

## Intervention de Diego Kuonen - Transcription

Quand on m'a demandé si je voulais intervenir, j'ai tout de suite dit "oui". Je veux parler de ce qui est important, je pense, à ce jour, dans le domaine de l'IA.

Voici une illustration. En allemand, mais la question c'est la notion "d'experts en IA" : on a l'impression depuis 2, 3 ans qu'il y a plus d'experts en IA que de citoyens en Suisse et c'est le problème dans le monde entier : aujourd'hui tout le monde est expert en IA.

Voici une autre illustration : quand on débat sur la thématique de l'IA il est important de débattre avec un langage commun. Un langage commun c'est aussi expliquer que l'IA existe depuis plus longtemps que ChatGPT : qu'il y a une histoire derrière, et qu'est-ce que l'IA et à quoi cela sert ? Et c'est cela que j'essaie de faire dans cette introduction avec vous.

Donc je commence avec une brève histoire de l'IA, qui débute dans les années 1943-1950. Le terme "intelligence artificielle" est apparu en 1956 : d'un point de vue marketing c'était un super terme. Mais, on ne peut pas parler "d'intelligence artificielle" sans définir ce qu'est l'intelligence. Et depuis là il y a des romans qui sont sortis, des histoires à la Frankenstein de tous côtés. C'est aussi pour cela que beaucoup d'expériences ont été menées. Mais à l'époque il manquait la puissance de calculs des machines et la disponibilité des données d'un point de vue numérique.

Puis il y a eu une grande évolution dans les années 80, où de grandes entreprises comme IBM ont investi dans cette recherche. D'un coup les ordinateurs ont commencé à battre des experts comme au jeu d'échecs par exemple. Là aussi c'est une étude d'application très concrète où l'IA fait du sens car dans les tâches répétitives, les règles sont clairement définies. Par exemple, aux échecs, vous aurez demain les mêmes règles qu'hier. Si demain, vous changiez les règles du jeu, par exemple, si le pion peut se comporter comme la tour, ce système ne va plus fonctionner. Il était clair, sur ce type d'exemples que l'on aurait très vite des succès. Là, on est dans les années 1996.

Après, il y a eu de plus en plus de développements, les grandes entreprises s'en sont mêlées. Quand OpenAI a lancé ChatGPT, c'était comme une phase de démocratisation de l'IA, mais ce n'était pas de l'IA : c'était des outils d'IA génératifs. Et cela, c'est un petit sous-ensemble de l'IA que j'essaie d'illustrer ici pour vous dans ce contexte.

Donc l'IA en elle-même, c'est un immense domaine où il s'agit à la base de systèmes informatiques, point final.

C'est une technologie où on essaie - et puis là il y a tous les aspects de l'intelligence, toutes les métaphores qui rentrent dedans -, les systèmes analysent leur

environnement, c'est-à-dire qu'ils collectent les données, qu'ils les analysent puis qu'ils réfléchissent comment réagir.

Puis à l'intérieur de tout ce domaine, il y a un petit sous-domaine qui s'appelle machine learning, qui est l'apprentissage automatique. Il fait ce jeu, en prenant des données, en créant un modèle puis en faisant des prévisions.

Par exemple, si je prends ici cette salle, je vous regarde, je regarde ici et là, et puis le modèle, en collectionne des données : je vous regarde, je collecte les données puis je développe le modèle : là, la majorité, c'est des hommes, ils ont des lunettes. Ce modèle, je l'applique de l'autre côté, et c'est ça que fait l'apprentissage automatique : c'est trouver des choses puis essayer de les appliquer quelque part ailleurs.

Donc ça, c'est l'apprentissage automatique. Pour le faire, il me faut des modèles statistiques, il me faut des modèles mathématiques. Et puis il y a une immense quantité de ça. Puis à l'intérieur de ceci, il y a des réseaux neurones. Puis à l'intérieur, là, il y a ce qu'on appelle les réseaux neurones profonds. Ça, c'est le *deep learning*, ce qu'on appelle *l'apprentissage profond*.

