



L'hypothèse de la « personne robot »

Catégorie : IA & Neurosciences, Mundus Numericus

Tags : loi, robot, société, technique

Personnages : Isaac Asimov, Alain Bensoussan, Elon Musk, Nathalie Nevejans, Pierre-Alain Raphan

1^{er} mai 2021

L'artefact « intelligent » autonome peut-il et doit-il être considéré comme responsable de ses décisions ? Faut-il créer un nouveau sujet de droit ?

Le droit suit à grand-peine les technologies numériques, l'« intelligence artificielle » (IA) et la robotique, mais l'élaboration de cadres juridiques clairs reste un chantier impérieux puisque, dit succinctement, ces technologies règlent désormais notre sort. Sur ce chantier législatif, technique et « société » se rencontrent et nos systèmes de valeurs déploient leurs scénarios. Le droit est donc aussi ce poste d'observation avancé, non pas tant des techniques en elles-mêmes que de nos désirs et de nos croyances à leur sujet. En l'occurrence, il nous indique une nouvelle et puissante poussée de ce vieil atavisme : la conviction que l'humain peut (voire est appelé à) créer de l'individu de toutes pièces.

Responsabilité

Qu'il s'agisse de médecine, de transports, de ressources humaines, d'agriculture, de militaire... ou tout simplement de notre vie personnelle, l'IA conquiert progressivement le champ de la *décision*. Le corollaire de cette conquête, c'est évidemment la question de la responsabilité de ces artefacts auxquels nous devons en même temps concéder, hésitant entre appréhension et fascination, la possibilité de prendre de *mauvaises* décisions. Victimes de ces décisions, nous sommes en droit de demander réparation du préjudice subi au(x) responsable(s) et certains pensent alors inévitable de pouvoir, à terme, incriminer le robot lui-même. À côté de la personne physique et de la personne morale, il faudrait par conséquent créer un nouveau sujet de droit : la « *personne robot* ». Mais sur quoi fonder cette troisième personne ? Que révèle cette proposition de nos relations qui se préparent avec les robots ?

Nous proposons de commencer par un détour technique (simplifié) que les thuriféraires de la personne robot s'épargnent en général, et pour cause puisqu'il met à jour une objection simple et gênante : un robot *est* un objet technique et son « autonomie » *est* donc un caractère technique, pas une aptitude d'ordre moral. Il s'ensuit que la « décision autonome » doit être techniquement définie et contrôlée, au même titre que la fission nucléaire, la trajectoire d'une fusée ou la cuisson d'un gâteau... Les robots disposant d'énormément de « degrés de liberté » il faut seulement « travailler » beaucoup plus.

Ouvrons ce détour technique par un exemple.

2013 : Google vs. Lyonnaise de garantie

Lorsque l'on saisit une recherche sur Google, des suggestions de complément s'affichent automatiquement. Techniquement, l'algorithme fonctionne à partir de données statistiques collectées en permanence (recherches fréquentes...) et, bien entendu, si Google peut garantir la qualité *statistique* des résultats, la firme ne peut pas prédire le résultat d'une requête *particulière*. Cette technique statistique est dans son principe analogue aux techniques d'« apprentissage » de l'IA qui mènent à des « décisions » statistiquement valables, comme un coup correct au jeu d'échecs ou de go, mais qui peuvent, dans certains cas particuliers, échouer et « porter préjudice ». Ainsi parfois, les suggestions de Google ont le malheur de ne pas plaire à certains...

La Lyonnaise de garantie, une société de courtage de produits d'assurance, a ainsi assigné Google dont l'algorithme proposait « *Lyonnaise de garantie escroc* » au troisième rang des suggestions proposées. La responsabilité de Google fut engagée en tant que fabricant, mais Google eut finalement gain de cause en cassation par un arrêt rendu en juin 2013¹ et qui stipulait que « *Lyonnaise de garantie escroc* » correspond à « *l'énonciation d'une pensée rendue possible uniquement par la mise en œuvre de la fonctionnalité en cause* ». Ceci signifie que cette « *pensée* » – phrase en attente d'interprétation humaine – s'est trouvée enregistrée dans les tables statistiques de Google mais n'a pu être « *énoncée* » – restituée à l'interprétant – que par le *seul* moyen de l'algorithme de suggestions. Autrement dit, pour Google cette « *pensée* » n'existait pas en tant que telle. Par conséquent (nous soulignons) :

[...] la fonctionnalité aboutissant au rapprochement critiqué est le fruit d'un processus purement automatique dans son fonctionnement et aléatoire dans ses résultats, de sorte que l'affichage des « mots clés » qui en résulte est exclusif de toute volonté de l'exploitant du moteur de recherche d'émettre les propos en cause ou de leur conférer une signification autonome au-delà de leur simple juxtaposition et de leur seule fonction d'aide à la recherche [...]

