

168. En ce qui concerne l'éthique de la recherche en robotique, Allistene, un consortium d'instituts de recherche français (CNRS, CEA, INRA, CDEFI, etc.) a formulé plusieurs propositions pratiques visant à renforcer la responsabilité des chercheurs en robotique sur le plan éthique et juridique. Ce document insiste sur la nécessité d'anticiper l'interaction homme-robot. Les chercheurs devraient aussi réfléchir aux limites de l'autonomie des robots et à leur capacité à imiter les relations affectives et sociales entre êtres humains (Allistene, 2016).

169. L'Organisation internationale de normalisation (ISO) a émis pendant les dernières années un certain nombre de normes relatives aux robots et à la robotique. La plupart de ces normes sont évidemment de nature technique et, par conséquent, axées principalement sur la fiabilité et la qualité des produits. Néanmoins, les normes de l'ISO se rapportant aux robots et à la robotique ont au moins une pertinence éthique et légale indirecte dans la mesure où

elles visent à favoriser une standardisation des appareils robotiques du point de vue de la sécurité des utilisateurs. On peut citer, en particulier, la norme ISO 13482: 2014 (Robots et composants robotiques – Exigences de sécurité pour les robots de soins personnels), la norme ISO 10218-2: 2011 (Robots et dispositifs robotiques – Exigences de sécurité pour les robots industriels – Partie 2 : Systèmes robots et intégration) et, plus généralement, bien qu'elle ne porte pas spécifiquement sur la robotique, la norme ISO 13850: 2015 (Sécurité des machines – Fonction d'arrêt d'urgence – Principes de conception).

170. Le *Projet de Rapport contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publié en 2016 par la Commission des affaires juridiques du Parlement européen, est l'une des initiatives les plus concrètes à ce jour en vue de la réglementation légale et éthique de la robotique et de l'industrie des robots. Ce document présente une proposition de résolution à adopter par le Parlement européen et inclut, en annexe, une Charte sur la robotique comprenant un Code de conduite éthique pour les ingénieurs en robotique, un Code de déontologie pour les comités d'éthique de la recherche, une Licence pour les concepteurs et une Licence pour les utilisateurs. Préoccupé par l'impact possible de la robotique sur la sécurité, la vie privée, l'intégrité, la dignité et l'autonomie des êtres humains, ce rapport aborde un certain nombre de questions éthiques et juridiques touchant à la robotique et à l'utilisation des robots, notamment : la protection des données et de la vie privée, les dommages qui pourraient être causés par la nouvelle génération de robots et la responsabilité des fabricants de robots, le principe de précaution, le test des robots en milieu réel, le consentement éclairé dans la recherche en robotique impliquant des êtres humains, les dispositifs d'interruption (coupe-circuit) et l'impact de la robotique sur l'emploi et d'éducation. Il souligne également l'importance de la traçabilité et de la mise en place d'un système d'immatriculation des robots avancés : « Il y a lieu de créer un système d'immatriculation des robots avancés sur la base de critères de classification des robots, à des fins de traçabilité » (JURI, 2016, p.13).

171. En 2017, le Rathenau Instituut a publié un rapport sur « Les droits de l'homme à l'ère des robots », commandité par l'Assemblée parlementaire du Conseil de l'Europe (APCE). Ce rapport part de l'idée qu'avec les nouvelles technologies, « les frontières se brouillent entre l'homme et la machine, entre les activités en ligne et hors ligne, entre le monde physique et le monde virtuel, entre le naturel et l'artificiel et entre la réalité et le virtuel » (Rathenau Instituut, 2017, p.10). Il examine de ce point de vue les impacts potentiellement négatifs de la robotique et des technologies apparentées – en particulier la convergence entre nanotechnologies, biotechnologies, technologies de l'information et sciences cognitives (NBIC) – dans un certain nombre de domaines relatifs aux droits de l'homme : le respect de la vie privée, la dignité humaine, la propriété, la sécurité et la responsabilité, la liberté d'expression, l'interdiction de la discrimination, l'accès à la justice et l'accès à un procès équitable. Il prête ensuite une attention particulière aux véhicules autonomes et aux robots de soins, en recommandant au Conseil de l'Europe de réfléchir à l'élaboration d'outils juridiques nouveaux et/ou plus pertinents pour réglementer l'utilisation de ce type d'appareils. Il recommande également l'introduction de deux nouveaux droits fondamentaux : le droit de ne pas faire l'objet de mesures, d'analyses ou de contrôles (en relation avec le mésusage possible de l'IA, de la collecte de données et de l'internet des objets) et le droit à un contact humain significatif (en relation avec les mésusages possibles, intentionnels ou non, des robots de soins).

172. Les approches théoriques et la recherche sur les enjeux juridiques associés à la robotique, aux robots et à l'industrie des robots semblent être en plein essor et il existe aujourd'hui nombre de contributions intéressantes en ce domaine (Matsuzaki and Lindemann, 2016 ; Calo, 2015 ; Asaro, 2015 ; Holder *et al.*, 2016 ; Leenes *et al.*, 2017 - pour ne citer que quelques-unes des plus récentes). Cependant, à un niveau plus concret, tant sur le plan national qu'international, la robotique, l'industrie des robots et le marché des robots restent en grande mesure non légalement réglementés. Selon une analyse par pays figurant dans le manuel de droit comparé intitulé *Comparative handbook: Robotic technologies law* (Bensoussan and Bensoussan, 2016), dans des pays comme l'Afrique du Sud, l'Allemagne,

la Belgique, le Brésil, la Chine, le Costa Rica, l'Espagne, les États-Unis, la France, la Grèce, Israël, l'Italie, le Japon, le Liban, le Portugal, le Royaume-Uni et la Suisse, il n'existe aucune législation spécifique sur les robots (mais, dans plusieurs pays, certaines technologies robotiques – par exemple les drones et les véhicules autonomes – sont effectivement réglementées).

173. La raison principale pour laquelle la robotique demeure légalement sous-réglémentée est probablement qu'il s'agit d'un domaine de recherche nouveau, qui évolue rapidement, et dont l'impact sur le monde réel et les conséquences juridiques sont difficiles à conceptualiser et à prévoir. Selon Asaro (2012), tous les dommages potentiels causés par les robots et la technologie robotique sont actuellement couverts par la législation civile sur la responsabilité des produits ; autrement dit, les robots sont traités comme n'importe quel autre produit technologique (jouets, voitures ou armes, par exemple). Le préjudice éventuel entraîné par l'utilisation d'un robot relève des normes sur la responsabilité des produits comme la « négligence », le « défaut de mise en garde » ou le « défaut de précaution ». Toutefois, l'application de ces normes de droit civil à l'industrie des robots risque de se révéler inadéquate au fur et à mesure qu'augmenteront la sophistication technique et la capacité d'autonomie des robots mis sur le marché, en brouillant la ligne de démarcation entre les responsabilités des fabricants de robots et celles des utilisateurs (Asaro, 2012).

174. La réglementation légale reposant sur la responsabilité des produits applicable à l'industrie robotique sera sans doute inadaptée aux futures technologies robotiques mais une réglementation plus spécifique de la robotique, en particulier si elle prenait une forme trop restrictive, pourrait avoir des incidences négatives sur le développement de la technologie robotique. Leenes *et al.* (2017) décrivent comme suit le problème que pose la réglementation européenne et américaine actuelle en matière de responsabilité : « le fait est qu'il n'est pas possible d'évaluer rigoureusement l'amélioration des normes de sécurité et ce type de réglementation est perçu comme ayant un effet paralysant sur le développement des technologies » (p.17). Il est probable, à leur avis, que les appareils robotiques de l'avenir, comme les véhicules autonomes, atteindront un niveau de sophistication et de fiabilité tel qu'ils pourront fonctionner sans intervention humaine, mais « imposer aux producteurs des normes de responsabilité rigoureuses avant que ce niveau de sophistication ait été atteint (...) pourrait les dissuader tout simplement de développer cette technologie, ces normes étant considérées comme posant un risque trop important et créant un degré d'incertitude trop élevé » (Leenes *et al.*, 2017, p.17).

175. La réglementation du secteur de la robotique est donc confrontée à quatre difficultés : (1) rester en phase avec l'évolution rapide des technologies ; (2) trouver un équilibre entre stimuler – ou, tout au moins, ne pas freiner – l'innovation et protéger les valeurs et droits humains fondamentaux ; (3) savoir quand affirmer les normes sociales existantes ou les pousser à évoluer dans de nouvelles directions ; et (4) mettre en balance l'efficacité et la légitimité de la réglementation des technologies (Leenes *et al.*, 2017).

V. ENJEUX ÉTHIQUES

176. Les robots constituent l'une des technologies révolutionnaires de notre temps, non seulement à cause de leur impact profond sur la société, comme on l'a vu au chapitre III, mais aussi par la manière dont ils remettent en cause les cadres existants de réflexion et d'évaluation éthique. Les machines robotiques brouillent les limites entre sujets humains et objets technologiques en manifestant une capacité d'agir qui ressemble, sous certains aspects, à celle des êtres humains, à savoir la capacité de prendre des décisions, d'« exprimer » et d'« interpréter » des sentiments, et de prendre en charge ou d'aider à la réalisation de tâches humaines telles que conduire une voiture, enseigner ou dispenser des soins. Ce faisant, ces technologies obligent à repenser des notions éthiques essentielles comme celles de « capacité d'agir » et de « responsabilité ». En outre, leur développement et leur acceptation rapides par la société pourraient entraîner une modification de nos cadres de

valeurs dans divers domaines de la vie comme les soins, le travail, l'éducation, l'amitié et l'amour.

V.1. Techno-pessimisme, techno-optimisme et au-delà

177. Le débat en cours au sein de la société au sujet de la robotique est délimité par deux points de vue extrêmes : le « techno-optimisme », dont l'utopie du transhumanisme peut être considérée comme une version extrême, et le « techno-pessimisme », qui repose sur la crainte d'un avenir dystopique et invoque la nécessité du « bioconservatisme » pour défendre l'humanité contre l'impact trop radical des technologies.

178. Le techno-optimisme est l'idée que les progrès de la technologie déboucheront nécessairement sur un avenir meilleur. Sous sa forme la plus modérée, il incarne une conception instrumentale de la technologie et pense que celle-ci fournit les moyens de résoudre les problèmes sociétaux et de préparer un avenir meilleur. Les véhicules autonomes permettront de réduire le nombre d'accidents, les robots chirurgicaux d'accroître la précision des interventions médicales, les robots de service d'améliorer la qualité de vie des personnes âgées et des personnes atteintes de maladies chroniques, et ainsi de suite. Sous sa forme la plus radicale, le techno-optimisme exprime un idéal transhumaniste reposant sur l'idée que l'innovation technologique pourra être mise au service de l'amélioration des capacités humaines. La robotisation des tâches et activités humaines et le développement maximum de l'hybridité homme-machine font partie du programme transhumaniste, qui s'appuie souvent sur des scénarios de science-fiction et un discours valorisant l'innovation technologique dans une économie capitaliste florissante, en promettant le bien-être général et le perfectionnement individuel (Ford, 2015 ; Hottois *et al.*, 2015). Le mouvement transhumaniste a été critiqué comme mettant en avant une conception trop restrictive de l'être humain axée sur l'autonomie, l'individualisme et la compétitivité et comme étant en fait plus « hyperhumaniste » que « transhumaniste ».

179. Le techno-pessimisme repose sur une vision radicalement différente : se méfiant des implications des innovations technologiques, il craint que la technologie ne conduise ultimement non pas à un avenir plus lumineux mais plus sombre, en affectant les relations sociales, nos capacités cognitives et la connaissance humaine en général. Le bioconservatisme, qui vise à assurer la préservation de l'être humain en défendant certains éléments essentiels de l'existence humaine comme la dignité humaine (Fukuyama, 2002), le caractère de don de la vie humaine (Sandel, 2009) et la capacité à devenir l'auteur de sa propre vie (Habermas, 2003), est l'expression la plus radicale de ce mouvement. Il est parfois animé par un profond pessimisme au sujet de l'avenir de l'humanité et de la planète, et par une hantise de l'effondrement de la civilisation occidentale moderne, qui serait déjà entrée dans une phase de déclin, sans aucune possibilité d'échapper à son destin.

180. Du fait de leur caractère abstrait et de leur radicalisme, les points de vue extrêmes représentés par le transhumanisme et le bioconservatisme ne sont guère d'utilité pour aborder concrètement les questions de développement, de mise en œuvre et d'utilisation des robots. Même si le caractère révolutionnaire des technologies robotiques conduit naturellement à s'interroger sur la question de savoir dans quelle mesure les êtres humains doivent être protégés contre ces technologies, ou fusionner avec elles, d'autres questions portant plus directement et plus concrètement sur les implications effectives des développements actuels de la robotique pour les pratiques sociales existantes, à divers niveaux, méritent aussi d'être examinées.

