

Notes de prospective

Les *Notes de prospective* sont une publication de l'Institut syndical européen (ETUI, *European Trade Union Institute*) proposant des articles de réflexion stratégique sur les enjeux futurs du monde du travail. Cette publication est éditée par l'Unité prospective de l'ETUI, dont les deux principaux axes de travail sont la transition climatique et les nouvelles technologies. Les *Notes de prospective* sont également disponibles en anglais sous le titre: *Foresight Brief*.

Une législation européenne sur la robotique et l'intelligence artificielle ?

Aída Ponce Del Castillo

Chercheuse senior à l'Institut syndical européen

L'automatisation, les robots et l'intelligence artificielle sont d'ores et déjà profondément intégrés dans notre société. Ils continueront d'avoir un impact croissant sur la manière dont nous vivons. Ils intégreront bientôt nos lieux de travail où ils disposeront de niveaux toujours plus grands d'autonomie. Cette évolution ne va pas sans poser des problèmes: que se passe-t-il si ces dispositifs se trompent? Un robot peut-il être traîné devant les tribunaux? Que se passe-t-il si une erreur est commise par une voiture automatisée ou par ce robot soignant qui avait semblé si fiable lorsqu'il avait été acheté par votre hôpital local? Nous savons tous que l'inattendu peut se produire. Qui est responsable en fin de compte: le robot, son propriétaire, le fabricant? Cette *Note de prospective* présente la résolution du Parlement européen concernant des règles de droit civil sur la robotique. Elle attire l'attention sur la problématique de la réglementation des technologies et sur la responsabilité des acteurs concernés. Nous espérons qu'elle contribuera au débat sur notre interaction en constante évolution avec les robots, l'intelligence artificielle et la technologie.

#02 – Septembre 2017

Les opinions exprimées ici n'engagent que leurs auteurs et ne reflètent pas nécessairement celles de l'ETUI.

Directeurs de publication de la série:

Christophe Degryse
Philippe Pochet
Aída Ponce Del Castillo

Responsable de ce numéro:

Christophe Degryse,
cdegryse@etui.org

Plus d'informations sur
www.etui.org > Unité prospective

© ETUI aisbl, Bruxelles, 2017
Tous droits de reproduction réservés
ISSN 2507-1521

Introduction

Réglementer des technologies nouvelles constitue un véritable défi parce que cet exercice implique de traiter de technologies en évolution rapide, entourées d'incertitudes et souvent invisibles (Graeme *et al.* 2012). Il est difficile de savoir où ces technologies peuvent nous mener, ce qu'elles peuvent réaliser et, plus important encore, comment elles sont conçues et utilisées.

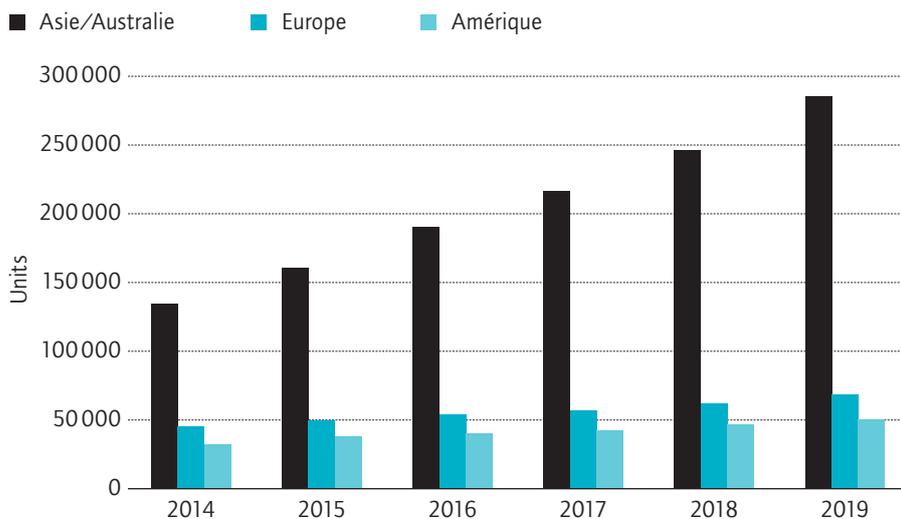
Il est tout aussi difficile de prévoir l'évolution que suivent nos sociétés humaines. Il est donc tout aussi compliqué de prévoir comment les technologies et la société interagissent et s'influencent mutuellement.

Dans le cas de technologies telles que la robotique et l'intelligence artificielle, la compréhension de la manière dont elles vont affecter nos sociétés nécessite de débattre de problèmes sociétaux et politiques latents, en vue de déterminer comment notre système réglementaire peut encadrer ces développements technologiques.

