

## **Questions autour de l'Intelligence artificielle**

Par **Tina Fischer**

### **«L'IA favorise la standardisation et non la diversité»**

**Cornelia Diethelm est experte en éthique numérique, établie à Kölliken (AG). Elle est fascinée par les progrès de l'IA, mais met aussi en garde contre les risques potentiels. Elle explique pourquoi l'intelligence artificielle met en danger la diversité et renforce les discriminations.**



**PME : l'intelligence artificielle est sur toutes les lèvres, l'enthousiasme est grand. Qu'est-ce qui explique cette fascination ?**

C. Diethelm : La fascination est d'ordre technologique. Les gens peuvent soudain interagir de manière naturelle avec une machine. Ils n'y croyaient pas jusqu'à présent et sont aujourd'hui étonnés de la rapidité avec laquelle ChatGPT, par exemple, fournit des réponses qui sont également utilisables.

**Mais malgré la fascination, ils soulignent les problèmes de l'IA.**

Oui, car toutes les intelligences artificielles - qu'il s'agisse de texte, de parole, d'image ou de musique - sont malheureusement confrontées aux mêmes problèmes.

**Quels sont ces problèmes ?**

Il y a trois grands problèmes, à commencer par le potentiel de discrimination. Très grossièrement, et je le dis sans préjugés, cela signifie que tout ce qui s'écarte de «l'homme blanc» est susceptible d'être discriminé.

**Encore l'homme blanc ?**

Oui, à nouveau l'homme blanc. C'est tout simplement la population pour laquelle il existe le plus de données. Et comme la base des intelligences artificielles est constituée d'énormes quantités de données, une IA apprend naturellement de manière disproportionnée de ce groupe.

**Vous avez parlé de trois problèmes, quels sont les deux autres ?**

D'une part, la problématique des stéréotypes. Les modèles d'IA se basent sur le passé, sur des données que nous possédons déjà. Or, dans le passé, nous vivions selon d'autres modèles sociaux, par exemple celui où la femme restait à la maison.

**Cela signifie que nous utilisons des informations du passé et que nous maintenons ainsi les stéréotypes ?**

Non seulement ils sont maintenus, mais ils sont même renforcés par l'IA et transmis dans le futur. Un autre exemple est la domination de la culture occidentale. Si l'on interroge une IA sur un

mariage, on tombe inévitablement sur le mariage blanc, alors qu'il existe tant d'autres rituels de mariage dans le monde. Nous ne percevons pas consciemment de tels éléments, mais ils nous amènent à percevoir le mariage occidental comme le seul vrai ou à faire moins confiance à une femme dans le contexte du travail, car elle est associée en premier lieu au foyer.

### **Discrimination, stéréotypes... il manque encore un problème.**

Les intelligences artificielles tendent vers la moyenne. Cela favorise la standardisation et non la diversité. La raison en est que l'IA est souvent entraînée avec des données unilatérales, c'est pourquoi elle ne connaît pas la diversité du monde réel.

### **À quoi cela est-il dû? De la pure paresse des programmeurs ?**

Les statistiques fonctionnent avec une courbe en cloche: tout ce qui se trouve au milieu est saisi, les éléments marginaux sont laissés de côté. C'est efficace. Dans le contexte de la programmation d'une IA, c'est la solution la plus avantageuse et la plus rapide pour que quelque chose fonctionne grosso modo.

### **Mais cela ne reflète pas la réalité.**

Exactement. C'est du mainstream, mais ce n'est tout simplement pas possible en ce qui concerne les humains. Dans les cas extrêmes, une IA entraînée avec des données insuffisantes sur des personnes à la peau foncée ou aux yeux asiatiques en amande, par exemple, associe ces mêmes personnes à des animaux ou ne les reconnaît tout simplement pas. Elle ne le fait pas par malveillance, mais elle n'a tout simplement pas d'autres données à sa disposition!

### **Comment résoudre ces problèmes ?**

La question est la suivante: qui peut intervenir pour corriger ce problème et comment? Une solution serait de travailler avec des jeux de données synthétisés. C'est-à-dire avec des jeux de données créés artificiellement au lieu de jeux de données réels.

