

« L'intelligence artificielle : qu'est-ce que c'est ? quels impacts sur le travail et le management ? »

Synthèse de la conférence introductive par **Dominique Boullier**, Professeur des Universités en sociologie - Digital Humanities Institute - École Polytechnique Fédérale de Lausanne (EPFL) et **Valérie Fert**, expert analyste du système d'Intelligence Artificielle Mileva, cofondateur et présidente de GMAP (<https://www.gmap-center.ch/>)

Mileva est un système d'intelligence artificielle avancé basé sur la théorie de l'information, l'approche quantique et la représentation de la complexité. Il révèle la trame des situations complexes et fournit prospectives et simulations qui mettent en lumière les proximités des différents enjeux et acteurs, les réseaux qu'ils forment, leurs comportements, leurs relations d'influence et les risques de rupture. Outre des **analyses**, **prospectives** et **simulations**, Mileva fournit également des **tableaux de bord** et des **newsletters intelligentes personnalisées**.

Valérie FERT : il n'y a pas une mais différentes intelligences artificielles, selon leur objet, leur usage, leur technicité. Cela va de l'assistant personnel en passant par le robot capable de relayer un chirurgien dans le cadre d'une opération à distance, ou encore la machine learning à laquelle on aura présenté mille images de chat afin qu'elle soit capable de reconnaître le 1001^{ème}. Mileva relève encore d'un autre registre, celui de la machine thinking, un système avancé d'intelligence artificielle basé sur une approche quantique de la théorie de l'information, un système capable de proposer des réponses aux questions qui lui ont été adressées : par exemple, qui sera le prochain pape ?

Dans tous les cas, la posture est la même : celle d'un outil au service de l'homme, et non l'inverse. L'intelligence artificielle est un outil comme un autre, outil qui permet au cerveau de se libérer de tâches fastidieuses et de se concentrer sur les tâches à valeur ajoutée : réfléchir, anticiper, analyser, décider, ...

Le principe de Mileva et des intelligences artificielles de type « machine thinking » est d'adopter une logique d'analyse plus systémique de la complexité, comparée à la logique cartésienne, base de notre enseignement et donc de notre approche du monde (une décomposition ou segmentation des problèmes, une pensée linéaire interrogeant causes et effets, le recours à la comparaison qui nous formate à ne prendre en compte les éléments que s'ils sont de même nature, ...). Mileva s'affranchit de ces contraintes : elle raisonne sur un nombre exponentiel d'informations et surtout, elle s'appuie sur des éléments hétérogènes, sur des interactions de nature différentes (boucles, rétroactions, échanges, ...), et sur le principe de recherche de stabilité du système pour apporter « sa réponse ».

Tout se joue au niveau du travail préparatoire : la requête (la question posée) et des différents critères qu'on lui demande de prendre en compte pour répondre à la question (les axes, décomposés en agents puis en mots clés). La qualité de ces informations est essentielle et relève bien à ce stade de l'intelligence humaine. Il en est de même pour la façon dont la machine a été « éduquée » (en l'occurrence par Valerie Fert, non ingénieur). Ensuite, la machine rassemble, tri, analyse, compile une somme importante d'informations qui peuvent être très disparates mais restent en lien avec les axes, agents et mots-clés paramétrés, pour donner une série de cartographies ou image d'ensemble : à

l'utilisateur, et donc à l'intelligence humaine, d'en déduire ou retenir le résultat. A lui de le refuser au motif que cela ne correspond pas à ce qu'il attend, ou de se demander pourquoi cette image d'ensemble lui est donnée, et alors d'en apprendre beaucoup sur la façon de penser la complexité ! C'est ainsi que Mileva, utilisée pour prédire qui serait le prochain Pape quelques jours avant l'élection du Pape François, avait fourni une cartographie qui mettait en avant une convergence vers le cardinal Jorge Mario Bergoglio... contre toute intuition humaine.

Dominique BOULLIER : je préfère parler de "machine learning" pour susciter moins de craintes ou d'attentes qu'avec le mot d'Intelligence Artificielle. C'est un domaine en plein évolution : on a déjà changé de génération et cela change les perspectives d'utilisation que l'on peut imaginer. Nous avons fait un saut considérable sur les cinq dernières années dans les capacités, les types de calculs et dans la baisse du taux d'erreur. Un exemple en est la progression des traducteurs automatiques.

