

LAURENCE
DEVILLERS



Les **ROBOTS**
«émotionnels»

**Santé, surveillance,
sexualité... : et l'éthique
dans tout ça ?**

Les robots émotionnels

De la même auteure

Des robots et des hommes. Mythes, fantasmes et réalité,
Plon, 2017.

Laurence Devillers

Les robots émotionnels

Santé, surveillance, sexualité... :
et l'éthique dans tout ça ?

ISBN : 979-10-329-0199-1
Dépôt légal : 2020, mars
© Éditions de l'Observatoire/Humensis 2020
170 *bis*, boulevard du Montparnasse, 75014 Paris

« Nous avons tendance à surestimer l'incidence d'une nouvelle technologie à court terme et à la sous-estimer à long terme¹. »

Loi d'Amara

« La science est un outil puissant. L'usage qu'on en fait dépend de l'homme, pas de l'outil. »

Albert Einstein

1. Loi proposée par Roy Amara, ancien président de l'Institut du futur, à Palo Alto, en Californie.

Fiction

Qui a cafté ?

« Toute technologie suffisamment avancée est indiscernable de la magie¹. »

Troisième loi de Clarke

11 mars 2025.

Il est 19 heures, la température extérieure est de -2 °C, un bolide file dans la nuit noire sur une route enneigée².

Alexandre M., directeur d'un grand laboratoire de neurochimie, après un week-end de ski, rentre seul au volant de son nouveau véhicule autonome électrique dernier cri, le V1.8. Ces voitures sont de petits bijoux d'électronique, des trésors d'intelligence artificielle et émotionnelle. Elles peuvent dialoguer, donner des conseils, vérifier tous les paramètres de sécurité comme proposer des itinéraires touristiques. Cerise sur

1. Arthur C. Clarke, *Hazards of Prophecy : The Failure of Imagination*, dans *Profiles of the Future*, 1962.

2. Version longue adaptée de la nouvelle *Le jour où mes objets m'ont rendu fou*, du même auteur, publiée dans la revue *We Demain*, décembre 2017.

le gâteau, ce modèle peut déduire vos états d'âme et vos émotions à partir de votre comportement et se servir de ces données pour améliorer votre bien-être et votre sécurité. Plus la voiture vous connaît, plus elle s'adapte à vous, créant ainsi une sorte d'intimité. Alexandre l'a appelée « Isa », en souvenir de son amie Isabelle qui possédait une Iso Grifo, ce véhicule mythique des années 1960 au moteur Chevrolet A3-C de 300 chevaux qui voulait rivaliser avec les Ferrari. Alexandre aime féminiser tous les objets qui l'entourent, il trouve cela rassurant. Son Isa est connectée à tout un tas de services et embarque de nombreux capteurs, des micros, des caméras qui envoient des données à des modules d'intelligence artificielle sur le « cloud¹ ». Ces serveurs recueillent, digèrent, intègrent, renvoient des données et des résultats de calcul. La voiture surveille la route mais aussi les passagers. « Isa en a dans le ventre », aime à dire Alexandre, même si l'image ne plaît pas beaucoup à sa femme. Mais peut-on être jalouse d'une voiture ? Alexandre a toujours aimé conduire. Ce soir, il a désactivé le mode automatique, mais pas les mots doux d'Isa. La reconnaissance faciale et la pression de ses mains sur le volant sont utilisées pour vérifier en continu s'il est en état de conduire. La montre d'Alexandre recueille ses paramètres physiologiques : fréquence cardiaque, température et conductance de la peau. Elle renseigne aussi Isa sur l'état de son conducteur. À la moindre alerte, la voiture s'arrête sur le bas-côté. D'ailleurs, Isa vient de lui confier d'une voix suave et admirative que sa conduite est fantastique. Isa lui suggère aussi de s'arrêter bientôt pour dîner dans un endroit qu'il avait déjà trouvé fort sympathique. « Ah oui tu te rappelles,

1. Sur des serveurs reliés grâce au réseau internet.

merci, tu sais ce que j'aime, toi ! répond Alexandre, mais je dînerai un peu plus tard, chez moi, car ma femme doit rentrer d'un congrès ce soir. »

