



Pénélope KOMITES
Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

AUDITIONS D'ACTEURS DU SECTEUR DE L'INTELLIGENCE ARTIFICIELLE (IA)

AUDITION #5
L'IA ET LA SECURITÉ

27 mars 2024

Intervenants :

- Renaud ALLIOUX, CTO et co-fondateur, Preligens
- Olivier BABEAU, Président, Institut Sapiens

Groupes politiques :

- Jean-Pierre GRYSO, Collaborateur, Groupe Les Ecologistes
- Jérôme GLEIZES, Conseiller de Paris, Groupe Les Ecologistes
- Emile MEUNIER, Conseiller de Paris, Groupe Les Ecologistes
- Lucas ESTAGNASIE, Conseiller technique, Groupe Indépendants et Progressistes
- Jean Noël AQUA, Conseiller de Paris, Groupe Communiste et Citoyen
- Jeanne BES DE BREC, Secrétaire Générale, Groupe Modem, Démocrates et Ecologistes
- Inès SLAMA, Collaboratrice, Groupe Communiste et Citoyen

Adjoints ou leurs cabinets :

- Pénélope KOMITES, Adjointe à la Maire de Paris chargée de l'innovation, de l'attractivité, de la prospective Paris 2030 et de la résilience - Conseillère de Paris et du 12^{ème} arrondissement
- Johanne KOUASSI, Conseillère de Paris déléguée auprès du premier adjoint, en charge des relations avec les arrondissements et les services publics de proximité.
- Kevin REVILLON, Directeur de cabinet de Pénélope KOMITES
- Alexandra MEDER, Collaboratrice de Pénélope KOMITES
- Honoré BOUSSAMBA, Chef de cabinet, cabinet de Nicolas NORDMAN

Administration :

- Ottavia DANINO, Chef de projet Innovation – bureau de l'Innovation, DAE
- Thierry WEIBEL, Inspection Générale



Pénélope KOMITES
Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

Intervention de Renaud ALLIOUX CTO et co-fondateur, Preligens

En introduction, ce que je vais dire dans cette conférence reflète mes avis et ma position. Preligens est une société que nous avons créée pour la défense et le renseignement extérieur. Nous travaillons beaucoup avec le ministère des Armées. Nous ne travaillons pas avec la police, la municipalité ou la gendarmerie. C'est à garder en tête, parce que nous avons beaucoup travaillé sur des sujets liés aux données satellites, aériennes, qui sont plus utiles pour du renseignement extérieur ou des opérations militaires, que pour de la police ou de la sécurité, où nous pouvons mettre des gens au sol.

Je ne suis pas un expert en sécurité intérieure. Cela fait huit ans que nous travaillons sur l'IA. Notre société compte environ 300 personnes et environ 30 millions d'euros de chiffre d'affaires, avec plusieurs solutions opérationnelles en France, mais aussi à l'étranger. Nous travaillons avec l'OTAN, les États-Unis, le Japon, pour ce que nous pouvons communiquer.

Dans le domaine du renseignement et dans le domaine militaire, il existe de plus en plus de données. Chaque blindé, hélicoptère, bateau ont une caméra et de nombreux moyens d'embarquer des données. En voyant les hélicoptères de la gendarmerie, la police municipale, les policiers eux-mêmes qui ont des caméras ou des téléphones, nous avons de plus en plus de flux de données. Cependant nous n'avons pas beaucoup de gens pour les traiter derrière. Mais, similarité entre le monde de la sécurité et le monde militaire, nous notons une accélération du tempo.

Dans le monde politique, je crois qu'il en est de même. Entre le moment où il se passe quelque chose et le moment de réaction, le temps est beaucoup plus court. En Ukraine, une frappe d'artillerie est demandée à telles coordonnées, le canon est tiré et l'obus va arriver sur l'adversaire. À la fin de la Seconde Guerre mondiale, il fallait une ou plusieurs heures entre l'ordre et le moment où l'obus arrive, aujourd'hui, en Ukraine, il faut quelques minutes.

Ces accélérations logistiques font que les données que nous allons récupérer ont besoin d'être traitées par une machine pour être synthétisées par des humains derrière. C'est tout le sens de l'IA que nous développons : de nombreuses données, des informations et l'humain doit prendre des décisions en étant mieux informé, de manière éclairée.