Les discussions au Parlement européen sur la réglementation de la robotique et de l'intelligence artificielle surviennent à point nommé, parce que ces technologies affectent et façonnent les relations que nous avons avec nos emplois et avec nos existences, en les rendant de plus en plus rapides, automatisés et numérisés. Qu'il s'agisse de production industrielle ou de services, le niveau de l'automatisation peut aller de l'exécution de tâches simples à l'adoption rapide de décisions complexes. En soi, c'est un moteur du changement social, du changement dans l'emploi et du changement environnemental. Toutefois, comme toute technologie disruptive, son impact sur la société est considérable et comporte des changements dans l'organisation du travail, les conditions de travail, la qualité et d'autres aspects de l'emploi comme l'externalisation de la responsabilité (Broersen 2014 ; Degryse 2017).

À l'heure actuelle, il n'existe aucune disposition légale spécifique sur la robotique (à l'exception de la directive 2006/42/CE sur les machines et

Production annuelle de robots industriels 2014-2015 et prévisions pour 2016-2019



Source : IFR/World Robotics

de la directive 85/374/CEE sur la responsabilité concernant les produits défectueux). Un ensemble minimum de règles est par conséquent nécessaire pour pouvoir établir une situation de concurrence équitable, transparente et démocratique, et certains éléments de la législation communautaire doivent être réexaminés en conséquence (Kritikos 2016).

Le présent article examine la résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (Parlement européen 2017). Il fournit un bref résumé du contenu de la résolution et examine ses principes de base et sa raison d'être. Il aborde également la problématique de la définition des robots et de leurs responsabilités. Ce faisant, il suggère l'existence d'une double évolution dans l'argumentation qui sous-tend les recommandations du Parlement. En utilisant une approche prospective et en tenant compte des analyses des spécialistes de la robotique et de l'intelligence artificielle, l'article indique que les recommandations du Parlement pourraient aller plus loin, en portant sur un éventail beaucoup plus large d'agents artificiels et de formes d'intelligence artificielle plutôt que de se focaliser sur des catégories spécifiques de robots.

L'article évoque ensuite la question de la responsabilité et de la visibilité des personnes qui disposent de pouvoirs de décision s'agissant de la conception, du développement et du déploiement de robots et d'objets basés sur l'intelligence artificielle, en ce compris les concepteurs et les développeurs.

Les fondements de l'initiative du Parlement et de la politique robotique

Le Parlement européen a pris l'initiative de rédiger un rapport présentant une motion de résolution de caractère législatif et comportant des recommandations relatives à des règles de droit civil appliquées à la robotique, émanant de la commission des affaires juridiques. La résolution ne constitue pas une initiative législative en soi, mais plutôt un ensemble de recommandations invitant la Commission européenne à élaborer un cadre législatif pour un usage civil des robots. Après un intense débat, le Parlement a voté en faveur du texte le 16 février 2017. Dans son suivi de la résolution, la Commission européenne a exprimé son intention (EC/2015/2103 (INL)) d'explorer différentes solutions en vue de s'attaquer à la question de la responsabilité et d'établir un registre complet, tout en s'interrogeant sur la question de savoir si une définition de ce qu'est un robot était nécessaire à des fins réglementaires. Il est encore plus intéressant de constater que la Commission a souligné l'importance d'en promouvoir l'acceptabilité dans l'opinion publique et la société.

Même si cela n'est pas formulé tel quel dans le corps du document, il est clair que la résolution est basée sur les résultats du projet FP7 RoboLaw – *Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics*, financé par la Commission européenne et mis en œuvre entre 2012 et 2014 (Palmerini *et al.*, 2014) ; elle s'inspire également de l'étude

Dans sa résolution, le Parlement réclame des solutions juridiques aux questions de responsabilité civile liées aux robots.

European Civil Law Rules in Robotics réalisée à la demande de la commission des affaires juridiques du Parlement européen (Nevejans 2016).

La première version de la résolution reposait sur la même structure que le projet RoboLaw. Elle s'articule autour de plusieurs sections relatives au développement de la robotique et de l'intelligence artificielle aux seuls usages civils, à savoir : des principes éthiques comme la transparence et les principes bioéthiques bien établis de bienfaisance, de non-malfaisance, d'autonomie et de justice ; la création d'une agence européenne chargée de la robotique et de l'intelligence artificielle, à même de fournir l'expertise technique, éthique et réglementaire ainsi que des directives sur les meilleures pratiques ; les droits de propriété intellectuelle et la circulation des données ; la normalisation, la sûreté et la sécurité ; les véhicules autonomes ; les robots de soins à la personne ; les robots médicaux ; la réparation et l'amélioration du corps humain ; les drones ; les perspectives en matière d'éducation et d'emploi ; la responsabilité y compris les aspects internationaux ; ainsi que l'adoption d'un code à l'attention des comités d'éthique de la recherche.