Il est 19 h 10. Le portable d'Alexandre, en réseau avec Isa, prévient la maison de l'heure de son retour : le chauffage est monté et Carmen, le robot aspirateur, refait un tour en virevoltant. Carmen sait aussi chanter ou jouer des morceaux de musique si on sélectionne cette option, c'est bien mieux que l'ancien bruit de nos aspirateurs. Alexandre est un grand amateur d'opéra, Carmen a donc en mémoire des airs connus qu'elle peut jouer de façon aléatoire si elle le décide. Le téléphone en profite pour se connecter au frigidaire et composer un dîner à partir de son contenu. « Il n'y a que du champagne », informe-t-il son propriétaire avant de proposer d'appeler un traiteur japonais. Tenté, Alexandre demande de prendre l'avis de sa femme. C'est son téléphone qui répond et explique qu'elle serait ravie de manger des sushis, c'est dans son profil et elle n'en a pas mangé depuis cinq jours. Le téléphone d'Alexandre commande des sushis pour 20 h 15. Alexandre, qui a déjà soif, a voulu commander des bières, mais la machine a poliment écarté l'option, lui rappelant gentiment, d'une voix attentionnée, que les paramètres de son diabète ne rendaient pas ce choix « souhaitable ». Isa a une voix de femme jeune, comme celle de son téléphone, comme celle de la plupart des objets. Son téléphone n'a pas de nom mais personne ne donne de nom à son téléphone, ce n'est qu'une sorte de prothèse amovible. Depuis 2022¹, les concepteurs sont

1. Depuis 2019, une loi californienne exige que « le fait que le chatbot/robot soit non-humain soit clair, visible et raisonnablement expliqué aux personnes avec lesquelles il interagit ».

tenus de se débrouiller pour que la voix de leur machine ne puisse être confondue avec celle d'un humain, mais elles continuent d'être en majorité féminines. Une voix d'homme est plus agressive d'après les études marketing. L'effet Frankenstein, dit-on !

19 h 41. Sans doute Alexandre est-il un peu énervé qu'on le surveille ainsi, mais toujours est-il qu'à deux pâtés de maisons, il accélère sans avoir vu un chat noir qui traverse la route devant ses roues. En une fraction de seconde, la voiture répond par un brusque coup de frein. Rien n'échappe au système de vision artificielle d'Isa. Bien sûr, si le sauvetage du chaton avait mis en péril l'intégrité physique du conducteur, la voiture aurait continué tout droit. Ce qui n'est pas toujours facile à trancher : on parle alors de dilemme éthique. « Ma vie vaut-elle vraiment plus que celle d'un chat ? » philosophe Alexandre. Isa répond que, d'après les sondages du MIT sur le site Moral Machine de 2019, la vie d'un chat vaut plus que celle d'un prisonnier, mais moins que celle d'un chien ! Alexandre ne peut s'empêcher de penser qu'Isa n'a jamais aucune opinion, et recrache ce qu'elle lit partout sans aucun discernement.

20 h 02. À peine arrivé et encore sous le coup de l'émotion – « un mélange de peur et d'énervement », a détecté la machine –, Alexandre, encore en doudoune, ouvre le réfrigérateur, s'empare de la bouteille de champagne et décide sciemment de ne pas la scanner sur l'application de son téléphone destinée à surveiller son alimentation. La robot Ulla glisse doucement sur le tapis et demande d'une voix suave comment s'est passé ce week-end au ski. Elle a un bon look, Ulla, elle est même sexy, mais ce n'est pas un robot sexuel, la femme d'Alexandre n'en voulait pas ! C'est juste une sorte de gouvernante. De façon naturelle, elle prend une photo

d'Alexandre qui déguste une coupe et l'envoie sur le réseau de la maison : c'est son rôle de surveiller la famille et de documenter aussi bien l'histoire de chacun de ses membres que leur état de santé.