Nous avons beaucoup travaillé sur des images satellites. Le volume d'images satellites acquises pour faire du renseignement explose, avec un facteur 100 000. Si bien que l'analyste arrive le matin, des milliers d'images ont été prises, il faut entre quelques heures et plusieurs jours pour regarder les images. Il faut donc savoir quelle image regarder. Parfois, la personne a des commandes : elle doit regarder telle base militaire, mais elle a aussi besoin de ne rien rater. Ce qu'il peut se passer dans certains conflits, il ne faut pas qu'elle le découvre sur Twitter. C'est là que l'IA peut lui dire : « Attention, va mettre ton intelligence humaine à tel endroit parce que nous avons vu des choses intéressantes, anormales : des avions ou des sous-marins qui ont bougé ».

La sécurité peut être du terrorisme, de la défense, ce genre de choses. Nous avons les caméras extérieures, les réseaux sociaux, tout ce qui est flux d'information en continu. Et ces flux, nous ne pouvons pas les regarder en permanence. C'est là que l'IA que nous développons permet d'avoir des informations de façon synthétisée.



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

L'important est que l'IA reste un outil pour que les humains décident. C'est l'humain qui décide, toujours. Cela est important dans la défense, même assez codifié, et encore plus important dans la sécurité avec la police, puisque nous n'allons pas vouloir que ce soit une machine qui fasse la politique de la ville quand nous allons vouloir décider.

Ces IA que nous développons sont un peu différentes de l'IA générative parce que nous mettons au point des IA qui détectent des choses sur des images : « Là : un avion, un bateau, un tank, etc. ». Pour faire cela, nous avons identifié les bateaux, les tanks, les avions et la machine finit par apprendre ce que c'est. Lorsque nous lui montrons un flux d'images, elle finit par comprendre. Cela est totalement différent des IA génératives qui vont générer des images, du texte, par rapport à ce que va demander l'utilisateur. Et là, le risque de biais ou le risque que la décision soit prise par la machine plus que par l'humain est plus important.

Nous avons été interrogés par le ministère des Affaires étrangères, le ministère de l'Intérieur sur différents sujets. Tout ce qui est police prédictive est un sujet que nous ne traitons pas aujourd'hui, sur lequel il faut être très prudent, parce que les IA génératives ou orientant la décision peuvent le faire de manière très précise ou totalement fausse, et cela est fait de la même manière. Aujourd'hui, nous demandons à l'IA générative de produire du code, cela marche très bien. Or, il lui arrive de faire du code qui n'existe pas. Et cela avec le même aplomb que si c'était un code qui existait. C'est là que pour des systèmes de défense, des systèmes critiques, nous faisons assez attention à ces IA génératives parce qu'il s'agit d'un domaine dans lequel l'humain prend une décision influencée par une machine.

Nous avons beaucoup de données, et savoir identifier un piéton, une voiture, peut générer des biais. Si la personne a des biais cognitifs, elle va les placer dans la machine et la machine va reproduire ces biais cognitifs.

Intervention d'Olivier BABEAU, Président, Institut Sapiens

Je ne suis pas sûr de ce que vous avez envie d'entendre par « IA et sécurité ». Évidemment nous pouvons parler de sécurité intérieure. Pour ma part, je parlerai de la souveraineté ou de la dépendance, parce que vous n'avez pas de sécurité si vous n'êtes pas souverain. J'ai une très forte sensibilité par rapport à la liberté publique et à la gouvernance de façon plus générale.

La première question est celle de la concentration des pouvoirs technologiques. L'ère des plateformes, qui a commencé il y a 20 ans, du fait de la loi de METCALFE¹ selon laquelle l'utilité d'un réseau est proportionnelle au carré du nombre de ses utilisateurs, a créé cette économie du gagnant qui prend tout. C'est l'économie *Highlander* : il ne doit en rester qu'un. Pourquoi tout le monde va-t-il sur Google ?

Parce que les résultats sont meilleurs. Pourquoi les résultats sont-ils meilleurs ? Parce que tout le monde va sur Google. C'est un cumulatif autoentretenu. Donc, le meilleur est meilleur et Bing est nul.

¹ La loi de Metcalfe est une loi théorique et empirique de l'effet de réseau énoncée par Robert Metcalfe au début des années 1980, fondateur de la société 3Com et à l'origine du protocole Ethernet. L'utilité d'un réseau est proportionnelle au carré du nombre de ses utilisateurs



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

Si vous êtes deuxième, vous n'existez même pas. Par conséquent, il existe une concentration extraordinaire de pouvoir. Nous allons tous devoir les utiliser ces technologies. Donc, vous allez devoir vous en servir auprès de gens qui sont ailleurs, en l'occurrence sur la côte ouest des États-Unis. Ils ont leur propre vision, leur propre agenda. D'autres sont sur la côte Pacifique, ils ont aussi leur agenda et leur vision.