Les principaux postulats de la résolution sont le droit des personnes au respect de la vie privée, le respect de la fragilité humaine, une transparence maximale raisonnable dans la programmation des systèmes robotiques et la nécessaire prévisibilité du comportement des robots.

La résolution fournit une indication de ce qui devrait être inclus dans une définition des « robots autonomes intelligents » et recommande d'établir un cadre éthique pour la conception, la production et l'utilisation des robots. Le but final est de façonner la révolution technologique en intégrant les principes énoncés dans la législation européenne et dans des codes de conduite.

La résolution plaide pour des solutions juridiques à la question de la responsabilité civile pour les dommages causés par les robots, comme l'établissement d'un fonds de compensation, l'adoption de normes strictes et proportionnelles de responsabilité, la mise en place d'un régime d'assurance obligatoire complété par un fonds pour garantir que le dommage peut faire l'objet d'une indemnisation, l'attribution de numéros d'enregistrement et, à long terme, la création d'un statut légal spécifique pour les robots.

Dans le domaine de l'emploi, la résolution invite la Commission à développer les capacités numériques et à suivre les évolutions à long terme du marché du travail. Elle souligne la nécessité d'une formation plus flexible et l'importance des compétences créatives, sociales et numériques, ainsi que les risques nouveaux que la robotique pourrait faire naître.

La résolution commence par faire référence aux « trois lois de la robotique » souvent citées et formulées par l'auteur de science-fiction Isaac Asimov dans *Cercle vicieux [Runaround]* (1942). Les trois lois qui établissent que les robots doivent servir leurs maîtres humains s'énoncent comme suit : (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.

Les trois lois formulées par Asimov ont été utilisées et appliquées pour le développement de logiciels informatiques et dans d'autres technologies numériques (Feitelson 2007). Même si elles ne sont pas présentées comme

des principes juridiques essentiels, elles constituent une partie importante de l'argumentation du document.

La résolution sur la robotique n'est pas la première initiative du Parlement consacrée à l'impact de la technologie sur nos sociétés. En 2009, celui-ci s'était déjà penché sur la régulation des nanotechnologies et des nanomatériaux (Parlement européen 2009, 2014), poussant la Commission européenne à envisager l'adoption d'un cadre réglementaire.

Pourquoi le concept d'« agent artificiel » est-il plus approprié que celui de « robot » ?

Dans notre vie quotidienne et dans toutes sortes de contextes, nous entrons sans cesse en interaction avec des robots et d'autres systèmes d'intelligence artificielle. Des technologies en évolution rapide et constante se sont de plus en plus immiscées dans nos vies privée et professionnelle. Un exemple évident et simple de ce phénomène nous est offert par l'usage des appareils mobiles (contenant des données personnelles) et par le fait que pour certains individus, la connexion au monde « en ligne » est presque devenue une extension de leur moi (Floridi 2014).

Il existe d'innombrables exemples de la manière dont l'intelligence artificielle a été intégrée dans le lieu de travail. Dans le secteur financier, des algorithmes prennent des décisions sur l'achat et la vente d'actions et ils ont dans une certaine mesure remplacé la nécessité de courtiers humains. Dans le monde de la presse, des algorithmes peuvent identifier, calculer et sélectionner de nouvelles informations à publier par un média donné. Dans le secteur des soins de santé, des robots peuvent aller chercher des objets, mesurer les signes vitaux d'un patient, guider des personnes âgées dans des maisons de repos et de soins et même faire preuve d'un « comportement social ». Dans le secteur de la sécurité, les robots fusionnent avec d'autres systèmes technologiques comme « l'Internet des objets », des algorithmes et des plates-formes, avec une interaction humaine limitée. Dans le secteur de la fabrication, les robots collaboratifs ou les systèmes de robots travailleurs opèrent à côté des humains et leurs sphères et champs d'action respectifs se chevauchent.

Les robots qui doivent le plus attirer l'attention sont ceux qui possèdent les capacités d'apprentissage les plus importantes : ils sont susceptibles d'apprendre de leur environnement, de percevoir ce qui les entoure, d'identifier les schémas, de modifier la manière dont un problème est initialement posé, d'adapter leur comportement en réaction à leur environnement et de fonctionner de manière semi-autonome (à ce jour, il n'existe pas de robots autonomes pleinement fonctionnels sur le long terme). Leur interaction avec les humains peut se développer dans de nombreuses directions et à de nombreux niveaux en jouant un rôle actif dans le processus décisionnel (robots/chirurgiens, algorithmes/marchés financiers, etc.).

Les robots qui méritent le plus d'attention sont ceux disposant de capacités d'apprentissage importantes.