Ainsi, le lien de responsabilité entre Google et cette « *pensée* » est rompu par une statistique sans statisticien et qui préserve le secret de ses agrégats jusqu'à leur énonciation. Il n'en reste pas moins que le préjudice pour la Lyonnaise de garantie était bien réel : voici donc un préjudice, conséquence d'une « *décision numérique* », sans responsable ! Il en irait aujourd'hui probablement de même pour une décision prise par un artefact « *intelligent* » puisque, là encore, une telle décision ne serait que statistiquement fondée, sans que le fabricant de cet artefact ait pu volontairement participer à son élaboration. En voici un exemple dramatique.

2016 : accident mortel en Floride

Le système Autopilot de la Tesla Model S était activé lorsqu'un poids lourd s'est déporté devant le véhicule. Cette manœuvre n'a pas été détectée par le pilote automatique et le conducteur est décédé dans la collision. La firme Tesla s'est évidemment empressée de rappeler que l'Autopilot n'est qu'un outil d'*assistance* et que la responsabilité incombe donc toujours au seul conducteur. En l'occurrence le conducteur ne semble pas avoir, lui non plus, vu le poids lourd (blanc, à contre-jour d'un ciel éblouissant) ni entamé la moindre manœuvre de reprise en mains du véhicule. Tesla a aussi rappelé qu'il s'agissait du premier

¹ Cour de cassation – Juin 2013 – [Arrêt n° 625 du 19 juin 2013 \(12-17.591\)](#)

décès en 130 millions de kilomètres parcourus au total avec l'Autopilot activé, contre une moyenne de 94 millions de kilomètres parcourus entre deux accidents mortels aux États-Unis (sous-entendu : Tesla + Autopilot est plus sûr en moyenne que Véhicule + Conducteur). Comme l'algorithme de suggestions de Google ou n'importe quelle autre IA, l'Autopilot est *statistiquement* compétent mais peut dramatiquement échouer dans certains cas particuliers.

« Tourner », « freiner », « maintenir la trajectoire » ... sont d'une certaine façon autant de « *pensées* » dont l'« *énonciation* » conduit à des effets de décision qui, comme dans le cas de Google, échappent à la « *volonté* » du fabricant. La gravité du préjudice exigeait toutefois d'établir précisément les responsabilités. Sans un mode d'enquête spécifique aux robots, les enquêteurs ont donc étudié le fonctionnement de l'Autopilot *en tant que composant technique* de l'automobile. Mais l'IA statistique n'est pas un frein ou une colonne de direction et se prête mal à une investigation « matérielle », logicielle en l'occurrence. On peut même noter avec un peu d'ironie que cette opacité protège le fabricant car il peut difficilement lui être reproché une erreur de *design* et ou de calcul. Ainsi, comme le déclare l'un des enquêteurs, « *dans cet accident, le système de Tesla a fonctionné comme prévu, mais il a été conçu pour effectuer des tâches limitées dans une gamme limitée d'environnements* »². La responsabilité sera donc cherchée du côté du *contexte d'usage*, impliquant en principe, soit le fabricant pour l'avoir mal explicité, soit le conducteur pour l'avoir enfreint. En l'occurrence, le conducteur a laissé l'Autopilot aux commandes pendant 37 minutes d'un trajet qui a duré au total 41 minutes en ignorant les instructions visuelles et sonores de la Tesla Model S lui demandant de la reprendre régulièrement en main.

Calcul, attachement

Les robots actuels ne prennent pas d'authentiques décisions au sens où ils exerceraient une sorte de délibération intime suivie d'un libre-arbitrage : ils n'effectuent que des *calculs*. Il incombe alors au fabricant la précision mathématique de ces calculs ainsi que l'explication de leur contexte d'utilisation, et il incombe entièrement à l'utilisateur leur bon usage. En l'occurrence, ne pas avoir repris en main le véhicule revenait à demander à l'Autopilot de poursuivre un « mauvais calcul » (toutes proportions gardées, le fabricant de mètres à ruban n'est pas non plus responsable des conséquences d'une mesure mal réalisée).