En tout cas, une chose est à peu près certaine : nous n'avons pas en Europe les technologies. À un moment, il va falloir acheter à l'extérieur. Nous avons eu un doute au moment de l'arrivée de l'IA générative, qui apparaissait comme un *big business*. Allons-nous pouvoir faire de l'IA générative dans notre garage ? La réponse est non : cela coûte plusieurs dizaines de milliards. Pour rappel, le CNRS a un budget de 4 milliards d'euros, celui d'Amazon est de 40 milliards de dollars. Plus d'une centaine de milliards sont jetés dans l'IA générative depuis quelques mois en investissement. Et nous ne pouvons pas rivaliser avec cela. C'est un grand *business*, avec des outils d'une puissance extraordinaire qui progressent, avec des effets de concentration, de goulot d'étranglement dans la production, dont la bonne illustration est Nvidia ; laquelle est passée devant Saudi Aramco (entreprise de pétrole saoudien) en capitalisation boursière. C'est un symbole.

Jusque-là, cette entreprise était une des dernières qui n'étaient pas *tech* et qui était encore dans le top des plus grandes entreprises en capitalisation boursière. Le fait que Nvidia, qui n'était rien, soit passé devant témoigne de façon symbolique de la concentration du pouvoir. Aujourd'hui le nouveau pétrole est le GPU², tout le monde a besoin du GPU pour les cartes graphiques, pour entraîner les IA, donc tout le monde doit utiliser pour l'instant Nvidia. Ma première idée est la concentration de pouvoir. Et le pouvoir n'est pas chez nous.

Deuxième chose : toutes les entreprises se posent la question : « Que faisons-nous ? », dans tous les secteurs. Banques, divertissement : « Quel est l'avenir, l'avenir du travail ? » La banque a beaucoup de questions de sécurité, les assurances, le transport logistique, etc. Tous se posent ces questions avec des outils que nous aimerions bien intégrer chez nous. Mais quand vous leur donnez les informations, vous ne savez que trop ce qu'elles deviennent. Ou bien vous ne le savez que trop, donc vous avez un grand problème de sécurité. Peut-être que vous avez l'appli Open AI, vous pouvez lui poser des questions comme moi, peu m'importe, je n'ai pas de question stratégique. Mais si vous êtes une grande banque ou naturellement une entreprise de défense, jamais vous n'allez lui poser de questions. Cela implique de développer des systèmes fermés, ce qui est plus long et plus coûteux.

Autre préoccupation : je ne vois pas malheureusement comment une société d'hyper sécurité n'est pas une société d'hyper surveillance ; donc une société où la liberté humaine s'amenuise. Le cas paroxystique est le phénomène chinois avec la cotation sociale qui existe. Cela est vrai. Je connais quelqu'un qui vit en Chine et qui a vu sa note baisser parce qu'il n'avait pas traversé au vert, avec ses enfants. Il était tout de suite identifié, le système est d'une efficacité redoutable. Derrière, vous ne pouvez pas mettre vos enfants à l'université, vous ne pouvez pas prendre l'avion parfois, vous êtes exclu. C'est un système extrêmement puissant d'emprise. Même Michel FOUCAULT n'aurait jamais pu imaginer que c'était possible, puisque vous avez une forme de centralisation.

² « Graphics processing unit » ou carte graphique.



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

Ma grande inquiétude porte sur les démocraties libérales, malgré les meilleures intentions. Il est vrai qu'il existe des menaces terroristes, personne n'a envie que ses enfants soient victimes d'un fou, nous avons envie que cela soit contrôlé. Nous aurions même envie d'aller vers *Minority Report*. Si nous pouvions arrêter les gens à l'avance, à la limite tant mieux, parce que rien n'est plus scandaleux qu'un attentat qui a lieu. Donc, comment articuler liberté publique et surveillance ou sécurité ? Je n'ai pas de réponse.

Autre préoccupation autour de la post-vérité et des capacités de manipulation. Cela a déjà commencé. Les croix gammées étaient russes, cela a été prouvé. S'agissant des fermes à trolls, si vous regardez des vidéos, vous avez l'impression qu'il s'agit de vraies personnes qui parlent, mais elles ont été entièrement créées par l'IA. Elles peuvent s'exprimer dans toutes les langues, quasiment à l'infini. Donc l'influence sur la façon dont l'opinion publique se forme est énorme.