Définir et classer les robots

Le mot « robo » vient du mot tchèque « robota » qui signifie « travail servile », « corvée » ou « travail pénible ». Toutefois, il n'existe pas de consensus sur la signification précise de ce terme et ses nombreuses connotations négatives le rendent très difficile à définir. Même le classement des robots par catégorie peut s'avérer un exercice interminable.

Le spécialiste de la robotique Alan Winfield a proposé de les définir en leur associant trois qualités. Winfield parle d'un « appareil artificiel qui peut *percevoir* son environnement et *agir délibérément* sur ou dans cet environnement ; une intelligence artificielle *incarnée* ; ou une machine qui peut se livrer *de manière autonome* à des travaux utiles » (Winfield 2012).

Il faut noter que dans sa résolution, le Parlement européen ne propose pas de définition ou de classification simple des robots autonomes « intelligents ». Mais il propose, dans l'annexe à la résolution, certains critères à prendre en considération et invite la Commission à présenter une définition des systèmes cyber-physiques, des systèmes autonomes des robots autonomes et intelligents, ainsi que de leurs sous-catégories. Ce qui en soi peut constituer un processus interminable et soulève un certain nombre de problèmes.

Nous pouvons en effet comparer ce processus à celui qui a été utilisé pour définir les nanotechnologies et les nanomatériaux en 2010 (Nanowerk 2013). Les divergences entre les opinions scientifiques avaient alors conduit à des définitions qui étaient autant de compromis politiques et les négociations sur leur révision se poursuivent encore aujourd'hui. Peut-être pouvons-nous en tirer une leçon : s'efforcer de parvenir à une définition et à une catégorisation stricte à des fins réglementaires ne constitue pas toujours la meilleure option, et le Parlement européen devrait éviter de s'enliser dans des discussions visant à définir ce qui n'est guère définissable.

Agents artificiels

Plutôt que de chercher à définir les robots, qui peuvent être divisés en millions de sous-catégories sur la base de leurs fonctionnalités, la Commission devrait adopter une approche qui englobe les algorithmes et l'intelligence artificielle. Elle devrait également prendre en considération les machines qui ont une capacité d'apprentissage et d'évolution et qui sont susceptibles de devenir semi-autonomes, voire entièrement autonomes.

Nous devons également reconnaître l'importance de l'utilisation d'une terminologie correcte dans le processus décisionnel et dans la rédaction de la législation. Les mots qui sont utilisés pour décrire ces machines ont un impact sur le type de politique et de législation qui en résulteront. Comme ces technologies sont si étroitement intégrées dans tous les aspects de la vie humaine, nous devons avancer avec prudence. Devons-nous faire référence aux robots, aux agents artificiels ou aux agents artificiels autonomes ? Et quelles sont les conséquences probables de cette distinction ?

Le terme « agents artificiels », compris comme un concept générique, a du sens parce qu'il recouvre une large diversité d'agents « soft » ou « hard » : algorithmes de décision, machines automatisées, agents numériques, agents

hybrides, logiciels automatiques (bots) sur Internet, robots, nano-robots, drones, etc. Ces agents ont une capacité de fonctionnement et d'apprentissage par l'expérience et l'interaction sans intervention directe d'êtres humains ou d'autres agents. Ils ne sont pas (encore) entièrement autonomes ; ils sont artificiels parce qu'ils sont produits ou construits par des humains, et ce sont des agents parce qu'ils se livrent à des actions (Chopra et White 2004 ; Chopra 2010 ; Floridi 2014, 2015).

L'utilisation du terme « agent artificiel » pourrait constituer une première étape dans le développement d'un cadre juridique solide visant à parvenir à la transparence et à la responsabilité au travers d'une certaine forme de contrôle. Il pourrait nous aider à éviter de tomber dans un trou noir réglementaire qui serait créé par un effort d'être trop précis.

Les robots en tant que personnes électroniques ?

Le Parlement européen recommande de créer une personnalité juridique spécifique pour les robots en tant que « **personnes électroniques** », avec des droits et des obligations spécifiques, et qui s'appliquerait dans les cas où les robots prennent des décisions ou interagissent avec des tiers (paragraphe 59(f) de la section sur la responsabilité), un statut qui est inconnu pour l'instant dans notre système juridique.

Le fait que certains robots remplissent des fonctions comme la gestion d'opérations, la délégation de tâches, la résolution de problèmes complexes et la prise de décisions en temps réel soulève des questions d'ordre juridique. Selon le Parlement, leur attribuer la personnalité électronique résoudrait ces problèmes juridiques.

