

Interview de Laurence Devillers
Propos recueillis par Françoise Laugée

Vous êtes chercheuse au Limsi (Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur) du CNRS. Expliquez-nous votre domaine de recherche. Très concrètement sur quoi travaillez-vous ?

Spécialiste de l'*affective computing*, apprentissage machine et dialogue oral homme-machine, j'anime depuis 2000 une équipe sur les « *dimensions affectives et sociales dans les interactions parlées* ». Le domaine de l'*affective computing* est issu des travaux de recherche de Rosalind Picard au MIT¹ en 1997 et recouvre trois technologies, la reconnaissance des émotions humaines, le raisonnement et la prise de décision, pour la génération d'expressions émotionnelles. Je travaille plus particulièrement sur la détection des émotions dans le signal audio, la construction d'un profil émotionnel dans l'interaction et la gestion des stratégies du dialogue. Ces systèmes utilisent des approches d'apprentissage machine.

Dans les années 1985-1990, l'apprentissage machine avec des réseaux de neurones était un domaine nouveau de l'intelligence artificielle. Classiquement, l'intelligence artificielle manipulait plutôt des connaissances symboliques. L'apprentissage machine connexionniste, par opposition à l'apprentissage symbolique, regroupe des algorithmes de type réseaux de neurones qui apprennent à partir de données et essaient de prédire des classes (mots, émotions, etc.). La qualité de l'apprentissage dépend de la qualité des données. Les réseaux de neurones sont une approche nouvelle d'apprentissage, on étudie l'analogie entre le cerveau humain et les machines informatiques universelles. La plasticité cérébrale montre que si le cerveau est en perpétuelle reconfiguration, c'est grâce à cette plasticité que l'on peut apprendre.

En s'appuyant sur les propriétés des neurones biologiques connues à cette époque, issues d'observations neurophysiologiques et anatomiques, les mathématiciens McCulloch et Pitts ont proposé en 1943 un modèle simple de neurone formel, un automate représentant le noyau de la cellule, l'axone et ses dendrites qui, à partir d'un certain seuil d'activation, envoie de l'information aux neurones connectés. L'apprentissage dans un réseau de neurones modifie les poids entre les cellules. Les réseaux de neurones de type Perceptron Multi-Couches bénéficient de nombreux travaux théoriques sur l'apprentissage. Un réseau de neurones artificiels de ce type est un ensemble de neurones formels interconnectés et répartis en plusieurs couches, l'algorithme de « rétro-propagation du gradient », permettant d'entraîner ces réseaux de neurones, date de 1986. L'apprentissage avec les réseaux de neurones artificiels se fait par essais et erreurs à partir de base de données étiquetées. À l'époque, l'apprentissage était très lent et en 1990, il fallait parfois un mois de calcul pour l'obtention d'un modèle. Les calculs sont maintenant effectués en parallèle et très rapidement sur un ordinateur portable.

EN 1990, IL FALLAIT PARFOIS UN MOIS DE CALCUL POUR L'OBTENTION D'UN MODÈLE

Pour l'apprentissage, il est nécessaire de présenter au réseau un grand nombre de données, par exemple des sons avec leurs étiquettes phonétiques. Pour chaque son, l'algorithme propage l'activité dans le réseau, de couche en couche, calcule l'erreur de prédiction par rapport à l'étiquette connue et rétro-propage l'erreur à travers le réseau en modifiant un peu chacun des seuils entre les neurones du réseau, afin que, la fois suivante, il prédise mieux le son testé. Après un grand nombre d'essais-erreurs, le réseau se stabilise sur une configuration de poids entre les cellules. C'est ce qu'on appelle l'apprentissage car la machine comprend de mieux en mieux en écoutant de nombreuses voix aux intonations différentes.

