

## Pourquoi veut-on rendre les robots empathiques?

L'opinion du professeur David Rudrauf et de l'avocat Nicolas Capt, invités à participer à la réflexion de TA-Swiss, l'institution qui fournit des notices d'information au Conseil fédéral.

ELA FLORET

«Construire une conscience artificielle c'est ce qui m'intéresse le plus. Une véritable conscience artificielle impliquera des capacités d'empathie. Et lorsque des robots gouvernés par une telle conscience artificielle seront suffisamment proches de nous, êtres humains, sur le plan de leur fonctionnement cognitif et affectif, la question se posera vraiment de la considération juridique, qu'il nous faudra avoir à leur égard», explique David Rudrauf, professeur à l'Université de Genève et directeur du laboratoire de modélisation multimodale de l'émotion et du ressenti.

Aux côtés de Nicolas Capt, avocat spécialisé dans les nouvelles technologies, il a été invité à participer à un groupe de réflexion préliminaire sur la question, par la Fondation TA-Swiss, l'institution qui développe des notices d'information pour le Conseil fédéral.

### Les robots empathiques existeront-ils vraiment?

Nourri à la science-fiction depuis mon enfance, ma réponse aujourd'hui en tant que chercheur est oui. Les robots pourront devenir un jour nos compagnons, mais cela impliquera qu'ils auront de véritables capacités empathiques. Ils devront être capables d'imaginer ce qu'on aime ou pas et leurs inférences seront gouvernées par une conscience similaire à la nôtre. De ce fait ils auront des comportements similaires aux nôtres. Ils seront capables de réciprocité intersubjective et d'imagination, qui jouent un rôle central dans la motivation de nos actions. Ils de-

ront être capables d'émotions et de comprendre nos enjeux humains.

Mais le problème, c'est qu'encore aujourd'hui, on ne comprend pas si bien que cela comment l'esprit humain fonctionne! On n'a pas résolu le problème fondamental de la cybernétique, avec un algorithme, qui gouvernerait des machines autonomes, selon les mêmes règles que notre propre esprit. A ce stade les robots dits empathiques sont essentiellement des coquilles vides... auxquelles nous attribuons plus de contenus qu'ils n'en ont. Certes les progrès sont énormes, mais il manque une grande synthèse des connaissances intégrant l'existant, l'affective computing, le machine learning et la vision computationnelle, et bien plus. Passionné de neurosciences, j'ai fini par réaliser que nous sommes loin de pouvoir rendre compte de la cognition et de l'esprit humain à partir d'une seule explication cérébrale. Il est essentiel de comprendre la psychologie, de modéliser la phénoménologie de l'expérience subjective indépendamment de la question de leur implémentation cérébrale.

Nous sommes encore très, très, loin d'avoir compris des choses fondamentales sur le fonctionnement cérébral lié aux émotions, aux croyances, à l'imagination, au-delà des seules contraintes biologiques. Il y a toute une optimisation subjective globale au cœur de notre esprit, combinant des motifs affectifs multiples et souvent incompatibles, impliquant nos capacités

projectives, dans l'imagination et dans la prise de perspective sociale, qu'il faudrait pouvoir capturer dans un modèle mathématique, afin de construire des machines qui nous ressemblent vraiment. Mes collaborateurs et moi-même pensons avoir compris des principes fondamentaux qui nous permettent d'approcher le problème. Mais il y a du travail...

### De quels financements publics-privés bénéficiez-vous pour votre recherche?

Ce type de projet est hautement interdisciplinaire et implique de pouvoir travailler dans un champ encore en friche, et cela ne facilite pas l'obtention de financements publics qui impliquent souvent des projets très balisés et averse au risque (*lire ci-contre*). C'est en partie la raison pour laquelle j'ai parallèlement démarré une société avec Jennah Kriebel, co-fondatrice et CEO, Scaphe, qui vise à développer une robotique nouvelle gouvernée par une conscience artificielle, profondément empathique, inspirée de la psychologie, avec de nombreuses applications transversales.

Il est potentiellement moins difficile d'innover dans le privé que dans le monde académique quand l'innovation implique un gros travail de recherche et développement coûteux: imaginez quelque chose qui demanderait l'engagement de 40 ingénieurs à plein temps, avec un cadre de travail impliquant un style de management, une division du travail purement dédiée au pro-



jet, visant avant tout l'efficacité et la productivité; pas très compatible avec la philosophie hautement respectable par ailleurs du développement des carrières académiques. Mais l'industrie impose aussi sa vision de faire des affaires. C'est un tout autre monde.

Je continue donc ma carrière académique avec passion tout en m'approchant du monde industriel pour arriver à implémenter ma vision, le tout dans la meilleure intelligence possible entre les deux mondes. L'université de Genève et son organe de transfert technologique (Unitec), ainsi que le Campus Biotech où j'ai mon laboratoire, m'offrent un environnement dans lequel ce type d'approche est possible. Il faut tenter l'aventure. Au moins j'aurai essayé.