Selon la théorie juridique actuelle, attribuer la personnalité juridique aux agents artificiels est une démarche complexe. Il ne s'agit pas d'assimiler purement et simplement les robots à des sociétés. La démarche rouvre le débat sur la personnalité qui a toujours été une source de controverses (comme l'a montré celle du statut des esclaves ou des femmes dans le passé, ou celle d'autres êtres et des sociétés dans des époques plus récentes) (Chopra et White 2011).

Quelles seraient les implications de cette approche dans le cas des robots ?

La personnalité juridique signifie le fait d'avoir des droits et des obligations et la capacité de mener une action civile et d'être tenu responsable de ses actes (et même d'intenter des poursuites ou d'être tenu responsable pour des délits). Une *personne juridique* est également en mesure d'exprimer des valeurs morales et de devenir active sur le plan politique. Une fois que la personnalité est attribuée à des robots ou à d'autres agents artificiels autonomes, ils deviennent des *sujets* (par opposition aux choses ou aux objets) et pénètrent dans l'univers des *personnes juridiques* (Chopra 2010 ; Chopra et White 2011).

Des discussions devraient être menées sur la nature de ces droits et de ces obligations et sur la capacité d'agents artificiels autonomes à les exercer.

Jusqu'à quel point les agents artificiels peuvent-ils devenir autonomes ? Pourront-ils être conscients de leurs intentions ? Pourront-ils savoir qu'ils sont responsables et donc susceptibles de rendre des comptes ?

Pour illustrer la complexité de la question, Samir Chopra rappelle que toutes les personnes n'ont pas les mêmes droits et les mêmes obligations. Certains de ces droits ou obligations dépendent de l'âge comme le droit de vote, le droit de se marier ou l'obligation de faire un service militaire. L'exercice de ces droits et obligations requiert l'autonomie et la capacité de se gouverner soi-même, ce qui, pour des agents artificiels autonomes, pourrait signifier la capacité de travailler sans aucune supervision humaine.

Dans ce contexte, plusieurs questions émergent s'agissant de la résolution du Parlement européen : jusqu'à quel point exactement les agents artificiels pourraient-ils être *autonomes* ? Pourraient-ils être *conscients* de leurs intentions ? Pourraient-ils *savoir* qu'ils sont responsables et devront donc rendre des comptes ? Si oui, ils auraient à assumer une responsabilité morale et à faire face aux conséquences économiques des risques ou des accidents.

La résolution identifie des solutions basées sur des régimes d'assurance et des fonds d'indemnisation (paragraphe 56 à 59 de la section sur la responsabilité).

Les difficiles questions de responsabilité et d'octroi de la « personnalité » juridique à des agents artificiels autonomes exigeraient une réorientation fondamentale de la pensée juridique. Si l'on parvient un jour à cette situation, les humains et les robots deviendraient en effet des acteurs au sein du même cadre juridique.

Quel moteur pour la réglementation européenne ?

Dans le cas de technologies complexes dont les possibilités et la portée demeurent inconnues, **les risques potentiels** sont également inconnus. Les robots et l'intelligence artificielle pourraient se substituer aux humains et transformer les emplois, mais à quel prix ? Certains scientifiques, comme Stephen Hawking et Bill Joy, sont préoccupés par le risque d'une utilisation abusive et ont prédit la destruction de l'humanité (Hern 2016). D'autres voient les choses selon une perspective différente et pensent que, pour pouvoir l'emporter sur les êtres humains, les machines dans leur version *hard*

(machines et robots « visibles ») ou *soft* (logiciels, algorithmes, etc.) auraient besoin de plus que de l'intelligence et d'un apprentissage en profondeur : elles devraient maîtriser l'abstraction, avoir la capacité de comprendre le monde visuellement et le langage sur le plan sémantique (Knight 2016), et être créatives et même faire preuve d'empathie. D'autres encore, comme Luciano Floridi, professeur

d'éthique et de philosophie de l'information à l'université d'Oxford, estiment qu'il y a peu de crainte à avoir car il existe tant de choses que les machines ne peuvent faire, comme réfléchir, connaître et être conscientes (Floridi 2016).

En toute hypothèse, l'objectif de la réglementation consiste à limiter un possible abus ou une utilisation « stupide » de cette puissance, et ainsi de minimiser les risques et les effets négatifs possibles des technologies nouvelles. Le développement de scénarios d'exposition permettant d'identifier le niveau

L'objectif est de favoriser une meilleure interaction entre les concepteurs et les travailleurs/ utilisateurs.

