



## INTERVIEW. "Il existe des limites intrinsèques aux techniques d'IA actuelles"



Par Arnaud Devillard le 19.07.2022 à 16h59 Lecture 8 min. Abonnés

Maître de conférences à Télécom Paris au laboratoire Traitement et communication de l'information, Jean-Louis Dessalles se passionne pour une approche scientifique de l'intelligence artificielle, en observant un certain recul sur les méthodes statistiques actuellement en vogue.

réagir

Pourquoi l'intelligence artificielle n'arrive pas à être à la hauteur des fantasmes ? Auteur de *"Des intelligences très artificielles"* en 2019 ([Odile Jacob](#)), Jean-Louis Dessalles n'est ni exagérément enthousiaste ni un sceptique forcené de l'IA. Il remet simplement en cause les approches techniques adoptées, ou plutôt leur manque de variété. Car si la puissance des réseaux de neurones, le "deep learning", a produit des résultats époustouflants, elle a généré un engouement un peu trop généralisé. Aujourd'hui, "deep learning" est quasiment devenu, à tort, synonyme d'intelligence artificielle. Avec le risque de buter sur des impasses. En complément du numéro d'été de *Sciences et Avenir* et de son dossier "Pourquoi ?", où le sujet était abordé, Jean-Louis Dessalles développe ici son propos.

**Sciences et Avenir : Vous faites partie de ces spécialistes de l'informatique qui relativisent les pouvoirs de l'intelligence artificielle et les exagérations sur le sujet. De quelles exagérations parle-t-on ?**

**Jean-Louis Dessalles :** Il plane cette idée que tout est apprenable, à coup de données, de statistiques et de fonction d'erreur. Rien de péjoratif dans mes propos : des trésors d'ingéniosité sont déployés pour faire des réseaux de neurones. J'ai suivi tous ces développements, même sans y croire au départ. Or je n'avais pas compris qu'après une montée en puissance quantitative, qu'avec l'usage des GPU, on passerait un seuil qui



permettrait de traiter des données brutes. Mais de là à extrapoler sur la prise de pouvoir des machines et la fin de l'humanité...

### **Vous ne croyez pas en une IA forte, omnipotente ?**

Ce n'est en tout cas pas pour demain. On a beaucoup dit que l'intelligence artificielle allait supprimer un tas de métiers. C'est une énorme erreur : elle supprime des tâches, pas des métiers. En 2019, IBM a présenté Debater : une vidéo montrait un système capable de mener un débat d'expert. Il parvenait à argumenter pendant 19 min et 30 secondes et à la fin, il a dit : "*J'ai terminé. Maintenant mon contradicteur va vous dire ceci, ceci et cela, ne l'écoutez pas*". Et le débateur humain a effectivement présenté les arguments prédits par la machine ! Cela m'a bluffé, j'ai pensé qu'ils avaient réussi le test de Turing.

Pourquoi l'intelligence artificielle n'arrive pas à être à la hauteur des fantasmes ? Auteur de "*Des intelligences très artificielles*" en 2019 (**Odile Jacob**), Jean-Louis Dessalles n'est ni exagérément enthousiaste ni un sceptique forcené de l'IA. Il remet simplement en cause les approches techniques adoptées, ou plutôt leur manque de variété. Car si la puissance des réseaux de neurones, le "deep learning", a produit des résultats époustouflants, elle a généré un engouement un peu trop généralisé. Aujourd'hui, "deep learning" est quasiment devenu, à tort, synonyme d'intelligence artificielle. Avec le risque de buter sur des impasses. En complément du numéro d'été de *Sciences et Avenir* et de son dossier "Pourquoi ?", où le sujet était abordé, Jean-Louis Dessalles développe ici son propos.

### **Sciences et Avenir : Vous faites partie de ces spécialistes de l'informatique qui relativisent les pouvoirs de l'intelligence artificielle et les exagérations sur le sujet. De quelles exagérations parle-t-on ?**

**Jean-Louis Dessalles** : Il plane cette idée que tout est apprenable, à coup de données, de statistiques et de fonction d'erreur. Rien de péjoratif dans mes propos : des trésors d'ingéniosité sont déployés pour faire des réseaux de neurones. J'ai suivi tous ces développements, même sans y croire au départ. Or je n'avais pas compris qu'après une montée en puissance quantitative, qu'avec l'usage des GPU, on passerait un seuil qui permettrait de traiter des données brutes. Mais de là à extrapoler sur la prise de pouvoir des machines et la fin de l'humanité...