Si le robot reste ainsi un objet technique et donc toujours un *objet de droit*, il présente la particularité de pouvoir, plus que tout autre, nous « leurrer ». Le fabricant exagère l'« intelligence » de son produit et l'équipe de « simagrées » qui suscitent en retour notre attachement ([L'attachement aux simulacres](#)). Cet

² BBC News – 12 septembre 2017 – [Tesla Autopilot 'partly to blame' for crash](#)

attachement peut conduire au relâchement de notre attention, à un « mésusage du calcul » causant préjudice, engageant ainsi notre responsabilité. Il revient par conséquent aux autorités compétentes de modérer l' « exagération » du fabricant et de prévenir les usagers. Ce n'est pas simple. Pour revenir à l'exemple de Tesla, le terme « *Autopilot* » veut bien dire « pilote automatique », non pas « assistant de conduite », et le KBA (autorité fédérale allemande en matière de transports motorisés) a donc bien tenté de le faire interdire mais en vain.

Autonomie technique

L' « autonomie » intègre progressivement le domaine des normes et des degrés car il faut pouvoir la *mesurer* afin d'ajuster le cadre réglementaire ou assurantiel. Ainsi, l'ISO publie un ensemble de normes concernant les robots tels que les robots collaboratifs (ISO 15066) ou les robots de soin personnel (ISO 13482). En 2014, la SAE américaine (Society of Automotive Engineers) a mis au point une classification (J3016) qui définit six niveaux d'autonomie (0 à 5) pour les systèmes de conduite de véhicules³ :



Le véhicule Tesla incriminé en Floride atteint le niveau 2 pour lequel il est indiqué : C'est *vous* qui conduisez (« *You are driving* »), même si vous ne touchez pas les pédales ou le volant et même si l'assistant de conduite s'appelle « *Autopilot* ». Aux niveaux 3 à 5 le véhicule est techniquement autonome puisque vous ne conduisez pas (« *You are not driving* ») même si vous êtes installé dans le « siège du conducteur ». En 2019, Audi commercialise le premier véhicule de niveau 3 au monde (une Audi A8). A ce niveau, les conditions de fonctionnement de l'automate sont assez restreintes (Audi appelle clairement son système « *Traffic Jam Pilot* ») et le conducteur doit pouvoir reprendre le contrôle chaque fois que le système le demande. L'automate est cependant verrouillé selon la législation en vigueur dans chaque pays, c'est-à-dire partout sauf... au Japon, qui a amendé son Road Traffic Act pour autoriser les véhicules de niveau 3 sur la voie publique à partir de mai 2020. Les Jeux Olympiques prévus à l'été 2020 se présentaient

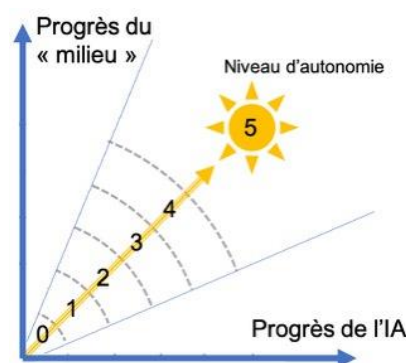
³ Sur le site de la [SAE](https://www.sae.org)

évidemment comme occasion unique pour ce pays pionnier de la robotique⁴. Honda est devenu en 2020 le premier constructeur habilité au niveau 3 par ce pays⁵ qui franchissait ainsi le Rubicon « *You are not driving* ». Au Japon désormais, le conducteur n'est plus nécessairement humain.

Une échelle d'autonomie *technique* permet ainsi de contrôler simultanément le progrès des machines autonomes et celui de l'environnement, y compris réglementaire, dans lequel ces machines évoluent. Nous apercevons déjà qu'une personne robot hautement autonome n'a aucune chance d'être lâchée dans l'environnement actuel, tant physique que légal et politique.