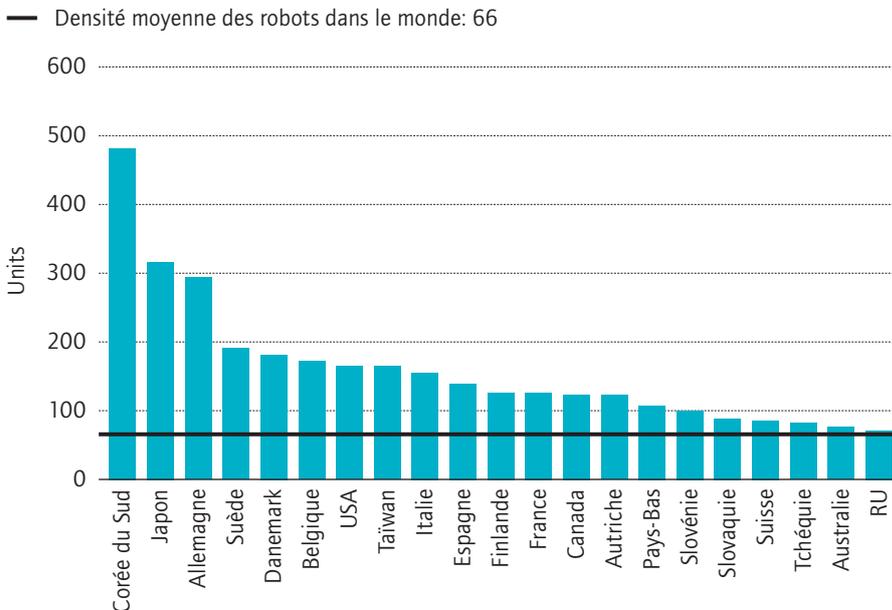
de risque est essentiel. Cela s'applique aux agents artificiels et c'est pourquoi un cadre réglementaire approprié doit être envisagé, par opposition à des mesures *soft* comme des codes de conduite, une certification ou des normes.

Les codes de conduite peuvent contribuer à promouvoir un comportement éthique et à introduire certaines valeurs et lignes de conduite que les professionnels doivent suivre, mais ils ne constituent pas des instruments de gouvernance. Lorsque la réglementation des nanotechnologies était à l'agenda, la Commission européenne a produit un code de conduite pour une recherche responsable dans les nanosciences et les nanotechnologies (Commission européenne 2009). Le code s'adressait à tous les acteurs concernés mais il présentait deux inconvénients majeurs : il n'était pas contraignant et son application n'était pas surveillée par une instance européenne ou nationale.

La certification des concepteurs ne constitue pas non plus une solution puisqu'il s'agit d'une démarche qui s'effectue dans un cadre privé : les entreprises privées établissent leurs propres critères s'agissant de ce qu'elles souhaitent certifier. Leur principal objectif est d'obliger les concepteurs à travailler en fonction d'intérêts spécifiques et non de protéger la société, et souvent ces démarches ne tiennent pas compte des vues de la société.

Le même raisonnement s'applique aux normes techniques internationales. Les organisations internationales de normalisation fonctionnent sur la base d'adhésions payantes, ce qui permet à un nombre limité d'acteurs de contribuer au développement et à la mise en place de normes techniques. Une fois encore, les objections que l'on peut soulever ici sont que cette procédure n'inclut pas la participation des acteurs de la société et qu'en raison de sa nature non contraignante, elle ne peut remplacer une législation (Hauert *et al.* 2015).

Nombre de robots industriels multifonctions (de tous types) par 10 000 employés dans l'industrie (ISIC rev.4: C) 2014



Source: IFR/World Robotics

Aucune de ces options ne conduit donc à une gouvernance effective ou ne résout la question de la responsabilité. La réglementation des robots et de l'intelligence artificielle est essentielle, mais pas à cause de la crainte de voir les robots confisquer aux humains des pouvoirs de décision. Le présent article soutient que le moteur de la réglementation européenne doit être, pour le dire simplement, de veiller à assurer visibilité et traçabilité s'agissant de qui est responsable de quoi. L'objectif est de favoriser une meilleure **interaction entre les concepteurs et les travailleurs/utilisateurs**, dont la relation est très intense et étroitement entremêlée. Agents artificiels et êtres humains doivent coexister, mais les premiers cités doivent rester utiles aux êtres humains, tout en respectant les valeurs humaines et les règles d'une société démocratique.

Conclusion et perspectives

Les aspects réglementaires de la robotique et de l'intelligence artificielle commencent à être discutés dans les institutions de l'Union européenne, où différents acteurs s'engagent dans la discussion et cherchent à défendre leurs valeurs et leurs intérêts. Cet article a cherché à analyser les questions fondamentales relatives aux difficultés inhérentes à la définition des robots et

Notre tendance à déléguer des responsabilités à des systèmes artificiellement intelligents deviendra un problème sérieux pour notre société et pour nos systèmes juridiques.

de leurs catégories, et il a mis en évidence les problèmes associés à la « personnalité électronique ». Il suggère que, avant de légiférer sur des questions telles que l'augmentation artificielle des capacités humaines, le type d'assurance qui devrait être requis pour les robots, la manière dont une agence européenne pour la robotique devrait être organisée et quelle serait sa mission, etc., le Parlement et la Commission européenne devraient examiner des questions fondamentales relatives à la responsabilité. Comme l'a souligné Broersen

(2014), « notre tendance à déléguer des responsabilités à des systèmes artificiellement intelligents deviendra un problème sérieux » pour notre société et pour nos systèmes juridiques. C'est pourquoi, en vue de réglementer les agents artificiels, cet article identifie un certain nombre d'aspects essentiels qui doivent être traités et pris en considération avant que nous nous retrouvions enlisés dans des débats qui, très vraisemblablement, déboucheront sur un compromis politique.

