

# Entretien avec Daniel Goossens, chercheur en IA et auteur de bandes dessinées reconnu !

1<sup>er</sup> octobre 2021

<https://carnetsdalerte.fr/entretien-avec-daniel-goossens-chercheur-en-ia-et-auteur-de-bandes-dessinees-reconnu/>



**Entretien avec Daniel Goossens : c'est étonnant car l'IA nous permet de gagner aux échecs, de démontrer des théorèmes mathématiques complexes, bref, de résoudre des problèmes hors de portée de notre intelligence naturelle mais elle n'a toujours pas de compréhension similaire à la nôtre.**

**Entretien avec Daniel Goossens**, chercheur en intelligence artificielle à l'université Paris 8 et auteur de bandes dessinées reconnu. Ses recherches portent sur la compréhension du langage naturel.

« Je suis suffisamment conscient pour savoir que je n'ai aucune chance d'appréhender comment fonctionne

notre compréhension. C'est un mystère que nous parviendrons peut-être à percer un jour... »

**Comment a commencé votre double carrière de chercheur en intelligence artificielle et dessinateur ?**

J'ai toujours mené en parallèle mes deux métiers. J'ai commencé à dessiner très tôt dans les marges de mes cahiers. J'ai découvert la recherche à l'université de Vincennes, où j'ai rencontré Patrick Greussay, qui m'a transmis sa passion pour l'intelligence artificielle. Je me suis spécialisé dans ce domaine dès le début. J'avais cette curiosité scientifique que je n'arrivais pas à exploiter dans le cadre des études usuelles. J'avais le profil de chercheur-explorateur plutôt que celui d'ingénieur. À l'université, j'ai trouvé un environnement, des ordinateurs, des objectifs ludiques avec tout mon temps pour pouvoir développer mes projets.

**Vous travaillez sur la compréhension du langage naturel en IA. Pourriez-vous nous l'expliquer ?**

J'ai travaillé un peu sur le signal sonore, l'enregistrement de la parole, avec des objectifs purement pédagogiques. C'est plus difficile que de traiter du texte tapé au clavier, où les lettres, les syllabes, les mots, sont déjà séparés. Pour un même locuteur, et à fortiori pour des locuteurs différents, les oscillogrammes d'un même son ont une énorme variabilité. On stabilise quelque peu avec des outils numériques comme « la transformée de Fourier »\*, qui fournissent des sonagrammes, moins variables, mais la vraie solution, c'est de mémoriser d'énormes quantités de façons de prononcer. C'est le succès du *Deep Learning* ou *apprentissage profond* (processus adapté aux tâches complexes s'appuyant sur les réseaux de neurones artificiels). C'est en quelque sorte de la statistique qui modélise un aspect de la perception humaine. Ça dépasse même la perception humaine, puisque la mémoire est quasi illimitée.

A partir du son ou du texte, il faut ensuite analyser la syntaxe puis vient l'énorme difficulté de la sémantique, de la pragmatique, etc. La compréhension

du langage naturel est donc un vaste continent, impossible à appréhender dans son ensemble.

« La compréhension du langage naturel est donc un vaste continent, impossible à appréhender dans son ensemble. »

Ma recherche spécifique vise une organisation logique des connaissances qui permette une forme simple de compréhension, par exemple une capacité à répondre correctement à des questions intuitives, ce qui est déjà excessivement difficile. Les systèmes de question-réponse accessibles sur Internet répondent à des questions factuelles, comme “À quel âge est mort Elvis Presley ? “, en se servant de sources comme Wikipédia. Ils n’ont pas de compréhension véritable, ils sont perdus dans les graphes de connaissances qu’ils parcourent et échouent sur des absurdités comme “À quel âge est né Elvis Presley ? “ à moins d’y être spécifiquement préparés. La compréhension humaine se déplace dans la connaissance comme un acrobate. Une étape intermédiaire, déjà inatteignable actuellement, serait de concevoir un programme capable de répondre comme le ferait un humain, à toutes les questions, censées ou absurdes, formulables à partir d’un dictionnaire de quelques dizaines de mots sur un domaine quelconque. On peut le faire partiellement pour ce qui est représentable en logique de Boole, la plus simple.

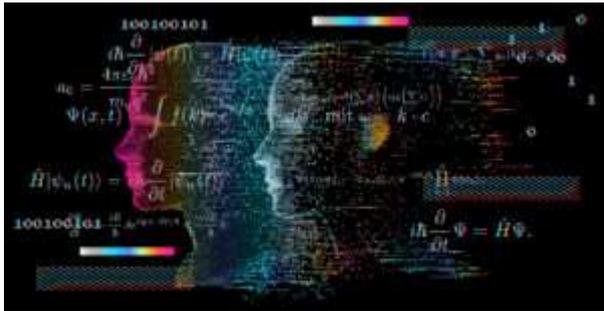
**Est-il possible que l’intelligence artificielle puisse un jour reproduire notre sens commun, c’est à dire l’ensemble de nos connaissances générales ?**

Oui mais ça me semble actuellement hors de portée. C’est encore un mystère entier qu’on percera sûrement un jour. C’est un domaine intéressant à étudier et explorer comme une jungle.

