

# REPORT ON AI AND SOCIETY

*Marion Pavaux*

*07/06/2021*

## SOMMAIRE:

- I. Introduction
- II. Impacts sociétaux de l'IA
  - a. Dans la santé
  - b. Dans la sécurité routière
  - c. Dans l'expérience client
  - d. Dans la publicité et le marketing
- III. Mon point de vue sur les problèmes d'éthique que peuvent poser cette technologie
  - a. Définition des valeurs et principes
  - b. Attaque ou support de ces valeurs et principes
- IV. Références

## I. Introduction

Depuis le congrès en 1956 qui introduisit l'intelligence artificielle, celle-ci s'est développée, et connaît une popularité grandissante, chez les chercheurs, les entreprises, et le public. Beaucoup de mythes entourent l'IA, jusqu'à penser que celle-ci pourrait développer une conscience humaine. Les producteurs inspirés créent des films où les robots remplacent les humains, ou développent des émotions. Parmi eux on peut citer les films *Ex-Machina*, *Intelligence Artificielle*, et la série *West World*.

De nouvelles avancées risquent encore d'alimenter les débats, notamment les technologies utilisant la reconnaissance des émotions, et l'IA conversationnelle. Il s'agit de technologies ayant la capacité de reconnaître les expressions faciales, l'humeur, et les émotions de leurs utilisateurs par la vue, ou par écoute comme la technologie développée par l'entreprise *Vokaturi*. Un second volet est de répondre par les bonnes émotions à un environnement.

Bien sûr, en combinant ces deux volets, on peut créer des chats bots, ou robots interactifs, reconnaissant les émotions de leurs interlocuteurs, les prenant en compte dans leurs plans d'actions, et créant eux-mêmes des émotions artificielles en réponse. Cette émotion peut être générée à l'aide d'un visage humanoïde ou de smileys par exemple. Leur champ d'application est très varié.

Par exemple, ces robots peuvent être utilisés dans les hôpitaux, pour aider les patients, en particulier les personnes âgées, en les accompagnant au quotidien grâce à la connaissance de leur état de bien être, humeur, et santé mentale. Lors d'enquêtes policières, une technologie de reconnaissance des émotions pourrait faire office de détecteur de mensonge par exemple. Enfin, elle peut permettre de statuer nos préférences au regard d'un contenu ; image, film, vidéo, œuvre d'art et bien sûr à propos d'une publicité ou d'un produit.

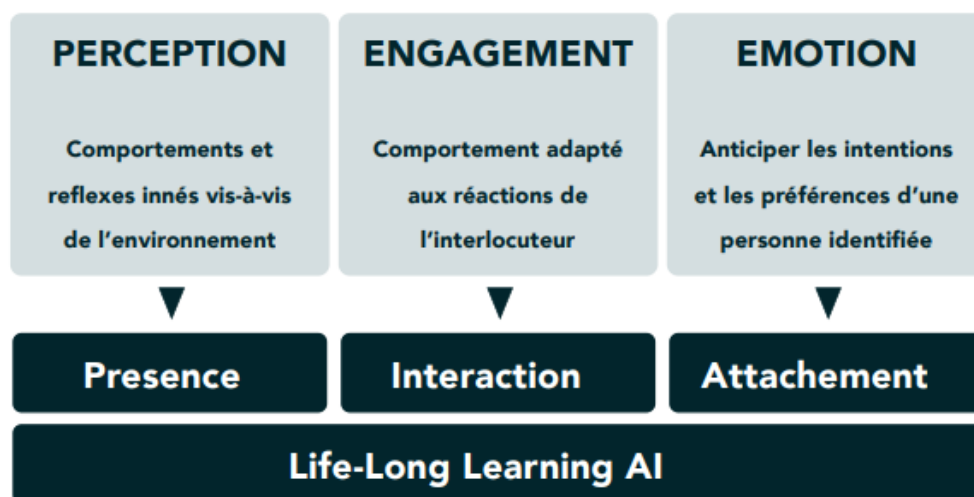


Figure 1: Schéma provenant de la société Hoomano<sup>IV</sup>

Cette technologie d'IA se base sur des connaissances de son environnement, et la façon dont il doit agir dessus, puis apprend comment interagir avec les humains, en fonction de ses erreurs de buts et d'émotion perçue chez l'autre. Elle devient ensuite capable de généraliser des lois sociales et psychologiques, et de spécialiser son apprentissage en fonction des personnes.

La technologie peut donc entre autres se servir de plusieurs réseaux de neurones, par exemple un dans la reconnaissance des émotions, bien que cette partie soit bien plus complexe que de la simple reconnaissance faciale. Un autre réseau de neurone peut servir à déterminer l'émotion du robot à afficher pour créer une

certaine réaction chez l'interlocuteur. Celui-ci utiliserait une fonction coût d'émotions, traduisant l'écart entre l'émotion ou la réaction souhaitée chez l'interlocuteur et l'émotion ou la réaction perçue par le robot.