### **Vous ne croyez pas en une IA forte, omnipotente ?**

Ce n'est en tout cas pas pour demain. On a beaucoup dit que l'intelligence artificielle allait supprimer un tas de métiers. C'est une énorme erreur : elle supprime des tâches, pas des métiers. En 2019, IBM a présenté Debater : une vidéo montrait un système capable de mener un débat d'expert. Il parvenait à argumenter pendant 19 min et 30 secondes et à la fin, il a dit : "*J'ai terminé. Maintenant mon contradicteur va vous dire ceci, ceci et cela, ne l'écoutez pas*". Et le débateur humain a effectivement présenté les arguments prédits par la machine ! Cela m'a bluffé, j'ai pensé qu'ils avaient réussi le test de Turing.

Puis j'ai visionné la séquence et j'ai isolé une phrase incroyable sur le financement des écoles aux Etats-Unis pour la saisir dans Google. J'ai découvert qu'elle provenait d'un article du *Santa Cruz Sentinel* dix ans auparavant. La prouesse de Debater n'était donc pas de générer cette phrase, mais de l'extraire de son contexte pour donner l'illusion d'une argumentation. On pourrait utiliser un tel système pour réaliser des dossiers de presse, des jurisprudences. Pas pour mener des raisonnements.

### **Votre constat est-il généralisé au sein du monde académique ?**

Les étudiants, les doctorants, ne comprennent même pas quand je leur dis qu'il existe des limites intrinsèques aux techniques actuelles. Attention, je n'ai pas d'opposition de





principe, je pense au contraire qu'une mécanisation de l'intelligence est possible. Le tout est de savoir avec quelle technique.

Mais pour un étudiant, ce genre d'objection ne tient pas. Il a à disposition des réseaux de neurones pré-entraînés et au bout de ses doigts, il y a la "magie" du résultat. C'est spectaculaire. Comment résister ? Donc non, nous ne sommes pas du tout en train de revenir de l'IA.

### **En quoi ces approches statistiques, sont-elles si puissantes ?**

Avant, pour caricaturer, il fallait des bases de données structurées. Si on voulait travailler sur des images de cancer du sein, il fallait les pré-décrire : dire que telle image a un contour irrégulier, que telle autre est comme ceci, comme cela. Ce prétraitement était manuel et coûtait très cher. Désormais, on traite des données brutes : les pixels, les pixels bruités... Et nous ne sommes pas au bout de nos surprises, vu la foule de chercheurs qui se sont lancés là-dedans. Mais mon devoir auprès de mes étudiants est de prévenir qu'il ne s'agit "que" de statistiques.

### **Quels en sont les limites, justement ?**

C'est de l'intelligence en boîte, du pré-calcul. On voit ce qui marche mais pas beaucoup ce qui ne marche pas. Le sens de verbes comme "poursuivre", "empêcher", "gêner", "rencontrer", un enfant le comprend après une, deux, trois occurrences. On sait reconnaître une situation de jalousie sans vraiment en avoir appris de définition. Mais on ne teste jamais les réseaux de neurones là-dessus. Ils reconnaissent les mots, les objets, les formes, pas des relations.

Prenez même la traduction automatique. Appliquée à des textes littéraires hors contexte, cela fonctionne. Or, je me suis amusé avec un exemple pour mes étudiants. Je sortais de la cantine, on nous avait servi du lieu, le poisson. Dans Google Translate et DeepL (l'outil de traduction développé par les équipes du dictionnaire bilingue en ligne Linguee, NDLR) j'écris donc : "*Aujourd'hui j'ai mangé du lieu*". Résultat : "*Today, I ate the place*" ! Lieu, le poisson, se dit "pollack" en anglais (lieu jaune) ; "place", c'est la notion topographique, géographique.