C’est étonnant car l’IA nous permet de gagner aux échecs, de démontrer des théorèmes mathématiques complexes, bref, de résoudre des problèmes hors

de portée de notre intelligence naturelle mais elle n'a toujours pas de compréhension similaire à la nôtre.

L'IA connexionniste actuelle est sûrement différente de la façon dont notre perception fonctionne. Un enfant modélise le réel à partir de très peu d'exemples. Les réseaux de neurones artificiels ont encore besoin de grandes quantités d'exemples pour obtenir leurs capacités de discrimination.



Pour différencier un chat et un chien sur une photo, un programme a besoin de centaines de milliers de photos de chats et de chiens. Il y a un exemple intéressant : une équipe a modifié des photos représentant des objets (un immeuble, une voiture, une grue, un arbre...) en ajoutant des pixels quasi invisibles à l'œil nu, qui ne modifient pas la photo mais trompent le réseau de neurones. Il le reconnaît comme un autre objet. L'expérience montre que les réseaux de neurones n'ont pas la "conscience" de ce qu'ils voient.

« Pour différencier un chat et un chien sur une photo, un programme a besoin de centaines de milliers de photos de chats et de chiens. »

**Il y a beaucoup de mythes sur l'IA. L'idée d'une machine autonome qui prendrait le pouvoir et remplacerait l'humain est-elle réaliste ?**

On est spontanément plus facilement attiré par le mythe d'une société de robots qui prend le contrôle et met l'humanité en esclavage que par le mythe de cerveaux artificiels collectifs qui résoudraient les problèmes hors de notre portée. Un film de science-fiction qui raconterait l'évolution d'une

administration de services publics qui se dotent de réseaux pensants aura forcément moins de succès que Terminator.

Il y a des mythes devenus réalité comme la voiture autonome, qui marche aussi bien que la reconnaissance de la parole, c'est-à-dire encore imparfaitement. Pour l'instant, la voiture autonome ne peut avoir de conscience comme celle que nous avons. La voiture reconnaît des situations, la route, la ligne jaune qu'il ne faut pas dépasser, le trottoir. Mais il peut y avoir une "hallucination", qui provoque une catastrophe.

Le mythe qui, pour moi, n'est pas crédible est celui du *Big Brother* ultime, c'est-à-dire un cerveau unifié qui aurait l'objectif de manipuler tout le monde. Actuellement, il existe plein d'organisations qui se servent de nos données sur Internet pour faire avancer leurs intérêts. C'est la rançon du capitalisme, de la concurrence, la loi des marchés, etc. Autrement dit, c'est l'absence d'une organisation globale. Quand on veut organiser, planifier l'économie, on finit régulièrement dans un marécage de problèmes insurmontables. Le réel est difficile à dompter. C'est ce qui fait la longévité des solutions qui contrôlent le moins possible et laissent à la concurrence sauvage le rôle de trancher sans pitié. Si dans un futur très lointain se profilait un cerveau collectif, il n'aurait pas de concurrence. Il ressemblerait à une administration. On pourrait même dire que les administrations actuelles sont un cerveau collectif qui se construit patiemment depuis des millénaires, un cerveau collectif qui a de la mémoire et de l'organisation mais pas la moindre trace d'une compréhension comparable à celle d'un individu humain.

### **Pourrions-nous nous passer de l'intelligence artificielle ?**

Les problèmes qui s'annoncent, les dérèglements climatiques causés par le développement technologique mais aussi les désordres sociaux violents qui en découlent, sont de plus en plus hors de notre portée. Ils sont hors de la portée de nos capacités intellectuelles individuelles mais aussi des capacités de toutes nos machines sociales. Nos machines sociales les plus sophistiquées peuvent tout juste percevoir les dangers, ce qui est déjà beaucoup parce que sans elles,

on foncerait dans le mur sans jamais en avoir conscience donc sans chercher de solution.

Arrêter tout progrès technologique source de pollution est illusoire. Les simples explosions démographiques sont déjà impossibles à maîtriser et elles sont la source de toutes les explosions de consommations. Tout le monde est capable d'imaginer sur le papier un planning familial international qui stabiliserait la population mondiale mais personne n'est en mesure de l'imposer où que ce soit.

Il faut juste espérer qu'Internet, qui est un embryon de cerveau collectif, se transforme progressivement en quelque chose qui soit capable de fournir des solutions qui nous sauveraient de quelques cataclysmes. Mais tout le monde sait que rien n'est garanti, bien sûr. En tous cas, pour moi, sans ces espoirs technologiques, c'est le cauchemar assuré. On est malgré nous dans une course sans fin. Les religieux devraient intégrer ça : Dieu est un beau sadique. Ou



alors, il n'a pas le choix ?