Il faudrait mettre en place rapidement non pas l'authentification (je suis plutôt pour l'anonymat en ligne, quand vous êtes en entreprise, vous avez envie de pouvoir parler, mais pas au nom de votre entreprise), en revanche il faudrait un certificat d'être humain. Il faudrait être sûr qu'il s'agit d'un être humain. Le test de Turing,³ c'est fini. Aujourd'hui, nous pouvons identifier qu'il s'agit d'une machine parce qu'elle est trop parfaite. Si elle sait répondre à toutes les questions dans toutes les langues, nous pouvons penser que c'est une machine. Mais il faudrait un moyen d'authentifier les images et les vidéos. Vous avez tous vu les images inventées *ad libitum*. Le cerveau humain adore croire des choses, il est très crédule, il adore croire des choses même s'il sait que c'est faux, cela lui fait beaucoup de bien. Concernant les *fake news*, les gens ne veulent pas être détrompés, cela ne les intéresse pas. Donc, une fois que vous avez enlevé une fausse nouvelle, vous aurez toujours quelqu'un pour imaginer que Madame MACRON est un homme, que la Terre est plate, que Kate est morte et a été remplacée par un *cyborg*, etc. Ce ne sont que des exemples.

Beaucoup de gens ont intérêt à déstabiliser la France en particulier et l'Europe, et sont très imaginatifs. Demain, il faudrait être capables de prouver que nous sommes des êtres humains. Cela limiterait beaucoup. J'aurais dû dire au début que le grand problème des réflexions sur l'IA est que le mot « intelligence », en France, est très lourd. Il fait peur. Or, il faut le comprendre au sens anglo-saxon, de « traitement de l'information », rien de plus. S'agissant de la question : « L'IA a-t-elle une conscience, va-t-elle vouloir nous contrôler ? », nous n'en sommes pas là. Peut-être que cela arrivera plus vite que prévu. En revanche, le code, c'est la loi.

Nous suivons de plus en plus des décisions d'algorithmes : quand le taxi suit le GPS qui parfois se trompe par exemple. Vous connaissez cette fameuse expérience menée chez les chauffeurs londoniens, montrant que la taille de leur hippocampe a diminué depuis qu'ils utilisent le GPS pour conduire à Londres. L'hippocampe est la partie du cerveau qui stocke la mémoire faciale. Le grand problème n'est pas la technologie, mais l'être humain. Nous adorons déléguer notre effort aux autres. Nous adorons qu'une machine puisse se souvenir à notre place. D'ailleurs une étude a montré

³ Le test de Turing vise à déterminer si une intelligence artificielle est capable de raisonner à la manière d'un être humain. Ce test, qui met en scène deux humains et un ordinateur, est à la base des recherches sur l'IA et trouve encore son application aujourd'hui



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

que les jeunes qui utilisent beaucoup les portables ont moins de mémoire et moins de capacités de concentration.

Jusque-là, avec Google, vous appuyiez sur le moteur : il vous montrait une liste où chercher la bonne réponse, et vous compariez. Maintenant, c'est la loi du prophète : nous sommes contents d'avoir la réponse à nos questions, nous allons faire copier-coller et nous allons nous laisser aller. C'est le sens de la plus grande pente. Parce que les machines vont vous proposer de décider à votre place : de votre consommation, de la bonne personne à élire. Nous avons de nombreuses bonnes façons d'orienter les décisions et les comportements.

De plus en plus, la responsabilité va nous échapper. 80 % de ce que vous regardez sur Netflix vient d'algorithmes de recommandation. Nous pouvons nous faire influencer sans même en avoir l'impression. La décision à laquelle nous tenons le plus fort est celle que nous avons l'impression d'avoir prise librement. L'illusion de la liberté est donc très puissante. Or, l'algorithme décide de plus en plus, que nous le voulions ou non. Ce n'est pas lisible donc vous avez un grand problème de compréhension de la logique de l'algorithme. Quelle logique me donnera Chat GPT ? Comment est-elle orientée ? Et est-ce auditable ? Pouvez-vous ouvrir le capot ? Chat GPT 3.5, c'était 175 milliards de paramètres. Chat GPT 4 que j'ai sur mon téléphone, c'est 1 700 milliards de paramètres et Nvidia a annoncé qu'ils allaient pouvoir faire marcher une architecture à 27 000 milliards de paramètres.

Il faut prendre en compte l'augmentation exponentielle des capacités informatiques : 1937, le premier ordinateur du monde faisait une opération par seconde, aujourd'hui le meilleur ordinateur du monde, premier à l'exaflop c'est Frontier, qui fait un milliard de milliards d'opérations par seconde. L'Aurora a deux milliards d'opérations par seconde, soit deux exaflops. Et Amazon a annoncé 400 exaflops, donc nous devrions arriver aux hetaflops bientôt.

