



HAL
open science

Les revendications de droits pour les robots : constructions et conflits autour d'une éthique de la robotique

Charles Corval

► To cite this version:

Charles Corval. Les revendications de droits pour les robots: constructions et conflits autour d'une éthique de la robotique. Implications philosophiques, 2023, <https://www.implications-philosophiques.org/les-revendications-de-droits-pour-les-robots-constructions-et-conflits-autour-dune-ethique-de-la-robotique/>. 10.5281/zenodo.8267808 . hal-04183760

HAL Id: hal-04183760

<https://sciencespo.hal.science/hal-04183760>

Submitted on 21 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

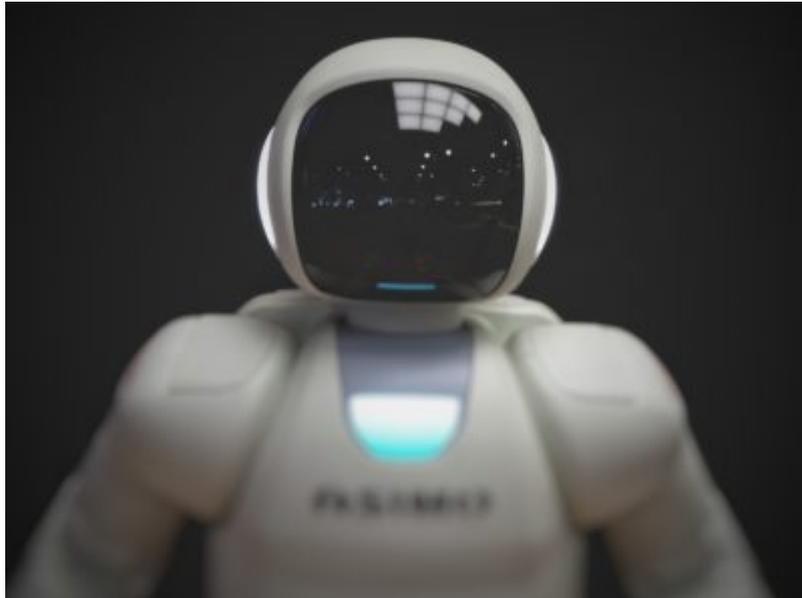


Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

Les revendications de droits pour les robots : constructions et conflits autour d'une éthique de la robotique

implications-philosophiques.org/les-revendications-de-droits-pour-les-robots-constructions-et-conflits-autour-dune-ethique-de-la-robotique/

19 juillet 2023



Charles Corval est doctorant en théorie politique au centre d'étude de la vie politique de Sciences Po (CEVIPOF). Dirigées par Frédéric Gros et Pierre-Antoine Chardel, ses recherches portent sur l'éthique et la gouvernementalité des nouvelles technologies. Il est actuellement lauréat de la Bourse Fulbright et chercheur-étudiant invité à l'Université de Berkeley afin d'effectuer une enquête en Californie. Il y organise des entretiens avec des producteurs d'objets connectés de la Silicon Valley afin de cerner leurs perceptions de l'avenir et de l'éthique.

Résumé

Ce travail examine les revendications contemporaines de droits pour les robots. Il présente les principales formes argumentatives qui ont été développées en faveur d'une considération éthique ou de droits positifs pour ces machines. Il met en relation ces argumentations avec un travail de recherche-action afin de produire un retour critique sur l'idée de droit des robots. Il montre enfin le rapport complexe entre les récits de la modernité et la revendication de droits pour les robots.

Mots-clés : robot, droit, revendication, modernité, robotique

Abstract

This article presents contemporary vindications of robot rights. The first part exposes main argumentations in favor of moral consideration or legal personhood for robots. The second part builds on cooperative action research with several French roboticists and experts in order to evaluate these arguments. These dialogues eventually show the complex relationships between the narratives of modernity and the vindication of robot rights.

Keywords: robot, rights, law, vindication, modernity, robotics

Introduction : une rapide histoire des revendications de droits pour les robots[1]

Le 16 février 2017, le Parlement européen vote une motion qui propose de lancer une réflexion sur la création d'une personnalité juridique spécifique aux robots[2], mesure qui ne sera finalement pas adoptée. Six mois plus tard, le 24 octobre 2017, lors du forum *Future Investment Initiative* qui s'est déroulé à Riyad, l'Arabie Saoudite accorde la nationalité au robot humanoïde Sophia[3]. Ces deux événements exemplifient dans le champ politique les premières concrétisations d'une revendication de droits pour les robots. Il s'agit de discuter politiquement la reconnaissance et la protection de certaines machines et plus seulement leur régulation technique.

La compréhension des revendications de droits pour des machines ne peut se faire qu'en rappelant sommairement la logique d'extension constante des droits personnels. Depuis la mise en avant par le libéralisme et le constitutionnalisme modernes de la valeur de l'individualité et de la rationalisation de ses rapports avec l'autorité politique, la question des droits s'est focalisée autour de l'homme. Le XIX^{ème} siècle a amorcé un ensemble de querelles encore présentes autour des droits humains. Comment peuvent-ils être promus et défendus ? Sont-ils efficaces ? Quelles fonctions jouent-ils dans l'ordre social ? Dans quelle mesure s'opposent-ils au choix démocratique ? Depuis les années 1960, un autre ensemble de problèmes est toutefois apparu autour de la question des droits. Il ne s'agit plus seulement de discuter de l'effectivité du droit et de son bien-fondé, mais de questionner la figure de son porteur, classiquement pensé comme l'individu masculin, majeur et propriétaire. Des réflexions nouvelles apparaissent d'abord pour étendre cette figure aux femmes, aux enfants et aux plus fragiles ; ensuite autour du patrimoine, des droits des animaux, de la nature, des générations futures et plus récemment des robots. L'approche critique de l'anthropocentrisme a ouvert des recherches de plus en plus nombreuses sur la possibilité de protéger des choses et des êtres non-humains. Ces divers phénomènes socio-politiques sont loin d'être homogènes mais engagent tous un certain rapport au droit et à la morale, dans une démarche de protection d'un ensemble de groupes et choses historiquement dominés.

La réclamation de droit pour les robots se logent entre ces deux pôles de la contrainte et de l'appel à des obligations pour les humains. La distinction du droit au sens moral (droit naturel) et du droit positif permet de réclamer parfois l'un sans l'autre : ici une considération morale mais sans dispositif de contrainte pour protéger des robots. D'un

point de vue stratégique, le droit positif est alors un moyen de porter l'attention sur une éthique. L'efficacité du droit positif peut se révéler presque nulle dans sa dimension punitive tout en ayant le pouvoir de diffuser la représentation d'une vulnérabilité. La réclamation de droit est alors proche du *soft law* : un support en faveur de l'autorégulation ou une arme pour blâmer les auteurs de comportements pensés comme nuisibles[4].