- a. Au niveau personnel : quels types de robots sommes-nous prêts à accepter dans notre sphère privée ou dans notre vie personnelle ? Qui ou quoi souhaitons-nous admettre dans notre *oïkos* (maison, communauté) et pour quels types de pratiques et de relations : l'éducation des enfants, les soins infirmiers, les soins des personnes âgées, le nettoyage et la cuisine, les relations affectives et l'amour ? Comment les robots vont-ils affecter l'identité humaine et la qualité des relations entre individus ?

- b. Au niveau sociétal : dans quelle mesure sommes-nous prêts à déléguer le travail à des robots ? Quelles en seront les retombées sociales et culturelles ? Quelles en seront les conséquences du point de vue de la dignité humaine, de l'égalité et de l'inclusion ? Quel type de dépendance à l'égard des robots pourrait-il en résulter ? (Carr, 2014)
- c. Au niveau politique : dans quelle mesure peut-on accepter que des drones armés ou des robots autonomes soient utilisés pour tuer des individus en temps de guerre ou blesser ou neutraliser des personnes à des fins de sécurité ? Peut-on et doit-on déléguer à des robots des décisions concernant la vie et la mort ? Les robots peuvent-ils et doivent-ils être programmés de manière à respecter les principes éthiques de la « guerre juste » ?

181. Le caractère révolutionnaire des technologies robotiques ressort clairement de toutes ces questions : les robots brouillent les frontières entre l'humain et le technologique, en soulevant des questions fondamentales quant à la responsabilité, à la capacité d'action et au statut moral des robots.

V.2. Robots et responsabilité

182. Étant donné la complexité des robots contemporains, ainsi que de leur conception, de leur fabrication et de leur programmation, la question se pose de savoir qui exactement est responsable sur le plan éthique et/ou juridique dans l'éventualité où le fonctionnement ou le dysfonctionnement d'un robot serait cause de dommages pour des êtres humains, des biens ou l'environnement. Cette question est d'autant plus urgente à résoudre que les robots sont de plus en plus autonomes, mobiles et présents dans un très grand nombre de domaines de la vie humaine.

183. Dans le cas, par exemple, où un véhicule robotique autonome provoquerait un accident entraînant des victimes humaines, à qui devra en incomber la responsabilité ? L'équipe de roboticiens qui a conçu le véhicule ? Le fabricant ? Le programmeur ? Le vendeur ? La personne ou l'entreprise qui a décidé d'acheter et d'utiliser le véhicule ? Le robot lui-même ? Des questions éthiques et juridiques semblables se posent à propos des robots médicaux (par ex. dans le cas d'une opération chirurgicale effectuée au moyen d'un robot entraînant des blessures ou le décès d'un patient), des robots militaires ou de sécurité (par ex. en cas de meurtre de civils par un robot militaire autonome), des robots de service (par ex. si un robot de cuisine prépare un repas ou une boisson dangereuse pour la santé de l'utilisateur) et des robots industriels (par ex. dans le cas de blessures provoquées par un robot ou d'atteinte à l'environnement due à des substances toxiques manipulées par un robot). Il n'est pas nécessaire que les dommages subis par des êtres humains soient des dommages physiques ; il pourra s'agir de dommages psychologiques (par ex. en cas de violation de la vie privée par un robot personnel, ou bien si l'interaction avec un robot, en raison de ses caractéristiques humanoïdes très marquées, provoque des sentiments et un puissant attachement affectif, chose qui a plus de chances de se produire avec des enfants ou des personnes âgées).

184. Dans toutes ces situations, la responsabilité semble « commune » ou « partagée » entre le concepteur, l'ingénieur, le programmeur, le fabricant, l'investisseur, le vendeur et l'utilisateur du robot. Aucun de ces acteurs, en effet, ne peut être désigné comme la source ultime de l'action. Cependant, une telle approche a pour effet de diluer tout à fait la notion de responsabilité : si tous ces acteurs ont une part de la responsabilité totale, aucun d'eux n'est entièrement responsable. Ce problème – à savoir la responsabilité d'un trop grand nombre de personnes – ne peut être résolu en déclarant que les roboticiens devraient être tenus pour responsables dans tous les cas de dommages qui résultent de l'utilisation de leurs produits. Cela nierait la possibilité d'un « double usage » : un robot, comme beaucoup d'autres technologies, peut être utilisé à de bonnes et à de mauvaises fins. C'est le même dilemme auquel sont confrontées la plupart des recherches scientifiques, par exemple, chimistes qu'appliquent les mêmes produits chimiques afin de fabriquer des produits pharmaceutiques précieux qui sont utilisés pour fabriquer d'armes chimiques. Un robot peut être utilisé aux fins

prévues par son concepteur mais aussi à diverses autres fins, en particulier si son « comportement » peut être « piraté » ou « reprogrammé » par l'utilisateur final. Une autre difficulté tient à la « multi-stabilité » associée à l'impact des technologies, qui dépend toujours d'un certain contexte (Ihde, 1990). De même que la machine à écrire, conçue à l'origine comme un outil pour personnes malvoyantes, a fini par avoir des implications majeures pour le travail de bureau, l'impact des robots pourrait excéder de loin les attentes de leurs concepteurs. Il est impossible pour les roboticiens de prédire exactement comment leurs créations affecteront la société.

185. Éviter l'effet potentiellement paralysant de cette difficulté à endosser ou assigner la responsabilité représente donc un enjeu majeur pour l'éthique de la conception, de la mise en œuvre et de l'utilisation des technologies robotiques, compte tenu en particulier du fait que les équipes robotiques sont en général très nombreuses et que chacun de leurs membres n'est pas nécessairement conscient des implications ultimes du segment qui correspond à son travail. Deux voies ont néanmoins été proposées pour résoudre la question de la responsabilité. La première serait de développer des techniques permettant d'anticiper dans la mesure du possible les impacts du développement robotique (Waelbers and Swierstra, 2014 ; Verbeek, 2013). La deuxième consisterait à réfléchir soigneusement à l'apparition inévitable d'effets imprévus, en envisageant l'introduction dans la société de technologies robotiques comme une « expérimentation sociale » qui demande à être conduite avec le plus grand soin (Van de Poel, 2013).

V.3. Une « capacité d'agir » non humaine

V.3.1. Les robots en tant qu'agents autonomes

186. Comme ils peuvent agir de façon autonome et entrer par eux-mêmes en interaction avec leur environnement, le comportement des robots se caractérise par une capacité d'action indépendante. C'est la raison principale pour laquelle les technologies robotiques sont des technologies révolutionnaires. La capacité d'agir, en effet, a toujours été considérée comme le propre de l'homme et non des machines (Coeckelbergh, 2012). Il existe cependant des différences très nettes entre la capacité d'agir d'un être humain et celle d'un robot : la capacité d'agir d'un robot n'est pas aussi autonome puisqu'elle est le produit du travail de concepteurs et de programmeurs et des processus d'apprentissage de systèmes robotiques cognitifs. Néanmoins, les robots finissent par « faire » des choses et cela est le résultat de leurs propres interactions et processus de décision et pas uniquement des instructions fournies par leurs concepteurs.

187. La question de la capacité d'agir des robots prend un tour plus urgent – et même perturbant – lorsqu'on l'aborde sous l'angle moral. Pour être un agent moral et pour être responsable de ses actes, un être quelconque doit pouvoir agir d'une manière libre et intentionnelle et non être dirigé ou contraint à faire accidentellement quelque chose. Les robots ne sont pas dotés d'une liberté et d'une intentionnalité comparables à celles des êtres humains. Néanmoins, ils disposent d'un degré de liberté plus important que celui des machines ordinaires, puisqu'ils peuvent prendre des décisions, et d'une certaine forme d'intentionnalité liée aux algorithmes qui régissent leur interaction spécifique avec l'environnement et leurs « actes » potentiels.

188. L'approche dite de « médiation technologique », développée en philosophie de la technique, offre un modèle intéressant pour aborder la question d'une capacité d'agir non humaine. S'appuyant sur les travaux de Don Ihde (1990), cette approche envisage les technologies non comme des objets non humains, par opposition aux sujets humains, mais comme des « médiateurs » entre les êtres humains et leur environnement (Rosenberger and Verbeek, 2015). En effet, les technologies en usage se positionnent généralement au niveau des relations entre les êtres humains et le monde extérieur : par exemple, les téléphones portables permettent aux individus de rester en contact, les appareils d'imagerie par résonance magnétique (IRM) permettent aux médecins d'obtenir une image du corps de leurs patients. En remplissant cette fonction de médiation, les technologies contribuent à modeler

l'expérience que les êtres humains ont du monde et la manière dont ils agissent dans ce monde. Cette manière d'envisager les choses permet d'échapper à l'alternative consistant à décrire les robots soit comme des « objets » technologiques, soit comme des « sujets » quasi-humains. Lorsque l'on introduit un robot dans une pratique spécifique – les soins infirmiers, l'enseignement, le nettoyage ou la guerre, par exemple –, ce robot commence immédiatement à jouer un rôle de médiation, en modifiant la pratique en question. Dans une telle perspective, au lieu d'« être » des agents à part entière, les robots servent à *médier* la capacité d'agir. La question essentielle, par conséquent, n'est pas de savoir s'ils peuvent être comparés aux êtres humains, ou même les dépasser, mais plutôt de déterminer de quelle façon ils modifient les pratiques humaines, en déplaçant ainsi le cœur de la réflexion vers l'analyse de la qualité des relations homme-robot et la prise en compte de cette analyse dans la conception, la mise en œuvre et l'utilisation des robots.

V.3.2. Les robots en tant qu'agents moraux

189. Les risques potentiels associés à l'utilisation de robots sont souvent liés au fait que de nombreux robots sont conçus pour exécuter des fonctions ou travailler en relation étroite avec des êtres humains. Dans ce type de situation, les erreurs de conception, de programmation ou de fabrication des robots, qui sont inévitables, peuvent entraîner des dommages pour des êtres humains. Le dysfonctionnement de certains robots sophistiqués actuels pourrait sans doute infliger des dommages importants à un très grand nombre de personnes (par ex. les robots militaires armés ou les voitures robotiques autonomes en cas de perte de contrôle). La question, par conséquent, n'est pas seulement de savoir si les roboticiens devraient respecter certaines normes éthiques mais aussi de déterminer si certaines normes éthiques devraient être inscrites dans la programmation même des robots. Cette nécessité existe déjà si l'on réfléchit aux dommages que pourraient infliger à des êtres humains les robots personnels (par ex. les robots de cuisine, de conduite, de protection contre l'incendie, d'achat ou de comptabilité, ainsi que les robots compagnons ou infirmiers) ou les véhicules autonomes qui doivent prendre des décisions pouvant entraîner la mort de leurs passagers et d'autres personnes en cas d'événement routier inattendu. L'autonomie des robots étant probablement destinée à s'accroître, la réglementation éthique les concernant devra de plus en plus être spécifiquement conçue pour prévenir les comportements dangereux.

190. La nouvelle discipline appelée « éthique des machines » réfléchit déjà aux moyens de « doter les machines de principes éthiques ou d'une procédure de résolution des problèmes éthiques qu'elles peuvent rencontrer, en leur permettant ainsi de fonctionner d'une manière éthiquement responsable sur la base d'un processus de décision éthique propre » (Anderson and Anderson 2011a, p.1). Ce nouveau domaine de recherche comporte déjà des disciplines spécialisées, par exemple l'« éthique des machines médicales » qui s'occupe tout particulièrement des machines médicales capables d'accomplir « des tâches qui exigent une aptitude à l'interaction et une sensibilité psychologique, des connaissances pratiques et toute une gamme de règles de conduite professionnelle, ainsi qu'intelligence, autonomie et responsabilité en général sur le plan éthique » (Van Rysewyk and Pontier, 2015, p.7).

191. Deux questions importantes sont à examiner dans ce contexte : est-il possible de construire des « agents moraux artificiels » ? Et, dans l'affirmative, quel code moral faut-il intégrer à leur programmation ? La plupart des chercheurs travaillant dans le domaine de l'éthique des machines reconnaissent que les robots sont encore loin d'être des « agents éthiques » comparables aux êtres humains et, en particulier, à l'aptitude de ces derniers à raisonner sur la base de principes moraux, en appliquant ces principes dans des situations concrètes diverses. Cependant, ils s'accordent à penser que l'aptitude à devenir un sujet moral est une affaire de degré ; c'est pourquoi ils distinguent les agents éthiques « implicites » et « explicites ».