## Quels sont pour vous les dangers liés à l'utilisation de l'IA ?

Avant qu'on ne puisse un jour planifier quelque peu une économie mondiale et garantir ainsi un confort minimal à chaque individu, l'IA va offrir des positions de domination à des potentats locaux. Elle fournit des capacités nouvelles pour profiter de grands volumes d'informations sur les adversaires économiques, elle permet de concevoir des engins de guerre redoutables, etc. Par ailleurs, c'est une technologie qui consomme de plus en plus d'énergie et qui donc pollue toujours plus. Il paraît que le cerveau d'un individu consomme près de 20 % de

l'énergie du corps. Il faut donc s'attendre à ce que le développement de cerveaux collectifs soit gourmand.

« C'est une technologie qui consomme de plus en plus d'énergie et qui donc pollue toujours plus. »

### **Pourrait-elle aider à lutter contre les enjeux climatiques ?**

Comme je disais, les problèmes d'épuisement des matières premières, la pollution, le réchauffement climatique, la surpopulation, etc, me semblent hors de portée à la fois de la pensée de tout individu mais aussi des capacités de nos machines sociales (institutions, administrations, gouvernements, organisations, clubs de poésie...) les plus développées.

Nos fonctionnements sociaux sont bridés par un goulot d'étranglement qui est cette nécessité d'avoir recours à un individu isolé pour prendre des décisions. Plus nous sommes nombreux à prendre des décisions et plus cela devient incohérent. C'est la cacophonie de la pensée collective. Le vote, par exemple, est probablement l'embryon le plus primitif de pensée collective. Pour moi, le seul espoir d'éviter cela serait d'avoir des pensées collectives artificielles avec une compréhension semblable à la nôtre mais une mémoire largement supérieure.

### **Comment expliquer le concept de pensée collective artificielle ?**

Ce serait comme un réseau Internet pensant. Je ne sais pas s'il pourrait être unique ou s'il faudra passer par une société de réseaux pensants, possiblement en concurrence les uns contre les autres. Il aurait une mémoire comparable à celle de l'Internet actuel et il aurait les mêmes mécanismes de compréhension que nous, ceux qui nous donnent l'impression qu'on est pleinement conscients plutôt que dans le cirage. Honnêtement, je ne sais pas ce que donnerait un tel cocktail mais c'est intimidant. Et bien sûr, pour l'instant, c'est de l'utopie.

Le fait que la pensée individuelle se prolonge en une pensée collective me semble dans la logique de la nature. Les individus s'organisent spontanément en sociétés structurées. Ces sociétés communiquent, mémorisent, gèrent de l'information en continu. Elles sont donc des embryons de cerveaux, composés d'individus. Cependant, les organisations collectives sont toujours plus rigides et frustrantes que la souplesse d'une pensée individuelle isolée. C'est la frustration fondamentale. Les pensées collectives de nos sociétés structurées étaient hier trop archaïques et cruelles, trop religieuses, trop normatives et aujourd'hui dans les démocraties les plus avancées, trop chaotiques. Ma conviction est que si les pensées collectives gagnaient en capacité de compréhension, la frustration de nos pensées individuelles diminuerait.

L'IA actuelle est encore loin d'une compréhension comparable à la nôtre. On ignore les problèmes théoriques à résoudre et je ne sais pas si les problèmes matériels qui nous en séparent (coûts en énergie, pollution, etc) peuvent se surmonter. Elle consomme déjà beaucoup pour des capacités qui s'apparentent plus à la simple perception, mémorisation et résolution de problèmes complexes qu'à la compréhension.

### **L'intelligence artificielle vous semble-t-il être un outil favorable à l'organisation d'un collectif plus juste ?**

Dans les échanges d'arguments (discussions politiques, débats houleux à la télé, etc), il n'y a pas de compréhension collective, pas de synthèse, rien qui récupère les hypothèses émises et tente d'en développer une théorie cohérente. Les haines entre les camps opposés peuvent être intenses et durables et il n'y a pas d'arbitre spontanément accepté.

Il n'y a pas de bouc-émissaire à désigner comme responsable. Personne, dans ces interactions, n'est en mesure d'imposer une vérité, bien que chacun le souhaite secrètement. Il faut espérer l'arrivée lente dans un futur lointain d'une IA dite "forte" qui serait capable de comprendre tous les points de vue et d'en faire des synthèses rassurantes et en accord avec le réel. Déjà aujourd'hui, qui est terrifié par le fait que Wikipédia en sait plus que lui ?

En attendant, il faut se contenter d'une IA qui résout des problèmes concrets (médecine, transports, finance, sécurité...) et fait planer des menaces (armes de guerre autonomes, fichages politiques...) le tout sans le moindre état d'âme car elle n'en est pas encore capable.

### **Juliette Duquesne et Marion Biosse Duplan**

\*La **transformée de Fourier** est une opération qui permet de représenter en fréquence des signaux qui ne sont pas périodiques.