Ces observations introductives mettent en perspective ce que sont les revendications de droits pour les robots : des tentatives pour encadrer les relations entre les humains et certaines machines imitant la vie. En premier lieu, elles ont été historiquement pensées dans la littérature, souvent sous l'angle d'un retournement montrant la cruauté d'un créateur envers sa créature robotique. C'est après le succès de ce genre de roman que le droit des robots a fait l'objet d'argumentations philosophiques et militantes. Depuis les années 1990, les robots sont devenus des entités dont l'encadrement moral et juridique peut faire l'objet de propositions effectives. Cette évolution est liée à une première période de robotisation dans l'industrie depuis les années 1960 puis aux annonces de grands projets pour la robotique de service depuis les années 2000[5]. Les revendications de droits pour les robots ont depuis lors pris des formes de plus en plus élaborées et concrètes. Pour certaines, elles concernent avant tout la préparation d'un avenir imminent où les robots rejoindront l'intelligence humaine, d'autres se conçoivent comme une opportunité pour rationaliser des questions juridiques, pour d'autres encore, le but est surtout de concevoir une discussion des interactions avec les robots.

Dans ce travail, je catégoriserai et présenterai les arguments en faveur du droit des robots en observant les principaux discours diffusés dans la sphère publique. L'ambition est également d'utiliser les pratiques sociales pour compléter la réflexion morale. Ce point sera particulièrement développé dans une deuxième partie où je proposerai de revenir sur les dialogues que j'ai pu établir au cours de l'année 2019 avec plusieurs experts de la robotique. Ces discussions donnent à voir les réactions des spécialistes aux divers arguments portés en faveur du droit des robots. La conclusion de ce travail permettra d'ouvrir la réflexion à un retour critique sur le droit des robots comme phénomène politique.

I. Les formes de la revendication d'un droit pour les robots

Sans rentrer dans l'étude de l'institutionnalisation des questionnements éthiques[6] concernant la robotique (nommés *machine ethics* et *robot ethics*), il est possible de distinguer quatre grandes formes argumentatives qui se sont construites autour de l'idée d'un droit des robots. Ces formes argumentatives mobilisent des foyers de sens parfois anciens et inscrivent la robotique dans des contextes politiques spécifiques. Le passage par la notion de *forme* ici doit permettre de réunir les diversités argumentatives dans des structures communes.

I.1. Une forme argumentative futuriste

La forme argumentative la plus ancienne renvoie à la mise en adéquation du droit avec un avenir hypothétique. Elle peut se qualifier comme « futuriste » dans la mesure où son projet est orienté par la prédiction de l'avenir sociotechnique. Elle se trouve principalement chez des universitaires et entrepreneurs de la robotique. La nécessité d'une préparation est justifiée par l'éminence du phénomène technique et par l'injustice qui se déploierait si ce phénomène ne pouvait être encadré comme il se doit. Cette approche se demande à quelles conditions un robot pourrait avoir des droits. Une telle interrogation présuppose trois enjeux distincts : premièrement, le fait que l'accès à un droit ou une considération morale soit directement lié à une condition ; deuxièmement, que cette condition soit identifiable par l'être humain ; troisièmement, la spécification de celle-ci. La première présupposition est partagée par l'ensemble des arguments futuristes : il faut une condition spécifique pour qu'un robot quitte le champ de l'objet et devienne digne de droits. La seconde présupposition fait l'objet de plusieurs débats à l'intérieur du groupe. Elle reprend la formulation du fameux test de Turing et plus globalement du protocole par lequel une compétence peut être identifiée. La troisième étape du raisonnement fait l'objet de débats. Certains avancent que la souffrance doit être le critère principal ; d'autres qu'il s'agit de la conscience ; d'autres encore, qu'il faut un acte de réclamation pour qu'un droit soit attribué. La formulation la plus simple de l'argument se repère dans les articles de presse : les robots vont devenir intelligents, donc il faut se préparer à leur accorder des droits. On trouve par exemple ce type de prédictions, souvent interrogatives et prudentes, chez Peter Asaro. Ce dernier envisage un critère pour attribuer des droits aux robots : le fait qu'ils soient capables d'en demander par un auto-apprentissage[Z]. Kevin Warwick, quant à lui, ayant travaillé sur l'usage d'une culture de cellules pour construire un « cerveau » dans un robot, posait la question quantitative du nombre de neurones nécessaire pour savoir à partir de quand un robot doit pouvoir obtenir des droits, a fortiori si son système neurologique dépasse la complexité matérielle de celui de l'Homme[8].

Historiquement, un des premiers arguments futuristes peut se trouver chez Laszlo Versenyi qui publie en 1974 un article intitulé « *Can Robots be Moral ?* ». Dans celui-ci, Versenyi s'interroge sur la manière dont un robot peut nous obliger moralement. Son travail se focalise sur la doctrine platonicienne et propose comme critère de valorisation des robots la connaissance rationnelle[9]. Eric Schwitzgebel et Mara Garza ont repéré l'archétype de ce genre de raisonnement. Respectivement professeur de philosophie et étudiante en philosophie, les deux auteurs résument l'argument comme il suit :

Prémisse 1 : Si l'entité A mérite quelque degré de considération morale et que l'entité B ne la mérite pas, il doit y avoir une différence pertinente entre les deux entités pour fonder cette différence de statut moral.

Prémisse 2 : Il y a des intelligences artificielles qui ne diffèrent pas de manière pertinente des êtres humains.

Conclusion : Donc, il y a des intelligences artificielles qui méritent un degré de considération morale similaire à celle des êtres humains.[10]

Cette présentation a l'originalité d'inverser l'approche de sorte que l'enjeu n'est pas de prouver la ressemblance entre l'humain et le robot, mais la différence. La charge de la preuve est alors du côté de ceux qui affirment une différence ontologique entre les hommes et les machines. Sur le fond, la structure de l'argument reste la même : la considération morale pour le robot dépend de la comparaison entre son statut ontologique et celui de l'humanité. Le critère ici n'est ni un seuil d'intelligence, de sensibilité ou de complexité biologique, mais la ressemblance avec ce qui est déjà un sujet de droit. La finalité argumentative est donc d'offrir un critère ontologique qui servirait à préparer la perception morale à accueillir de nouveaux agents artificiels.

Du fait de sa structure, cette argumentation est la plus diffusée sur la question[11]. Cependant, cette revendication n'est que rarement explicite dans ses concrétisations pratiques. Le droit est souvent conçu dans son versant moral, comme considération pour le patient d'une action qui mériterait de se voir protéger de certains comportements. La dimension futuriste repose dans une large mesure sur cette indétermination pratique de ce que produirait comme rapports entre humains et robots l'acquisition effective de ces nouveaux droits.

I.2. Une forme argumentative juridique

Une autre forme argumentative assez populaire comble cet enjeu pratique. Il s'agit des approches juridiques qui superposent à l'aspect éthique les problèmes de la pratique du droit. Cette forme argumentative se trouve non seulement chez des universitaires du droit, mais également chez des praticiens.