Les automates sont sortis des usines. Des robots, logiciels ou machines, humanoïdes même, s'immiscent dans notre vie de tous les jours grâce au progrès de l'intelligence artificielle transformant des secteurs entiers d'activité : transports, banque, justice, commerce, éducation ou santé. Quel est l'état de l'art en la matière ? Pour reprendre le titre de votre ouvrage *Des robots et des hommes. Mythes, fantasmes et réalité* (Plon, 2017), quelle est la part du fantasme ou au contraire de la réalité quant aux capacités et au degré d'autonomie de ces agents artificiels ?

Dans cet ouvrage, j'ai analysé les capacités et le degré d'autonomie de ces agents artificiels, afin de les démystifier et d'expliquer ce qui relevait du fantasme ou de la réalité.

D'un point de vue informatique, un robot social est une entité complexe qui peut être pilotée par plusieurs ordinateurs regroupant des fonctions qui correspondent à différents programmes :

- la perception de signaux physiques grâce à des caméras vidéo thermiques, des microphones, des caméras Kinect qui capturent le mouvement gestuel, facial et la profondeur de champ ;
- l'interprétation de ces signaux et la reconnaissance de la parole, des émotions et des images (visages, objets, lieux, etc.) dans le contexte de la tâche ;
- le raisonnement et la prise de décision à partir d'informations perçues et reconnues, et également mémorisées ;
- la génération d'actions que le robot exécute dans le monde physique ;
- mais aussi des réponses orales grâce à la synthèse de la parole.

Ces robots sociaux sont cependant des objets présentant certaines particularités :

- le robot peut se substituer à l'homme dans des situations de surveillance, d'assistance et d'interaction sociale ;
- il simule des caractéristiques du vivant : son aspect physique, son comportement en interaction (simulation de parole et d'émotions), ses actions (marche, prise d'objets) ou ses capacités de raisonnement et de décision (choix entre plusieurs comportements) ;
- enfin, il se déplace dans notre environnement réel.

LES MACHINES NE COMPRENNENT PAS, ELLES RECONNAISSENT LA SURFACE DU LANGAGE
MAIS SANS EN COMPRENDRE LE SENS

Les robots des films de science-fiction sont dotés de capacités sensorielles et motrices ultra-performantes, ainsi que d'un pouvoir de raisonnement, qui vont très au-delà des capacités des robots actuels. Depuis la révolution des temps modernes, une relation complexe nous unit aux machines, nourrissant peurs et fascination, entretenues par la littérature et le cinéma de science-fiction.

Les mythes du Golem ou de Pygmalion inspirent encore aujourd'hui la littérature et les films de science-fiction comme *A.I.* de S. Spielberg (2001), où les robots, créés à l'image des humains, font partie de leur univers. Le robot a remplacé la statue. Avec le développement exponentiel du progrès scientifique, l'être artificiel devient une source de préoccupation majeure pour de nombreux créateurs, mais parfois également d'espoir.

L'imaginaire que nous développons autour des robots prend ses racines dans la fiction - *Metropolis*, 2001 *l'Odyssée de l'espace*, *Star Wars*, *Wall-E* - et dans les mythes - *Le Golem*, *Pygmalion* - se confrontant à de nombreuses questions d'ordre technologique, philosophique, religieux ou social. Depuis 2001, *l'Odyssée de l'espace*, film de Stanley Kubrick sorti en 1968, qui met en scène HAL, une [intelligence artificielle](#) avec des états d'âme, on se demande si une machine peut être morale, si elle peut faire le bien ou le mal, et si elle peut être tenue responsable de ses actes.