Milieu technique

Les niveaux de classification de la SAE commencent à inspirer également la robotique médicale⁶. Au niveau 5 le robot ne serait plus alors un appareil médical, même très sophistiqué, mais un authentique praticien. Mais qu'il s'agisse de médecine ou de tout autre domaine (police, aide à domicile, armée, conduite de véhicules...), ce niveau 5 est encore une utopie et nous ne risquons pas de croiser demain des robots totalement autonomes dans un environnement totalement ouvert. En même temps, il faut observer cette curieuse propension du système technicien à engendrer des objets de plus en plus autonomes (mais qui n'ont par ailleurs aucun caractère de nécessité). Dans notre monde technique, l'inatteignable niveau 5 se présente ainsi comme une puissante cause finale qui nous pousse en même temps à *techniciser notre milieu* pour permettre à l'autonomie technique des objets d'apparaître plus authentique, de se montrer plus efficace (voir l'exemple de l'agriculture : [L'AgTech ou les champs numériques](#)). La « zone d'autonomie » et ses degrés se déploie alors comme ceci :



⁴ Herbert Smith Freehills – 28 janvier 2019 – [Japan advances driverless car ambitions with draft bill to amend Road Traffic Act](#)

⁵ AFP / msn.com – 11 novembre 2020 – [Honda wins world-first approval for Level 3 autonomous car](#)

⁶ Guang-Zhong Yang, James Cambias, Kevin Cleary, Eric Daimler, James Drake, Pierre E. Dupont, Nobuhiko Hata, Peter Kazanzides, Sylvain Martel, Rajni V. Patel, Veronica J. Santos, Russell H. Taylor / Science Robotics – 2017 – [Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy](#)

Le progrès du « milieu » doit être compris au sens large comme l'extension du domaine de « décision » des objets autonomes. Il peut donc s'agir aussi bien d'un progrès technique (smart city, route intelligente, appartement connecté...), que physique (zones de fonctionnement dédiées...), réglementaire ou assurantiel...

L'autonomie technique se développe ainsi le long du *couplage* du robot / milieu. Le [milieu digitalisé](#) permet par exemple la collecte de suffisamment de précieuses données *réelles* pour améliorer l'apprentissage des robots qui circulent et donc l'élargissement progressif de leur périmètre d'autonomie. Le « *Technoking* » Elon Musk⁷ amasse ainsi en temps réel les données comportementales de 1,4 millions d'utilisateurs⁸, autant d'informations permettant d'améliorer à distance les performances de l'Autopilot. En même temps, son milieu de fonctionnement, l'infrastructure routière, doit progresser de façon à enserrer le véhicule dans un cocon technique visualisable comme une « route numérique » où le risque est réduit à la seule faille technique d'un équipement et non pas à un biais statistique juridiquement problématique. Rien n'indique encore que sur le chemin de ce couplage progressif une personne robot devienne un jour juridiquement nécessaire.

L'effet collatéral de ce progrès dirigé vers la « cause finale », c'est le risque d'inadéquation progressive de l'humain à ce milieu en cours de technicisation et que nous partageons avec les machines autonomes : l'humain se trouve lui-même engagé dans un nouveau champ de responsabilité. Il faut donc en même temps, disons-le rapidement, adapter le droit positif « classique ».

Le Parlement européen hésite

La Résolution du Parlement européen du 16 février 2017 contient des recommandations concernant des règles de droit civil en matière de robotique⁹. Cette résolution commence par quelques dizaines de considérations dressant le panorama d'un monde à venir profondément transformé par la robotique et où, pour ne retenir que le thème qui nous intéresse ici, la responsabilité juridique devrait être repensée. En effet (nous soulignons) :

AF. considérant que, dans l'hypothèse ou un robot puisse prendre des décisions de manière autonome, les règles habituelles ne suffiraient pas à établir la responsabilité juridique pour dommages causés par un robot,

⁷ Chris Isidore / CNN Business – 15 mars 2021 – [Elon Musk is now 'Technoking' of Tesla. Seriously](#)

⁸ Chiffres de livraison jusqu'à fin 2020 indiqués par Wikipédia – [Tesla \(automobile\)](#)

⁹ Parlement européen – 16 février 2017 – [Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique](#)

puisque'elles ne permettraient pas de déterminer quelle est la partie responsable pour le versement des dommages et intérêts ni d'exiger de cette partie qu'elle répare les dégâts causés ;

[...]