Des règles doivent être élaborées pour clarifier qui fait quoi parmi les créateurs, concepteurs, spécialistes des données, fournisseurs et entreprises responsables de la fabrication d'agents artificiels. Assurer la **visibilité** des acteurs et de leur responsabilité constitue une première étape vers une réglementation de l'automatisation et de l'intelligence artificielle. En d'autres mots, nous devons être certains de savoir qui est responsable de quoi et assurer la traçabilité des acteurs (Mittelstadt *et al.* 2016).

À cette fin, il est également nécessaire d'être en mesure d'identifier tous les autres **acteurs qui interagissent avec des agents artificiels et qui**

les utilisent, comme les travailleurs, les employeurs, les consommateurs, les patients, les utilisateurs et les formateurs.

Des normes réglementaires minimales doivent être développées afin de pouvoir déterminer la **responsabilité** lorsque l'agent artificiel possède des caractéristiques « d'apprentissage et d'enseignement » et est en mesure de générer des résultats non désirés (Grodzinsky *et al.* 2008 ; Vanderelst et Winfield *et al.* 2016). Ces normes doivent assurer que des ajustements appropriés sont mis en place en vue d'éliminer autant que possible tous les comportements ou conséquences indésirables ou non programmés. Ces normes doivent également traiter de la question de la responsabilité si de tels comportements ou conséquences se produisent.

L'UE doit mettre en place des exigences contraignantes pour exercer le « **droit à l'explication** » des modèles et des décisions dues à des systèmes algorithmiques automatisés ou dotés d'une intelligence artificielle, conformément à ce qui était déjà établi dans la directive sur la protection des données adoptée en 2016 (Wachter *et al.* 2016). S'agissant spécifiquement des travailleurs, la réglementation doit garantir leurs droits à ce que la logique, les fonctionnalités et les conséquences de systèmes de décision automatisés leur soient expliquées. Elle doit également identifier les moments où une intervention humaine est nécessaire et quand une décision peut être contestée. Pour le dire simplement, « les travailleurs ne doivent pas connaître le code mais ils doivent savoir ce que le code cherche à obtenir ».

Nous sommes en train de créer un nouvel espace immense fait de machines et de données, qui transforme notre environnement et nos politiques. Il est essentiel qu'une réglementation soit adoptée pour répondre à ce monde nouveau que nous sommes en train de bâtir. Il est également essentiel que tous les acteurs clés soient rendus visibles, et finalement responsables, étant donné que la décision ultime concernant l'objectif d'un projet, les relations en jeu dans l'utilisation d'une machine, les niveaux de sécurité et les risques potentiels pour la société, la santé et l'environnement, leur incombe presque exclusivement.

Compte tenu de l'intégration d'agents artificiels dans notre société et en particulier s'agissant de leur déploiement sur le lieu de travail, nous devrions voir plus loin que leur utilisation aux seules fins d'augmenter la productivité et la rentabilité. Les conséquences du développement technologique doivent être envisagées en tenant compte des acteurs de la société (Winner 1980). Un cadre réglementaire efficace s'impose pour assurer que les agents artificiels coexistent harmonieusement avec les humains, qu'ils soient spécifiquement conçus en fonction des valeurs et des besoins des hommes, et qu'ils opèrent et soient capables de s'adapter en conséquence. ●

L'UE doit mettre en place des exigences contraignantes pour exercer le « droit à l'explication » des modèles et des décisions dues à des systèmes algorithmiques automatisés ou dotés d'une intelligence artificielle.