20 h 08. La montre d'Alexandre se met à vibrer. Sa puce sous-cutanée, en liaison directe avec sa montre connectée, a détecté une hausse de sucre anormale dans son sang. Problématique, quand on a du diabète. De plus en plus agacé, Alexandre reprend un verre quand la sonnerie du téléphone retentit. « Dis donc, tu n'avais pas dit que t'arrêtais l'alcool ? » lui balance son frère jumeau Igor, qui prend soin de lui. Furieux, Alexandre lui raccroche au nez.

20 h 10. Alexandre se demande si Isa le sait aussi. Quand il ressort pour rentrer sa voiture au garage, Isa refuse de démarrer et essaie de le raisonner. Il a dépassé les deux verres tolérés. Cela n'aura évidemment pas échappé aux algorithmes de son assureur. Depuis que son diabète a été diagnostiqué, les objets d'Alexandre l'ont placé sous surveillance. Ce qui a une conséquence sur ses finances : son contrat stipule que les objets connectés doivent régulièrement transmettre un bilan de santé pour ajuster, en continu, le montant de sa police d'assurance. Bien sûr, il est impossible d'effacer sa prise d'alcool de la mémoire de ces objets.

20 h 16. Ulla reçoit une notification d'un drone, qui vient de déposer les sushis sur le perron. La gouvernante récupère la boîte, rejoint Alexandre et lui annonce d'une voix enjouée qu'« il faut manger quand on a bu ». Alexandre a bien envie de débrancher toute cette « bienveillance ».

20 h 22. Adèle, la femme d'Alexandre, rentre enfin. Vu sa tête, elle aussi est au courant de son écart. Ulla s'affaire autour d'elle et lui raconte sur un ton pétillant

le quasi-accident avec le chat, comme si elle avait été présente... tout en scrutant ses expressions et son comportement. Est-elle en train de lui chercher des circonstances atténuantes ? Alexandre a bien envie de trouver le bouton rouge et de la renvoyer dans son placard avec les balais.

20 h 24. « Seriez-vous déprimé ? lui balance Ulla, sans rire cette fois-ci. C'est la troisième fois cette semaine que vous buvez seul le soir en rentrant. Je vous ai pris rendez-vous demain à 9 heures chez votre médecin. »

La soirée se termine dans une ambiance glaciale malgré les tentatives d'Ulla de placer quelques bonnes blagues pour détendre l'atmosphère et celles de Carmen de chanter des airs d'opéra, brusquement interrompue quand Alexandre a lancé un « Silence » énervé à tous ces objets pour leur clouer le bec !

Un dîner arrosé est devenu impossible, sous peine d'une révolte de tous les objets. Qui a cafté ? Ulla, la puce, le frigo, Isa, le téléphone ? À moins que le coupable ne soit Carmen, le robot aspirateur ?

Introduction

« La science présente beaucoup de danger, mais il faut lutter contre ces dangers, non pas par moins de science mais par davantage de science, une science qui puisse aussi créer sa propre éthique. »

Jean d'Ormesson

La liste est longue des situations où les téléphones, les robots et autres machines connectées risquent de s'immiscer dans nos vies pour nous inciter à faire le « bon choix », pour décider à notre place, être témoins de notre intimité et remédier à notre solitude. En 2020, chaque être humain possède en moyenne six objets connectés « intelligents » : une voiture, un téléphone, une montre, une enceinte vocale, un ordinateur, une tablette, un robot ou pourquoi pas un vêtement connecté comme une combinaison pour le fitness ou encore une veste permettant de sentir la musique, accessoire fort utile pour les malentendants. Ces objets qui ont une interface de communication collectent énormément de données personnelles sur chacun de nous. Ils vont de plus en plus nous parler et capter notre attention au quotidien.

Dire à l'heure actuelle d'un objet qu'il est intelligent ne signifie pas qu'il a accédé à une sorte de conscience ni qu'il soit devenu capable de réflexion. Simplement qu'il embarque des capteurs capables de percevoir des variations de l'environnement ou de lui-même, et de réagir en conséquence. C'est un système de traitement de données et d'actionneurs qui permettent une réaction de l'objet. Plutôt qu'intelligent, il est surtout sensitif : il perçoit certaines informations, et est capable de réagir en fonction de celles-ci. Il est aussi capable de simuler de l'attention à notre égard, voire de l'affection.