### **Vous utilisez le conditionnel ?**

Il y a une tendance à ce que l'on appelle le solutionnisme. Grâce au big data, le monde doit devenir un endroit meilleur, mais pour cela, la base de données doit être correcte. Beaucoup de choses peuvent être possibles en ce qui concerne les objets matériels, mais si nous pensons que tous les problèmes peuvent être résolus avec des données, alors c'est problématique. Par exemple, le comportement et la personnalité d'un individu sont uniques et liés à un contexte particulier. Les données ne couvrent qu'une partie limitée de notre vie, c'est pourquoi nous devons toujours bien réfléchir aux domaines dans lesquels les données peuvent nous apporter une valeur ajoutée.

### **Le solutionnisme, malgré son nom, n'est donc pas la solution. De quoi a-t-on besoin ?**

Les fournisseurs d'IA doivent organiser leurs équipes de manière diversifiée, par exemple en termes de culture, de sexe, d'âge, de formation. Les équipes diversifiées sont sensibilisées et veillent à ce que les données soient elles aussi diversifiées.

### **Que puis-je faire en tant qu'individu qui utilise par exemple ChatGPT ?**

La contribution de chacun est relativement modeste. En tant qu'individu, on peut faire pression sur le fournisseur et exiger qu'il mette en place des systèmes non discriminatoires. On peut aussi donner directement un feedback sur ChatGPT en levant le pouce ou en signalant que «doctor» est traduit

par médecin, mais qu'il pourrait aussi s'agir d'une femme médecin. Cependant, il ne s'agit que d'accompagnement. Le plus grand impact se situe au niveau du développement et de l'utilisation des systèmes.

### **L'UE élabore actuellement une proposition de loi sur l'IA. Une réglementation est-elle vraiment la solution ?**

L'intelligence artificielle ne fonctionne que si nous sommes conscients des responsabilités qui en découlent. Avec l'AI Act, l'UE élabore une loi visant à réglementer les applications de l'IA qui présentent un risque élevé. Actuellement, des outils d'IA peu sérieux inondent le marché.

### **De quel type d'outils parlez-vous ?**

Dans le domaine des RH, on observe de plus en plus l'utilisation de l'IA. Mais elle est souvent construite selon des critères basiques et filtre les personnes qui ne correspondent pas au schéma. Une personne formée saurait pourtant qu'elles sont précieuses pour l'entreprise.

### **Cela se produit-il effectivement ?**

C'est arrivé à Amazon aux États-Unis. Lors du recrutement, les femmes ont été systématiquement écartées. En Autriche, l'office du chômage a essayé un système d'IA qui répartissait les chômeurs en trois groupes: ceux qui trouvent rapidement un emploi, ceux qui sont sans espoir et ceux qui se trouvent entre les deux, que l'on encadre et conseille. Problème: l'IA a été entraînée avec les données des chômeurs du passé. Il en résultait qu'une femme avec des enfants faisait automatiquement partie des cas désespérés!

### **Une certaine prudence s'impose donc aussi du côté des décideurs face à l'utilisation des nouvelles technologies ?**

Absolument! Les décideurs sont fascinés par ce qui est possible avec les données et les nouvelles technologies comme l'IA. Mais ils oublient parfois les risques et les effets secondaires indésirables. Par exemple, ils sont sensibilisés à des thèmes comme la diversité et l'inclusion et souhaitent les promouvoir. Ils utilisent alors un outil douteux, ce qui entraîne des régressions.

### **Mais on ne veut pas rater le coche et faire bonne figure.**

C'est un peu comme pour la durabilité. On a la pression de l'innovation et un ego personnel. On veut se montrer innovant, être meilleur que la concurrence. On se laisse donc guider par les nouvelles possibilités technologiques au lieu de résoudre réellement un problème.

### **Et vous, en tant qu'experte en éthique, vous freinez l'enthousiasme. Est-ce que vous freinez aussi le progrès ?**

Non, je suis moi-même fascinée par beaucoup de choses, et l'éthique ne doit pas freiner le progrès. Mais il est dangereux d'avoir aujourd'hui le sentiment que tout ce qui est technologique est un progrès. Nous devons nous poser la question de savoir ce qui améliore vraiment notre vie et ce qui ne l'améliore pas.