Il y a déjà plusieurs approches possibles du Machine Learning : les spécialistes "évolutionnistes" travaillent sur les algorithmes qui font progresser les algorithmes, tandis que les spécialistes "bayésiens" travaillent à établir des probabilités dans des environnements incertains : ces derniers sont partout !

Une intelligence artificielle... artificielle

Quel que soit le modèle conceptuel (modèle symboliste qui s'appuie sur les principes de déduction, l'analogiste sur la recherche de similarité pour prédire des situations nouvelles, ou le connexionniste sur la pondération des connexions pour obtenir un résultat espéré), il y a derrière le « machine learning » de nombreuses tâches humaines pour classer, noter, corriger, de nombreux clics pour faire apprendre les algorithmes à la machine. On peut parler de "tâcherons" humains qui travaillent en micro-tâches élémentaires, qui n'ont pas de profils d'experts, mais dont le travail en amont sur le « set d'apprentissage » est nécessaire : l'IA est en fait une intelligence artificielle... artificielle !

Alors que l'une des ambitions de l'intelligence artificielle est d'éliminer les biais cognitifs humains, ce qui se joue au niveau du set d'apprentissage (finalement une forme de supervision de la machine élément par élément) est susceptible d'en apporter de nouveaux : que se passe-t-il dans la boîte noire de la « machine learning », qu'est-ce qui a été validé et comment au moment de l'éducation de la machine ?

Une intelligence qui n'est pas sans limite

Les capacités de l'IA dépendent fortement de la façon dont la problématique a été posée en amont, qui détermine le niveau de complexité, voire la possibilité de résolution du problème. Car même si les capacités de calculs et de traitement de l'information sont démesurées par rapport aux capacités humaines, et toujours en expansion, elles ont une limite donnée par la capacité des ordinateurs et par le panel des éléments qui ont contribué à l'éducation de la machine : trop ou trop peu d'observations pour un même critère, trop ou trop peu de critères discriminants. Dans un cas, le « machine learning » est confronté à une trop forte quantité de calculs à opérer, - ce qui incite à aller vers une réduction de dimension des critères et observations ; à l'opposé, trop peu d'observations rend difficile l'automatisation tandis que trop peu de critères peut introduire des biais.

D'autres limites existent du fait que le système se nourrit d'observations, donc de réalités passées. Pour ces raisons, la capacité du « machine learning » à traiter beaucoup d'informations est utile pour des décisions de court terme, mais perd de sa pertinence pour des décisions de long terme. « En effet, le passé, seule source de l'IA, obstrue les devenirs en en faisant des futurs probabilisables ». De même, le « machine learning » n'est pas fait pour prédire les événements exceptionnels. Si on l'oublie, on s'aveugle nous-mêmes, comme l'oie qui a été habituée à être nourrie tous les jours à 9h, jusqu'au 24 décembre où elle passe à la casserole !

Dans tous les cas, le « machine learning » doit laisser toute sa place à l'humain, à l'image de ce qui se joue dans la radiographie médicale : 80% des examens relèvent de diagnostics de routine et sont effectivement aisément automatisables ; si le problème est complexe et laisse place au doute et à l'interprétation, rien ne remplace une équipe pluridisciplinaire pour discuter et décider.

En termes de management, il est essentiel d'avoir en tête l'ensemble de ces éléments pour mesurer avec discernement la puissance et les limites du « machine learning », pour au final donner la place qui revient à l'humain dans les processus de décision et penser en conséquence le design organisationnel adapté.

En synthèse, deux nouvelles rassurantes qui donnent toute sa place à l'humain (non expert et expert) ...

- La pratique du machine learning consiste en un assemblage de méthodes, très empirique et non dogmatique, pilotées par l'humain
- Pour rendre les problèmes "tractables", il faut réduire les dimensions, nettoyer, rendre compatibles, piloter les pondérations, penser les questions pour qu'elles soient « calculables »

... Et deux problèmes sérieux :

- Le « deep learning » génère une opacité pour les développeurs eux-mêmes (heureusement de plus en plus travaillent à du Machine Learning interprétable)
- L'immensité des données qui peut faire croire à la toute-connaissance, la non prise en compte que le passé gouverne le présent, peuvent conduire à l'ignorance de l'ignorance !