Ce sont deux avocats qui ont diffusé en France un texte en faveur du droit des robots. Celui-ci se conçoit à la fois comme l'exposition des problèmes juridiques posés par la diffusion de la robotique et comme la proposition de solutions pour les traiter. Alain Bensoussan et JérémY Bensoussan proposent ainsi d'utiliser une personnalité « robotique » à la manière de la personnalité morale pour organiser un nouvel espace de la responsabilité. Bruno Maisonnier, fondateur de l'entreprise *Aldebaran Robotics* résume ainsi le cœur de l'argument juridique en question :

[Les robots] auront été dotés d'une capacité d'apprentissage, et sans doute, d'auto-apprentissage : certes très limitées et contraintes, par les temps qui courent, l'état de la technologie actuellement ne sachant pas vraiment faire mieux, mais qui va croître avec le temps ; jusqu'à devenir exponentielles (...) Les comportements d'un robot à un moment donné résulteront donc d'un mélange entre les capacités dont le fabricant l'aura doté, des expériences qu'il aura vécues en propre, incluant les interactions avec ses différents interlocuteurs, dont principalement ses « propriétaires », et les leçons qu'il aura choisi de retenir issues d'expériences de d'autres robots. Qui pourra être identifié comme « moralement » responsable d'un comportement à un moment donné ? Sans doute personne.[12]

Le développement d'objets capables d'acquiescer une autonomie décisionnelle poserait la question de l'impossible traçabilité de la décision et par-là exigerait un régime de droit qui définit clairement la responsabilité. Alain et Jérémy Bensoussan se reposent principalement sur cette prémisse pour dérouler la suite de leur raisonnement : si le comportement du robot ne peut être pleinement prédit par son producteur, opérateur, ou propriétaire, alors on doit considérer qu'il dispose d'une marge d'indétermination analogue à la liberté des sujets. L'établissement de la responsabilité avec des moyens de traçabilité limités produit l'idée qu'il faut organiser la robotique selon des règles spécifiques. La volonté de passer par une production juridique nouvelle vise la création d'un registre attribuant une identité originale au robot – pour faciliter son identification. Un capital lié à cette identité constituerait un fond d'indemnisation en cas de défaillance. Le vocabulaire de l'acceptabilité et de la gestion sociale des risques est omniprésent derrière cette approche juridique et débouche sur une charte des droits des robots[13]. Ses différents articles encadrent une protection assez vague et symbolique de la machine en question[14]. Ils organisent également une structure juridique capable de prendre en charge le déploiement économique des robots[15].

Cette approche n'est toutefois pas exclusive aux praticiens du droit français ; aux Etats-Unis, dès 1992, Lawrence Solum, à la suite des travaux de Christopher Stone, avait produit un long article sur la personnalité juridique pour les formes d'intelligence artificielle[16]. Il défendait qu'en plusieurs situations la personnalité juridique pouvait se révéler efficace pour encadrer le déploiement d'intelligences artificielles, notamment pour la gestion d'actifs économiques par des machines. La possibilité de faire jouer des droits constitutionnels pour les machines n'en restait cependant qu'à l'état d'expérience de pensée. C'est plus récemment qu'une juriste nord-américaine a proposé une approche de la question de la personnalité juridique sous un angle novateur. Christina Mulligan reprend l'idée d'une impossible traçabilité, mais ajoute la question de la punition et de son rôle social. L'argument se soucie de l'anthropomorphisation du robot par les êtres humains. Lorsqu'un humain se trouve victime d'un robot, il ne peut imputer la responsabilité à un être biologique spécifique et la cause même du tort reste invisible. Mulligan considère que les victimes ont tout à gagner et les robots rien à perdre à ce qu'une punition soit possible pour un être artificiel. La personnalité du robot ici est reconnue dans la mesure où il est l'agent auquel on peut imputer un acte, quand bien même le motif de l'acte repose sur d'autres agents. Cette imputation doit justifier que cet agent artificiel puisse être puni au nom de la victime. Mulligan mélange ici les considérations psycho-sociales concernant la satisfaction vis-à-vis de la peine et l'enjeu d'une efficacité juridique dans la compensation des torts. Elle achève ainsi sa proposition comme il suit :

Mais finalement, la conséquence la plus satisfaisante, pour une personne flouée par un robot, pourrait être la pratique du haut Moyen Âge, nommé « noxal surrender », où « les animaux ou objets causant de sérieux dommages ou la mort », désignés comme « banes », étaient données directement à la victime ou à sa famille. L'opportunité de

prendre le contrôle d'un robot pour pouvoir le détruire, combinée avec la reconnaissance sociale de recevoir l'appareil par la loi, pourraient ensemble servir au mieux la satisfaction des victimes.[17]

La spécificité de cette argumentation tient à ce que la définition ontologique du robot passe au second plan dans la considération juridique. Si les propos de Alain et Jérémy Bensoussan ainsi que ceux de Solum, reposent largement sur des prédictions de ce que sera le futur du robot – mobilisant des formes d'arguments futuristes au secours de leurs ambitions juridiques –, Mulligan se préoccupe avant tout de la manière dont l'humain conçoit ses interactions avec le robot.

I.3. Une forme argumentative déontologique

Cette absence de préoccupation pour la définition ontologique se retrouve dans les approches déontologiques du droit des robots. Les deux principaux représentants de cette forme d'argumentation sont Blay Whitby et Kate Darling. Sa structure est la suivante : étant admis qu'un robot ressemble à un être humain ou un être vivant, un abus qui est commis en raison de cette ressemblance pose un problème moral et social. Ainsi, peu importe que l'objet dispose de caractéristiques complexes ou pas, un abus envers lui outrepassé une loi morale qui implique de respecter la dignité humaine, quand bien même celle-ci serait-elle simulée grossièrement dans un objet. En 2008, Blay Whitby publie ainsi un article en forme d'appel pour soutenir cette réflexion[18]. Il développe une argumentation qui ne débouche pas directement sur la demande d'un droit positif. Sa construction est construite en deux temps. Premièrement, il défend qu'un robot, même sommaire, ne peut pas être assimilé à un simple bien de consommation. Whitby avance ainsi que le robot dispose d'une spécificité parce qu'il ressemble aux êtres vivants dans trois aspects : premièrement, il peut avoir une apparence physique humaine ou animalisée ; deuxièmement, il dispose d'un agir dans le monde réel, ce qui tend à lui accorder une apparence de vie ; troisièmement, il a un but à réaliser, souvent en remplacement de l'humain ou de l'animal. De cette ressemblance, Whitby remarque que l'habitude de maltraiter ce qui ressemble à l'humain pourrait habituer ou entraîner à commettre des sévices envers des êtres vivants[19]. Cette approche conséquentialiste est doublée d'une approche de principe :

Une seconde revendication peut être subsumée par l'argument que maltraiter les entités humanoïdes est mal. Cela revient à dire que certaines activités sont simplement inacceptables, même si elles sont faites en privées à un artefact incapable de souffrir.[20]

Une telle proposition revient à considérer que des contraintes doivent s'appliquer hors du cadre du consentement. La question posée par la théorie politique, qui cherche à savoir si une relation librement consentie entre individus est toujours acceptable, quand bien même elle implique une subordination, une alinéation, est ici déplacée dans le champ des relations entre humains et objets. Certains objets, en raison d'une ressemblance particulière avec l'être humain, devraient ainsi être considérés dignement.