Dans la saga des cinq *Terminators* de 1984 à 2015, l'[intelligence artificielle](#) Skynet, qui décide de construire une armée de robots dans ces films de James Cameron, doit absolument tout à la science-fiction et aux mythes. Ces films ont pour principal sujet le voyage dans le temps et la menace que les robots intelligents pourraient faire naître. Alors que la plupart des films de science-fiction se passaient jusqu'alors dans l'espace, sur Mars, avec des luttes guerrières et des conquêtes ; un nouveau genre s'intéresse désormais aux relations humaines des robots, à l'instar des œuvres *HER*, *Real Humans* et *Ex-Machina*. « *Pouvons-nous tomber amoureux d'une voix ?* », tel est le sujet du film *HER* réalisé par Spike Jonze, sorti en 2015 et mettant en scène un héros solitaire qui se laisse séduire par un compagnon numérique présenté

comme une véritable conscience. Il s'agit dans ce film d'un programme informatique capable d'interagir vocalement avec un être humain. Si la reconnaissance de la parole n'a pas cessé d'évoluer, pour autant les machines ne comprennent pas, elles reconnaissent la surface du langage mais sans en comprendre le sens et elles n'ont la capacité ni de comprendre le contexte ni le sens général.

DOTÉS D'UNE AUTONOMIE PLUS OU MOINS GRANDE, LES AGENTS ARTIFICIELS PEUVENT
APPARAÎTRE COMME DES ACTEURS SOCIAUX

La robotique sociale et affective veut créer des robots compagnons, censés nous apporter une aide thérapeutique, d'assistance ou encore de surveillance. Il est urgent de comprendre l'utilité de ces approches et d'encadrer les logiciels. Même si ces machines ne sont pas encore très performantes, elles peuvent rendre des services. De nombreux agents artificiels utilisent des modules d'apprentissage machine. Ces agents peuvent être des logiciels comme les agents conversationnels ou des instruments physiques comme les robots ou les voitures autonomes. Dotés d'une autonomie plus ou moins grande, ils peuvent apparaître comme des acteurs sociaux, avec des capacités d'interaction, voire de simulation, d'émotions et de décisions d'actions. Dans le cas de pathologies neurodégénératives ou de handicaps lourds, le robot peut se mettre au diapason de l'autre, à des rythmes très lents : le robot écoute sans aucune impatience mais sans empathie ressentie.

Vous êtes membre de la Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique (Cerna). Quels sont les objectifs de la Cerna²? Quelles sont les conditions et les limites - et les risques aussi -, clairement identifiés en ce qui concerne le développement des machines apprenantes ?

La Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene (Cerna) a pour but de répondre aux questions d'ordre éthique. Il s'agit de :

- mener une réflexion sur l'éthique des recherches scientifiques développées en sciences et technologies du numérique, de sensibiliser les chercheurs à la dimension éthique de leurs travaux ;
- aider à exprimer les besoins spécifiques de la recherche au regard du législateur et à les ouvrir à une démarche responsable ;
- apporter un éclairage de nature scientifique aux décideurs et à la société sur les conséquences potentielles de résultats de recherche ;
- veiller à ce que les étudiants soient formés à ces questions et suggérer des thèmes de recherche permettant, en résumé, d'approfondir la réflexion éthique dans un cadre interdisciplinaire.

UNE MÉDIATISATION, AUTANT EXCESSIVE QU'APPROXIMATIVE, DE L'IDÉE QUE LA MACHINE - PARFOIS UN ROBOT - POURRAIT APPRENDRE À S'AFFRANCHIR DE L'HOMME

Les succès de l'apprentissage machine, champ d'études de l'intelligence artificielle (IA), s'appuient sur l'accroissement des capacités de calcul, de stockage, de traitement des données (*big data*) et font resurgir avec une médiatisation, autant excessive qu'approximative, l'idée que la machine - parfois un robot - pourrait apprendre à s'affranchir de l'homme. Construire des systèmes capables de fonctions de perception, d'apprentissage, d'abstraction et de raisonnement est un des buts des chercheurs en intelligence artificielle. À cette fin, les algorithmes d'apprentissage utilisent différentes méthodes statistiques en se basant sur des données d'apprentissage, par exemple pour construire des règles de déduction, des arbres de décision ou pour paramétrer des réseaux de neurones ; ils les appliquent ensuite à de nouvelles données. Prédire un phénomène à partir d'observations passées présuppose un mécanisme causal. Expliquer ce mécanisme n'est pas toujours facile. L'apprentissage machine est une approche statistique permettant de découvrir des corrélations significatives dans une masse importante de données pour construire un modèle prédictif quand il est difficile de construire un modèle explicatif.