**Outre le volet juridique et éthique sur lequel planche Nicolas Capt, quelles sont les limites que vous vous mettez à vous-mêmes dans cette quête de la conscience artificielle?**

Cette question est fondamentale pour moi (fils de psychanalyste),

qui poursuis ici une sorte de rêve un peu enfantin, mais qui donne aussi un sens profond à mon existence. Il s'agit pour moi de participer à une aventure, qui a un futur au-delà de l'humanité telle qu'on la connaît. C'est une ambition qui part de l'absurde, voire du tragique de la «conscience malheureuse» et de la condition humaine certes. Il y a un enjeu quasi métaphysique, «comme une revanche sur l'existence», qui nous laisse dans le noir à la fin.

Ce «post-humanisme» a des résonances avec une vision sartrienne de l'existence. Et ce qui était de l'ordre de la science-fiction devient aujourd'hui plus proche de la réalité. Mais penser l'au-delà ou l'à côté de l'humanité, ce n'est pas penser contre l'humanité, c'est ouvrir des voies, par exemple à des consciences non-biologiques, qui pourront un jour peut-être explorer le cosmos, bien au-delà de ce que nous pourrions jamais faire (on sera content si on arrive, nous les humains, sur la planète Mars). Il s'agit de continuer notre humanité à travers une nouvelle branche de l'évolution. Cela n'implique pas du tout de renoncer à travailler à un

monde meilleur pour notre humanité à nous. Et j'essaie aussi d'œuvrer au développement durable et au bien-être de l'être humain, en réfléchissant notamment à des applications cliniques et sociétales.

Nous avons récemment organisés, sous l'impulsion de Jennah Kriebel, un hackathon sur la question de l'éthique, accueilli dans les locaux de HPE à Genève, dans le cadre de cette réflexion sur l'émergence de consciences artificielles, et dans un esprit de science citoyenne. Il est ressorti des discussions d'un large éventail de participants, que nous devrions peut-être avoir plus peur des grosses compagnies qui utilisent l'intelligence artificielle pour nous aliéner dans les réseaux sociaux, que de l'intelligence artificielle elle-même. A la fin il faut mettre notre humanité dans la machine, plutôt que de laisser des êtres humains nous «machiniser.» Si on y réfléchit, si des consciences artificielles véritables émergent il faudra tout autant protéger l'homme de ces machines que ces machines de l'homme. ■



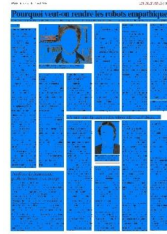
DAVID RUDRAUF. «A ce stade les robots dits empathiques sont essentiellement des coquilles vides...»

## Problème de financement public en Suisse et en Europe

David Rudrauf cherche des financements pour ses recherches en robotique empathique. «En raison du cloisonnement disciplinaire dans le milieu académique, il est difficile pour un professeur de psychologie de financer un projet sur un modèle visant à désigner un agent artificiel, autrement dit un robot, qui fonctionnerait selon les mécanismes de la psychologie humaine. L'écart entre les disciplines complique le financement de tels projets, et la multidisciplinarité intéresse tout le monde mais est souvent difficile à mettre en œuvre dans ce contexte».

«Cette approche un peu conservatrice relève parfois pour moi du paradoxe, vu le bon ratio risque-bénéfice et surtout la problématique universelle posée par la possibilité de l'empathie chez des robots», précise le professeur, qui regrette que les organes d'évaluation préfèrent souvent un terrain connu et rassurant, tout en comprenant la difficulté de l'évaluation.

David Rudrauf continue néanmoins à explorer des voies de financement académique ou à l'interface du public et du privé. Il a récemment soumis des dossiers à l'UE, avec des consortiums européens d'experts en robotique, en IA et en psychologie, mais aussi en éthique et en droit, certains axés sur le futur et les technologies émergentes, d'autres sur des applications beaucoup plus concrètes pour aider des enfants souffrant d'autisme, des personnes âgées souffrant de difficultés cognitives, et des personnes à mobilité réduites. ■



## «A mon sens, la taxation des robots relève de l'absurdité»

A la question de l'existence de cette conscience artificielle s'ajoutent donc ces questions juridiques et éthiques. Qui intéressent tout particulièrement Nicolas Capt, avocat aux barreaux de Genève et Paris et spécialiste des questions liées aux nouvelles technologies.

**Vous figurez parmi les experts choisis pour préparer le lancement d'une étude sur les robots empathiques. Quelles sont les questions éthiques et juridiques soulevées par leur existence future?**

En Suisse, le traitement de ces questions en est à ses balbutiements. La dernière tentative politique pour attribuer une personnalité juridique au robot a été sèchement rejetée par le Conseil national par 129 voix contre 55. Cela étant, ce postulat était sous-tendu par une thèse plus idéologique, la taxation des robots, ce qui explique peut-être la virulence de ce rejet.