Le propos de Kate Darling approfondit ce point via un travail psychologique. Darling explique que les observations des relations entre humains et robots permettent de saisir que ceux-ci se distinguent des autres objets par l'intensité des relations narratives dont ils sont le support. Le robot est un objet dont la finalité sociale est l'anthropomorphisation ou l'animalisation[21]. Un robot imite le mouvement et les fonctions de l'humain ou de l'animal ; il implique donc nécessairement une confusion entre le vivant et la machine. Darling se repose en outre sur les résultats des travaux de Sherry Turkle au sujet du « *caregiver effect* », c'est-à-dire l'amplification de l'importance attribuée à certains objets technologiques parce qu'on doit s'occuper d'eux[22]. La dernière considération empirique de Darling repose sur l'observation de nos intuitions morales : la maltraitance de robots produit chez la plupart des sujets humains une condamnation morale intuitive.

Ce que nous voyons à partir de ces exemples et de nos études actuelles, c'est que les comportements violents envers les objets robotiques paraissent mauvais à beaucoup d'entre-nous, même si nous savons que le robot abusé ne ressent rien. Nous pourrions être câblés pour répondre instinctivement aux signaux comme la peine ou le besoin simulé. Si cet inconfort devient plus fort avec les technologies de plus en plus sophistiquées, comment devrions-nous y faire face ? Devrions nous l'encourager ou le décourager ? Pourrions-nous même atteindre un point où nous voudrions des règles pour gouverner le traitement des robots compagnons ?[23]

Ces trois considérations sur les relations avec les robots permettent de reconnaître ces derniers comme des objets exceptionnels dont on peut affirmer la spécificité morale. Darling utilise ensuite un raisonnement kantien bien connu à propos des animaux : si ces derniers ne peuvent être reconnus comme des êtres rationnels, l'homme doit s'autocontraindre dans les actions qu'il exerce sur eux[24]. C'est pourquoi le modèle proposé par Darling admet une protection des robots sans leur accorder une personnalité : il suffit de produire des lois restrictives concernant certains comportements avec des créatures anthropomorphiques. Le robot ne dispose donc pas en toute rigueur de droits ; mais les humains disposent de devoirs envers lui. Et ces protections ne préjugent nullement des intérêts ou de la liberté du robot. Elles ne se basent que sur la présupposition qu'on attribue au robot une valeur morale lors d'une interaction, tant et si bien que c'est pour éviter que les hommes s'habituent à l'immoralité dans les comportements avec ceux-ci que le droit doit intervenir. Le droit des robots en tant que protection légale des robots vise ainsi à protéger des principes dans la communauté humaine.

I.4. Une forme argumentative relationnelle

La valeur de l'interaction entre l'humanité et les machines complexes se trouve au cœur d'une dernière forme argumentative que l'on peut qualifier de relationnelle. La spécificité de cette argumentation tient à son ancrage dans une tradition philosophique « continentale » ; elle mobilise les phénoménologues allemands et les philosophes français de l'après-guerre. Elle cherche à prendre un recul critique vis-à-vis des

arguments basés sur des principes ontologiques qui impliquent de définir la machine et la dignité de l'Homme. Les deux représentants principaux de cette approche sont Mark Coeckelbergh et David Gunkel.

L'argument de Coeckelbergh peut se résumer avec deux thèses principales : premièrement, l'idée de donner des droits à des êtres en partant de la description de caractéristiques est vouée à l'échec, épistémologiquement et moralement ; deuxièmement, la considération morale d'une chose ne peut donc reposer que sur une ontologie fondée dans les relations. La première proposition reprend des arguments classiques sur l'impossibilité de tirer une description absolument fiable de ce qu'un être est, et encore moins de passer de l'être au devoir-être[25]. La seconde proposition mobilise une alternative à la discussion morale basée sur une ontologie stable, pour partir des relations qui peuvent être tissées entre les êtres. Autrement dit, Coeckelbergh inverse le raisonnement : plutôt que de penser que les relations doivent être encadrées par le droit pour être justes, et que ce droit doit découler de ce que sont les êtres, il propose de penser ce qu'est la moralité à partir de la relation[26]. Coeckelbergh remarque que cette approche dispose de nombreux avantages : elle évite les généralisations excessives, interdit de parler de droits des robots, au bénéfice d'une considération morale pour une singularité robotique ; elle ne réduit pas l'être à une information ; elle permet une flexibilité dans les perceptions technologiques et naturelles ; et enfin, elle explique pourquoi malgré les droits de l'homme, des considérations morales variables peuvent se construire entre les humains. Peut-elle pour autant produire une éthique positive des non-humains ? Selon Coeckelbergh, oui, à condition de la concevoir cumulativement : en démultipliant le résultat des relations que les individus et les groupes peuvent entretenir et développer avec des entités, on obtient une vision globale de toutes les considérations morales qui se composent dans une société déterminée. D'un point de vue pratique, la conclusion de Coeckelbergh renvoie la question du droit des robots à une détermination collective, dans la mesure où un ensemble de coutumes se diffusent ou non dans la société.

Une seconde version de cette argumentation se trouve dans l'ouvrage *Robot Rights*, de David Gunkel. Ce dernier intègre la critique de Coeckelbergh et la reprend contre l'argument futuriste ou juridique. Tout son ouvrage se base sur l'impossibilité de passer du *is* au *ought to*[27] ; de l'être au devoir-être. Gunkel propose ainsi une typologie capable de distinguer quatre possibilités théoriques. Ces quatre possibilités se constituent autour de deux questions : « les robots peuvent-ils avoir des droits ? » et « les robots doivent-ils avoir des droits ? ». Ces questions ouvrent quatre possibilités :

1. Les robots peuvent avoir des droits et doivent avoir des droits.
2. Les robots peuvent avoir des droits mais ne doivent pas en avoir.
3. Les robots ne peuvent avoir de droits, mais devraient en avoir.
4. Les robots ne peuvent avoir de droits et ne devraient pas en avoir.

Pour Gunkel, ces quatre possibilités tombent toutes dans l'idée que l'ontologie devrait dicter la règle morale. Une seule position paraît en effet se jouer du rapport de conditionnalité entre l'être et le droit : la troisième position, où la moralité dépasserait le

fait. Cette position est ramenée par Gunkel à l'argument de Darling et à une approche déontologique. Il cache en vérité une conditionnalité implicite : les robots ne peuvent pas avoir de droits, mais devraient en avoir, pour protéger les principes humains[28]. Il ne se préoccupe pas, par conséquent, de la valeur morale propre du robot, mais de la valeur morale que la relation avec le robot peut produire. Pour cette raison, Gunkel rejoint directement la conclusion de Coeckelbergh : le droit des robots n'a pour l'instant pas été produit en dehors d'une ontologie classique et humaniste.