## II. Impacts sociétaux de l'IA

### a. Dans la santé

Avec une population âgée grandissante, et une population active en baisse, il deviendra de plus en plus difficile d'accorder à de nombreux patients, et de moindre effectifs, le temps nécessaire à leur bon rétablissement. Surtout les places dans les hôpitaux seront peut-être bien plus limitées qu'elles ne le sont déjà aujourd'hui.

Une intelligence émotionnelle pourrait permettre d'accompagner la personne âgée à la maison et à l'hôpital dans ses tâches quotidiennes et la prise de médicaments par exemple, avec tact, et fournir au médecin un condensé des informations cruciales sur le bien-être de la personne. Le fait d'interagir avec des émotions rend la technologie plus sympathique, et plus facilement intégrable dans le quotidien de personnes âgées, qui sont effectivement moins familières avec elle, et ont besoin d'un contact humain plus important.

Dans le cas d'études cliniques sur les maladies mentales et des troubles neurologiques, une technologie de reconnaissance des émotions et des expressions pourrait permettre d'améliorer grandement la compréhension des mécanismes de ces maladies. Avec cette compréhension de nouveaux traitements pourraient apparaître, donc l'efficacité pourrait continuer d'être suivie par la technologie.

L'impact sociétal de ces technologies dans le domaine de la santé serait par conséquent très positif. Malheureusement celles-ci ne sont encore qu'à l'étape de recherche, peu d'entreprises s'intéressent réellement à l'apport de cette technologie d'IA en santé.

### b. Dans la sécurité routière

Cette technologie d'IA est déjà utilisée dans la sécurité des personnes, notamment la sécurité routière. L'entreprise *Affectiva* a donc développé une technologie pour les voitures, qui détecte le niveau de fatigue de ses conducteurs, et prévient ainsi l'endormissement au volant et les accidents. Plus que pour la sécurité, la technologie est également capable grâce à la détection de ses expressions faciales, de personnaliser l'expérience du conducteur au sein du véhicule.

### c. Dans l'expérience client

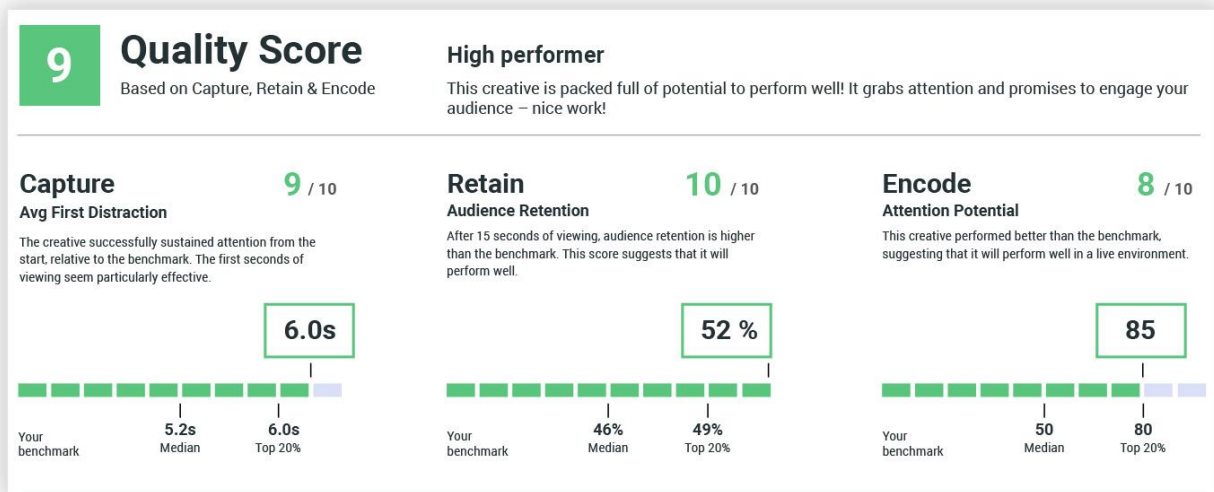
En parlant d'expérience client, la reconnaissance des émotions peut permettre de proposer à ces utilisateurs du contenu personnalisé, comme des musiques ou des films, correspondant à l'humeur du moment. Netflix utilise déjà des miniatures différenciées des utilisateurs pour un même film ou une même série. Ici, c'est l'humeur qui déterminerait la miniature, et proposerait le contenu à regarder par exemple.