192. Selon Moor (2011), une machine est un agent éthique « implicite » dans la mesure où elle est équipée d'un logiciel qui empêche ou restreint les comportements non éthiques. Autrement dit, il est possible de programmer de manière limitée une machine en un sens

éthique afin d'éviter les conséquences moralement non souhaitables de l'exécution de certaines tâches spécifiques (par exemple, les distributeurs automatiques de billets sont programmés pour rendre exactement la monnaie aux clients, les pilotes automatiques sont programmés pour ne pas mettre en danger la sécurité des passagers). Certains philosophes (Savulescu and Maslen, 2015) considèrent aussi que les « agents éthiques artificiels », grâce à leur vitesse de calcul et au fait qu'ils sont dépourvus des imperfections morales humaines typiques (partialité, égoïsme, biais individuels ou préjugés affectifs, faiblesse de la volonté), pourront en fait aider ou même remplacer les êtres humains en cas de décision morale difficile dans certains contextes précis et limités (de la même façon qu'un travail humain physique ou mental imparfait a pu être remplacé par un homologue robotique ou IA), en soutenant ainsi selon des modalités nouvelles la capacité d'action morale des êtres humains.

193. Une question plus difficile à résoudre est celle de savoir si des robots pourraient un jour devenir des agents éthiques « explicites » à part entière comme les êtres humains. Est-il possible de programmer des machines ou des robots de façon à ce qu'ils puissent « agir conformément à des principes », « manifester certaines vertus » ou « justifier moralement leurs actes » ? Pour disposer de ces capacités, comme on le rappelle souvent, le libre arbitre est nécessaire et celui-ci, dit-on, est la prérogative des êtres humains. Des robots pourraient-ils avoir des « connaissances morales » qu'ils appliqueraient à résoudre toute une gamme de problèmes moraux différents et peut-être même très complexes ? D'après Moor (2011), bien que des recherches intéressantes se poursuivent en ce domaine (en particulier dans les domaines avancés de la logique déontique, de la logique épistémique et de la logique de l'action), pour le moment, « les exemples clairs de machines agissant comme des agents éthiques explicites demeurent élusifs » et « de nouvelles recherches seront nécessaires avant qu'un agent éthique explicite solide puisse exister dans une machine » (p.17). D'autres auteurs sont plus optimistes. Whitby (2011), par exemple, soutient qu'un agent éthique explicite ne doit pas être conçu d'une manière essentiellement différente des programmes informatiques actuels de jeu d'échecs, compte tenu en particulier du fait que, dans ce type de programmes, tous les mouvements de pions possibles ne sont pas préprogrammés car ils s'appuient en fait sur des procédures générales de décision (et même des conjectures ou des intuitions au sujet des meilleurs mouvements possibles, qui peuvent être affinées et modifiées pendant le jeu).

194. Le fait demeure, toutefois, que « programmer un ordinateur à se comporter de manière éthique est beaucoup plus difficile que de programmer un ordinateur à jouer aux échecs comme un champion mondial » car « les échecs constituent un domaine simple où les mouvements autorisés sont bien définis », tandis que « l'éthique englobe un domaine complexe où certains mouvements souhaitables sont mal définis » (Moor, 2011, p.19). Autrement dit, prendre des décisions morales et agir sur la base de ces décisions en temps réel dans le monde réel nécessite non seulement la connaissance de principes moraux complexes mais aussi l'aptitude à reconnaître et évaluer un large éventail de faits concernant les êtres humains, d'autres êtres sensibles et leurs environnements. Pour ne citer qu'un aspect du problème : les décisions moralement pertinentes (par exemple, le choix de poursuivre différents avantages et d'éviter divers dommages) varient non seulement entre des groupes d'individus différents (jeunes/vieux, hommes/femmes, personnes en bonne santé/malades) mais aussi à l'intérieur d'un même individu au cours de son existence (au fur et à mesure de sa croissance, de sa maturation et de l'acquisition de nouvelles connaissances et de nouvelles expériences). Il ne semble guère probable qu'une machine dépourvue de sentiments comme l'empathie (qui est important pour évaluer les avantages et dommages physiques et psychologiques possibles) réussisse à traiter la diversité des faits et des choix moralement pertinents.

195. L'utilisation de robots dans la plupart des domaines de la vie est généralement considérée comme acceptable dès lors qu'ils exécutent leurs tâches mieux que les êtres humains et en commettant moins d'erreurs. Le même raisonnement pourrait sans doute s'appliquer à l'utilisation de robots en tant qu'agents éthiques « implicites » ou « conseillers

moraux artificiels » pour assister les êtres humains, qui prendraient la décision finale dans les situations morales complexes (par exemple, en cas de catastrophe naturelle lorsqu'il faut décider la répartition optimale de ressources limitées). Toutefois, même cet « usage éthique » limité des robots n'est pas sans présenter certains aspects potentiellement inquiétants. Il pourrait en effet amener les êtres humains placés dans de telles situations à éviter de réfléchir eux-mêmes au choix moral à effectuer et à déléguer trop vite et trop facilement la décision à leurs « conseillers artificiels », ce qui n'est guère souhaitable. Un risque connexe est qu'avec le temps, divers professionnels (médecins ou officiers militaires, par exemple) soient conduits à perdre toute sensibilité morale ou aptitude à la réflexion morale sérieuse, comme on a pu l'observer pour l'aptitude à l'écriture et au calcul après la diffusion des calculatrices de poche et des ordinateurs individuels.

196. La création de robots qui seraient des agents éthiques « explicites » (c'est-à-dire des agents capables de prendre des décisions morales générales et d'appliquer eux-mêmes ces décisions), bien que cette possibilité reste de l'ordre de la spéculation, requiert une attention éthique particulière. L'existence de tels robots ne serait hypothétiquement souhaitable que dans les situations où ils pourraient « dépasser les performances éthiques » des êtres humains, par exemple en parvenant à analyser des problèmes éthiques complexes sans manifester certaines faiblesses morales caractéristiques des êtres humains. Toutefois, cela présupposerait l'existence d'une norme de mesure des performances éthiques, norme aujourd'hui inexistante, en particulier au niveau mondial. La définition d'un « bon agent moral » varie selon les époques historiques et les contextes culturels.

197. Nonobstant les doutes que l'on peut avoir sur les chances qu'un tel agent éthique « explicite » devienne un jour techniquement possible, on voit mal à partir de quel moment une machine pourrait être considérée comme ayant atteint ce but. Le test de Turing, c'est-à-dire la meilleure proposition théorique qui existe aujourd'hui pour tester l'intelligence artificielle, demeurant lui-même controversé, il paraît difficile d'imaginer qu'un test analogue – le « test de Turing moral » – puisse être mis au point pour tester la création éventuelle d'agents éthiques « explicites » (Allen *et al.*, 2000 ; Allen *et al.*, 2011). Toute nouvelle technologie devant être rigoureusement testée avant de pouvoir être mise sur le marché, on voit mal comme un « agent éthique explicite » pourrait être testé – et surtout testé avec succès.

198. La difficulté à tester et à vérifier l'existence d'agents éthiques artificiels « explicites » nous ramène à la deuxième question importante au sujet de ces agents : si la fabrication d'agents éthiques « explicites » devenait possible, avec quel code éthique devraient-ils être programmés ? Les trois lois de la robotique d'Asimov sont fréquemment considérées comme la réponse type à cette question. Cependant, les chercheurs en « éthique des machines » s'accordent généralement à reconnaître que ces lois sont trop générales, potentiellement contradictoires et non applicables dans le monde réel (ce que montrent en fait les nouvelles mêmes d'Asimov). En outre, selon Abney (2012), la programmation de robots futurs au moyen d'une quelconque théorie morale descendante comme la déontologie ou le conséquentialisme est difficilement plausible à cause des conflits potentiels entre les obligations et/ou l'incapacité des robots à calculer les conséquences à long terme de leurs actions (problème de la traçabilité computationnelle). C'est pourquoi il propose de développer une « éthique de la vertu des robots » qui serait axée sur « la recherche des vertus qu'un bon robot (c'est-à-dire fonctionnant comme il faut) doit manifester, compte tenu de ses fonctions spécifiques ».

199. Parmi les autres propositions actuellement discutées, on peut citer le kantisme (Powers, 2011), la théorie pluraliste des obligations *prima facie* (Anderson and Anderson, 2011b), l'éthique bouddhiste (Hughes, 2012) et la théorie du commandement divin (Bringsjord and Taylor, 2012). La diversité de ces approches montre clairement que l'on est encore loin d'un consensus sur la question de savoir quelle théorie devrait être programmée dans les futurs « agents éthiques artificiels », si jamais cela devient techniquement possible. Il n'existe toujours pas de réponse philosophique définitive à la question de la nature même de la moralité, non plus qu'au sujet de l'objectivité ou de la subjectivité des cadres éthiques. Dans la discussion sur la capacité des robots à agir de manière morale réapparaissent les différents

points de vue présents dans le champ de la réflexion éthique en général, ce qui soulève la question de savoir s'il est nécessaire de développer dans la programmation des machines une capacité à accepter et respecter la diversité des cadres et orientations éthiques identique à celle qui existe à propos des cadres éthiques humains.

V.4. Statut moral des robots

200. Une question intrigante au sujet des technologies robotiques dotées d'une autonomie et d'une capacité de décision plus grandes – et peut-être même d'une capacité de décision morale – est celle de leur statut moral. De tels robots deviendront-ils non seulement des agents moraux capables d'action morale mais aussi des entités moralement utiles, et pas seulement utiles en un sens instrumental comme de simples appareils fabriqués pour exécuter des tâches spécifiques ? Mériteront-ils un respect moral et une protection contre les dommages identiques à ceux actuellement accordés aux êtres humains et à certains animaux non humains ? Ces robots pourront-ils en définitive être dotés non seulement d'obligations et de devoirs moraux mais aussi de droits moraux ?

201. La manière dont s'acquiert ou se perd le statut moral est une vieille question philosophique. Certains philosophes considèrent que le statut moral se résume à certaines propriétés psychologiques et/ou biologiques. D'un point de vue déontologique, avoir un statut moral signifie être une personne et être une personne signifie être doué de raison ou être capable de délibération rationnelle et morale. Dans la mesure où ils sont capables d'effectuer par eux-mêmes de nombreuses tâches cognitives difficiles, on peut dire des robots qu'ils possèdent une certaine forme de rationalité. Cependant, il serait absurde de les qualifier de « personnes » puisqu'ils sont dépourvus de certaines autres qualités généralement associées aux êtres humains comme le libre arbitre, l'intentionnalité, la conscience de soi, le sens moral et le sentiment de l'identité personnelle. On notera toutefois à ce propos que la Commission des affaires juridiques du Parlement européen, dans son *Projet de Rapport contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique* de 2016 envisage la possibilité de « la création d'une personnalité juridique spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques dotées de droits et de devoirs bien précis, y compris celui de réparer tout dommage causé à un tiers ; serait considéré comme une personne électronique tout robot qui prend des décisions autonomes de manière intelligente ou qui interagit de manière indépendante avec des tiers » (JURI, 2016, section 59.f).

202. Dans une perspective utilitariste, le statut moral n'est pas lié à la rationalité mais à la sensibilité ou à la capacité d'éprouver le plaisir et la peine (au sens large) et les sentiments correspondants. Dans cette conception, les êtres humains et de nombreux animaux non humains ont un statut moral, mais non les robots car ils ne sont pas doués de sensibilité et n'éprouvent pas de sentiments. Selon certains auteurs (par exemple, Torrance, 2011), seuls les êtres organiques et non les robots sont dotés d'une sensibilité véritable. Il n'existe pas encore de robots sensibles et/ou capables d'éprouver des sentiments mais la recherche se développe sur les sentiments « artificiels » ou « synthétiques » qui pourraient être programmés dans de futurs « robots sociables » (Valverde and Casacuberta, 2009). En attendant d'éventuels progrès dans ce domaine de recherche, on ne peut exclure la possibilité que des robots futurs acquièrent une sensibilité, des sentiments et, par voie de conséquence, un statut moral ou au moins un certain type de droits moraux.

203. Dans la discussion sur le statut moral des robots (Floridi, 2011), il est utile de distinguer à des fins d'analyse les *agents moraux* (êtres capables d'agir par eux-mêmes de manière morale et de traiter autrui de façon moralement juste ou injuste) de ce qu'on pourrait appeler les *patients moraux* (êtres incapables d'agir par eux-mêmes de manière morale mais pouvant être traités de façon moralement juste ou injuste). La plupart des êtres humains, par exemple, sont à la fois des agents moraux et des patients moraux, tandis que certains êtres humains (comme les personnes dans le coma ou les nouveau-nés) et certains animaux sont uniquement des patients moraux. On admet généralement que certains animaux non humains

disposent effectivement de certains droits moraux (comme le droit de ne pas être soumis à la torture), bien qu'ils soient incapables d'obligations morales. Le développement de robots autonomes et capables de décision compliquera sans doute cette classification, qui est en fait assez imprécise. Si des robots deviennent un jour des agents éthiques « explicites », ils entreront clairement dans la catégorie des agents moraux mais on ne voit pas clairement s'ils entreront aussi dans la catégorie des patients moraux, car on ignore ce que serait effectivement un « tort moral » ou un « préjudice » infligé à un robot (louer ou blâmer moralement un robot dépourvu de sentiments, par exemple, serait sans aucun effet, tout comme le fait de le menacer d'une sanction morale ou juridique ou d'une peine physique).