Toutefois Gunkel ajoute à l'approche de Coeckelbergh la question de l'altérité que constitue le robot. Pour Gunkel, une fois retiré le cadrage ontologique comme base de la réflexion éthique, le robot apparaît comme une étrange altérité. La position anthropocentrée se trouve dans une situation intenable et empêche de considérer des arguments éthiques pour exclure *a priori* du cercle moral[29]. Mais, pour autant, Gunkel reconnaît qu'aucune base ne permet non plus de tirer une norme morale spécifique pour protéger les robots. Sa conclusion est donc que le problème du droit des robots constitue pour l'heure une question indécidable, mais que cette indécidabilité n'est pas négative[30]. Au contraire, la « hantise » d'un droit des robots permet de ramener une inquiétude morale vis-à-vis de nos actes. Ici, Gunkel se dit donc militant moins en faveur du droit des robots que de sa discussion et de ce qu'il peut apporter à l'espace public. Cette position amène à faire du droit des robots moins une question à laquelle il faudrait donner une solution pratique et technique, qu'une expérience de l'éthique démocratique face au développement de son pouvoir d'agir. Les travaux de Coeckelbergh et Gunkel sont davantage des appels à l'activisme, à la société civile et à l'inventivité locale, que l'imposition d'un droit positif vertical. Dit autrement, le droit des robots ne doit pas être *imposé*, mais *construit* relativement à des situations et des contextes propres à des communautés politiques en relation directe avec les machines.

II. Contre la revendication, la réactivation du programme de la modernité

Le passage en revue de ces diverses formes argumentatives donne à voir à la fois une grande diversité des propos et dans le même temps de nombreuses fragilités théoriques. Sans rentrer dans une critique rigoureuse de chaque argument, on peut néanmoins remarquer que beaucoup d'entre eux reposent sur une dynamique prospective pour le moins incertaine – beaucoup citent Ray Kurzweil, dont les prédictions sont encore loin de se réaliser[31] –, ou sur des constructions juridiques dont la redondance avec les dispositifs déjà existants laisse sceptique[32]. Les approches déontologiques, quant à elles, se basent sur la présupposition d'une anthropomorphisation naturelle et indépassable dans l'interaction entre humains et robots, sans tenir compte des questions de croyances personnelles impliquées dans ce type de perception. La dénonciation sociale et quotidienne de l'anthropomorphisation – que cela concerne les animaux ou les choses – comme une pratique enfantine et risible suffit à monter qu'un recul existe sur ce genre de comportements. Par ailleurs, pourquoi les robots se verraient-ils protégés de maltraitance et pas des personnages virtuels dans un jeu vidéo ? Il faudrait pouvoir prouver la naturalité d'une intensité relationnelle plus forte avec un robot physique qu'avec une image. Enfin, les arguments relationnels renvoient la question du droit des

robots à l'espace public et aux pratiques, mais ne discutent que peu des rapports de force. La promotion d'un droit pour les robots par un individu n'est pas un fait abstrait, mais dépend concrètement de ressources sociales et économiques, des relations qu'il peut entretenir avec le pouvoir et de l'ouverture de fenêtres d'opportunités auprès des élites politiques. L'espace du dialogue vis-à-vis d'une machine, qui est également une marchandise au cœur de divers intérêts industriels et politiques contradictoires, doit aussi se concevoir comme un problème et pas seulement comme une solution.

Ces quelques critiques sont apparues à de nombreuses reprises dans un travail de terrain que j'ai pu entreprendre auprès de plusieurs roboticiens et théoriciens de la robotique. Le but de ces huit entretiens était d'explorer les problèmes posés par les argumentations en faveur du droit des robots. Ces entretiens peu directifs devaient permettre de tester les réactions de mes interlocuteurs aux arguments précédemment exposés et de travailler avec eux, dans une perspective collaborative, sur la valeur du débat autour de la revendication de droits pour les robots[33].

Nom	Activités	Date de l'entretien
Dominique Lestel	Philosophe	27/03/2019
Raja Chatila	Roboticien universitaire	30/03/2019
Natalie Nevejans	Juriste	05/04/2019
Jérôme Monceaux	Roboticien et entrepreneur	05/04/2019
Denis Vidal	Anthropologue	09/04/2019
Sophie Sakka	Roboticienne universitaire	11/04/2019
Catherine Simon	Consultante en robotique	19/04/2019
Amélie Cordier	Roboticienne dans une entreprise privée	23/04/2019

II.1. Contre le droit des robots, la mobilisation de trois frontières

Les résultats de ces principaux échanges se résument en trois points renvoyant à trois frontières théoriques dans nos perceptions éthiques. La première frontière est constituée par la distinction ontologique. L'ensemble de nos entretiens ont vu s'activer le rappel de la différence entre les hommes et les machines. L'homme doit disposer d'une dignité qu'une machine ne peut revendiquer ; car la machine n'a pas les capacités de l'humain, mais aussi parce qu'elle est un outil, et parce qu'elle dispose d'une dangerosité spécifique – tant par sa puissance mécanique, que par l'empathie artificielle qu'elle peut susciter. Tous les intervenants ont ainsi défendu d'une manière ou d'une autre l'idée que cette frontière était indépassable, et même l'approche plus ouverte de Jérôme Monceau, qui voyait dans les robots des collectivités humaines en construction, refuse la remise en cause de cette segmentation des droits. Le danger pointé par la plupart des intervenants concerne le rabaissement de la dignité humaine en raison d'une valorisation du robot. Ce point vise directement les argumentations futuristes, juridiques et relationnelles, qui toutes pourraient laisser penser une équivalence à venir entre un être humain et une

entité artificielle. Il s'oppose directement aux approches post-humanistes qui ont cherché à remettre en cause les approches éthiques anthropocentrées[34]. La seconde frontière relève de la sphère politique et en particulier de la distinction entre le collectif et le particulier. Si les arguments en faveur du droit des robots ne discutent que secondairement de la notion de propriété et d'intérêt, celle-ci est apparue au centre de plusieurs entretiens. Les préoccupations politiques étaient liées aux récents événements au Parlement européen, qui discutait explicitement de la possibilité d'offrir aux robots une personnalité juridique. La députée Mady Delvaux avait remis en 2017 un rapport qui envisageait :

La création, à terme, d'une personnalité juridique spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques responsables de réparer tout dommage causé à un tiers ; il serait envisageable de considérer comme une personne électronique tout robot qui prend des décisions autonomes ou qui interagit de manière indépendante avec des tiers.[35]

Plusieurs personnes ont vu là l'influence des travaux de messieurs Bensoussan et une victoire de la stratégie d'influence des partisans de l'argument juridique. Nathalie Nevejans et Raja Chatila ont été directement impliqués dans la tournure politique que le débat a prise, en signant, avec d'autres roboticiens et juristes, une lettre qui s'opposait et s'inquiétait de la concrétisation de cette option juridique[36]. Cette prise de position publique se comprend en regard de trois questions politiques qui ont traversé nos entretiens. La première concerne le pouvoir dans sa dimension économique et médiatique : à qui l'idée du droit des robots peut-elle servir ? D'un point de vue sociologique, le droit des robots est un thème médiatiquement utile pour obtenir de la visibilité. Économiquement, l'idée peut offrir des perspectives de déresponsabilisation des entreprises productrices de robots. Politiquement enfin, il permet à certains acteurs de se faire reconnaître par les institutions comme experts. Ces éléments qui matérialisent la pratique argumentative sont finalement des critiques de la dimension abstraite de la revendication de droits pour les robots. Une deuxième question politique se pose à propos de la libre propriété et de la justification des régulations. Plusieurs entretiens ont souligné le problème normatif d'encadrer la relation à un bien comme si elle était une relation de personnes. Donner des droits à un robot revient logiquement à restreindre les libertés humaines, en posant de nouvelles contraintes dans le domaine privé. La libre jouissance de son bien est pourtant une valeur centrale des sociétés libérales contemporaines et sa remise en cause ne pourrait se justifier qu'en raison d'un gain collectif clair et d'une définition rigoureuse. Or la définition d'un robot étant loin de faire consensus, la limitation des libertés reviendrait à ouvrir un champ de revendications plus vastes ; pourquoi ainsi ne pas accorder des droits à des objets religieux, ou artistiques, ou identitaires ? Une dernière question politique concerne l'éducation. La plupart des intervenants ont pointé les nombreuses illusions qui entourent les robots. La science-fiction, la complexité de ces machines et un traitement médiatique parfois défaillant, diffusent un anthropomorphisme qui ne se confirme nullement dans la pratique technique. Face à cette tendance à l'anthropomorphisation ou à l'animalisation, le devoir des experts n'est pas tant de réguler que de diffuser le savoir pour saisir ce que peuvent et sont ces