De même, lorsque nous sommes sur les réseaux sociaux, un algorithme étudie en même temps le temps que nous passons à lire chaque publication, ce que nous aimons, partageons, suivons, et nos données personnelles pour nous proposer du contenu toujours plus pertinent, qui va nous faire rester sur la plateforme. Si cet algorithme pouvait également prendre en compte les expressions de notre visage, il serait à même de collecter des informations sur notre humeur. Il serait également à même de détecter plus facilement si un contenu que nous aimons est corrélé à une émotion particulière, comme la peur, la joie, ou la surprise par exemple. Cela lui permettrait de proposer du contenu correspondant à notre humeur du moment, et lié à des émotions que nous souhaitons ressentir. Il pourrait s'en servir pour nous faire changer d'émotion de sorte que l'émotion corresponde mieux à un temps élevé passé sur la plateforme de contenu.

## d. Dans la publicité et le marketing

Ces mêmes réseaux sociaux pourraient transmettre et vendre ces données sur ce que nous aimons et nos émotions aux annonceurs, pour qu'ils soient d'autant plus à même de toucher le client. Ainsi ces annonceurs ne se baseraient plus seulement sur nos centres d'intérêt, à la manière des cookies, mais à une version beaucoup plus actuelle de nous-même.

Des agences de publicité peuvent également, à la manière des études de santé, mener des études sur les réactions d'une audience à une annonce, vidéo, etc... et ainsi les adapter de manière à toucher leur audience le plus possible. Beaucoup d'entreprises se sont intéressées à l'application de l'IA dans ce domaine.



## III. Mon point de vue sur les problèmes d'éthique que peuvent poser cette technologie

### a. Définition des valeurs et principes

Puisque des données seront enregistrées, les utilisateurs devraient avoir la capacité d'exprimer clairement leur accord avant d'être analysés par une technologie de ce type. Les données récoltées devraient également appartenir aux utilisateurs avec la possibilité d'en contrôler l'accès, leur utilisation, et de les supprimer définitivement. Les données devraient être anonymisées pour éviter de faire une correspondance entre les données et l'identité d'une personne. Ces mesures visent à préserver la vie privée de la personne et ses intérêts.

Aujourd'hui, la consommation énergétique est un aspect de protection environnementale majeur. Les demandes en ressources énergétiques qu'a une technologie d'IA, ainsi que les serveurs sur lesquels sont stockés les données des utilisateurs, peuvent être conséquentes. Cette technologie se doit de limiter ses consommations, voire d'être plus bénéfique que son empreinte environnementale.

Pour certaines applications, il peut être important que l'IA s'explique, c'est-à-dire qu'elle donne les éléments de raisonnement pertinents qui l'ont amené à une analyse, action ou décision. Cela peut être important dans le médical où l'IA remplace des tâches de spécialistes, et où cela peut mettre en danger la vie des patients si de mauvaises conclusions étaient tirées.

Dans cette même optique, il faut avoir une base de données d'entraînement conséquente et de bonne qualité, évitant autant que possible les biais pour permettre les conclusions les plus exactes et objectives possibles.

De manière générale, la technologie doit impacter la société de manière positive, c'est-à-dire être ne pas être néfaste pour une partie de la société (en prenant en compte aussi son environnement), ou si elle peut l'être, être bénéfique au plus grand nombre de personnes sans discrimination. Elle doit évidemment avoir des applications légales.

Dans le cas où l'IA est aussi conversationnelle se pose la question de l'éthique de sa parole, puisque, celle-ci peut alors mentir, rendre une personne heureuse ou triste. Il serait très délicat d'interdire à une technologie d'IA de mentir, ou même d'avoir pour but de rendre quelqu'un heureux à tout prix, même en autorisant le mensonge. Cela se fait au cas par cas selon le domaine d'application de la technologie.

## b. Attaque ou support de ces valeurs et principes

Dans le domaine de la santé, nous avons vu que cette technologie pouvait avoir des conséquences très positives sur la société. Cependant, dans le cas d'une IA conversationnelle, par exemple avec comme interlocuteur une personne âgée, on peut envisager des dilemmes moraux à certaines demandes.

Si la personne prise en charge demande si elle va mourir bientôt, ou de l'assister au suicide en ne l'en empêchant pas, ou encore de mentir à son médecin sur sa prise de médicament ou habitudes alimentaires. Il est très difficile pour un humain d'envisager quelle est la bonne manière de répondre à ces questions. Une personne a le choix de mentir à son médecin et une technologie ne peut pas non plus se transformer en surveillance forcée, mais qu'en est-il de la responsabilité de la technologie si de mauvaises conséquences venaient à advenir ? Même si on pourrait dire qu'il s'agit de la responsabilité de la personne elle-même, il reste certaines situations à envisager. Peut-être la personne a moins d'informations sur les conséquences qu'une décision pourrait avoir que l'IA. Peut-être cette personne n'a plus les capacités cognitives pour être indépendante et prendre de bonnes décisions.

C'est pour cela que la technologie doit bien cibler avec quel interlocuteur elle interagit pour prendre les meilleures décisions.