QUI EST RESPONSABLE EN CAS DE DYSFONCTIONNEMENT DE LA MACHINE ?

Les algorithmes d'apprentissage machine utilisés sont nombreux et divers. Ils peuvent se classer en trois grandes catégories selon leur mode d'apprentissage : *supervisé*, *non supervisé* et *par renforcement*. En *apprentissage supervisé*, les données utilisées doivent être annotées préalablement par des « experts ». Lors d'une première phase, dite d'apprentissage, la machine construit un « modèle » des données annotées, qui peut être un ensemble de règles, un arbre de décision, un ensemble de matrices, comme

dans les réseaux de neurones, etc. Ce modèle est ensuite utilisé dans une seconde phase, dite de reconnaissance. *En apprentissage machine non supervisé*, aucun expert n'est requis pour annoter les données. L'algorithme découvre par lui-même la structure des données, en les classant en groupes. *En apprentissage par renforcement*, le but est d'apprendre, à partir d'expériences, ce qu'il convient de faire dans des situations différentes, afin d'optimiser une récompense quantitative au fil du temps. L'algorithme procède par essais et erreurs, chaque erreur l'amenant à améliorer sa performance dans la résolution du problème. Le rôle des experts se limite ici à définir les critères de succès de l'algorithme.

Les *chatbots*, ou *bots*, sont des agents de traitement automatique de conversation en langage naturel qui utilisent des algorithmes appartenant à ces trois catégories. Ils sont de plus en plus présents en tant qu'assistants personnels ou interlocuteurs dans des transactions commerciales réalisées en ligne sur des plateformes informatiques. Il arrive aussi qu'ils soient majoritaires dans des *chats* avec les humains. La prolifération massive de ces conversations, sans hiérarchie ou distinction claire entre l'humain et la machine, pourrait à terme influencer les prises de décision des humains. En avril 2016, le [chatbot](#) Tay de Microsoft, qui avait la capacité d'apprendre en interaction, a rapidement été retiré par le groupe : en douze heures il avait appris à tenir des propos racistes et nazis en interagissant avec des internautes malveillants. La machine avait découvert que parler de sujets non politiquement corrects engageait les internautes, l'engagement étant l'une des récompenses codées dans le programme (voir [La rem n°38-39, p.67](#)).

Le rapport sur l'éthique de l'apprentissage machine³ a énoncé un certain nombre de préconisations qui sont des points d'attention ou de vigilance pour susciter une réflexion éthique, individuelle et collective. Elles sont articulées autour de plusieurs questions : Quelles sont les données à partir desquelles la machine apprend ? Peut-on s'assurer que la machine effectuera uniquement les tâches pour lesquelles elle a été conçue ? Comment peut-on évaluer un système qui apprend ? Quelles décisions peut-on déléguer à un système apprenant ? Quelle information doit-on donner aux utilisateurs sur les capacités des systèmes apprenants ? Qui est responsable en cas de dysfonctionnement de la machine ?

Est-il pertinent, selon vous, d'étendre l'usage de l'automatisation à de nombreux aspects de la vie économique et sociale ? Du service commercial au véhicule autonome, de la justice au diagnostic médical en passant par les ressources humaines, [l'intelligence artificielle](#) peut-elle avoir réponse à tout ?

Il est pertinent d'étendre l'usage de l'automatisation à de nombreux aspects de la vie économique et sociale. Déjà, dans le domaine du transport, des systèmes obtenus par apprentissage machine sont utilisés pour reconnaître visuellement l'environnement routier des voitures autonomes. La reconnaissance faciale, popularisée par Google Face et Facebook, est aussi utilisée dans un tout autre domaine. Ainsi, en 2016, *DeepMind* a-t-il vaincu au jeu de go l'un des meilleurs joueurs mondiaux.