204. Une troisième manière d'attribuer un statut moral aux robots, sans prendre en compte aucune propriété psychologique ou biologique particulière, serait d'adopter un « point de vue relationnel » en vertu duquel les robots seraient considérés comme dotés d'un statut moral dans la mesure où ils entrent dans des relations singulières, parfois significatives d'un point de vue existentiel, avec des êtres humains. Cette approche relationnelle de la question du statut moral des robots est intuitivement attrayante car les êtres humains sont des êtres sociaux et affectifs qui sont naturellement en relation avec d'autres êtres humains et se soucient d'eux. Une approche similaire est parfois utilisée en éthique de l'environnement, un écosystème entier étant considéré comme méritant une protection morale, c'est-à-dire un certain type de statut moral, en raison de son « interconnectivité » et des relations d'une importance vitale qui existent entre ses éléments vivants et non vivants.

205. Toutefois, dans le cas des robots, l'approche relationnelle se heurte à un problème : la tendance psychologique des êtres humains à anthropomorphiser des objets inanimés ou des artefacts et à projeter sur eux des propriétés humaines. Les êtres humains étant capables de développer un attachement puissant et durable à une voiture, un bateau ou une guitare, par exemple, il est fort probable qu'ils soient conduits à développer une forme d'attachement comparable aux robots. On connaît des exemples de soldats ayant développé un lien si étroit avec un robot de déminage qu'ils ont pleuré lors de la destruction du robot au cours d'une intervention (Lin, 2012). Le développement de liens entre êtres humains et robots n'est d'ailleurs pas nécessairement dû à l'anthropomorphisme : bien que n'ayant aucune apparence humaine, des artefacts technologiques, y compris des robots, peuvent devenir si précieux et importants que l'on considère normal de les protéger, comme le montrent de nombreux sites du patrimoine culturel de l'UNESCO.

206. Le développement rapide de robots extrêmement intelligents et autonomes conduira sans doute à remettre en cause la classification actuelle des êtres en fonction de leur statut moral, d'une façon similaire – ou peut-être même plus radicale – à ce qui s'est produit dans le cas des animaux non humains sous l'effet du mouvement pour la défense des droits des animaux. Il pourrait même modifier la manière dont est aujourd'hui perçu le statut moral des êtres humains. Bien qu'ayant souvent l'air de spéculations futuristes, ces questions ne devraient pas être prises à la légère, compte tenu en particulier du fait que l'« écart homme-machine » se comble peu à peu (Romportl *et al.*, 2015), et de l'apparition probable à l'avenir d'hybrides homme-machine ou animal-machine ou de cyborgs (robots intégrés à un organisme biologique ou contenant au moins certaines composantes biologiques).

V.5. Dynamisme des valeurs

207. Les technologies robotiques ont encore un autre effet révolutionnaire sur les sociétés : leur impact sur les systèmes de valeurs morales. Les technologies n'ont pas seulement des effets sociétaux que l'on peut évaluer d'un point de vue éthique, elles affectent directement les cadres éthiques utilisés pour les évaluer. Les technologies font évoluer les valeurs humaines et les orientations normatives. L'introduction de la pilule contraceptive, par exemple, a entraîné une modification des valeurs en matière de sexualité, en distendant le lien entre sexualité et reproduction et en ouvrant la voie à une évaluation nouvelle des pratiques hétéro- et homosexuelles (Mol, 1997). L'introduction d'une technologie de réalité augmentée comme Google Glass semble aussi modifier le sens de ce que l'on entend par « vie privée » dans nos

sociétés (Kudina and Bas, 2017), en suscitant la formulation de nouvelles définitions aptes à prendre en compte la manière spécifique dont ce type de technologie redessine les frontières entre sphère publique et sphère privée. Ce type d'influence crée par conséquent un enjeu spécifique pour la réflexion éthique sur les technologies : les valeurs que nous utilisons pour évaluer les technologies ne peuvent être considérées comme des normes données à l'avance car elles changent avec le temps, en interaction avec les technologies mêmes que nous cherchons à évaluer.

208. Cette plasticité ou ce dynamisme des valeurs occupe une place centrale dans l'approche contemporaine du « changement techno-moral ». En développant des scénarios d'avenirs technologiques possibles, cette approche cherche à anticiper le changement moral lié aux évolutions de la technologie, afin de promouvoir le développement de politiques et de pratiques dans le domaine technologique (Swierstra *et al.*, 2009). L'approche dite de « médiation technologique » envisage l'évolution des valeurs sous l'angle de la « médiation morale » exercée par les technologies : en médiant les pratiques, les perceptions et les interprétations humaines, les technologies contribuent à modeler les actions et les décisions morales, et elles ont donc des incidences sur les cadres moraux existants (Verbeek, 2011).

209. Il n'existe pas encore, dans le domaine de la robotique, de travaux explicites sur l'impact effectif des robots du point de vue des valeurs éthiques et des cadres normatifs. Cependant, les approches axées sur le changement techno-moral et la médiation morale permettent tout à fait d'imaginer les formes que prendra cet impact. Les robots de soins risquent de modifier les aspects des soins qui paraissent les plus importants aux êtres humains, tant sur le plan physique qu'affectif. Les robots d'enseignement modifieront probablement notre idée de ce qui constitue une bonne éducation et notre appréciation du rôle des enseignants et des élèves. Les voitures autonomes conduiront sans doute à l'apparition de nouveaux critères de l'« aptitude à conduire » des êtres humains et des robots. Les robots sexuels pourraient modifier la relation entre amour et rapports sexuels et ce que les êtres humains valorisent dans leurs relations avec d'autres individus.

210. Pour prendre en compte ces effets normatifs de manière responsable, il sera nécessaire de mettre soigneusement en balance l'anticipation et l'expérimentation. Grâce au développement de scénarios techno-moraux, l'approche axée sur le changement techno-moral permet de fournir dès aujourd'hui aux utilisateurs, aux concepteurs et aux responsables de l'élaboration des politiques des éléments aptes à guider la prise de décision sur les effets sociétaux potentiels d'une technologie à l'avenir. Dans une autre optique, Van De Poel (2013) a proposé de traiter le développement d'une technologie émergente révolutionnaire comme une « expérimentation sociale ». L'anticipation risque à son avis de « passer à côté de certaines conséquences sociales réelles importantes » et de « nous aveugler sur les effets imprévus » d'une technologie (Van de Poel 2016, p.667). Il déclare que les innovations sont en fait des expérimentations sociales que nous devons mener de manière responsable. Le dynamisme des valeurs exige par conséquent d'évaluer de près, à petite échelle et dans un contexte expérimental, les incidences des technologies robotiques sur les systèmes de valeurs, afin de pouvoir prendre en compte ces incidences dans les pratiques de conception des robots, dans les discussions publiques et dans l'élaboration des politiques.

VI. RECOMMANDATIONS

VI.1. Un cadre éthique fondé sur la technologie

211. Comme indiqué dans l'introduction générale de ce rapport, les premiers robots étaient programmés pour exécuter des tâches clairement définies. Ces robots peuvent être décrits comme des robots *déterministes* car leurs actions sont contrôlées par une série d'algorithmes dont l'enchaînement est prévisible.

212. Avec les progrès de l'informatique, la notion d'*intelligence artificielle* a commencé à être appliquée à des ordinateurs. Bien que cette notion fasse l'objet d'interprétations différentes, certains auteurs allant jusqu'à considérer qu'elle implique une intelligence réelle

identique à celle des êtres humains, les machines reposant sur l'intelligence artificielle manifestent des capacités sensorielles, linguistiques et interactives comparables à celle des êtres humains. En outre, ces machines présentent des capacités d'apprentissage qui sont peu à peu améliorées – et continueront à être améliorées – à l'aide des techniques d'apprentissage profond.

213. Ces développements ouvrent la voie à ce qu'on appelle la *robotique cognitive*. Il s'agit d'un domaine technologique visant à la mise au point de robots reposant sur des systèmes de computation cognitive et capables, par conséquent, d'apprendre à partir de l'expérience, non seulement en interaction avec des formateurs humains mais par leurs propres moyens, en développant ainsi la capacité de réagir à leur environnement sur la base de ce qu'ils ont appris. Comparés aux robots classiques ou déterministes, les robots cognitifs sont capables de prendre des décisions dans des situations complexes, décisions qui ne peuvent être prévues par un programmeur.

214. La distinction entre robots déterministes et robots cognitifs est importante aux fins de l'élaboration de recommandations concernant l'éthique de la robotique. En effet, le comportement d'un robot déterministe est dicté par le programme qui contrôle ses actes. La responsabilité de ces actes est donc clairement identifiable et peut faire l'objet, pour l'essentiel, d'une réglementation légale. En revanche, les décisions et les actes d'un robot cognitif peuvent donner lieu uniquement à des estimations statistiques et sont donc imprévisibles. Par conséquent, la responsabilité des actes de ce type de robot n'est pas claire et son comportement dans un environnement autre que celui dans lequel il a été formé pendant la phase d'apprentissage – c'est-à-dire un environnement « aléatoire » – peut être catastrophique. Assigner la responsabilité des actes d'une machine en partie *stochastique* ne va pas de soi. Cependant, si nous devons vivre avec ces machines et réglementer leur utilisation, le problème de leur responsabilité devra être résolu.

215. En outre, pour être raisonnablement durable dans un domaine comme celui-ci, qui évolue rapidement, une série de recommandations pertinentes doit prendre en compte, outre le niveau de développement actuel des robots, les tendances et les directions de ce développement.

216. C'est pourquoi la COMEST propose d'élaborer des recommandations prenant en compte la distinction susmentionnée. À un premier niveau, celui des machines déterministes dont on peut assigner la responsabilité du comportement, les recommandations de la Commission porteront essentiellement sur les outils juridiques nécessaires pour réglementer leur utilisation. À un second niveau, celui des machines cognitives dont le comportement n'est pas prévisible à 100 % et est donc pour une part importante stochastique, les recommandations envisageront, outre l'adoption d'outils juridiques, l'élaboration de codes de pratique et de directives éthiques à l'intention des producteurs et des utilisateurs. Enfin, s'agissant des machines stochastiques utilisées dans un contexte où le risque de dommages existe (voiture ou arme autonome, par exemple), il conviendra d'examiner le degré d'autonomie qui peut raisonnablement être accordé à ces machines et les modalités de maintien d'un niveau significatif de contrôle humain.

217. Le tableau ci-dessous présente le modèle proposé sous une forme schématique. La structure envisagée est simple mais sa mise en œuvre sous forme de délimitation des responsabilités et de réglementation de l'utilisation sera complexe et difficile, à la fois pour les scientifiques, les ingénieurs, les responsables de l'élaboration des politiques et les spécialistes des questions éthiques. Cependant, si nous voulons vivre dans un monde où la robotique prendra de plus en plus en charge à l'avenir les tâches actuellement exécutées par des êtres humains, ces difficultés devront être surmontées.

Décision du robot	Implication humaine	Technologie	Responsabilité	Réglementation
Effectuée à partir d'une série limitée d'options, sur la base de critères stricts prédéfinis	Critères mis en œuvre dans un cadre légal	Machine uniquement : algorithmes/robots déterministes	Fabricant du robot	Légale (normes, législation nationale ou internationale)
Effectuée à partir d'une gamme d'options, avec une certaine flexibilité, sur la base d'une politique prédéfinie	Décision déléguée au robot	Machine uniquement : algorithmes d'intelligence artificielle, robots cognitifs	Concepteur, fabricant, vendeur et utilisateur du robot	Codes de pratique des ingénieurs et des utilisateurs ; principe de précaution
Effectuée via l' interaction homme-robot	Contrôle humain des décisions du robot	Capacité des êtres humains à prendre contrôle du robot dans le cas où ses actes pourraient entraîner des dommages graves ou la mort d'un individu	Êtres humains	Éthique

VI.2. Valeurs et principes éthiques pertinents

218. Étant donné la diversité et la complexité des robots, établir un cadre de valeurs et de principes éthiques peut être utile en vue de la définition de la réglementation à tous les niveaux – conception, fabrication et utilisation – et d'une manière cohérente, sous une forme allant de codes de conduite des ingénieurs à des textes de loi nationaux et à des conventions internationales.

219. Le principe de la responsabilité humaine est le fil conducteur reliant les différentes valeurs mentionnées ci-dessous.

VI.2.1. Dignité humaine

220. La **dignité humaine** est une valeur fondamentale affirmée dans la *Déclaration universelle des droits de l'homme* (UN, 1948), qui dispose que, naissant libres et égaux, les êtres humains « sont doués de raison et de conscience et doivent agir les uns envers les autres dans un esprit de fraternité » (art. 1).

