

Daniel Nissille le 22.11.2025

Retours de la discussion avec Madame Kerstin Preuschoff en personne:

La discussion orale n'a pas pu être enregistrée et est synthétisée ci-dessous le plus fidèlement possible.

Dans un premier temps, j'ai présenté au mieux la problématique que nous avions à aborder.

Madame K.Preuschoff a expliqué que la problématique était trop vaste et qu'elle faisait partie d'un enseignement dispensé sur au moins une année entière.

J'ai alors expliqué la démarche d'un pipeline général pour simplifier l'accès au risque:

Son retour fut alors d'expliquer que nous étions un peu trop optimistes. Pour étudier ou comprendre une problématique de risque il faut d'abord se concentrer sur un seul domaine d'étude et la catégorisation claire du risque. "Il faut toujours commencer par: quelle est votre définition du risque, de quel risque vous parlez?"

Je pense qu'à ce niveau il y a eu une discordance sur notre idée de simplifier la représentation des données d'une étude sur un risque X et le fait de concevoir un programme permettant de l'évaluer. Effectivement nous n'avons pas la connaissance pour évaluer un moindre risque présenté par une étude. En plus une démarche IA ne peut fonctionner que si elle est entraînée et les thématiques de risques étant diverses et de complexité différente il est probablement impossible d'entraîner un modèle global. La recommandation était donc de se concentrer sur un cas d'étude précis et donc un modèle spécifique. C'est-à-dire que chaque LLM doit être entraîné sur des données d'un domaine précis, ce qui impose en réalité de construire autant de pipelines qu'il y a de domaines d'études du risque. Mais cela dépassait de loin ce que nous avions entrepris.

J'ai donc abordé la thématique des démarches de risk literacy visant à représenter les données scientifiques de manière intuitive comme avec les "facts box" de Gerd Gigerenzer.

Madame Preuschoff a reconnu ne pas avoir entendu parler de ce modèle, mais la discussion fut toutefois encore plus enrichissante à ce moment précis alors que nous fûmes rejoints par les deux experts en IA.

J'ai présenté à nouveau la thématique et expliqué l'idée derrière les facts box. J'ai insisté sur le fait qu'elles ne sont aucunement un moyen de valider la qualité des données, mais permettent de les présenter avec le plus de clarté possible.

A ce point et avec le peu de temps restant, l'idée conductrice fut de nous focaliser sur un exemple le plus parlant possible afin d'illustrer notre démarche lors de la présentation. Il a aussi été évoqué qu'une bonne manière pour dédramatiser une impression de risque est de comparer ce risque avec un autre cas. L'exemple très clair mentionné fut de comparer les risques en avion par rapport à ceux de la voiture. En fait, il s'agit de contrecarrer un biais où la personne surestime des risques face à ce qui lui semble à tort moins risqué. Les facts box

permettent justement de faire cela avec une très grande efficacité et démontreront à la personne à quel point elle a surestimé les risques de prendre l'avion par rapport à ceux de monter dans une voiture. "La partie la plus risquée lors de la prise de l'avion, c'est la route entre chez vous et l'aéroport."

Le cas du scraping de données a aussi ultimement été traité, là encore il n'est pas évident de définir un protocole global, les publications scientifiques utilisant une pléthore de représentations différentes. J'ai abordé l'exemple d'utiliser des données pour connaître le taux d'emploi à partir d'un cours, donc globalement une étude de risque déguisée. Mais nous en sommes revenu à la première étape, comment obtenir et synthétiser les données. Ce dernier point étant en réalité le facteur limitant de notre "problématique".

En conclusion, nous avions peu de temps pour aborder une problématique vraiment complexe et trouver une direction solide n'a pas été chose aisée. Cet échange très constructif bien qu'un peu tardif dans la journée explique notamment pourquoi nous avons parlé de voitures et d'avions et pas vraiment de l'apprentissage du risque permis par l'accès facilité aux données qui pourrait être très enrichissant pour l'enseignement. En outre, nous avons pu pleinement comprendre quels étaient les facteurs limitants et nous concentrer sur un cas précis et concret bien que peut être un peu hors sujet en définitif.