Donc vous voyez, c'est très grand. Maintenant, on vient à ce petit morceau-là que vous êtes en train de voir maintenant avec l'IA générative. Et puis l'IA générative, ça fait encore plus, ça va sur le traitement du texte. Donc vous avez la chose plus bas, vous avez encore certaines interactions sur le traitement du langage naturel, et puis c'est l'apprentissage profond.

Et puis là, vous avez ce qu'on appelle de *grands modèles de langage*. Et puis ça, si vous voulez, c'est l'engin qui est derrière un ChatGPT. Des modèles comme ChatGPT 4 ou des trucs comme ça. Donc vous voyez, ça c'est un petit sous-domaine. Et qu'est-ce qu'ils font, ces modèles ? Ils font une prévision de mot après mot.

Maintenant, la différence par rapport aux autres succès : historiquement, on pourrait voir si c'était faux ou pas. Si vous jouez aux échecs, vous pouvez voir si vous avez gagné la partie ou pas. L'ordinateur, vous avez "oui" ou "non". Si vous cherchez un criminel, vous pouvez savoir, une fois que vous l'avez, si vous avez le faux ou le vrai. Donc vous pouvez voir si vous avez fait juste.

Avec tous ces outils d'IA génératifs, il n'y a pas ça. Cela fait du sens pour nous, ça a l'air intelligent, mais vous n'avez pas une preuve de si c'est vrai ou faux.

Il y a maintenant cette démocratisation avec une société, des décideurs, même dans le milieu académique, avec des gens qui n'ont jamais appris à travailler avec de telles technologies. Donc, ce qui est important, les messages à faire passer avec l'IA, c'est :

Premièrement, que c'est un processus de résolution de problèmes, ça peut être un problème sociétal, un problème environnemental, un problème à résoudre : à un certain moment on utilise une technologie, et l'IA fait du sens.

En effet de nos jours, vous avez beaucoup de données, vous avez une puissance du calcul, vous avez ce qu'on appelle une stabilité, vous avez les mêmes règles demain, ou hier que demain, aussi dans le futur, et puis vous avez ces étages répétitifs.

Et puis ces technologies, on a toujours eu, je pense tout le monde : il y avait le vélo, il y a eu maintenant le e-vélo, on peut toujours aller à pied. Mais si on a un e-vélo, on prend l'e-vélo. Donc on utilise la technologie. Ceux qui font le gâteau à la maison, vous pouvez toujours faire votre gâteau à la main. Vous prenez un mixeur ou un truc comme ça, on s'est habitués.

Donc c'est exactement la même chose. C'est clair, la vitesse de croisière va très vite maintenant parce qu'il y a cette démocratisation.

Ce qui est important, c'est que ce n'est pas artificiel et que ce n'est pas intelligent. N'oubliez pas, c'est inspiré par les humains. Ça va être créé par les humains. Ça influence les humains.

Donc, sur les humains, ça ne va pas fonctionner. Et puis, ça ne consiste pas à remplacer l'humain par un robot. Il s'agit de faire sortir le robot de l'humain. Ça ne va pas détruire des emplois. Mais chaque emploi, c'est un processus qui a certaines tâches. Et puis, certaines tâches, de nos jours, on peut les faire de manière plus efficace, plus rapide. C'est pour ça qu'on peut toujours s'écrire des lettres par papier. De nos jours, on s'envoie un email, on s'envoie un WhatsApp, mais c'est juste passer un message de A à B.

Et puis, ça fournit, et ça, c'est très important, des outils d'aide à la décision.

Pour qui ? Pour l'être humain qui doit réfléchir si ça fait du sens ou pas. Et puis maintenant le message clé que je veux faire passer, et que je souhaite que vous reteniez, c'est que ce qui sort c'est probablement faux.

Nous, comme êtres humains, on le sait, comment on peut contrôler. Tout le monde connaît les poissons du 1er avril. Je pense que c'est le seul jour de l'année où on ne croit pas à l'information qu'on lit.

Prenez cette approche, quand vous utilisez un ChatGPT ou un autre outil : ce qui sort est probablement faux.