221. La dignité est intrinsèque aux êtres humains mais non aux machines et aux robots. Il n'est pas possible, par conséquent, d'assimiler les robots à des êtres humains, même dans le cas de robots androïdes ayant l'apparence séductrice d'un être humain ou de robots cognitifs puissants dotés de capacités d'apprentissage supérieures aux capacités cognitives humaines individuelles. Les robots ne sont pas des êtres humains ; ils sont le produit de la créativité humaine et ont besoin, pour être efficaces et demeurer des outils ou des médiateurs efficaces, d'un système de soutien technique.

VI.2.2. Valeur de l'autonomie

222. La reconnaissance de la dignité humaine implique que l'**autonomie**, en tant que valeur, ne concerne pas uniquement le respect de l'autonomie individuelle, qui peut aller jusqu'au

refus de toute forme de dépendance à l'égard d'un robot. L'autonomie exprime en tant que valeur la reconnaissance de l'*interdépendance des relations* entre êtres humains, entre êtres humains et animaux, et entre êtres humains et environnement. Dans quelle mesure les robots sociaux pourront-ils enrichir ces relations, ou bien les réduire et les standardiser ? Cela demande à être évalué de manière scientifique dans le cadre des pratiques médicales et éducatives où les robots peuvent être utilisés, en particulier lorsque sont concernés des groupes vulnérables tels que les enfants et les personnes âgées. L'utilisation systématique de robots pourrait aggraver la rupture des liens sociaux dans certaines sociétés.

223. L'interdépendance signifie que les robots font partie des créations techniques de l'humanité (et du « technocosme » qu'elle construit) et qu'ils ont aussi un *impact environnemental* (déchets électroniques, consommation d'énergie, émission de CO₂, empreinte écologique), qui doit être pris en compte et évalué en mettant en balance les avantages et les risques.

VI.2.3. Respect de la vie privée

224. Le **respect de la vie privée** découle de l'article 12 de la *Déclaration universelle des droits de l'homme* (UN, 1948), qui stipule :

Nul ne sera l'objet d'immixtions arbitraires dans sa vie privée, sa famille, son domicile ou sa correspondance, ni d'atteintes à son honneur et à sa réputation. Toute personne a droit à la protection de la loi contre de telles immixtions ou de telles atteintes.

225. Divers systèmes de protection et réglementations ont été introduits dans nombre de pays pour limiter l'accès aux données à caractère personnel et protéger la vie privée des individus. Cependant, le développement des méga-données modifie la manière dont les données sont collectées et traitées (utilisation d'algorithmes de profilage) L'échelle des pratiques de collecte et de traitement s'accroît énormément et les utilisations des données se multiplient (par exemple, dans les domaines du commerce, de la sécurité nationale, de la surveillance publique et de la recherche), ainsi que les formes d'intrusion. Les robots sont des appareils capables de recueillir des données au moyen de capteurs et d'exploiter des méga-données au moyen de l'apprentissage profond. Par conséquent, la collecte et l'utilisation des données sont des aspects à examiner de près dans la conception des robots, en mettant en balance la finalité du robot et la protection de la vie privée des individus. Certaines catégories de données sont plus sensibles que d'autres ; c'est pourquoi diverses approches – législation, réglementation professionnelle, gouvernance, surveillance publique – doivent être envisagées conjointement pour maintenir la confiance du public à l'égard des robots et leur bonne utilisation.

VI.2.4. Principe d'innocuité

226. Le **principe d'innocuité** désigne la ligne à ne pas franchir par les robots. Comme nombre de technologies, les robots ont potentiellement un « double usage ». Ils sont généralement conçus à des fins utiles ou positives (par exemple, réduire les risques sur le lieu de travail), pour aider les êtres humains et non pour leur porter préjudice ou pour porter atteinte à leur vie. Les trois lois formulées par Isaac Asimov (voir paragraphe 18) demeurent pertinentes à cet égard. Si l'on prend ce principe éthique au sérieux, il est nécessaire de se demander si les armes autonomes et les drones armés devraient être interdits.

VI.2.5. Principe de responsabilité

227. Le rapport de la COMEST sur « Le principe de précaution » (UNESCO, 2005b) déclare que « la responsabilité éthique implique un minimum de liberté de choix dans l'action » (p.17). S'agissant de l'utilisation des robots, « l'idée selon laquelle les individus (ou bien les entreprises, ou bien les États) sont moralement responsables des choix qu'ils font » est extrêmement importante. C'est parce que les êtres humains sont responsables de leurs actes qu'ils peuvent être poursuivis et faire l'objet d'une reconnaissance légale de responsabilité lorsque leurs actes entraînent un préjudice ou des dommages. De ce point de vue, la responsabilité éthique humaine ne peut être déléguée aux robots.

228. Les robots déterministes, et même les robots cognitifs sophistiqués, ne peuvent endosser une quelconque responsabilité éthique ; cette responsabilité est celle du concepteur, du fabricant, du vendeur, de l'utilisateur et de l'État. Les êtres humains, par conséquent, doivent toujours être dans le circuit et avoir la possibilité de contrôler un robot par différents moyens (traçabilité, coupe-circuit), afin de maintenir la responsabilité humaine sur les plans moral et légal.

229. Dans le développement de la robotique, trois aspects demandent à être éclaircis et précisés à cet égard : **la responsabilité civile**, la **transparence** et l'**imputabilité**. La question de la responsabilité civile est particulièrement importante dans le cas des véhicules présentant un certain degré d'automatisation. La mise en place de systèmes de positionnement et d'enregistrement pourrait aider à établir ce type de responsabilité mais risque de poser des problèmes du point de vue du droit au respect de la vie privée et de la protection des données à caractère personnel.

230. Il existe un lien fort entre responsabilité éthique et **principe de précaution**. La COMEST a défini ce principe comme suit :

Lorsque des activités humaines risquent d'aboutir à un danger moralement inacceptable, qui est scientifiquement plausible mais incertain, des mesures doivent être prises pour éviter ou diminuer ce danger. Le danger *moralement inacceptable* est un danger pour les humains ou pour l'environnement qui est :

- menaçant pour la vie ou la santé humaine, ou bien
- grave et réellement irréversible, ou bien
- inéquitable pour les générations présentes ou futures, ou bien
- imposé sans qu'aient été pris dûment en compte les droits humains de ceux qui le subissent. (UNESCO, 2005b, p.14)

L'application du principe de précaution renforce la nécessité du contrôle et d'études empiriques systématiques faisant appel à un large éventail de disciplines scientifiques.

231. Ce principe est particulièrement pertinent au regard du développement de robots autonomes (robots cognitifs) dotés de capacités d'apprentissage profond et dont certains processus comportementaux ou décisionnels ne peuvent être programmés de la même façon que pour les robots déterministes. Un certain degré d'incertitude est inhérent aux robots cognitifs et les roboticiens ne peuvent se fier uniquement à leur capacité à évoluer sans prendre en compte les risques et effets négatifs cumulés sur une longue échelle de temps.

232. Les chercheurs en robotique, l'industrie et les gouvernements sont des partenaires dans l'optique de la **recherche** et de l'**innovation responsables**. Il est donc nécessaire de ne pas envisager les perspectives de développement de la robotique uniquement du point de vue de la productivité et de l'efficacité économiques. Un engagement public est requis afin d'opter pour des voies de développement scientifique qui contribuent au bien commun et sont adaptées au contexte culturel.

VI.2.6. Valeur de la bienfaisance

233. Les robots sont utiles pour améliorer la sécurité, l'efficacité et les performances dans l'exécution de nombreuses tâches humaines physiquement difficiles. Les robots industriels, les robots de recherche en cas de catastrophe et les robots de déminage peuvent être utilisés pour remplacer les êtres humains dans des environnements dangereux. Cependant, la question de la bienfaisance des robots conçus pour intervenir dans un contexte social tel que l'éducation, les soins de santé, la surveillance ou le travail de police de l'État doit faire l'objet d'une discussion et d'une réflexion plus approfondies.

234. La bienfaisance d'un robot, en tant que valeur, doit être évaluée à l'aide du **principe de proportionnalité**, en tenant compte des choix technologiques existants. Plusieurs questions importantes se posent ici : quel est le but final poursuivi ? Le contexte social et

culturel a-t-il été pris en compte lors du processus d'évaluation et de mise en œuvre ? Un type particulier de robot a-t-il été imposé aux utilisateurs ou bien le robot a-t-il été conçu à l'intention des utilisateurs et éventuellement en collaboration avec eux ? Dans les pays en développement plus encore que dans les pays développés, l'utilisation de robots doit être mise en balance avec d'autres priorités économiques et sociales.

235. Cela nous amène à la promotion de **la diversité culturelle** comme valeur. En 2005, l'UNESCO a adopté la *Convention sur la protection et la promotion de la diversité des expressions culturelles*, qui définit la diversité culturelle comme :

[...] la multiplicité des formes par lesquelles les cultures des groupes et des sociétés trouvent leur expression. Ces expressions se transmettent au sein des groupes et des sociétés et entre eux. La diversité culturelle se manifeste non seulement dans les formes variées à travers lesquelles le patrimoine culturel de l'humanité est exprimé, enrichi et transmis grâce à la variété des expressions culturelles, mais aussi à travers divers modes de création artistique, de production, de diffusion, de distribution et de jouissance des expressions culturelles, quels que soient les moyens et les technologies utilisés (UNESCO, 2005a, p.4).

La recherche et l'innovation en robotique devraient prêter une plus grande attention aux questions culturelles et de genre. Du fait de la diversité des cultures, les robots – en particulier les robots sociaux – pourront être acceptés dans certains contextes et non dans d'autres.

VI.2.7. Valeur de la justice

236. La justice, en tant que valeur, doit conduire à examiner la question de l'inégalité. L'utilisation systématique de robots industriels et de robots de service contribuera à accroître le chômage de certains segments de la force de travail. Si l'on ne parvient pas à compenser de quelque façon cette évolution, à fournir du travail aux personnes concernées ou à organiser le lieu de travail d'une manière différente, l'augmentation des inégalités au sein de la société restera une source de préoccupations. Le travail, en effet, demeure un élément essentiel de l'identité individuelle et de la reconnaissance sociale.

237. La justice, en tant que valeur, exige aussi d'aborder la question de la non-discrimination. Les roboticiens devraient être sensibilisés au risque de reproduction des préjugés sexistes et des stéréotypes de genre dans la conception des robots. Le risque de discrimination et de stigmatisation par le biais de l'exploitation des données collectées par les robots n'est pas un problème trivial. Les États devraient prendre des mesures adéquates à ce sujet.

VI.3. Recommandations spécifiques de la COMEST sur l'éthique de la robotique

VI.3.1. Recommandation sur le développement de codes d'éthique de la robotique et des roboticiens

238. Il est recommandé de poursuivre, tant au niveau national qu'international et selon des modalités pluridisciplinaires, le développement, la mise en œuvre, la révision et l'actualisation de codes d'éthique des roboticiens, en prenant en compte les avancées futures possibles de la robotique et leur impact sur la vie humaine et l'environnement (énergie, déchets électroniques, empreinte écologique). Il est également recommandé que les disciplines et les professions qui contribuent de façon importante à la robotique ou peuvent y recourir – notamment l'ingénierie électronique, l'intelligence artificielle, la médecine, les sciences animales, la psychologie et les sciences physiques – révisent, de préférence de manière coordonnée, leurs codes d'éthique propres, en cherchant à anticiper les défis qui pourront résulter de leurs liens avec la robotique et l'industrie des robots. Enfin, il est recommandé que l'éthique - y compris les codes d'éthique, les codes de conduite, et d'autres documents pertinents – devienne une partie intégrante des programmes d'études pour tous les professionnels impliqués dans la conception et la fabrication des robots.

VI.3.2. Recommandation sur la conception éthique des robots

239. Les considérations éthiques devraient être prises en compte dans la conception des technologies robotiques. Les robots utilisent, pour prendre des décisions, des algorithmes qui incarnent des valeurs et des cadres éthiques. En outre, l'introduction de robots a des implications éthiques pour les pratiques concernées, qu'il s'agisse des soins de santé, de l'éducation ou des relations sociales. Pour prendre effectivement en compte la dimension éthique des robots, il est nécessaire que le processus de développement de ces machines accorde une place à l'éthique, en s'appuyant sur une approche telle que la « conception éthique » (Value Sensitive Design). Cette approche devrait aussi être adaptée pour prendre en compte le bien-être des animaux.

VI.3.3. Recommandation sur l'expérimentation

240. Les implications sociales des nouvelles technologies robotiques sont souvent difficiles à prévoir. Pour assurer l'utilisation responsable des robots dans la société, une soigneuse et transparente expérimentation est nécessaire. L'introduction préalable des nouvelles technologies robotiques à petite échelle, dans des situations contrôlées, devrait permettre d'étudier ouvertement leur impact sur les pratiques et activités humaines, ainsi que sur les valeurs et les cadres d'interprétation. Les résultats de telles expérimentations pourraient ensuite être utilisés pour adapter la conception des robots, informer l'élaboration des politiques et de la réglementation, et fournir aux utilisateurs un point de vue critique sur l'emploi de ces technologies.