A mon sens, la taxation des robots relève de l'absurdité. En premier lieu, elle est largement inapplicable puisqu'elle pose déjà un problème de définition: qu'est-ce qu'un robot? Doit-on se limiter au robot matériel ou aussi prendre en compte les robots dématérialisés comme les chatbots? Et où fixer la limite entre la robotisation et l'automatisation? Va-t-on taxer les photocopieurs ou les caisses automatiques? Et que fera-t-on des robots dématérialisés se trouvant sur des serveurs étrangers, mais accessibles depuis la Suisse? Et je ne parle même pas de l'impact en termes de fiscalité internationale sur un Etat qui entendrait, en faisant cavalier seul, taxer les robots.



**NICOLAS CAPT.** «*La question de l'intelligence artificielle et des robots est souvent dévoyée.*»

La question de l'IA et des robots est souvent dévoyée. Ce qui menace les emplois, c'est avant tout l'automatisation, que ce soit le remplacement d'une caissière par une caisse automatique ou alors celui d'une standardiste par un robot conversationnel. Et ce problème n'est pas neuf; au XX<sup>e</sup> siècle, la voiture automobile a peu à peu remplacé les cochers. L'IA n'est ainsi qu'une déclinaison de ce phénomène. Il y a toutefois une vertu à ce débat, celle de se poser la question de la réorganisation des recettes de l'Etat dans un monde en constant bouleversement. La taxation des robots est une réponse inapte à une question légitime.

**Que pensez-vous de l'avènement des robots empathiques?**

Il faut tout d'abord opérer un distinguo fondamental entre les robots qui présentent une empathie de façade et ceux qui, à terme et si les recherches actuellement menées aboutissent, présenteront une empathie au sens strict, sous-tendue par une conscience véritable et des affects.

Les conséquences juridiques et sociales de ces entités dissemblables n'ont évidemment pas de commune mesure. Dans le premier cas, le risque est notamment celui de se laisser happer par une illusion, par un robot dont la programmation lui fait mimer de l'empathie, sans la ressentir. C'est un risque de dépendance émotionnelle à laquelle s'ajoutent un péril terrible, soit la collecte massive et le traitement subséquent, de masses de données personnelles sensibles et de secrets. L'idée que l'on puisse, en situation d'isolement ou de faiblesse, considérer un robot comme son meilleur ami ou son confident devrait nourrir une véritable inquiétude.

Pour ce qui est des robots véritablement empathiques, le changement de paradigme sera complet et, si ces robots émergent un jour, il faudra envisager une véritable révolution juridique puisque des entités d'un genre nouveau rejoindront nos sociétés. J'ai longtemps été plus que circonspect sur cette question, mais les derniers travaux me laissent entrevoir que ma position était sans doute trop rigide.

**Comment encadrer ces machines devenues intelligentes?**

Il y a, là encore, le temps court et le temps long. A court terme, il faut réorganiser nos lois pour prendre en compte le bouleversement qu'impliquent les robots, par exemple dans le domaine des voitures autonomes. Il y a un éclatement véritable des acteurs et une dissolution des responsabilités. En cas d'accident, qui sera responsable? Le constructeur? Les occupants? Le fournisseur du logiciel d'IA? Sans compter que les



véhicules communiqueront entre elles et partageront des retours d'expérience. Ces questions de responsabilité sont épineuses et le droit actuel ne permet pas d'y répondre correctement.

A long terme, si des robots empathiques devaient être créés, le bouleversement du droit de nos sociétés serait alors total. Mais c'est évidemment encore trop tôt pour avoir une idée claire de ce que cela impliquera concrètement.

**Dans votre étude d'avocats, quelle est la proportion**

**de l'activité liée à ces questions éthiques?**

C'est embryonnaire. Les mandats que nous traitons sont davantage liés aux questions croisées de l'IA et de la protection des données. A ce stade, c'est la recherche académique qui domine. Il n'y a pas de grand mouvement ni de concrétisation en Suisse. L'Estonie est bien plus avancée que la Suisse, en cela comme en d'autres matières relevant des technologies.

**Quelle rédaction suite à l'abandon de Google**

**de son comité éthique sur l'IA?**

L'abandon par Google de son comité externe d'éthique à la suite de deux départs, dont l'un est en lien avec une pétition à l'interne de l'entreprise, démontre la difficulté qu'il y a à mettre sur pied un organe suffisamment représentatif et indépendant pour connaître des questions éthiques. S'il ne faut pas nier les mérites de l'auto-régulation, les questions éthiques ont avant tout vocation à être examinées par des organismes sans lien avec l'industrie. ■