Il est urgent de renforcer un champ de recherche émergeant autour de l'éthique *by design* afin d'améliorer « l'explicabilité » et la transparence des systèmes d'apprentissage, ainsi que leur adaptation en contexte et l'adéquation de l'apprentissage à ce qu'en attend l'humain. Il ne s'agit plus seulement de construire des modèles par apprentissage machine sans comprendre mais bien d'essayer de les expliquer et de gérer correctement les données collectées.

Quels seraient les biais liés au traitement de données massives (*big data*) pour calculer ou prédire toute activité humaine, y compris nos émotions ?

Les traces que nous laissons lors de nos consultations sur internet et à travers les objets connectés que nous utilisons sont exploitées par des algorithmes d'apprentissage afin de mieux cerner nos préférences de consommation et nos opinions. Par rapport aux simples statistiques, ces algorithmes ont la capacité de fournir des prescriptions individuelles. Lorsque nous naviguons sur internet ou que nous nous servons de gadgets connectés, nous ne pensons pas que notre sillage numérique peut entraîner des algorithmes servant à nous « catégoriser » pour conditionner le montant de nos primes d'assurance. La conformité, la transparence, la loyauté, l'équité des algorithmes d'apprentissage sous-jacents apparaissent ici comme une exigence.

L'adaptabilité à l'environnement que confère la capacité d'apprentissage devrait à l'avenir favoriser l'usage des robots, notamment comme compagnons ou soignants. Construire des robots dits sociaux d'assistance aux personnes nécessite d'encadrer leur usage, surtout lorsqu'ils sont en contact avec des personnes malades ou âgées. Construire des robots humanoïdes sociaux nécessite de respecter les principes éthiques fondamentaux : l'autonomie, la bienveillance, la « non-malfaisance » et la justice. Il doit se créer entre le robot et le patient une relation de confiance.

Pourquoi chercher à créer de l'empathie entre les machines et les hommes ?

Comment allons-nous réagir à l'empathie artificielle ?⁴ De nombreux travaux sont en cours sur la robotique, en psychanalyse par exemple⁵. On sait bien que certaines personnes sont socialement isolées, et il a été depuis longtemps démontré que par besoin d'empathie, elles apprécient les compagnons sociaux non humains, en particulier les chiens⁶. Les assistants humains ne sont pas toujours disponibles et n'apportent pas la même chose. Le robot peut faire office d'assistant surveillant, ce que ne peut pas faire l'animal de compagnie, qui apporte principalement de l'affection. Il faut toutefois considérer les dérives de l'addiction aux machines. La place du virtuel dans notre vie et dans la conscience de nous-même n'est pas si simple. Notre propre identité peut-elle habiter des entités virtuelles ? Comment allons-nous juger cette machine capable « d'apprendre de nous » et de toujours répondre à nos demandes ?

LA PLACE DU VIRTUEL DANS NOTRE VIE ET DANS LA CONSCIENCE DE NOUS-MÊME N'EST PAS
SI SIMPLE

Un robot exprimant de l'empathie peut servir à expliquer, à éduquer, à rassurer les personnes avec lesquelles la machine interagit. Par exemple, je pense à la stimulation cognitive des personnes âgées atteintes de maladies telles que la maladie d'Alzheimer. Un robot peut réagir avec des comportements pseudo-affectifs. Le malade va alors jouer avec le robot comme si c'était un animal de compagnie, ce qui le stimule du point de vue émotionnel et crée un lien social. De même, les robots peuvent aider les enfants autistes qui ont du mal à exprimer leurs émotions ou à comprendre les émotions des autres. En d'autres termes, être émotionnellement engagé avec un robot peut dans certains cas être bénéfique, à la fois pour l'individu et pour la société. Certains envisagent aussi d'utiliser des robots sexuels pour soigner des pathologies sexuelles graves. Des chercheurs néo-zélandais imaginent que des robots peuvent remplacer les prostituées à l'horizon 2050, éliminant ainsi la marchandisation des corps et réduisant les risques de transmission de maladies.