VI.3.4. Recommandation sur la discussion publique

241. L'introduction de robots aura des effets profonds sur la société et la vie quotidienne des individus. Pour faire face à ces effets de manière responsable, les citoyens devront disposer de concepts, de connaissances et de cadres d'interprétation adéquats. À cette fin, des discussions publiques devraient être organisées au sujet des implications des nouvelles technologies robotiques eu égard à divers aspects de la vie sociale et de la vie quotidienne, y compris l'impact environnemental sur l'ensemble du cycle de production du robot. De telles discussions aideraient les individus à développer une attitude réflexive à ce sujet et contribueraient à mieux sensibiliser les concepteurs et les décideurs.

VI.3.5. Recommandation sur la reconversion et la réadaptation de la force de travail

242. Les robots seront amenés à prendre la place des êtres humains dans un large éventail de domaines. Ils entraîneront une réduction significative des opportunités d'emploi et susciteront aussi la création d'emplois nouveaux dans certains secteurs. Les États, les organisations professionnelles et les établissements d'enseignement devraient donc réfléchir aux conséquences de cette évolution, en prêtant une attention particulière aux secteurs de la société plus susceptibles d'être les plus vulnérables aux changements, et prendre des mesures adéquates en vue de la reconversion et de la réadaptation de la force de travail pour permettre la réalisation d'avantages potentiels.

VI.3.6. Recommandations relatives aux transports et aux véhicules autonomes

243. Les véhicules autonomes (VA) reposent sur deux types d'algorithmes : des algorithmes déterministes, qui contrôlent le comportement du véhicule dans une situation donnée de manière prévisible, et des algorithmes IA cognitifs avec lesquels le comportement du véhicule n'est pas entièrement prévisible. La COMEST formule les recommandations suivantes au sujet des véhicules autonomes.

244. Les fonctions d'un VA qui sont prévisibles (par ex. la protection de la vie privée des utilisateurs) devraient être identifiées et leur mise en œuvre reposer sur des algorithmes déterministes. Eu égard à ces fonctions, les VA pourront être considérés comme une technologie conventionnelle, comparable aux ordinateurs par exemple, et les cadres juridiques et éthiques existants pourront donc leur être appliqués.

245. Cependant, la caractéristique spécifique des VA est leur capacité à fonctionner et à prendre des décisions sur la base d'algorithmes cognitifs d'apprentissage automatique. Même si les critères d'optimisation des algorithmes cognitifs sont transparents, les décisions effectives prises par un VA seront imprévisibles car elles dépendront de l'expérience particulière, fondamentalement aléatoire, du véhicule. La question de savoir à qui incombera la responsabilité des conséquences de ces décisions imprévisibles pose de graves problèmes éthiques.

246. Les cadres juridiques et éthiques existants prévoient en général certains outils pour traiter les situations non déterministes (aléatoires). Un exemple est celui des catastrophes naturelles pour lesquelles des outils d'assurance ont été mis au point afin de répondre à des dommages imprévus. La COMEST recommande de développer un cadre similaire pour traiter les conséquences imprévisibles des décisions de VA.

247. Cependant, on ne peut imaginer que toutes les conséquences imprévisibles des décisions d'un VA puissent être considérées comme des « actes divins » comparables aux catastrophes naturelles. Un être humain devrait conserver le contrôle de certaines décisions des VA car il est impossible de programmer des critères moraux. La COMEST recommande, par conséquent, de recenser et de définir les situations dans lesquelles la responsabilité des résultats de l'action d'un VA incombe uniquement à un être humain (« le conducteur »). L'autonomie du véhicule devrait être limitée eu égard aux décisions susceptibles d'avoir des conséquences mortelles – par exemple, dans les situations où un accident est impossible à éviter – et la responsabilité de ces décisions graves laissée à un être humain. Celui-ci pourrait prendre la décision de façon autonome, suivre l'avis du VA ou même déléguer la décision au VA. Cependant, il resterait responsable dans tous les cas.

VI.3.7. Recommandations sur les Systemes robotiques militaires armés (« drones armés »)

248. Les drones armés donnent à l'humanité la possibilité de faire la guerre à distance. Cette possibilité, superficiellement attrayante car elle réduit la nécessité d'envoyer des hommes sur le terrain, menace de modifier fondamentalement la nature des conflits armés. Ces nouvelles armes soulèvent, par conséquent, de graves questions éthiques et juridiques auxquelles les Etats semblent ne pas avoir répondu jusqu'à présent.

249. Les questions éthiques soulevées par les drones armés ne concernent pas seulement le droit international humanitaire. La guerre à distance que permettent ces nouvelles armes est massivement asymétrique, l'attaquant ayant la possibilité de tuer un adversaire sans être aucunement menacé. On voit mal comment une telle situation pourrait être justifiée d'un point de vue éthique. L'assassinat à distance va aussi à l'encontre du principe de respect de la dignité humaine. Enfin, la possibilité d'entrer en guerre sans exposer ses propres soldats à une menace directe abaisse le coût perçu – et donc le seuil d'activation – d'un conflit armé. Cela ouvre la perspective inquiétante d'une guerre continue à faible coût.

250. L'utilisation de drones armés contre des acteurs non étatiques soupçonnés dans une situation d'insurrection soulève également des questions éthiques et juridiques. L'assassinat ciblé par un drone armé ne supprime pas seulement le droit à la justice des individus directement pris pour cible, qui sont privés de leur droit à une procédure légale. Cette méthode d'assassinat, lorsqu'elle est employée fréquemment en situation de contre-insurrection, a des conséquences négatives sur le plan économique, social et psychologique pour la population civile – en particulier les enfants – et suscite de graves préoccupations éthiques au regard des droits de l'homme.

251. La COMEST considère par conséquent qu'outre les problèmes juridiques que cela pose, l'assassinat d'un être humain par un robot armé est moralement inacceptable. La Commission constate que certains pays utilisent de plus en plus fréquemment des drones armés dans des situations de conflit. Néanmoins, les arguments éthiques et juridiques contre une telle pratique sont suffisamment convaincants pour que la COMEST recommande aux États de la réexaminer, comme ils l'ont fait pour d'autres armes dont l'utilisation a été limitée

ou rendue illégale, par exemple les mines antipersonnel et les armes chimiques et biologiques. Le processus envisagé à l'article 36 – même appliqué rétrospectivement – fournit peut-être un mécanisme pertinent à cette fin. Les États pourraient ainsi, dans la transparence et sur la base des normes internationalement reconnues, non seulement examiner la question de savoir si ces armes peuvent être utilisées de façon légale et éthique dans certains cas, mais aussi identifier les situations où leur emploi serait compatible avec les critères du droit humanitaire et du droit des droits de l'homme. D'un point de vue éthique, il nous paraît douteux que de telles situations existent et nous craignons que, si aucune mesure n'est prise pour empêcher la prolifération de ces armes, se poursuivent des conflits continus à distance, ainsi que le déni de justice que constituent les assassinats ciblés.

VI.3.8. Recommandations sur les armes autonomes

252. Deux remarques s'imposent au sujet de l'utilisation éventuelle d'armes autonomes. Premièrement, d'un point de vue légal, le déploiement de telles armes constituerait une violation du DIH. Deuxièmement, d'un point de vue éthique, ces armes sont contraires au principe directeur selon lequel les décisions de vie et de mort concernant des êtres humains ne doivent en aucun cas être prises par des machines.

253. Au vu de leurs capacités techniques, les armes robotiques autonomes ne disposent pas des composantes essentielles requises pour assurer le respect des principes de distinction et de proportionnalité. Même si l'on peut soutenir qu'elles seront capables de respecter ces principes à l'avenir, de telles spéculations sont dangereuses s'agissant de machines à tuer dont le comportement dans des circonstances particulières est stochastique et donc intrinsèquement imprévisible.

254. Le principe moral selon lequel le pouvoir de recourir à la force létale ne peut légitimement être délégué à une machine – quel qu'en soit le degré d'efficacité – est inscrit dans le droit international : l'acte de tuer doit rester la responsabilité d'un être humain capable de rendre des comptes et tenu de prendre sa décision de manière réfléchie.

255. Notre vigoureuse et seule recommandation est donc qu'il soit maintenu, pour des raisons juridiques, éthiques et militaires-opérationnelles, le contrôle humain sur les systèmes d'armement et l'utilisation de la force. Compte tenu de la rapidité potentielle du développement des armes autonomes, il est urgent, comme le préconise le CICR, « de déterminer le type et le degré de contrôle humain sur l'utilisation de systèmes d'armement qui est jugé nécessaire pour satisfaire aux obligations légales et répondre à certains critères éthiques et sociétaux » (ICRC, 2016).

VI.3.9. Recommandations concernant la surveillance et le travail de police

256. Les États devraient établir des politiques sur l'utilisation de drones dans les activités de surveillance, de police et dans d'autres contextes non militaires. Les modalités d'utilisation de drones par la police devraient être décidées par les représentants du public et non par les forces de police. Les politiques en la matière devraient être claires, écrites et accessibles au public. Elles devraient veiller, au minimum, à ce que le déploiement d'un drone soit autorisé uniquement sur la base d'une commission rogatoire, dans une situation d'urgence ou lorsqu'il existe des raisons spécifiques et clairement définies de penser qu'un drone permettra de recueillir des éléments de preuve relatifs à une infraction pénale particulière. La conservation des images ne devrait être autorisée que s'il existe des raisons de soupçonner qu'elles contiennent des éléments de preuve au sujet d'un délit ou sont pertinentes aux fins d'une enquête ou d'un procès en cours. L'utilisation de drones devrait faire l'objet d'un contrôle ouvert et d'une surveillance adéquate afin de prévenir toute utilisation abusive.

257. Les drones utilisés par la police ne devraient pas être équipés d'armes létales ou non létales.

258. L'utilisation d'armes autonomes par la police ou les services de sécurité ne devrait pas être autorisée.

VI.3.10. Recommandation concernant l'utilisation privée et commerciale des drones

259. L'utilisation privée de drones devrait être soumise à l'octroi d'une licence et leur aire d'utilisation devrait faire l'objet d'un contrôle strict pour des raisons juridiques, de sécurité et de respect de la vie privée. Équiper des drones privés d'armes létales ou non létales devrait être interdit.

VI.3.11. Recommandation sur l'égalité des genres

260. Une attention particulière devrait être accordée aux questions de genre et aux stéréotypes concernant tous les types de robots décrits dans ce rapport, et en particulier, les robots de jouets, les compagnons de sexe et les remplacements de travail.

VI.3.12. Recommandation sur l'évaluation de l'impact sur l'environnement

261. Tout comme d'autres technologies de pointe, l'impact environnemental devrait être considéré comme faisant partie de l'analyse du cycle de vie, permettant d'évaluer de manière plus globale si une utilisation spécifique de la robotique donnera à la société plus de bénéfices que des préjudices. Cela devrait tenir compte des impacts négatifs possibles de la production, de la consommation et du gaspillage (par exemple, l'exploitation minière des terres rares, les déchets électroniques, la consommation d'énergie), ainsi que les avantages environnementaux potentiels. Lors de la construction de robots (nano, micro ou macro), des efforts devraient être faits pour utiliser des matériaux dégradables et des technologies respectueuses de l'environnement, ainsi que pour améliorer le recyclage des matériaux.

VI.3.13. Recommandations sur l'internet des objets

262. L'internet des objets (IdO) est une technologie en plein développement qui permet l'interconnexion d'appareils intelligents de la vie quotidienne, y compris des appareils ménagers. Il rend possible l'utilisation comme capteurs d'appareils de types divers dans le but de recueillir des données massives pouvant être utilisées à de nombreuses fins.

263. Ces réseaux reliant tout individu et toute chose créent des possibilités entièrement nouvelles dans de nombreux domaines comme la réalité augmentée, l'interconnexion tactile, la fabrication distribuée et les villes intelligentes.

264. L'IdO soulève clairement certaines questions éthiques, notamment en matière de respect de la vie privée et de sécurité, mais la prochaine génération de l'IdO, parfois appelée IdO++, présente des enjeux encore plus graves. L'IdO++, en effet, s'appuiera sur l'intelligence artificielle (IA) pour traiter les données recueillies. Les algorithmes cognitifs utilisés, qui agiront comme des apprenants autonomes, pourront aboutir à des résultats imprévisibles.

265. De même, les technologies émergentes créent des robots de petite¹ taille qui peuvent servir de capteurs mobiles, collectant des informations dans des endroits ciblés - élargissant la portée des IdO même au-delà des « choses » existantes.

266. Considérant que les enjeux éthiques liés à l'IdO sont, bien que non identiques, comparables à ceux associés à la robotique cognitive, la Commission recommande d'élargir son travail en ce domaine afin d'étudier les questions liées à l'IdO et de formuler des recommandations pertinentes à ce sujet.