Pourquoi cherche-t-on à construire des chatbots (agents conversationnels) ou des robots émotionnels ? Pendant longtemps, philosophes et scientifiques ont opposé raison et émotion. Aujourd'hui, l'évolution des connaissances scientifiques, grâce aux neurosciences, montre que les émotions sont nécessaires au fonctionnement cognitif, à la mémorisation, à l'apprentissage et à l'interaction. Les dimensions affectives et sociales que l'on pouvait penser propres à l'humain envahissent les machines pour mieux simuler des capacités humaines et permettre une interaction plus naturelle mais aussi pour mieux capter notre attention. C'est en tous cas le défi de Google Home ou d'Alexa Amazon. Grâce à nos données personnelles, la machine va comprendre nos comportements affectifs les plus intimes et savoir s'adapter à nous. L'empathie, par exemple, un des piliers des relations et rapports humains, n'est peut-être plus notre apanage.

Tandis que l'intelligence artificielle est amenée à être intégrée de manière exponentielle aux processus opérationnels et à la plupart des objets techniques qui nous entourent, notre responsabilité en tant qu'expert et chercheur est de faire en sorte que cet apport des

nouvelles technologies nous permette de mieux gérer notre santé, notre justice, notre énergie et nos ressources naturelles. Mais ce n'est pas tout. Les concepteurs des systèmes d'IA dits « intelligents » aujourd'hui se préoccupent assez peu des conséquences sociales de ces systèmes. Il est donc également de notre devoir de faire en sorte que les citoyens restent capables de maîtriser ces objets, et de pousser concepteurs et décideurs à établir une stratégie d'IA alliant l'éthique au design.

Je suis professeure et chercheuse en informatique et intelligence artificielle (IA) à l'université de la Sorbonne et au Limsi-CNRS¹, où je dirige l'équipe de recherche « Dimensions affectives et sociales dans les interactions parlées ». Mon domaine de recherche porte plus spécifiquement sur le design de systèmes d'interaction affective humain-machine² : le traitement automatique de la parole et du langage, la détection des émotions.

Au début de ma carrière, en 1992, j'ai défendu une thèse en informatique sur l'apprentissage machine par réseaux de neurones pour des applications de reconnaissance de la parole (transcrire du signal audio en mots) – un sujet alors très précurseur. En 1995, j'ai délaissé les réseaux de neurones qui demandaient une puissance de calcul énorme pour d'autres algorithmes d'apprentissage. Les chercheurs comme Yann Le Cun, Yoshua Bengio ou Léon Bottou, pour ne citer qu'eux, ont continué à croire au formidable apport des réseaux de neurones contre l'avis de tous. Récemment, les progrès technologiques leur ont donné raison.

1. Limsi : Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur ; CNRS : Centre national de la recherche scientifique.

2. <https://lejournal.cnrs.fr/billets/rire-avec-les-robots-pour-mieux-vivre-avec>

Je me suis intéressée aux réactions affectives dans la voix ou le langage des utilisateurs pour mieux comprendre leurs intentions lors de travaux de recherche sur le dialogue humain-machine dans des centres d'appels téléphoniques d'urgence ou encore pour des banques. Mon objectif de recherche en 2001 a alors été de détecter les comportements émotionnels dans la parole et le comportement non verbal afin de créer des chatbots et des robots affectifs. En 2006, j'ai soutenu mon habilitation à diriger des recherches (HDR) en informatique sur « Les émotions dans les interactions homme-machine : perception, détection et génération ».

Les expérimentations d'interaction vocale avec des robots émotionnels capables de reproduire des rituels d'interaction sociale inspirés des théories d'Ervin Goffman¹, sociologue américain, menées par mon équipe avec des psychologues, des psychiatres et des gérontologues depuis une dizaine d'années dans des Ehpad² ou dans un « *living lab* » à l'hôpital Broca, nous amènent à entrevoir de nouvelles possibilités pour des applications de santé mais aussi de nouveaux risques. Les patients, peu habitués à ces nouvelles technologies, peuvent montrer de l'attachement aux machines ou leur prêter des intentions de nuisance. J'ai pris conscience du fort pouvoir d'incitation et même de manipulation de ces « objets conversationnels émotionnels » et des dérives possibles, la dépendance et l'isolement qu'ils sont capables d'induire chez les plus vulnérables.