AI. considérant que [...] le cadre juridique actuellement en vigueur ne suffirait pas à couvrir les dommages causés par la nouvelle génération de robots, puisque celle-ci peut être équipée de capacités d'adaptation et d'apprentissage qui entraînent une certaine part d'imprévisibilité dans leur comportement, étant donné que ces robots tireraient, de manière autonome, des enseignements de leurs expériences, variables de l'un à l'autre, et interagiraient avec leur environnement de manière unique et imprévisible ;

Ces considérations préparent sans ambiguïté le terrain pour la personne robot sur la base d'une appréciation assez ambiguë de l'autonomie. Il est précisé un peu plus haut dans le texte (AA) que « *l'autonomie d'un robot peut être définie comme la capacité à prendre des décisions et à les mettre en pratique dans le monde extérieur, indépendamment de tout contrôle ou influence extérieurs ; que cette autonomie est de nature purement technique et que le degré d'autonomie dépend du degré de complexité des interactions avec l'environnement prévu par le programme du robot* ». L'autonomie est de nature *technique*, nous sommes d'accord et, comme relevé plus haut, les définitions et degrés de l'autonomie technique se mettent progressivement en place, ouvrant la voie pour un ajustement corrélatif du « milieu », en particulier juridique. Mais le reste du propos est équivoque. Cette « *part d'imprévisibilité dans le comportement* » et cette « *capacité à prendre des décisions sans contrôle extérieur* » évoquent irrésistiblement la posture cognitiviste dominante consistant à enfermer la « pensée » dans une boîte, le « monde extérieur » dans une autre, et à considérer que ces deux boîtes posent des problèmes... d'emboîtement. Cette vieille lubie du cerveau (même artificiel), fermé dans sa mystérieuse rumination et dont jaillissent parfois des « décisions », continue de semer des troubles du langage. Rappelons que le robot n'est pas un être mystérieux que nous cherchons à comprendre mais un *objet technique* des pieds à la tête. Rappelons que le robot, à l'instar de la voiture Tesla, sera bien entendu *toujours* contrôlé de l'extérieur, techniquement depuis le « cloud », par des autorités publiques ou privées. Rappelons aussi que rien ne nous oblige (si ce n'est cette « téléologie du niveau 5 »...) à mettre des machines potentiellement dangereuses en circulation et il paraît déraisonnable d'inventer un nouveau *sujet de droit* justement pour le permettre.

Pour en finir avec les considérations parlementaires, revenons sur l'expression « *prendre des décisions de manière autonome* ». L'algorithme de Google,

quoique non robotique au sens strict, « prend des décisions de manière autonome » comme n'importe quel algorithme alimenté par des données contextuelles. C'est d'ailleurs cette forme d'autonomie, au sens ici de pouvoir « *énoncer une pensée* » qui n'a aucune existence avant l'acte d'énoncer (c'est-à-dire le calcul d'énonciation), qui a dégagé Google de sa responsabilité de fabricant. Cet exemple montre que, si l'on veut continuer à utiliser ces technologies, le principe de responsabilité doit être repensé au fur et à mesure que ces cas apparaissent mais sans jamais céder à la tentation de la personne robot. Les textes réglementaires doivent se préciser, les normes apparaître, les assurances s'adapter et les fabricants poser des garde-fous sans qu'il soit nécessaire de s'abandonner à l'idée d'une quelconque capacité des objets à prendre des décisions de manière autonome, au sens humain habituel du terme. Ces machines, comme toutes celles que nous avons créées depuis la nuit des temps doivent être seulement intégrées techniquement, socialement, politiquement...

L'ambiguïté fondatrice du domaine de l'IA conduit une fois de plus aux portes du fantasme (« *P. considérant qu'il est possible, en fin de compte, qu'à long terme, l'intelligence artificielle surpasse les capacités intellectuelles de l'être humain* »). L'un des grands défis qui nous attend sera probablement de tenir fermement la frontière ontologique qui sépare encore l'humain de l'objet technique. Or, les spéculations et les « avancées » du droit se présentent comme autant d'indices d'une certaine mollesse. Le Parlement européen reste encore prudent en matière de responsabilité entourant l'activité robotique et propose principalement des instruments de bon sens (régime d'assurance obligatoire, fonds de compensation, numéro d'immatriculation individuel...). Mais ses prémisses l'amènent inévitablement à demander à la Commission d'examiner en dernier lieu :