machines. La troisième frontière que nos discussions ont révélé renvoie à l'espace et au temps. Le Japon et les pays asiatiques plus globalement sont apparus dans les entretiens comme un pôle de différence pour comprendre cet attachement occidental à la notion de dignité. Ainsi, plusieurs intervenants ont mis en avant l'importance accordée à la liberté et à la rationalité matérialiste en Occident par rapport à d'autres régions du monde, où l'animisme et les cultes religieux peuvent expliquer une approche différente de la robotique. Cette frontière culturelle est ainsi consciente et même théorisée par les anthropologues et juristes, mais elle n'implique pas pour autant un abandon des valeurs européennes ; penser à l'altérité peut même donner d'autant plus de valeur à la spécificité de nos pratiques. Cette frontière se module ensuite sur la relation entre le présent et l'avenir. La projection dans le futur fait apparaître de manière récurrente le transhumanisme et les risques technologiques. La transformation de la culture européenne dans sa dimension humaniste est ainsi largement dénoncée et crainte ; le robot est certes une machine extraordinaire, mais il porte en lui la peur d'un changement social où la valeur exceptionnelle accordée à l'homme est remise en cause.

Ces divers dialogues ont ainsi permis de constituer une réflexion critique non seulement sur le droit des robots, mais plus généralement sur les problèmes adjacents au développement des interactions entre humains et robots. L'humanisme qui a fait l'objet de plusieurs mises en question dans le champ philosophique semble conserver sa place chez les théoriciens et praticiens de la robotique contemporains.

II.2. Le rejet du droit des robots comme discussion de la modernité

Au cours de cette enquête française, il est ainsi apparu progressivement que la discussion du droit des robots dans diverses instances médiatiques se constituait de manière asymétrique. Le droit des robots en tant que revendication ne peut pas reposer sur un mouvement social institutionnalisé et ne dispose que de peu de partisans. Au contraire, nombre de spécialistes en position de pouvoir ne voient dans ce problème qu'un enjeu secondaire, ou une mécompréhension des enjeux actuels par les populations[37]. La question en suspens était donc de savoir si ces polémiques universitaires, juridiques et politiques pouvaient disposer d'une fonction autre qu'intellectuelle. Les entretiens avaient largement mis en avant le danger d'une instrumentalisation de la thématique par l'industrie de la robotique. Toutefois, ce soupçon paraît difficile à établir empiriquement. Si les productions de messieurs Bensoussan sont bien soutenues par divers cadres de l'entreprise *Aldebaran robotics*, ceux-ci restent très prudents sur la solution régulatrice proposée. Les quelques associations de promotion du droit des robots ont eu une existence aussi éphémère qu'une activité réduite ; elles étaient liées souvent à Alain Bensoussan. Les interventions des assureurs concernant l'intérêt d'une personnalité juridique pour les robots sont, elles aussi, peu engagées et tranchées[38]. Le droit des robots comme droit positif paraît donc pour le moment être un thème dont les groupes d'influence ne se sont pas encore emparés ou dont la manipulation a été un échec. En revanche, le traitement éthique des robots semble bien porter en lui des enjeux qui croisent les problèmes de collecte massive de données (*big data*) et d'influence, que ce soit via les questions d'économie de l'attention ou d'empathie artificielle[39]. Sous cet aspect, la robotique n'apparaît plus tant dans des formes

humanoïdes ou animalisées, que comme des objets connectés – enceintes, aspirateurs, réfrigérateurs, etc. – qui se trouvent semi-robotisés, au sens où ils disposent d'une agentivité minimale et d'une capacité organisationnelle. La question des relations avec ces nouveaux objets joue parfois sur des confusions anthropomorphiques qui empêchent de saisir les relations de pouvoir qu'ils recomposent.

Dans un précédent numéro d'*Implications philosophiques*, Olivier Sarre mettait en relation la question de la modernité et du droit des robots[40]. Sa conclusion amenait à penser que le droit des robots continuait la modernité tout en étant susceptible de s'inscrire dans des enjeux post-modernes de remise en cause de la science. Notre enquête montre que les critiques du droit des robots réactivent un ensemble de discours classiques de la modernité. L'importance accordée à la dignité humaine, à la définition scientifique des concepts, à la propriété privée, à l'éducation, à la distanciation entre un soi et une altérité, constituent en effet des thèmes récurrents de la période classique jusqu'aux Lumières. Cette relation entre le discours moderne et la lutte contre le droit des robots offre une autre explication de la place qu'a pu occuper ce discours dans l'espace public. De nombreuses rencontres et articles ont mis en scène une confrontation où les thèses classiques de la modernité s'exposent en pratique. Lors de nos entretiens, plusieurs intervenants ont concédé que le droit des robots permettait au moins de parler directement au public d'enjeux cruciaux de la robotique. Une explication du succès de ce thème dans l'ère contemporaine peut alors se concevoir sur le mode de la performance énonciative. Le thème offre une opportunité de réitérer un ensemble de discours de la modernité sur les distinctions ontologiques à la base de notre droit. Ce genre de performance se trouve autant dans la presse que dans des rencontres directes avec le public, sur des modes ludiques[41]. Les regroupements dans des rencontres organisées par les journaux, dans des conférences, et autres activités à la frontière de l'académie, donnent à voir de véritables mises en scène qui jouent sur la compatibilité entre modernité et robotique affective. Elles apparaissent comme des clarifications normatives concernant le bon comportement face aux robots. La volonté d'accorder un droit à des robots joue ainsi comme une hypothèse pédagogique devant être dépassée par le discours rationnel.