Je recommence en corrigeant le tir : "*Aujourd'hui j'ai mangé du lieu jaune*" et le résultat est bien : "*Today I ate pollack*". Je me dis qu'il fallait juste utiliser le bon terme anglophone. Sauf que j'ai prolongé l'expérience en écrivant : "*Si tu vas dans ce restaurant, prends le lieu jaune à la carte*". Je pensais que Google Translate tomberait dans le piège de traduire "carte" par "map" au lieu de "menu". Or, sa proposition a été : "*If you go to this restaurant, have the yellow place on the menu*" !

Autrement dit, l'algorithme a bien compris qu'on était au restaurant et qu'il fallait traduire "carte" par "menu" mais il remet "place" et non "pollack" pour le poisson !

C'est révélateur de l'absence de compréhension de la part de ces systèmes. Et c'est une bonne illustration du théorème de Bayes en théorie des probabilités selon lequel il existe un compromis entre la corrélation et l'a priori. Dans le cas qui nous occupe, il y a un énorme a priori pour traduire "lieu" par "place" et une corrélation entre les termes "manger" et "menu". Mais la corrélation n'arrive pas à compenser l'a priori. C'est toute la malédiction des systèmes statistiques.

Bien sûr, on va me rétorquer qu'il suffit de corriger les données et de faire apprendre à l'algorithme plus de cas où "lieu" désigne le poisson. Sauf que corriger un biais, c'est inévitablement en créer d'autres, un problème propre aux systèmes statistiques.

### **Pour vous, cela n'a donc rien à voir avec l'intelligence ?**

Cela relève d'un stimulus-réponse. Les mathématiques qui sous-tendent ces développements relèvent d'une ingénierie de haut niveau. Mais en terme scientifique,



cela ne permet pas de comprendre les mécanismes de l'intelligence humaine.

**La formule même d'intelligence artificielle vous paraît-elle, comme à certains chercheurs, abusive ?**

Non, je l'aime bien. Seulement le sens évolue. Maintenant, IA a été réduit à "apprentissage automatique" et "apprentissage automatique" à "apprentissage automatique statistique". Tout le monde a oublié qu'il existait un apprentissage automatique non statistique !

**Quelles approches privilégiez-vous ?**

Celles qui vont se rapprocher des mécanismes cognitifs. Les phrases "*Elle n'est pas toujours malade*" et "*Elle n'est toujours pas malade*" contiennent les mêmes mots mais ils n'ont pas le même sens. Je n'ose pas imaginer combien d'exemples il faudrait donner à un réseau de neurones pour qu'il réussisse à faire le distinguo. Je ne saurais même pas quels types de données lui montrer. Il aurait besoin à la fois des phrases et du sens, alors que n'importe quel humain comprend.

Même chose avec "*Elle est encore malade*" : c'est ambigu. Cela veut dire soit "*Elle n'est pas guérie*", soit "*Elle est malade une fois de plus*". Ce n'est pas le mot "encore" qui a deux sens, mais "malade", qui correspond à plusieurs durées. Ou : "*Elle a éternué pendant le spectacle*". Cela peut vouloir dire qu'elle a éternué de manière répétitive ou une fois mais tout le monde a entendu ! Dans les deux cas, il faut comprendre qu'il s'agit d'une situation anormale.

Tout cela, l'intelligence humaine le calcule à la volée. On n'a pas appris ce genre de nuance par cœur quand on était petit.

**On touche là aux notions de sens commun et d'intuition. C'est ce qui échappe à l'intelligence artificielle ?**

Pas exactement. Il suffirait d'extraire 1% des efforts de recherche actuellement focalisés sur les réseaux de neurones et de le concentrer sur ce genre de sujets, et on avancerait beaucoup plus vite. Mais ce n'est pas du tout à la mode.

J'enseigne dans une école d'ingénieurs. Et le critère de validité du travail d'un ingénieur, c'est l'efficacité. Pas la compréhension de phénomènes existants indépendamment de leurs recherches. Les premiers fabricants de machines à vapeur laissaient de côté la physique sous-jacente, ils n'en prenaient que ce qui était utile. Les premiers physiciens à avoir réfléchi ensuite à la machine à vapeur ont créé la thermodynamique, mais ce ne sont pas les mêmes personnes. Pour moi, en matière d'intelligence artificielle, nous sommes dans la phase machine à vapeur.