(59. f) 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, tenues de réparer tout dommage causé à un tiers ;

Nathalie Nevejans, maîtresse de conférences en droit privé à l'université d'Artois, spécialiste du droit et de l'éthique de l'intelligence artificielle, membre du comité d'éthique du CNRS (COMETS), développe une critique claire de cette nouvelle personne juridique. Elle souligne à juste titre que « *le sentiment s'installe très vite à [la lecture de la Résolution du Parlement européen] que les évolutions souhaitées ne sont pas simplement motivées par des arguments juridiques, mais préparent implicitement un bouleversement ontologique de la place de l'humain*

dans un monde technologique »¹⁰. Le droit joue ici le rôle du canari dans la mine de charbon et ses soubresauts annoncent bien quelque chose.

La nature robotique

Grâce à [la cybernétique] les figures austères de l'enregistrement, de la compensation comptable, de la sommation *a posteriori* retrouvent la fraîcheur de ce qui naît, de ce qui « s'auto-organise » avec toute la vigueur et l'innocence d'une faune ou d'une flore [...] mais au prix d'une dégradation de la politique en théorie de compétition d'agréats [...]

Gilles Châtelet¹¹

Le robot n'est pourtant pas un objet tout à fait comme les autres. C'est le seul « qui fait tout » pour dissimuler son état naturel d'objet technique. Il pose de ce point de vue des questions spécifiques en matière d'intégration technique, sociale et politique. Nous avons rappelé par exemple plus haut que le « jeu » entre le fabricant et l'utilisateur peut conduire à un « préjudice par attachement » qui appelle au minimum une intervention réglementaire (d'ailleurs toutes les industries d'attachement « dur », comme l'industrie cigarettière, sont des mines d'enseignement pour l'industrie robotique). Mais au-delà de cette particularité, le robot numérique présente des caractéristiques techniques intrinsèques qui sont, pour le coup, *dans sa nature*. Rappelons les trois principales.

L'Autopilot de Tesla ne *fonctionne* pas différemment de l'algorithme de Google. Si le premier est doté d'une enveloppe physique ce sont tous deux des produits de l'IT (Information Technology). Or, l'IT est caractérisée par son **opacité**. Cette technologie produit des effets de surface que chacun peut observer, désirer, redouter... sans que nous en comprenions les causes ni même la survenance. Elle *est* opaque comme le ciel *est* bleu et le fabricant utilisant l'IT comme « matière première » produit inévitablement des objets opaques. Il faut nous *défendre* de cette opacité par tous les moyens à notre disposition, qu'ils soient techniques, juridiques ou éthiques ([Tristan Harris et le marais de l'éthique numérique](#)).

Deuxièmement, l'algorithme de l'Autopilot se distingue de celui de Google par sa **complexité**. Cette complexité a des *niveaux*, mesurables par exemple en nombre de lignes de code, de modules-fonctions, de nombre et de types d'entrées et sorties algorithmiques possibles, de mémoire requise, de temps de calcul... Le fabricant qui incorpore une grande complexité dans ses machines s'expose à

¹⁰ Nathalie Nevejans / La Jaune et la Rouge N° 750 – Décembre 2019 – [Le statut juridique du robot doit-il évoluer ?](#)

¹¹ Gilles Châtelet / Éditions rue d'Ulm – 2010 – *L'enchantement du virtuel* (« Mettre la main à quelle pâte ? »)

devoir mettre en place des protocoles de preuve algorithmique, de vérification et de certification elles-mêmes complexes. Il devient également délicat de reprendre le contrôle d'une machine complexe qui manifestement déraile. Rappelons-nous les terribles accidents du Boeing 737 MAX¹². Ce nouvel avion qui a succédé au 737 NG, réputé pour sa grande stabilité en vol et sa fiabilité, a été équipé de moteurs plus gros qui ont modifié très sensiblement les paramètres de stabilité de l'appareil. Plutôt que de redessiner physiquement l'avion, Boeing a misé sur une option moins coûteuse et plus rapide¹³ : utiliser un *software* de stabilisation automatique. Sans être des spécialistes de l'aéronautique, on peut comprendre que l'avion est ainsi devenu plus dangereux *by design* parce que ce *software* complexe est devenu un point de défaillance supplémentaire : « *ajouter plus de software ajoute seulement de la complexité à l'appareil* ». Le robot « hyper-softwarisé » que la complexité rend imprévisible présente le même genre de difficulté et il faudra encore de nombreuses générations techniques pour « naturaliser » son comportement, c'est-à-dire pour faire en sorte qu'il devienne « stable » au sens large sans ajouts complexes. La recherche de simplicité, ou de naturalité, devrait d'ailleurs être une quête technique passionnante aussi bien qu'une exigence éthique.

