

En bref

> Nous interagissons de plus en plus avec des machines, ou des robots, qui nous parlent.	> Ces «agents conversationnels» simuleront des capacités qui auront l'air d'être élaborées sur le plan cognitif, mais aussi sur le plan émotionnel, alors qu'ils ne comprennent rien.	> Il est essentiel d'évaluer les capacités de tels systèmes, voire d'en réaliser un suivi : étant donné la faculté d'apprentissage de ces machines, leur comportement pourrait changer avec le temps.	> Il faut aussi étudier le comportement des humains vis-à-vis des « robots bavards », afin d'éviter certains abus et effets pervers, l'isolement social par exemple.
--	---	---	--

96

«Si vous souhaitez joindre le service commercial, tapez 1. Le service technique, tapez 2.» Cela vous dit quelque chose? Eh bien, l'époque où une demande de renseignements exigeait de taper sur les touches d'un ordinateur ou d'un téléphone avant d'obtenir une information ou de joindre le bon interlocuteur touche à sa fin. De plus en plus, nous utilisons des programmes informatiques auxquels on s'adresse en parlant naturellement, et qui nous renseignent aussi en langage naturel, que ce soit à l'écrit ou à l'oral. Connectés sur internet ou à une base de données, embarqués sur un ordinateur, un téléphone ou encore un robot, ces programmes, invisibles aux yeux des utilisateurs, sont ce qu'on nomme des «agents conversationnels», ou encore robots bavards (*chatbots*, en anglais). Le premier d'entre eux était Eliza, élaboré à l'institut de technologie du Massachusetts (MIT) vers 1964-1966 par Joseph Weizenbaum : ce système répondait à l'écrit en simulant un psychothérapeute dont la stratégie était de répéter les propos du patient.

Si les *chatbots* ont longtemps été développés de manière à répondre à un sujet de discussion précis ou une tâche particulière, les *chatbots* dits à domaine «ouvert» sont conçus pour parler de tout. Ce qui sous-entend certaines capacités élaborées de la machine, ou plus exactement des logiciels dont elle est munie. Les *chatbots* étaient jusqu'à présent composés de plusieurs modules : de reconnaissance de la parole, de compréhension, de dialogue, de production et de synthèse de la parole. Les nouvelles générations de *chatbots* sont de plus en plus performantes grâce à l'évolution des techniques d'apprentissage des

machines, à la puissance des processeurs et à la taille des bases de données. Les modèles de traitement automatique de la langue les plus récents, appelés des «transformeurs», sont des réseaux de neurones qui, à partir de vastes corpus linguistiques, apprennent les régularités les plus saillantes, sans être influencés, par l'ordre des mots. Des *chatbots* sont maintenant capables d'aborder n'importe quel sujet sans contrainte, par exemple Blenderbot, le modèle conversationnel de Facebook. Ou encore Meena, système fondé sur l'apprentissage profond de bout en bout entraîné sur des dialogues entre humains, qui montre des performances dépassant tous les *chatbots* existants.

LE TEST DE TURING, INSUFFISANT POUR PLUSIEURS RAISONS

Au-delà des avancées scientifiques et techniques mises en œuvre dans les *chatbots*, l'interaction croissante que nous avons avec ces derniers soulève des questions plus fondamentales, voire stratégiques et éthiques, sur les capacités des machines et sur la relation intersubjective qu'instaurent avec celles-ci leurs utilisateurs humains. En particulier, si l'on veut vivre harmonieusement avec des machines qui nous assistent et avec lesquelles nous dialoguons, il apparaît de plus en plus important d'évaluer de façon pertinente et continue, par des batteries de tests reproductibles, ces capacités ainsi que leur impact sur la relation entre humain et machine.

Que faut-il tester plus précisément, et comment? On s'est longtemps focalisé sur l'«intelligence» des machines. D'ailleurs, le domaine