¹ Y compris les micro-robots et les nano-robots

BIBLIOGRAPHIE

- Abney, K. 2012. Robotics, ethical theory, and metaethics: a guide for the perplexed. In: Lin, P., Abney, K. and Bekey, G. A. eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. London, MIT Press, pp. 35-52.
- Allen, C., Varner, G. and Zinser, J. 2000. Prolegomena to any future artificial moral agent, *Experimental and Theoretical Artificial Intelligence*, Vol. 12, No. 3, pp. 251-261.
- Allen, C., Wallach, W. and Smit, I. 2011. Why machine ethics. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 51-61.
- Alliance des sciences et technologies de numérique (Allistene). 2016. *Éthique de la recherche en robotique* [Ethics of research in robotics]. Paris, Allistene.
- Alonso, E., Sherman, A. M., Wallington, T. J., Everson, M. P., Field, F. R., Roth, R., and Kirchain, R. E. 2012. Evaluating rare-earth element availability: a case with revolutionary demand from clean technologies, *Environmental Science and Technology*, Vol. 46, pp. 3406-3414.
- Anderson, M. and Anderson, S. L. 2011a. General introduction. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 1-4.
- Anderson, M. and Anderson, S. L. 2011b. A *prima facie* duty approach to machine ethics: machine learning of features of ethical dilemmas, *prima facie* duties, and decision principles through a dialogue with ethicists. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 476-492.
- Angelo, J. A. 2007. *Robotics: A Reference Guide to the New Technology*. Westport, Greenwood Press.
- Asaro, P. M. 2012. A body to kick, but still no soul to damn: legal perspectives on robotics. In: Lin, P., Abney, K. and Bekey, G. A. eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. London, MIT Press, pp. 169-186.
- Asaro, P. M. 2015. The liability problem for autonomous artificial agents, *2016 AAAI Spring Symposium Series*. Available at: <https://www.aaai.org/ocs/index.php/SSS/SSS16/paper/download/12699/11949>
- Asimov, I. 1942. *Runaround*. New York, Street & Smith.
- Asimov, I. 1950. *I, Robot*. New York, Gnome Press.
- Asimov, I. 1985. *Robots and Empire*. New York, Doubleday.
- Balding, C. 2016. 'Will Robots Ravage the Developing World?', *Bloomberg View*, 25 July. London, Bloomberg. Available at: <http://www.bloomberg.com/view/articles/2016-07-25/will-robots-ravage-the-developing-world>
- Bar-Cohen, Y. and Hanson, D. 2009. *The Coming Robot Revolution: Expectations and Fears About Emerging Intelligent, Humanlike Machines*. New York, Springer-Verlag.
- Bekey, G. A. 2012. Current trends in robotics: technology and ethics. In: Lin, P., Abney, K. and Bekey, G. A. eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. London, MIT Press, pp. 17-34.

Bekey, G., Ambrose, R., Kumar, V., Lavery, D., Sanderson, A., Wilcox, B., Yuh, J., Zheng, Y. 2008. *Robotics: State of the Art and Future Challenges*. London, Imperial College Press.

Benetti, F. B. V. 2012. Exploring the educational potential of robotics in schools: A Systematic review, *Computers & Education*, Vol. 58, No. 3, pp. 978-988.

Bensoussan, J. and Bensoussan, A. eds. 2016. *Comparative Handbook: Robotic Technologies Law*. Brussels, Editions Larcier.

Beran, T. N., Ramirez-Serrano, A., Kuzyk, R., Fior, M. and Nugent, S. 2011. Understanding how children understand robots: Perceived animism in child-robot interaction, *International Journal Human-Computer Studies*, Vol. 69, No. 7-8, pp. 539-550.

Bertoncello, M. and Wee, D. 2015. Ten ways autonomous driving could redefine the automotive world, *Automotive & Assembly*. Available at: <http://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world>

Bonnefon, J.-F., Shariff, A. and Rahwan, I. 2016. The social dilemma of autonomous vehicles, *Science*, Vol. 352, No. 6293, pp. 1573-1576.

Bringsjord, S. and Taylor, J. 2012. The divine-command approach to robot ethics. In: Lin, P., Abney, K. and Bekey, G. A. eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. London, MIT Press, pp. 85-108.

British Standard Institution (BSI). 2016. *Robots and robotic devices: Guide to the ethical design and application of robots and robotic systems*. London, BSI Standards Limited.

Broadbent, E., Stafford, R., MacDonald, B. 2009. Acceptance of Healthcare Robots for the Older Population: Review and Future Directions, *International Journal of Social Robotics*, Vol. 1, pp. 319–330.

Butterfield, A., Ngondi, G. E. and Kerr, A. eds. 2016. *A Dictionary of Computer Science*. Oxford, Oxford University Press.

Calo, R. 2015. Robotics and the lessons of cyberlaw, *California Law Review*, Vol. 103, pp. 513-563.

Carr, N. 2014. *The Glass Cage: Automation and Us*. New York, W.W. Norton and Company.

Chamayou, G. 2015. *A Theory of the Drone*. New York, The New Press.

Chang, C. W., Lee, J. H., Chao, P. Y., Wang, C. Y. and Chen, G. D. 2010. Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school, *Educational Technology & Society*, Vol. 13, No. 2, pp. 13-24.

Coeckelbergh, M. 2012. Can We Trust Robots?, *Ethics and Information Technology*, Vol. 14, pp. 53.

Comitato Nazionale per la Bioetica (CNB, Italy) and Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita (CNBBSV, Italy). 2017. *Developments of Robotics and Roboethics*. Joint Opinion. Rome, CNB and CNBBSV.

Committee of Legal Affairs of the European Parliament (JURI). 2016. *Draft Report with recommendations to the to the Commission on Civil Law Rules on Robotics (2015/2103(INL))*. Brussels, European Parliament. Available at: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML%2BCOMPARL%2BPE-582.443%2B01%2BDOC%2BPDF%2BV0//EN>

Copeland, J. 1993. *Artificial Intelligence: Philosophical Introduction*. New Jersey, Wiley-Blackwell.

Crandall, J. and Armitage, J. 2005. Envisioning the Homefront: Militarization, Tracking and Security Culture, *Journal of Visual Culture*, Vol. 4, No. 1, pp. 17-38.

Diel, J. J., Schmitt, L. M., Villano and Crowell, C. R. 2012. The clinical use of robots for individuals with autism spectrum disorders: A critical review, *Research in Autism Spectrum Disorders*, Vol. 6, No. 1, pp. 249-262.

Driessen, C. P. G. and Heutinck, L. F. M. 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 No. 1, pp. 3-20.

Eguchi, A. 2012. Educational Robotics Theories and Practice: Tips for how to do it Right. In: Barker, S. B. ed. *Robots in K-12 Education*. Hershey, IGI Global, pp. 1-30.

European Commission (EC). 2012. *Public Attitudes Towards Robots*. Report, Special Eurobarometer 382. Brussels, EC. Available at: http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_382_en.pdf

Floridi, L. 2011. On the morality of artificial agents. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 184-212.

Ford, M. 2015. *Rise of the Robots: Technology and the Threat of Jobless Future*. New York, Basic Books.

Frankish, K. and Ramsey, W. M. eds. 2014. *The Cambridge Handbook of Artificial Intelligence*. Cambridge, Cambridge University Press.

Franklin, S. 2014. History, motivations, and core themes. In: Frankish, K. and Ramsey, W. M. eds. *The Cambridge Handbook of Artificial Intelligence*. Cambridge, Cambridge University Press, pp. 15-33.

Frey, C. B. and Osborne, M. 2013. *The Future of Employment*. Oxford, University of Oxford.

Fukuyama, F. 2002. *Our Posthuman Future: Consequences of the Biotechnology Revolution*. New York, Farrar, Straus, and Giroux

Gagnon, J.-A. 2018. Des enfants et des robots [Children and robots]. In: Parizeau, M.-H. ed. *Éthique des robots et transformations sociales* [Ethics of robots and social transformations]. Québec, Les Presses de l'Université Laval.

Gibilisco, S. 2003. *Concise Encyclopedia of Robotics*. New York, McGraw-Hill.

Habermas, J. 2003. *The Future of Human Nature*. Cambridge, Polity Press.

Heacock, M., Kelly, B. C., Asante, K. A., Birnbaum, L. S., Bergman, A. L., Brune, M. N., Buka, I., Carpenter, D. O., Chen, A., Huo, X., Kamel, M., Landrigan, P. J., Magalini, F., Diaz-Barriga,

F., Neira, M., Omar, M., Pascale, A., Ruchirawat, M., Sly, L., Sly, P. D., Van de Berg, M. and Suk, W. A. 2016. E-waste and harm to vulnerable populations: a growing global problem, *Environmental Health Perspectives*, Vol. 124, pp 550-555.

Holder, C., Khurana, V., Harrison, F. and Jacobs, L. 2016. Robotics and law: key legal and regulatory implications of the robotics age (part I of II), *Computer Law & Security Review*, Vol. 32, No. 3, pp. 383-402.

Holloway, L., Bear, C. and Wilkinson, K. 2014. Robotic milking technologies and renegotiating situated ethical relationships on U.K. dairy farms, *Agriculture and Human Values*, Vol. 31, No. 2, pp. 185-199.

Hottois, G., Missa, J-N., Perbal, L. eds. 2015. *Encyclopédie du trans/posthumanisme* [Encyclopedia of trans/posthumanism]. Paris, Vrin.

Hughes, J. 2012. Compassionate IA and selfless robots: a Buddhist approach. In: Lin, P., Abney, K. and Bekey, G. A. eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. London, MIT Press, pp. 69-84.

Husbands, P. 2014. Robotics. In: Frankish, K. and Ramsey, W. M. eds. *The Cambridge Handbook of Artificial Intelligence*. Cambridge, Cambridge University Press, pp. 269-295.

Ihde, D. 1990. *Technology and the Lifeworld*. Bloomington, Indiana University Press.

Institute of Electrical and Electronic Engineers (IEEE) Standards Association. n.d. *The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems*. New Jersey, IEEE Standards Association. Available at: http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html

International Federation of Robotics (IFR). 2016a. *Executive Summary World Robotics 2016: Industrial Robots*. Frankfurt, IFR. Available at: https://ifr.org/img/uploads/Executive_Summary_WR_Industrial_Robots_20161.pdf

IFR. 2016b. *Executive Summary World Robotics 2016: Service Robots*. Frankfurt, IFR. Available at: https://ifr.org/downloads/press/02_2016/Executive_Summary_Service_Robots_2016.pdf

Ingram, B., Jones, D., Lewis, A., Richards, M., Rich, C. and Schachterle, L. 2010. A code of ethics for robotics engineers, *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp.103-104.

International Committee of the Red Cross (ICRC). 1949a. *Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field*. Geneva, ICRC. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/365?OpenDocument>

ICRC. 1949b. *Convention (II) for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea*. Geneva, ICRC. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/370?OpenDocument>

ICRC. 1949c. *Convention (III) relative to the Treatment of Prisoners of War*. Geneva, ICRC. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/375?OpenDocument>

ICRC. 1949d. *Convention (IV) relative to the Protection of Civilian Persons in Time of War*. Geneva, ICRC. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/380>

ICRC. 1977a. *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I)*. Geneva, ICRC. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/470>

ICRC. 1977b. *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II)*. Geneva, ICRC. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/475?OpenDocument>

ICRC. 2007. *Distinction: Protecting Civilians in Armed Conflict*. Geneva, ICRC. Available at: <https://www.icrc.org/en/publication/0904-distinction-protecting-civilians-armed-conflict>

ICRC. 2014. *What is international humanitarian law?* Geneva, ICRC. Available at: <https://www.icrc.org/en/download/file/4541/what-is-ihl-factsheet.pdf>

ICRC. 2016. *Autonomous weapons: Decisions to kill and destroy are a human responsibility*. Published online 11 April 2016, Geneva, ICRC. Available at: <https://www.icrc.org/en/document/statement-icrc-lethal-autonomous-weapons-systems>

International Peace Conference. 1899. *Convention (II) with Respect to the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land*. The Hague, International Peace Conferences. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/INTRO/615?OpenDocument>

Kamga, R., Romero, M., Komis, V. and Mirsili, A. 2016. Design Requirements for Educational Robotics Activities for Sustaining Collaborative Problem Solving. In: Alimisis, D., Moro, M. and Menegatti, E. eds. *Educational Robotics in the Makers Era*. Edurobotics 2016. Advances in Intelligent Systems and Computing, Vol. 560. Cham, Springer.

Kappor, A. 2014. L'invasion robotique au Canada [The invasion of robotics in Canada]. *Canadian Urology Association Journal*, Vol. 8, pp. 5-6.