Et puis, ce qui n'est dit nulle part, c'est que pour le faire fonctionner, il faut des données. Ce ne sont pas des technologies qui tombent du ciel et puis ça fonctionne. Ce qui rentre, ce sont des données.

Comme vous voyez ici [sur la diapositive] au milieu c'est la magie qui est faite avec les données, entrées et ce qui sort, ce sont des données. Si vous prenez un ChatGPT, ce que vous rentrez, c'est du texte. Magie. Ce qui sort, c'est un texte : les données rentrent, données-magie, les données sortent.

Donc, d'abord, parler d'une IA digne de confiance, c'est une méthode mathématique. Donc, par rapport à la question, il faut avoir des données dignes de confiance. De nos jours, tout le monde utilise l'IA. Si on a n'importe quel type de données, qui ne sont pas dignes de confiance, qui n'ont pas la qualité suffisante, c'est un problème. Personne ne saute dans une piscine si l'eau est sale. Cela, personne ne le fait.

Mais de nos jours, avec ce souhait d'utiliser la technologie, cette réflexion sur les données n'est pas faite. Donc, ça m'amène à la question, avec un langage commun, c'est quoi les données ?

Les données, c'est ce qu'on appelle le carburant de l'IA. Données plus puissance calcul plus algorithmes, c'est l'intelligence artificielle, plus ou moins la technologie.

Et puis qu'est-ce que c'est les données ? Depuis mai de l'année dernière, il y a un langage commun en Suisse. L'Académie suisse des sciences avait défini les données de cette manière ici. Les données sont disponibles sous divers formats, sous forme de chiffres, textes, images, vidéos ou sons.

Donc ce que je fais ici, je produis des données. Ce que vous faites en me regardant, vous consommez des données. Mon image, mon son. Et puis, il y a toujours cet aspect de consommation, de production de données. Si je raconte une blague et puis vous riez, vous consommez ma blague principalement, vous produisez des données si vous riez, moi je vois, et puis si vous riez, ça montre que c'est une bonne blague, une blague de qualité. Puis là, on a cette notion de qualité qui est fondamentale. Il faut demander à celui qui consomme ce dont il a besoin comme données.

Ça c'est un truc de base, la qualité, qui s'applique aussi à la donnée. Donc ne réinventez pas la roue, je pense qu'en Suisse il y a un langage commun sur ce que ça veut dire, "donnée" Ici il y a celle de l'Académie de Suisse des Sciences, j'avais le plaisir, je suis co-initiateur et co-auteur de la charte suisse de la littératie des données, ça c'est un langage suisse. Donc avec ça vous pouvez commencer.

De la même manière, il y a cette discussion du jour : il faut former tous nos collaborateurs en IA. Après, quand vous regardez ce que vous allez faire comme cours, prompt engineering, cool. Produire des meilleures données. Pourquoi ? Parce que ce qui sort, ce n'est pas bon. Et puis, l'IA, principalement, c'est l'utilisation d'outils IA comme ChatGPT. L'IA, c'est beaucoup plus.

Donc, comment vous voulez former les gens en ce qu'on appelle la littératie en IA, s'ils n'ont pas la littératie des données ? Ça ne va pas fonctionner comme ça.

Et puis là, aussi sans littératie des données, il n'y a pas de littératie de l'IA, c'est aussi pour ça qu'il y a cette charte suisse de littératie des données où vous avez une définition qu'est-ce que c'est la littératie des données, mais ça prend aussi en compte les principes éthiques et la protection des données. Et aussi, le plus important dans la charte, c'est aussi clairement écrit, il faut aussi tenir compte des aspects économiques, sociétaux et écologiques.

Ici, je parle de l'IA écologique, c'est-à-dire là aussi sociétaux. Si vous saviez les histoires, comment ces modèles comme ChatGPT sont faits avec des contrats au Kenya, des gens qui étiquettent les données (data labelling) etc. L'aspirateur vous l'amène à la maison, s'ils n'arrive pas à tourner, l'image sera envoyée au Kenya. Là, il y a quelqu'un qui va recevoir un centime par heure ou un truc comme ça. Il va dire, c'est une chaise ou une table. Donc, il faut faire cette réflexion-ci. Cela a un impact environnemental, sociétal énorme.