Cette conclusion rejoint paradoxalement les appels de Coeckelbergh et Gunkel à tisser des relations éthiques avec les robots. Cependant cette éthique, bien loin de s'accorder à vouloir mettre en doute la qualification du robot comme outil inerte, cherche à lutter contre toute forme de mystification ou d'affaiblissement de l'humanisme. Contre les appels à faire du robot un objet dont la valeur éthique serait particulière, les opposants aux droits des robots s'engagent dans des démonstrations de la transparence du robot comme marionnette ou outil, certes élaboré, mais pas fondamentalement distinct de la cafetière ou de l'aspirateur[42]. De ce point de vue, la manipulation du droit des robots dans l'espace public s'oppose directement aux propositions élaborées par Bruno Latour afin de dépasser la modernité[43]. Latour est d'ailleurs cité plusieurs fois par les défenseurs du droit des robots, ce qui tend à renforcer l'idée d'une querelle qui, plus qu'elle ne parle des robots, revient à un enjeu éthique de la perception des machines. Contre les invitations à percevoir le robot sur un mode anthropomorphique ou

sentimental, les conflits publics concernant le droit des robots invitent à mobiliser des passions pour la rationalité, via la clarté, la discussion publique et la planification des technologies. En reprenant l'approche de Latour, il semble donc que ces débats structurent en fait un conflit de représentation concernant le « non-humain ». Le droit des robots apparaît alors comme un moyen de renouveler « le travail de purification » moderne^[44] afin de séparer clairement la science et la politique, la nature et la culture, les êtres humains et les « non-humains ». Cependant, ce travail est clairement assumé sous le vocable de « défense de l'humanisme » et dispose pour lui de sérieuses propositions éthiques favorables à un contrôle de la technologie, en vue de la liberté humaine.

L'analyse de ce débat ouvre par conséquent des réflexions sur les enjeux de la concrétisation de l'éthique de la robotique. Cette concrétisation d'un débat théorique, dans un champ économique et social où les rapports de force et les projections stratégiques concernant l'avenir sont prégnants, doit permettre de questionner les phénomènes technologiques, non seulement dans leurs apparitions locales, mais aussi dans la manière dont ils indiquent un mode de perception des objets. La représentation du droit des robots appartient ainsi à un cadrage social du robot dans des conflits philosophiques sur l'avenir social et économique. Bien loin d'être dépassés, les discours de la modernité continuent de répondre et de mobiliser une éthique face à des enjeux lourds de conséquences pour l'établissement des responsabilités juridiques, la dignité humaine, ou encore l'éducation populaire.

[1] Ce travail est la synthèse d'une partie d'un mémoire de recherche produit dans le cadre d'un master 2 en théorie politique pour l'École doctorale de Sciences Po Paris. Je remercie Philippe Portier et Frédéric Gros pour leurs conseils concernant sa production ; ainsi que les critiques qui m'ont été adressées par Annabelle Levers et Astrid Von Busekist. Toutes les erreurs qui pourraient être contenues ici sont toutefois de mon unique responsabilité.

[2] La décision est notamment rapportée dans BISMUTH Yves, *Petit guide juridique pratique de la robotique*, Paris, L'Harmattan, 2018, p. 35-40.

[3] JOSSET Christophe, « *Le Robot Sophia, citoyenne saoudienne a plus de droits que les femmes de son pays* », *L'Express*, 28/10/2017, en ligne : https://www.lexpress.fr/actualite/monde/proche-moyen-orient/le-robot-sophia-citoyenne-saoudienne-a-plus-de-droits-que-les-femmes-de-son-pays_1955877.html.

[4] CAZALA Julien, « *Le Soft Law international entre inspiration et aspiration* », *Revue interdisciplinaire d'études juridiques*, vol. 66, no 1, 2011, p. 41-84.

[5] FORTUNATI L., ESPOSITO A., LUGANO G., "Introduction to the special issue "beyond industrial robotics: social robots entering public and domestic spheres.", *The Information Society*, 31(3), 2015, p. 229-236.

[6] Sur ces enjeux, voir ASARO Peter et WALLACH Wendell, *Machine Ethics and Robot Ethics*, New York, Routledge, 2017.

[7] ASARO Peter M., "What should we want from a robot ethic", *International Review of Information Ethics*, vol. 6, no 12, 2006, p. 12.

[8] WARWICK Kevin, "Robots with biological brains", in LIN Patrick, ABNEY Keith, and BEKEY George A., *Robot Ethics*, Cambridge, MIT Press, 2012, p. 329: « Le robot aura plus de cellules cérébrales qu'un chat, chien ou un chimpanzé, et possiblement plus que beaucoup d'humains. (...) Nous devons considérer quels droits un tel robot devrait avoir. Avons-nous besoin d'aller aussi loin que l'attribution d'une forme de citoyenneté ? » (Nous traduisons.)

[9] VERSENYI Laszlo, "Can Robots be Moral ?", *Ethics*, vol. 84, no 3, 1974, p. 253 : « En supposant que certains futurs robots soient sophistiqués au point que des générations de robots en aient construites d'autres, et qu'ils connaissent mieux leurs propres besoins et capacités que nous, il serait logique de leur donner des pouvoirs de décision suffisant pour assurer leur fonctionnement. Et austade, quoiqu'improbable, où ces machines connaîtraient plus à propos des hommes que les hommes eux-mêmes, il serait alors logique de leur transférer le contrôle des hommes – pour le bien de l'accomplissement humain. »

[10] SCHWITZGEBEL Eric et GARZA Mara. "A defense of the rights of artificial intelligences", *Midwest studies in philosophy*, 2015, vol. 39, no 1, p. 98.

[11] Nous ne pourrions pas passer en revue toutes les figures qui peuvent être catégorisées dans ce courant. Nous renvoyons aux noms suivants : Ben Goertzel, Nick Bostrom, Eric Scheitzgebél, Mara Garza, Hutan Ashrafian, David Levy, Patrick Lin, Rob Sparrow, Kevin Warrick, Erica Neely ; David Calverley, Frank Wells Sudia, Amedeo Santosuosso, et Patrick Hubbard.

[12] BENSOUSSAN Alain et BENSOUSSAN Jérémy, *Droit des robots*, Bruxelles, Larcier, 2015, p. 2. On trouve le même raisonnement sur la non-prédictibilité p. 43 : « Si le comportement du robot ne peut être entièrement prédit, c'est qu'une indétermination, qu'il lui appartient de combler à vu le jour. Il se trouve placé, peu ou prou, en situation de liberté et ce, indifféremment du périmètre dans lequel cette liberté peut trouver à s'exercer. Même dans un périmètre limité, le robot, parce qu'il est libre, a une emprise sur le monde qu'il façonne nécessairement par son action ».

[13] *Ibid.*, p. 51-54.

[14] *Ibid.* p. 51 : Article 2 : « le robot dispose d'un nom » ; article 3 : « un robot a le droit au respect de sa dignité limitée aux données à caractère personnel conservées ».

[15] Article 2 : « le robot (...) dispose d'un capital et d'un représentant légal » ; article 5 « Chaque opération effectuée par un robot doit être conservée pendant une période d'un mois à compter de sa réalisation afin de pouvoir reconstituer chacune des opérations en

cas de sinistre » ; article 8 : « Les robots doivent être assurés pour les conséquences dommageables dont ils pourraient être à l'origine », p. 51-53.

[16] SOLUM Lawrence B., "Legal Personhood for Artificial Intelligences", *North California Law Review*, vol. 70, no 4, 1992, p. 1232-1287.

[17] MULLIGAN Christina, "Revenge Against Robots", *SCL Review*, 2017, vol. 69, p. 595.

[18] WHITBY, Blay, "Sometimes it's hard to be a robot: A call for action on the ethics of abusing artificial agents", *Interacting with Computers*, vol. 20, no 3, 2008, p. 326-333.

[19] *Ibid.*, p. 330.