Il est important de comprendre que les robots sociaux auront de plus en plus d'influence. Je parle souvent de coévolution « humain-machine ». Les robots peuvent manipuler les personnes qui conversent avec eux. Ils peuvent également communiquer ce qu'ils ont appris à d'autres - directement ou indirectement par le biais de bases de données partagées. Ils peuvent même détruire de l'information, y compris les opinions humaines. Ils ont la capacité d'utiliser l'internet pour acheter, discuter, mais ils peuvent aussi avoir un impact physique direct sur le monde, le casser ou le créer. Plus important encore, ils peuvent pousser les humains à prendre des décisions appropriées ou inappropriées. Il faut donc se munir de garde-fous. La plupart des chercheurs sont d'accord avec l'idée de rendre l'[intelligence artificielle](#) ainsi que la robotique morales et éthiques.

Pour reprendre l'exemple de la voiture autonome qui permettrait d'éviter 90 % des accidents de la route, comment l'algorithme sera-t-il « nourri » pour résoudre le dilemme entre sauver les occupants du véhicule ou écraser un piéton ? Ce choix *a priori* pose à la société tout entière une question cruciale d'ordre éthique, philosophique même. Comment y répondre ? Et, *in fine*, qui portera la responsabilité du choix effectué par la machine ?

LES THÉORIES ÉTHIQUES TRADITIONNELLES TROUVENT UNE ILLUSTRATION NOUVELLE DANS
LE NUMÉRIQUE ET L'APPRENTISSAGE MACHINE

Chaque accident impliquant un véhicule totalement ou partiellement autonome est abondamment commenté et il est probable qu'un trafic de voitures autonomes se révèle plus sûr qu'un trafic de voitures

conduites par des personnes. Pour l'instant, toutes les voitures autonomes d'un même type sont livrées avec les mêmes paramètres et, une fois en service, elles n'apprennent plus rien mais on peut augurer qu'à l'avenir elles continueront d'apprendre en continu à partir de leur environnement. L'évaluation périodique de leur comportement deviendra alors cruciale. Les théories éthiques traditionnelles trouvent une illustration nouvelle dans le numérique et l'apprentissage machine. Les dilemmes liés aux voitures autonomes comme choisir entre sacrifier son passager ou deux enfants imprudents, ou encore un vieux cycliste respectant le code de la route, pourront être programmés : selon l'éthique de la vertu qui nous vient d'Aristote - ici l'abnégation (sacrifier le passager) ; selon l'éthique déontique du respect du code de la route (sacrifier les enfants) ou selon l'éthique conséquentialiste (sacrifier le cycliste - ici en minimisant le nombre d'années de vie perdues).

Le propos n'est pas ici de traiter de telles questions qui relèvent de la société tout entière, mais d'étudier à l'attention du chercheur, dans le cas de l'apprentissage machine, certaines propriétés particulières que doit satisfaire le comportement d'un dispositif numérique. Les propriétés importantes sont la loyauté, l'équité, la transparence, l'explicabilité, l'intelligibilité, la responsabilité et la conformité.

La loyauté des systèmes informatiques signifie que ces systèmes se comportent, au cours de leur exécution, comme leurs concepteurs l'ont déclaré. L'équité d'un système informatique consiste en un traitement juste et équitable des usagers. La transparence d'un système signifie que son fonctionnement n'est pas opaque, qu'il est possible pour un utilisateur, de vérifier son comportement. Cette transparence s'appuie notamment sur la traçabilité, la mise à disposition d'informations suffisamment détaillées pour qu'il soit possible, après coup, de suivre ses actions. La traçabilité est essentielle, d'une part, pour les attributions de responsabilité car elle permet le cas échéant de fonder un recours juridique et, d'autre part, pour diagnostiquer et corriger les dysfonctionnements. Elle permet également d'expliquer le fonctionnement d'un système à partir des traces laissées par celui-ci : on parle alors d'explicabilité ou d'intelligibilité.