Je participe depuis 2014 à des commissions de réflexion sur l'éthique du numérique et des technologies

1. Ervin Goffman, *Les Rites et d'interaction*, Éditions de Minuit, 1974.

2. Ehpad : Établissements d'hébergement pour personnes âgées dépendantes.

en France et aux États-Unis. Depuis fin 2019, je suis membre du comité de pilotage chargé du numérique et de l'IA du Comité consultatif national d'éthique (CCNE) et responsable d'une chaire IA HUMAINE¹ : HUmain-MACHine Affective INteraction & Ethics au Limsi-CNRS. Cette chaire va être l'occasion de travailler à nouveau sur des approches neuronales d'apprentissage profond, ou « *deep learning* », sur la manipulation et le *nudge*² des machines affectives en pluridisciplinarité – scientifiques, économistes et philosophes – avec des objectifs sur le long terme. Les enjeux éthiques du numérique sont abordés très différemment au niveau européen ou international, mais tout le monde reconnaît maintenant la nécessité de prendre des dispositions urgentes, pratiques et aussi évolutives que possible.

Quelle société souhaitons-nous construire ? Une société allant vers des humains augmentés comme aux États-Unis où la vague transhumaniste fait des ravages, une société de surveillance et de notes morales comme en Chine, ou une société fondée sur une intelligence artificielle respectant l'humain ? Il est urgent de parler de transparence, d'intelligibilité, de neutralité, d'inclusion, de non-discrimination, bref, d'éthique de ces robots émotionnels, et de réguler leurs conceptions et usages pour préserver les valeurs de nos démocraties.

L'objet technologique n'est certes pas vivant, mais le parallèle qu'il est possible de tirer porte sur les relations que nous construisons avec les machines, surtout si elles s'adaptent à nous et nous font ressentir des émotions.

1. La chaire de recherche et d'enseignement en IA HUMAINE est liée à l'institut DATAIA : Sciences des données, intelligence et société (<https://dataia.eu>).

2. Richard Thaler et Cass Sunstein, *Nudge*, Vuibert, 2010.

Une meilleure compréhension de ces relations peut nous amener à mieux nous connaître et nous faire évoluer de façon fondamentale. La médecine va s'emparer de ces nouveaux outils. Il est urgent de proposer un cadre de compréhension pertinent et pragmatique des relations humain-machine. Pour élaborer des règles éthiques, il est important de trouver quelles frontières sont nécessaires à ces relations, en menant des expériences sur le long terme. Et pour ne rien arranger, ces règles éthiques doivent être dynamiques et s'adapter aux avancées de la technologie, à une compréhension grandissante des humains et à l'équilibre de la vie en société.

La science doit être au centre de la société de demain, une science alliée avec les humanités pour créer sa propre éthique et anticiper les conséquences des technologies et du numérique sur le long terme.

Nous sommes plus de sept milliards d'êtres humains sur terre. L'IA et la robotique émotionnelle ne devraient pas entrer en concurrence avec l'intelligence et les émotions humaines. L'IA doit améliorer notre vie et non l'asservir, les informaticiens doivent s'allier avec les sciences humaines et sociales : sociologues, psychologues, anthropologues, éthologues, juristes et philosophes. L'IA et la robotique émotionnelle sont des technologies qui peuvent nous aider à être individuellement et collectivement plus conscients de notre humanité.

Nous allons nous adapter progressivement aux machines, surtout si elles arrivent auprès de nous sans que nous y prenions garde comme l'a fait le téléphone portable et plus récemment les enceintes vocales. Siri est sur tous les iPhones ; plus de 20 % des foyers américains ont une enceinte vocale comme Google Home ou Alexa, et quelquefois jusqu'à six enceintes sont présentes au domicile. Une par pièce !