[20] *Ibid.*, p. 331.

[21] DARLING, Kate, "Who's Johnny? Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy", *Robot Ethics*, vol. 2., 2015.

[22] TURKLE Sherry, "In Good Company?: On the Threshold of Robotic Companions", in WILKS Yorick (dir.), *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, New York, John Benjamins Publishing, 2010.

[23] DARLING Kate, "Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects", in CALO Ryan, FROOMKIN Michael, et KERR Ian (éd.), *Robot Law*, Norphampton, Edward Elgar, 2016, p. 223.

[24] On trouve ce point dans deux textes fameux : KANT Emmanuel, *Leçons d'éthique [1775-1780]*, tr.fr. Luc Langlois, Paris, Le Livre de poche, 1997, p. 391-393 ; KANT Emmanuel, *Métaphysique des mœurs, deuxième partie, Doctrine de la vertu [1797]*, tr.fr. Alexis Philonenko, Paris, Vrin, 1985, §17.

[25] COECKELBERGH, Mark, "Robot rights? Towards a social-relational justification of moral consideration", *Ethics and Information Technology*, 2010, vol. 12, no 3, pp. 209-221.

[26] COCKELBERGH Mark, *Growing Moral Relations: Critique of Moral Status Ascription*, New York, Palgrave Macmillan, 2012 ; p. 43-46 pour l'argument phénoménologique de l'être comme relation ; p. 95-115 pour l'ancrage de l'être dans un monde qui lui permet de construire sa signification.

[27] GUNKEL David J., *Robot rights*, Cambridge, MIT Press, 2018, p. 2-12.

[28] *Ibid.*, p. 175 : « Pour Darling, la raison pour laquelle nous devrions bien traiter les robots s'explique parce que nous percevons quelque chose de nous-mêmes en eux. »

[29] *Ibid.*, p. 177-180.

[30] *Ibid.*, p. 180-186. Cette approche s'inscrit explicitement dans la lignée de Derrida et Levinas.

[31] GANASCIA Jean-Gabriel, *Le Mythe de la Singularité. Faut-il craindre l'intelligence artificielle ?*, Paris, Seuil, 2017, p. 25-33.

[32] Voir par exemple la critique établie par LACROIX Marieve, « Portrait-Robot de la Responsabilité du Robot », *Canadian Bar Review*, vol. 99, no 1, 2021, p. 1-27.

[33] Sur les questions méthodologiques, je renvoie aux approches de la recherche-action élaborée par HERON John, *Co-operative inquiry: Research into the human condition*, London, Sage, 1996 ; et en France aux questions liées à l'entretien compréhensif, KAUFMANN Jean-Claude, *L'entretien compréhensif*, Paris, Armand Colin, 2011. Dans ma perspective, ces dialogues devaient permettre de construire un discours philosophique collectif dont l'enjeu est une progression de la réflexivité.

[34] L'idée d'un post-humanisme est liée aux textes fameux de Foucault et Derrida sur la question. Par exemple : FOUCAULT Michel, *Les mots et les choses*, Paris, Gallimard, 1966, p. 398 : « L'homme est une invention dont l'archéologie de notre pensée montre aisément la date récente. Et peut-être la fin prochaine. » ; DERRIDA Jacques, *Marges de la philosophie*, Paris, Minuit, 1972, p. 140-141 : « Après la vague humaniste et anthropologiste qui a recouvert la philosophie française, on aurait donc pu croire que le reflux anti-humaniste et anti-anthropologiste qui suivit, et dans lequel nous sommes, allait redécouvrir l'héritage des pensées qui avaient ainsi été défigurées ou plutôt dans lesquelles on avait trop vite reconnu la figure de l'homme. ». Certains discours portés par Latour peuvent permettre de l'inscrire dans ce mouvement. Voir par exemple ce texte : HACHE, Émilie et LATOUR Bruno, « Morale ou moralisme ? Un exercice de sensibilisation », *Raisons politiques*, vol. 34, no 2, 2009, p. 143-165.

[35] Parlement Européen, *Résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL))*, par Mady Delvaux, 2017.

[36] Disponible à l'adresse suivante : <http://www.robotics-openletter.eu/> (consulté le 26 avril 2019).

[37] Le Parlement arrive à la même conclusion via l'OPECST. Voir Office parlementaire d'évaluation des choix scientifiques et technologiques, *Les robots et la loi, Compte rendu de l'audition publique du 10 décembre 2015 et de la présentation des conclusions du 3 mars 2016*, M. Jean-Yves LE DÉAUT, député, et M. Bruno SIDO, sénateur, Paris, 2016, disponible en ligne : <http://www.senat.fr/notice-rapport/2015/r15-570-notice.html> [consulté le 20/02/2022], p. 71-72 : « La majorité des intervenants s'accorde sur le fait que le robot n'est pas sujet de droit et n'a pas à avoir de personnalité juridique. Néanmoins, il faut pouvoir régler le problème de la responsabilité entre l'acquéreur, le revendeur, l'intégrateur, le fabricant, etc. L'un d'entre vous a proposé une charte constitutionnelle de la robotique. En tout cas, il faut une charte et des règles générales, même si celles-ci ne doivent pas forcément être inscrites dans la Constitution. »

[38] Voir ce même rapport du Sénat, *ibid.*, p. 60-63.

[39] TISSERON Serge, *Le jour où mon robot m'aimera, Vers l'empathie artificielle*, Paris, Albin Michel, 2015 ; DAMANIO Luisa et DUMOUCHEL Paul, *Vivre avec les robots : essai sur l'empathie artificielle*, Paris, Seuil, 2016.

[40] SARRE Olivier, « Intelligences et organismes artificiels », Implications Philosophiques. Consultable en ligne : <https://www.implications-philosophiques.org/recherches/le-droit-des-robots/des-intelligences-et-organismes-artificiels/>.

[41] On peut penser, par exemple, à la rencontre du 7 octobre 2018 organisée par le journal *Le Monde* au Studio de l'Opéra Bastille. Ou encore ce faux procès qui a été organisé dans le cadre de l'initiative *TechnoLex*, le 28 novembre 2018. Voir : https://www.lemonde.fr/festival/article/2018/07/02/un-amour-de-robot_5324650_4415198.html [consulté le 20/02/2022] ; <https://www.archimag.com/univers-data/2018/12/10/assiste-proces-fictif-deux-robots-policiers-voici-faut-retenir> [consulté le 20/02/2022].

[42] C'est la comparaison qui a souvent été mise en avant par Sophie Sakka lors de notre entretien : « Remplacez le mot robot quand vous l'utilisez par aspirateur, cafetière, machine à laver, si ça a du sens, c'est que vous l'avez bien utilisé. » Elle est aussi revenue dans d'autres entretiens pour ramener le robot à sa dimension d'objet.

[43] LATOUR Bruno, *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique*, Paris, La Découverte, 1997, p. 43-70. Dans cet ouvrage, Latour défend que la modernité repose sur une séparation entre les choses (objets naturels et artificielles), qualifiés de « non-humains », et les êtres humains. Il critique cette distinction comme une mystification qui n'a pas de réalité sociale effective.

[44] *Ibid.*, p. 47.