Kudina, O. and Bas, M. 2017. "The end of privacy as we know it": Reconsidering Public Space in the Age of Google Glass. In: Newell, B. C., Timan, T., Koops, B. J. eds. *Surveillance, Privacy, and Public Space*. United Kingdom, Taylor and Francis.

Leenes, R., Palmerini, E., Koops, B.-J., Bertolini, A., Salvini, P., Lucivero, F. 2017. Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues, *Law, Innovation and Technology*, Vol. 9, No. 1, pp. 1-44.

Leopold, A. 1949. *A Sand County Almanac and Sketches Here and There*. New York, Ballantine Books.

Liberati, N. 2017. Teledildonics and New Ways of "Being in Touch": A Phenomenological Analysis of the Use of Haptic Devices for Intimate Relations, *Science and Engineering Ethics*, Vol. 23, No. 3, pp. 801-823.

Lin, P. 2012. Introduction to robot ethics. In: Lin, P., Abney, K. and Bekey, G. A. eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. London, MIT Press, pp. 3-16.

Matsuzaki, H. and Lindemann, G. 2015. The autonomy-safety-paradox of service robotics in Europe and Japan: a comparative analysis, *IA & Society*, Vol. 31, No. 4, pp. 501-517.

Mavroidis, C. and Dubey, A. 2003. From pulses to motors, *Nature Materials*, Vol. 2, pp. 573-574.

Mavroidis, C., Dubey, A., Yarmush, M. L. 2004. Molecular machines, *Annual Review of Biomedical Engineering*, Vol. 6, pp. 10.1-10.33.

Miller, D. P., Nourbakhsh, I. R. and Sigwart, R. 2008. Robots for education. In: Siciliano, B. and Khatib, O. eds. *Springer handbook of robotics*. New York, Springer.

Mol, A. 1997. *Wat is kiezen? Een empirisch-filosofische verkenning*. Inaugural lecture, Enschede, Universiteit Twente. Available at:
<http://www.stichtingsocrates.nl/tekstenpdf/Wat%20is%20kiezen.pdf>

Moor, J. 2011. The nature, importance, and difficulty of machine ethics. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 13-20.

Mordoch, E., Osterreicher, A., Guse, L., Roger, K. and Thompson, G. 2013. Use of social commitment robots in the care of elderly people with dementia: A literature review, *Maturitas*, Vol. 74, No. 1, pp. 14-20.

Murashov, V., Hearl, F. and Howard, J. 2015. *A Robot may Not Injure a Worker: Working safely with robots*. NIOSH Science Blog, posted 20 November. Available at:
<http://blogs.cdc.gov/niosh-science-blog/2015/11/20/working-with-robots/>

Murphy, R. R. 2000. *Introduction to IA Robotics*. Cambridge, The MIT Press.

OECD Insights (n.a.) 2016. *The rise of the robots – friend or foe for developing countries?* Available at: <http://oecdinsights.org/2016/03/02/the-rise-of-the-robots-friend-or-foe-for-developing-countries/>

Oost, E. and Reed, D. 2010. Towards a Sociological Understanding of Robots as Companions. In: Lamers, M. H., Verbeek, F. J. eds. *Human-Robot Personal Relationships*. Heidelberg, Berlin, Springer.

Pearson, Y. and Borenstein, J. 2014. Creating “companions” for children: the ethics of designing esthetic features for robots, *IA & Society*, Vol. 29, pp. 23-31.

Peláez, L. 2014. *The Robotics Divide. A New Frontier in the 21st Century?* London, Springer-Verlag.

Pilkington, E. 2015. ‘Life as a drone operator: ‘Ever step on ants and never give it another thought?’’, *The Guardian*, 19 November. London, The Guardian. Available at:
<https://www.theguardian.com/world/2015/nov/18/life-as-a-drone-pilot-creech-air-force-base-nevada>

Powers, T. M. 2011. Prospects for a Kantian machine. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 464-475.

Rathenau Instituut. 2017. *Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality*. The Hague, Rathenau Instituut. Available at: <https://www.rathenau.nl/en/publication/human-rights-robot-age-challenges-arising-use-robotics-artificial-intelligence-and>

Riek, L. D. and Howard, D. 2014. A code of ethics for the human-robot interaction profession, *Proceedings of We Robot 2014*. Available at: <http://robots.law.miami.edu/2014/wp-content/uploads/2014/03/a-code-of-ethics-for-the-human-robot-interaction-profession-riek-howard.pdf>

Rifkin, J. 1995. *The End of Work*. New York, Putnam.

Romportl, J., Zackova, E. and Kelemen, J. eds. 2015. *Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*. Switzerland, Springer International Publishing.

The Royal Society. 2017. *Machine learning: the power and promise of computers that learn by example*. London, The Royal Society. Available at: <https://royalsociety.org/~media/policy/projects/machine-learning/publications/machine-learning-report.pdf>

Rosenberg, J. M. 1986. *Dictionary of Artificial Intelligence and Robotics*. New York, John Wiley & Sons.

Rosenberger, R. and Verbeek, P. P. 2015. A Field Guide to Postphenomenology. In: Rosenberger, R., Verbeek, P. P. eds. *Postphenomenological Investigations: Essays on Human-Technology Relations*. London, Lexington Books, pp. 9-41.

Sandel, M. 2009. *The Case Against Perfection: Ethics in the Age of Genetic Engineering*. Cambridge, Harvard University Press.

Savulescu, J. and Maslen, H. 2015. Moral enhancement and artificial intelligence: moral IA? In: Romportl, J., Zackova, E. and Kelemen, J. eds. *Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*. Switzerland, Springer International Publishing, pp. 79-96.

Scassellati, B., Admoni, H., and Mataric, M. 2012. Robots for autism research, *Annual Review of Biomedical Engineering*, Vol. 14, pp. 275-294.

Sharkey, A. and Sharkey, N. 2012. Granny and the robot: Ethical issues in robot care for elderly, *Ethics and Information Technology*, Vol. 14, No. 1, pp. 27-40.

Simut, R. E., Vanderfaeillie, J., Peca, A., Van de Perre, G. and Vanderborght, B. 2016. Children with Autism Spectrum Disorders Make a Fruit Salad with Probo, the Social Robot: An Interaction Study, *Journal of Autism Development Disorder*, Vol. 46, pp. 113-126.

Smith, C., Villanueva, A., Priya, S. 2012. Aurelia aurita bio-inspired tilt sensor, *Smart Materials and Structures*, Vol. 21, No. 10.

Stone, W. L. 2005. The history of robotics. In: Kurfess, T. R. ed. *Robotics and Automation Handbook*. Boca Raton, CRC Press, pp. 1.

Swierstra, T., Stemerding, D. and Boenink, M. 2009. Exploring Techno-Moral Change: The Case of the ObesityPill. In: Sollie, P. and Düwell, M. eds. *Evaluating New Technologies*. Dordrecht, Springer, pp. 119-138.

Ticehurst, R. 1997. The Martens Clause and the Laws of Armed Conflict, *International Review of the Red Cross*, No. 317, pp. 125-134.

Toh, L. P. E., Causo, A., Tzuo, P. W., Chen, I. M. and Yeo, S. H. 2016. A Review on the Use of Robots in Education and Young Children, *Educational Technology & Society*, Vol. 19 No. 2, pp. 148-163.

Torrance, S. 2011. Machine ethics and the idea of a more-than-human moral world. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 115-137.

Tzafestas, S. G. 2016a. *Roboethics. A Navigating Overview*. Switzerland, Springer.

Tzafestas, S. G. 2016b. *Sociorobot World: A Guided Tour for All*. London, Springer.

Ummat, A., Dubey, A. and Mavroidis, C. 2004. Bionanorobotics: a field inspired by nature. In: Bar-Cohen, Y. ed. *CRC Handbook on Biomimetics: Mimicking and Inspiration of Biology*. Boca Raton, CRC Press, pp. 201-226.

United Nations (UN). 1945. *Charter of the United Nations*. New York, UN. Available at: <http://www.un.org/en/sections/un-charter/un-charter-full-text/>

United Nations (UN). 1948. *Universal Declaration of Human Rights*. New York, UN. Available at: <http://www.un.org/en/universal-declaration-human-rights/index.html>

UN. 2010. *Study on targeted killings, Addendum 6 of the Report of the Special Rapporteur on extrajudicial, summary and arbitrary executions, Philip Alston*. New York, UN. Available at: <http://www2.ohchr.org/english/bodies/hrcouncil/docs/14session/A.HRC.14.24.Add6.pdf>

UN. 2015a. *Addis Ababa Action Agenda of the Third International Conference on Financing for Development*. New York, UN. Available at: http://www.un.org/esa/ffd/wp-content/uploads/2015/08/AAAA_Outcome.pdf

UN. 2015b. *Transforming Our World: the 2030 Agenda for Sustainable Development*. Resolution A/RES/70/1. New York, UN. Available at: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

United Nations Educational, Cultural and Scientific Organization (UNESCO). 2005a. *Convention on the Protection and Promotion of the Diversity of Cultural Expressions*. Paris, UNESCO.

UNESCO. 2005b. *The Precautionary Principle: Report of COMEST*. Paris, UNESCO. Available at: <http://unesdoc.unesco.org/images/0013/001395/139578e.pdf>

UNESCO. 2013. *Report to COMEST of Workshop on Ethics of Modern Robotics in Surveillance, Policing and Warfare (University of Birmingham, United Kingdom of Great Britain and Northern Ireland, 20-22 March 2013)*. Working Document. Paris, UNESCO. Available at: <http://unesdoc.unesco.org/images/0022/002264/226478E.pdf>

Valverdu, J. and Casacuberta, D. eds. 2009. *Handbook of Research on Synthetics Emotions and Sociable Robots: New Applications in Affective Computing and Artificial Intelligence*. Hershey, IGI Publishing.

Van de Poel, I. 2013. Why New Technologies Should Be Conceived as Social Experiments, *Ethics, Policy & Environment*, Vol. 16, No. 3, pp. 352-55.

Van de Poel, I. 2016. An Ethical Framework for Evaluating Experimental Technology, *Science and Engineering Ethics*, Vol. 22, No. 3, pp. 667-686.

Verbeek, P. P. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago, University of Chicago Press.

Van Rysewyk, S. P. and Pontier, M. 2015. Preface. In: Van Rysewyk, S. P. and Pontier, M. eds. *Machine Medical Ethics*. London, Springer.

Vandemeulebroucke, T., De Casterle, B. D. and Gastmans, C. 2017. How do older adults experience and perceive socially assistive robots in aged care: a systematic review of qualitative evidence, *Aging and Mental Health*, published online 9 February 2017, pp. 1-19.

Verbeek, P. P. 2013. Technology Design as Experimental Ethics. In: Van den Burg, S. and Swierstra, Tsj. eds. *Ethics on the Laboratory Floor*. Basingstoke, Palgrave Macmillan, pp. 83-100.

Veruggio, G. 2002. Views and visions in robotics. Hearing at the Italian Senate's 7th Permanent Commission (Rome).

Veruggio, G. and Operto, F. 2008. Roboethics: social and ethical implications of robotics. In: Siciliano, B. and Khatib, O. eds. *Springer Handbook of Robotics*. London, Springer, pp. 1499-1524.

Waelbers, K. and Swierstra, Tsj. 2014. The family of the future: how technologies can lead to moral change. In: Hoven, J. V. D., Doorn, N., Swierstra, T., Koops, B-J. and Romijn, H. eds. *Innovative Solutions for Global Issues*. Dordrecht, Springer, pp. 219-236.

Wallach, W. and Allen, C. 2009. *Moral Machines: Teaching Robots Right from Wrong*. Oxford, Oxford University Press.

Warwick, K. 2012. *Artificial Intelligence: The Basics*. New York, Routledge.

Weir, N. A., Sierra, D. P., and Jones, J. F. 2005. *A Review of Research in the Field of Nanorobotics*. SAND2005-6808, Unlimited Release.

Whitby, B. 2011. On computable morality: an examination of machines as moral advisors. In: Anderson, M. and Anderson, S. L. eds. *Machine Ethics*. Cambridge, Cambridge University Press, pp. 138-150.

Wildmer, G., Oswald-Krapf, H., Sinha-Khetriwal, D., Schnellmann, M. and Boni, H. 2005. Global perspectives on e-waste, *Environmental Impact Assessment Review*, Vol. 25, No. 5, pp. 436-458.

Wise, E. 2005. *Robotics Demystified*. New York, McGraw-Hill.

Wu, Y.-H., Faucounau, V., Boulay, M., Maestrutti, M. and Rigaud, A. S. 2010. Robotic agents for supporting community-dwelling elderly people with memory complaints: Perceived needs and preferences, *Health Informatics Journal*, Vol. 17, No. 1, pp. 33-40.

Zhang, M., Davis, C. T. and Xie, S. 2013. Effectiveness of robot-assisted therapy on ankle rehabilitation – a systematic review, *Journal of Neuroengineering and Rehabilitation*, Vol. 10, No. 30, pp. 1-16.