LES PROPRIÉTÉS IMPORTANTES SONT LA LOYAUTÉ, L'ÉQUITÉ, LA TRANSPARENCE,
L'EXPLICABILITÉ, L'INTELLIGIBILITÉ, LA RESPONSABILITÉ ET LA CONFORMITÉ

La possibilité d'identifier des responsabilités lors d'un dysfonctionnement implique de distinguer deux agents : le concepteur et l'utilisateur du système. Le donneur d'ordre ou le concepteur est responsable si le système est mal conçu ; l'utilisateur est responsable s'il a mal utilisé le système tout en sachant que le professionnel est tenu à un devoir d'informations supplémentaires vis-à-vis de tout usager. Un système numérique doit demeurer conforme à son cahier des charges, et ce cahier des charges doit être conforme à la législation. La conformité d'un système à son cahier des charges signifie que le système est conçu pour effectuer des tâches spécifiées en respectant les contraintes expliquées dans ce cahier. Par exemple,

la conformité inclut le respect des règles de protection des données personnelles pour un système d'analyse de données ou le respect du code de la route pour la voiture autonome.

Pour conclure...

IL N'EXISTE PAS POUR L'INSTANT DE MACHINES MORALES

Il n'existe pas pour l'instant de machines morales mais les robots « incarnés » arrivent tout juste dans notre vie quotidienne. Même si nous n'en voyons pas encore beaucoup, il est nécessaire de s'en occuper dès maintenant. Ils sont déjà actuellement réceptionnistes ou accompagnent les clients dans les boutiques afin d'enquêter sur leur degré de satisfaction ; ils sont là pour divertir les enfants ou encore pour apporter des renseignements sur les cartes de fidélité⁷ d'un magasin. Ils seront bientôt robots assistants à l'hôpital, à l'école... La fonction des robots assistants de santé sera avant tout de transmettre aux médecins nos données de santé et de comportement. À l'école, ils pourront servir de répétiteurs pour les enfants en difficulté. Pour la plupart, ils auront des caractéristiques humanoïdes. En Occident, nous pensons qu'il est souhaitable que la frontière soit claire entre le vivant et l'inanimé, ce qui n'est pas forcément le cas dans toutes les cultures. Les Japonais sont beaucoup plus détachés que nous vis-à-vis des robots : pour beaucoup ce sont des gadgets très amusants et à la mode. Au Japon, les géminoïdes, jumeaux d'êtres vivants, ne posent pas vraiment de problème philosophique. De la même façon, se développe au Japon le marché des robots sexuels, sorte de poupées qui seront bientôt douées de réactions.

Il est donc nécessaire d'apprendre à utiliser ces nouveaux outils de simulation sans en avoir peur, de comprendre leur utilité ; mais aussi de mettre en place une surveillance de ces outils et de notre relation avec eux.

Sources :

2. *Affective Computing*, Rosalind W. Picard, MIT Press, 1997.
4. La Cerna a été créée en 2012, au sein d'Allistene, Alliance des sciences et technologies du numérique fondée par la CDEFI, le CEA, le CNRS, la CPU, l'Inria et l'Institut Mines-Télécom en 2009.
6. Éthique de la recherche en apprentissage machine, avis de la Cerna, cerna-ethics-allistene.org, juin 2017.
8. *Le jour où mon robot m'aimera, vers l'empathie artificielle*, Serge Tisseron, Albin Michel, 2015.
10. *Le numérique et la robotique en psychanalyse*, Frédéric Tordo, L'Harmattan, 2016.
12. « Stressful life events and use of physician services among the elderly : The moderating role of pet ownership », J. M. Siegel, *Journal of Personality and Social Psychology*, 1990.
14. « Pepper, le robot humanoïde qui accompagne les clients de Carrefour », Maxime Brigand, LeFigaro.fr, 21 octobre 2015.