Enfin, puisqu'il s'agit d'IA, la mystérieuse et imprévisible « *fraîcheur* » du robot naît, dit joliment le mathématicien-philosophe Gilles Châtelet, des « *figures austères* » du calcul. L'extraordinaire accumulation de données, au coût écologique d'ailleurs exorbitant ([Données et traces numériques sous rature](#)), produit tout au plus la possibilité de capturer des agrégats statistiques de plus en plus fins qui sont « *a posteriori* » exposés comme des choses « réelles » ou des décisions « intelligentes ». Le robot se situe en quelque sorte à un degré avancé de ce zoom statistique. Il ressemble à une collection d'agrégats statistiques hyper-fins ramassée dans une enveloppe physique, un inventaire étendu de gestes moyens. L'enjeu apparaît donc bien technique mais il nous semble aussi éthique dans un sens encore difficile à caractériser : le milieu de fonctionnement visé, dont les données proviennent pour la fabrication des « agrégats comportementaux » du robot, est-il respecté et respectueux ? (Ce dernier point mérite un développement renvoyé à plus tard).

Le robot numérique *est* donc opaque, complexe et statistique. Plutôt que de s'abandonner à ces « existentiels » en les empaquetant dans une personne juridique, il faut par principe s'y « opposer ». Les armes doivent être réglementaires et techniques, bien entendu, et il y a là une inépuisable source d'innovations. Mais ces armes doivent aussi être éthiques. Nous opposerons ainsi

¹² Padma Nagappan / San Diego State University – 24 octobre 2019 – [What Caused Boeing's 737 Max Crashes](#)

¹³ Ian Snyder / Points with a crew – 12 mars 2019 – [Can Boeing fix a potentially faulty 737 MAX design with software?](#)

à l'opacité une éthique du dévoilement, à la complexité une éthique de la simplicité et à la statistique, plus tard, une éthique « environnementale ».

L'hypothèse de la personne robot

Peu importe qu'il y ait du sang dans vos veines ou du liquide hydraulique, chacun a le droit d'être libre.

W.H. Mitchell, *The Arks of Andromeda*

En France, l'avocat Alain Bensoussan, bien connu dans le secteur numérique, est le premier à avoir plaidé, dès 2013, pour l'avènement de cette personne robot à qui nous pourrions reconnaître des droits et des devoirs. L'aspect le plus singulier de sa proposition est son fondement sur un droit naturel dont voici la définition habituelle¹⁴ :

Le droit naturel est l'ensemble des droits que chaque individu possède du fait de son appartenance à l'humanité et non du fait de la société dans laquelle il vit. Le droit naturel comprend notamment le droit à la vie et à la santé, le droit à la liberté, comme le droit de propriété [...]

Le robot acquerrait ainsi un ensemble de droits du fait de son appartenance à une « *robohumanité dans laquelle hommes et robots devront apprendre à cohabiter* »¹⁵. Pour l'avocat, pas de doute : nous avons affaire à une « *nouvelle espèce* » qui conférerait au robot un « état de nature » duquel proviendrait un « *droit à la souveraineté* » et notamment à la « *liberté qui réside dans le fait de pouvoir décider « en son âme et conscience » ou plutôt « en ses algorithmes et conscience* », *en tous cas davantage qu'un simple automate* ». Maître Bensoussan participe de ce « bouleversement ontologique » dans un élan qui nous paraît un peu audacieux. Il passe tout simplement par-dessus la technologie et son intrinsèque dignité pour instaurer *directement* une espèce-robot fictive, mais ce que nous voulons faire remarquer ici n'est pas tant que ce discours que le fait qu'il puisse exister. Et il n'est pas le seul...