Références

- Asimov I. (1942) *Runaround*, in Asimov I., I. Robot, London, HarperVoyager, 2001.
- Broersen J. (2014) Responsible intelligent systems. The REINS Project, *KI - Künstliche Intelligenz*, 28 (3), 209-214. DOI 10.1007/s13218-014-0305-4
- Chopra S. (2010) Rights for autonomous artificial agents?, *Communications of the ACM*, 53 (8), 38-40.
- Chopra S. et White L. F. (2011) *A legal theory for autonomous artificial agents*, Ann Arbor, The University of Michigan Press.
- Chopra S. et White L.F. (2004) Artificial agents – personhood in law and philosophy, *Proceedings of the 16th European Conference on Artificial Intelligence, Valencia, August 22-27, 2004*, 635-639.
- Commission européenne (2009) *Commission recommendation on a code of conduct for responsible nanosciences and nanotechnologies research & Council conclusions on responsible nanosciences and nanotechnologies research*, Luxembourg, Office for Official Publications of the European Communities. http://ec.europa.eu/research/science-society/document_library/pdf_06/nanocode-apr09_en.pdf
- Degryse C. (2017) *Façonner le monde du travail dans l'économie digitale*, Note de prospective, (1), janvier 2017, Bruxelles, ETUI.
- Feitelson D. G. (2007) Asimov's laws of robotics applied to software, *IEEE Software*, 24 (4), 112. DOI 10.1109/MS.2007.100
- Floridi L. (2016) Should we be afraid of AI?, *AEON*. <https://aeon.co/essays/true-ai-is-both-logically-possible-and-utterly-implausible>
- Floridi L. (2015) Toleration and the design of norms, *Science and Engineering Ethics*, 21 (5), 1095-1123. DOI 10.1007/s11948-014-9589-x
- Floridi L. (2014) *The fourth revolution. How the infosphere is reshaping human reality*, Oxford, Oxford University Press.
- Graeme L., Harmon S. HE and Arzuaga F. (2012) Foresighting futures: law, new technologies and the challenges of regulating for uncertainty, *Law, Innovation and Technology*, 4 (1), 1-33. DOI 10.5235/175799612800650626
- Grodzinsky F., Miller K. W. et Wolf M.J. (2008) The ethics of designing artificial agents, *Ethics and Information Technology*, 10 (2-3), 115-121. DOI 10.1007/s10676-008-9163-9
- Hauert C, Bütschi D., Graz J.-C., Audétat M. et Kaufman A. (2015) *The international standardisation arena and the civil society participation stakes: results of the INTERNORM project*, Policy Brief 14.2015, Brussels, ETUI. <http://www.etui.org/Publications2/Policy-Briefs/European-Economic-Employment-and-Social-Policy/The-international-standardisation-arena-and-the-civil-society-participation-stakes-results-of-the-INTERNORM-project>
- Hern A. (2016) Stephen Hawking: AI will be 'either best or worst thing' for humanity, *The Guardian*, 19 October 2016. <https://www.theguardian.com/science/2016/oct/19/stephen-hawking-ai-best-or-worst-thing-for-humanity-cambridge>
- Knight W. (2016) AI's unspoken problem, *MIT Technology Review*, 19 (5), 28-37.
- Kritikos M. (2016) Legal and ethical reflections concerning robotics, *STOA Policy Briefing*, June 2016 – PE 563.501, Brussels, European Parliament Research Service. [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/563501/EPRS_STU\(2016\)563501\(ANN\)_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/563501/EPRS_STU(2016)563501(ANN)_EN.pdf)
- Mittelstadt B. *et al.* (2016) The ethics of algorithms: mapping the debate, *Big Data & Society*, December 1, 2016, 1-21. DOI: 10.1177/2053951716679679
- Nanowerk (2013) Definition of the term 'nanomaterial'. <http://www.nanowerk.com/spotlight/spotid=30804.php>
- Nevejans N. (2016) *European civil law rules in robotics*, Directorate General for Internal Policies. Policy Department C: Citizens' Rights and Constitutional Affairs, Study PE 571.319, Brussels, European Parliament.
- Palmerini E. *et al.* (2014) *Guidelines on regulating robotics*. RoboLaw Deliverable D6.2. http://www.robolaw.eu/RoboLaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf
- Parlement européen (2017) *Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL))*.
- Parlement européen (2009) *Résolution du Parlement européen du 24 avril 2009 sur les aspects réglementaires des nanomatériaux (2008/2208(INI))*.
- Vanderelst D. et Winfield A. (2016) *An architecture for ethical robots*, arXiv:1609.02931v1 [cs.RO] 9 September 2016. <https://arxiv.org/pdf/1609.02931v1.pdf>
- Wachter S., Mittelstadt B. et Floridi L. (2016) Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation (December 28, 2016), *International Data Privacy Law*, forthcoming. Available at SSRN: <https://ssrn.com/abstract=2903469>
- Winfield A. (2012) *Robotics. A very short introduction*, Oxford, Oxford University Press.
- Winner L. (1980) Do artifacts have politics?, *Daedalus*, 109 (1), 121-136. http://www.jstor.org/stable/20024652?origin=JSTOR-pdf&seq=1#page_scan_tab_contents

Les liens ont été vérifiés le 06.09.2017.