Et puis énergétique aussi. Moi je suis dans un canton où ils disent, "il faut faire attention". Je suis valaisan, il faut faire attention avec l'eau. Mais de l'autre côté, ils disent, utilisez ces outils. Donc là, il y a aussi, le bilan Energie qui n'est pas terrible actuellement sur les outils d'IA génératifs.

Donc, sans littératie des données, il n'y a pas de littératie en IA. Ça m'amène à la question de la confiance. Ce souhait d'une entreprise, d'une société, même d'un gouvernement de se dire "confiance", oui, il y a des labels, mais pour ça, vous ne pouvez pas acheter. Il faut démontrer que vous êtes digne de confiance. Et puis ça, c'est quelque chose qui est important, qu'une entreprise, un gouvernement, quand ils le font, oui, ils font l'IA, ils appliquent des algorithmes. L'Hospice a aussi des idées, tout le monde a des idées, mais il faut qu'en IA on puisse démontrer qu'on est digne de confiance. C'est-à-dire, on démontre quasiment, on explique aux gens de le faire.

Donc on a quelque part un côté bonnes pratiques à disposition, et puis ça c'est ce qui était mentionné en l'intro : j'ai eu le plaisir pendant longtemps au plus haut niveau de faire des choses pour le Conseil fédéral, là c'était le chef du projet, quand j'étais du côté bonnes pratiques, la Confédération pour les sciences de l'énergie, centré sur l'être humain et digne de confiance, où c'est vraiment l'idée, si on veut être digne de confiance, il faut démontrer qu'on l'est. Si on le fait, on le fait d'une manière en respectant des principes de base.

Et puis ces principes de base, ça doit être la gouvernance de l'entreprise. Et puis à la Confédération, c'était protection des données, sécurité de l'information, sécurité des

données, gouvernance des données, non-discrimination, explicabilité, traçabilité, transparence, neutralité, objectivité, traitement éthique.

Et puis selon l'application, une de ces dimensions est plus importante qu'une autre. Vous n'allez jamais avoir tous au maximum, c'est impossible. Donc il faut trouver un compromis.

Et puis là, on est de nouveau dans un monde de qualité. Qualité, il faut faire cette réflexion, à quelle fin, pour quel but précis on le fait. Selon ça, des cadres comme les cours de bonne pratique, aussi des choses qui sont au canton de Genève, aident l'utilisateur à avoir une sorte d'orientation dans laquelle on peut naviguer.

Et puis je pense qu'avec ça, on peut montrer, on peut se faire gagner la confiance en montrant qu'on est digne de confiance. Et puis là aussi, si vous regardez ce code des bonnes pratiques, c'est un document, je pense, de 30 pages. Il y a des définitions, qu'est-ce que ça veut dire "traçabilité", qu'est-ce que ça veut dire "transparence", "explicabilité" et tout. Tout ça, c'est documenté pour vraiment créer ce langage commun.

Là, il y avait deux possibilités. Si je regarde les représentants du canton, vous pouvez réinventer la roue. Je connais des cantons qui le font. Ou bien vous vous inspirez de ce qu'a fait la Confédération.

Où le futur va aller, je n'ai aucune idée, je ne suis pas futuriste. Pourtant avec l'IA on peut faire très vite beaucoup d'argent si on veut faire des futuristes, mais je pense qu'il ne faut pas oublier que l'IA ne va pas partir. Cela va être dans notre quotidien, comme l'énergie et tout.

Par contre, nous, on n'a jamais appris à utiliser ça. On nous a appris à l'école à lire, à écrire, mais jamais à avoir cette littératie des données. Donc c'est ça qui manque actuellement. On ne nous a pas appris, notre génération, on a peut-être une idée. Donc il faut préparer la future génération pour avoir vraiment une société qui est aussi inclusive, une société qui fait confiance aux entités qui sont dignes de confiance de fait.