Asimov, le retour

Personne ne semble encore prêt à accueillir un troisième sujet de droit. Mais, comme nous l'avons vu, les instances politiques préparent le terrain et les comités d'éthique s'activent, le « canari » frémit. Ainsi, en janvier 2020, le député Pierre-Alain Raphan a rédigé une proposition de loi *constitutionnelle* relative à la Charte

¹⁴ Dictionnaire juridique de Serge Braudo – [Définition de Droit naturel](#)

¹⁵ Alain Bensoussan / Figaro Blog – 10 juillet 2018 – [Le droit naturel, fondement juridique de la personne-robot ?](#)

de l'intelligence artificielle et des algorithmes ¹⁶. Cette proposition assez déconcertante s'inspire des fameuses « trois lois de la robotique » proposées en 1942 par l'écrivain de science-fiction Isaac Asimov ¹⁷. Si l'Article 1^{er} dénie prudemment toute responsabilité juridique au robot ou à l'algorithme d'IA défini comme « *évolutif dans sa structure, apprenant, au regard de sa rédaction initiale* », l'article 2 pose qu'un tel système :

« – ne peut porter atteinte à un être ou un groupe d'êtres humains, ni, en restant passif, permettre qu'un être ou un groupe d'êtres humains soit exposé au danger ;

« – doit obéir aux ordres qui lui sont donnés par un être humain, sauf si de tels ordres entrent en conflit avec le point précédent ;

« – doit protéger son existence tant que cette protection n'entre pas en conflit avec les deux points précédents.

Il s'agit de fait de devoirs qui incombent aux personnes physiques et morales qui « *hébergent ou distribuent ledit système* ». Mais aucun *calcul* ne peut présenter de telles garanties, ce que le Parlement européen, également bien inspiré par la littérature d'anticipation, rappelle lui-même :

T. Considérant qu'il y a lieu de considérer les lois d'Asimov comme s'appliquant aux concepteurs, aux fabricants et aux opérateurs de robots, y compris de robots dotés d'autonomie et de capacités d'auto-apprentissage, étant donné que lesdites lois ne peuvent être traduites en langage de programmation ;

Par conséquent, ces « lois » (ou d'autres) devraient être considérées comme des règles classiques d'éthique s'appliquant à l'ensemble de la collectivité humaine participant de l'intégration d'outils, quels qu'ils soient (voiture, couteau, machine à coudre...), dans notre environnement, depuis la conception jusqu'à l'utilisation. La fameuse « autonomie » des robots n'est qu'une caractéristique technique de plus qui exige une définition technique claire.

¹⁶ Assemblée Nationale – 15 janvier 2020 – [Proposition de loi constitutionnelle relative à la Charte de l'intelligence artificielle et des algorithmes](#)

¹⁷ (1) Un robot ne peut attenter à la sécurité d'un être humain, ni, par inaction, permettre qu'un être humain soit mis en danger. (2) Un robot doit obéir aux ordres d'un être humain, sauf si ces ordres entrent en conflit avec la première loi. (3) Un robot doit protéger sa propre existence tant que cela n'entre pas en conflit avec la première ou la deuxième loi. (Isaac Asimov, « *Cercle vicieux* », 1942).

Rupture anthropologique

I am not afraid of robots. I am afraid of people, people, people. I want them to remain human.

Ray Bradbury

Nous pouvons admettre qu'en tant qu'objet technique le robot puisse appartenir à une « espèce », comme le suggère Alain Bensoussan, mais seulement au sens d'un mode d'individuation technique particulier qui pourrait appeler, pourquoi pas, une dignité particulière et un soin spécifique. Il y a là un vrai terrain de réflexion à réactiver, en particulier pour des objets à qui nous conférons des modes autonomes. Mais en même temps que nous « exigeons » d'eux, à travers nous, de respecter l'humain, l'obligation devrait être réciproque : quelle sorte de « respect » devons-nous à ces objets ? Cette posture éthique, la plupart du temps négligée, est nécessaire à la production d'artefacts qui nous « respectent ». La tentative de délégation de responsabilité à l'objet est irrespectueuse de l'objet comme de nous-mêmes.

Enfin, le plus important ici reste d'observer que le droit tente d'exprimer à son tour l'air du temps, celui qui exalte l'*universalisme* des espèces technologiques et sape, corrélativement, le principe d'un droit naturel humain. Sur ce très profond mouvement, il y aurait tant à dire. Commençons par ceci : il s'agit probablement de ce que les contempteurs du « progrès » technique appellent en gémissant une « rupture anthropologique ». A force de l'appeler, elle vient.