Dans le cadre de la psychologie, et de la sociologie, on pourrait imaginer que des conclusions surprenantes sont tirées. En effet, des expériences, comme celle de la prison de Stanford par Philip Zimbardo en 1971 nous ont déjà montré que les humains peuvent avoir des comportements tyranniques lorsqu'ils en ont le pouvoir. Est-ce que ces conclusions, pourraient être trop extrapolées, et expliquées par l'IA, et ainsi justifier certains actes.

En sécurité routière, l'IA ne pourrait peut-être pas satisfaire une expérience utilisateur des plus complètes car cela pourrait ensuite être dangereux pour la conduite. Également, à quel point une IA peut empêcher un conducteur de conduire s'il montre des signes de fatigue ?

En ce qui concerne l'expérience client, nous avons déjà un avertissement de la dangerosité des données que nous possédons en termes de manipulation. Cependant, comme cette technologie va encore plus loin dans notre analyse, en particulier celle de nos émotions, son utilisation peut poser de véritables problèmes d'éthique. En effet, l'IA va vouloir nous proposer du contenu adapté à notre humeur. Si on est triste elle peut nous proposer du contenu associé à la tristesse, ou au contraire du contenu associé au bonheur. Dans ce cas si elle propose des contenus associés à la tristesse cela peut être néfaste pour la personne, mais d'un autre côté elle ne sera peut-être pas intéressée par des contenus à caractère joyeux ou l'inverse. Si l'IA a de mauvaises données elle pourrait être amenée à faire de mauvaises conclusions sur notre état, jusqu'à encourager des biais néfastes déjà présents dans la société à force de répétition.

A la même manière des technologies actuelles qui nous enferme dans une bulle en ne nous montrant que des contenus qui pourraient nous intéresser, cette technologie pourrait nous enfermer dans une bulle d'émotion

choisie. Cela revient à de la manipulation sur les émotions. Pire, la technologie pourrait nous enfermer dans une émotion ou une humeur choisie propice à l'achat. Cela peut-être en nous rendant tristes, en amoindrissant notre confiance en nous pour certains produits, ou en nous donnant une humeur aventureuse pour d'autres, en suivant la volonté des annonceurs. Cette manipulation pourrait à un certain point être préjudiciable au niveau juridique.

Sur une plateforme comme YouTube par exemple, l'IA pourrait choisir les annonces qui correspondent à notre humeur en même temps que de nous proposer du contenu pour essayer de la modifier. Également les annonceurs en utilisant cette technologie, augmentent les chances d'achat chez leur cible. Ces deux aspects de la publicité, création et diffusion qui travaillent ensemble avec cette technologie pourraient augmenter significativement les ventes d'un produit. Les personnes pourraient acheter des produits qui ne leur sont pas utiles ou ne leur correspondent pas, et qui seront jetés rapidement. Finalement, cette consommation excessive nuira dans une grande partie des cas à l'environnement et la biodiversité, par les ressources et énergie qu'elle demande pour exister, et les déchets massifs qu'elle va engendrer.

Cela supposerait cependant que l'utilisateur donne son autorisation pour être filmé ou entendu. Dans le cas d'un régime totalitaire le déploiement de cette technologie pourra servir, de la même manière qu'en publicité, la propagande, et inciter au vote de certains candidats et candidates.

En conclusion, cette technologie d'IA peut avoir de réels impacts positifs sur la société. Elle a des conflits d'éthique de gravité variée selon les applications, allant des simples questions éthiques que peuvent avoir l'interaction, aussi présente d'humain à humain, à la possibilité d'une manipulation de masse. La technologie reste de mon point de vue très discutable utilisée dans la suggestion de contenu et dans la publicité ciblée. Mais si elle peut tout de même être contrôlée dans ce contexte, on ne pourrait empêcher son utilisation dans le cas de régimes totalitaires par exemple, et dans le contrôle des personnes et de l'idéologie. Elle reste dangereuse dans la mesure où elle peut détecter toutes nos émotions en temps réel et peut transmettre nos données, et ce parfois de manière insidieuse, sans que l'utilisateur en ait réellement conscience.

## IV. Références

Technologies :

<https://www.hoomano.com/>

<https://vokaturi.com/>

<https://aws.amazon.com/fr/rekognition/>

<https://www.microsoft.com/fr-fr/ai/ai-lab-conversational-ai>

<https://azure.microsoft.com/fr-fr/services/cognitive-services/face/>

Sécurité :

<http://go.affectiva.com/in-cabin-sensing>

<https://www.datakalab.com/>

Santé :

<https://imotions.com/>

Marketing :

<https://imotions.com/>

<https://www.kairos.com/blog/using-emotion-analysis-for-marketing-consumer-insights>

<https://www.realeyesit.com/>