La loi de MOORE⁴ était empirique et s'est vérifiée : tout est allé beaucoup plus vite que prévu. Le premier génome humain, en 2000 ou 2001, coûtait 100 millions, aujourd'hui cela coûte un jean. Cette chute du traitement de l'information va de pair avec l'explosion des capacités à traiter l'information. D'ailleurs le *deep learning* est fils de cette capacité de traitement. Les réseaux de neurones existaient depuis un moment, mais nous n'avions pas les capacités informatiques pour les faire fonctionner avec la puissance qu'ils ont aujourd'hui : réseau de neurones, *deep learning*, *machine learning* sur des données non structurées. Il y a 15 ans, il ne pouvait pas reconnaître une photo de chat, aujourd'hui il peut le faire. Voilà le monde un peu inquiétant dans lequel nous vivons aujourd'hui du point de vue de la sécurité, de la liberté humaine, de la gouvernance. Cela pose des questions fondamentales.

Questions et remarques de Jérôme GLEIZES Conseiller de Paris, Groupe Les Ecologistes

Merci. Nous avons l'impression, aujourd'hui, que tous les traitements informatiques sont une IA. Or, ce n'est pas forcément une IA, même si cela aide à la décision. Ce qui fait le passage à une intelligence artificielle est une analyse un peu plus profonde qui n'est pas faite par des êtres humains. Sinon, tout le monde ferait des IA depuis très longtemps. J'aimerais aller un peu plus sur les conséquences de

⁴ Cofondateur de la société Intel Gordon Moore avait affirmé dès 1965 que le nombre de transistors par circuit de même taille allait doubler, à prix constants, tous les ans. Il rectifia par la suite en portant à dix-huit mois le rythme de doublement. Il en déduisit que la puissance des ordinateurs allait croître de manière exponentielle, et ce pour des années.



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

cette distinction, notamment par rapport à la deuxième intervention : vous avez parlé des IA qui ont modélisé la conscience. Les réseaux de neurones, au départ, étaient destinés à cela : la question de la conscience, avec des théories physiques sur la conscience. Il existe beaucoup de travaux à ce sujet en recherche fondamentale. Donc il existe plusieurs niveaux d'IA que nous ne distinguons pas assez. Vous parliez des conséquences sur les données publiques. Le système chinois peut marcher sans IA. Ce peut être uniquement du traitement d'informations brutes, avec beaucoup d'informations.

Ma question est la suivante : quelles seraient les préconisations juridiques sur la question du contrôle ou du contrôle du contrôle ? Au niveau de l'État, notre capacité à intervenir est relativement faible. La seule chose que nous puissions faire est le plaidoyer. Mais dans le système chinois, les tests ont commencé par des métropoles. D'ailleurs, le surcontrôle lié au Covid a fait imploser ce système, ce qui a entraîné une relance du Covid par ailleurs. Quelles préconisations pourrions-nous formuler vis-à-vis de cela ?

Questions et remarques de Jean Noël AQUA, Conseiller de Paris, Groupe Communiste et Citoyen

Merci pour vos interventions et vos réflexions qui nous nourrissent, nous en avons besoin pour affiner une politique publique qui soit affûtée et en phase avec la réalité du développement de ce sujet. J'ai une question à propos de l'erreur. Monsieur, vous disiez que finalement, la décision prise était toujours humaine, orientée par votre système d'analyse. Vous parlez de l'ancienne génération d'IA. Comment traitez-vous cela, comment l'analysez-vous, pour arriver à corriger ? Même si, finalement, j'entends bien que la décision n'est pas liée à la machine pour l'instant.

Par ailleurs, vous ne l'avez pas trop abordé, mais je vous pose la question à tous les deux de façon générique : en parlant de sécurité et d'intelligence artificielle, nous pensons notamment à la vidéosurveillance et à l'analyse comportementale, où désormais nous n'allons pas analyser que des photos satellites, mais des photos en temps réel, non pas pour reconnaître les individus, mais pour détecter des comportements à risque. La loi l'a autorisé à titre provisoire, par rapport aux JOP à Paris pour quelques mois. Mais nous savons bien qu'à partir du moment où nous commençons une expérimentation provisoire, il est difficile de l'arrêter. Avez-vous des éléments qui pourraient nous éclairer par rapport à ce point et comment devons-nous légiférer ?

Questions et remarques de Pénélope KOMITES, Adjointe à la Maire de Paris

Nous voyons bien les dangers potentiels d'une utilisation abusive de l'IA sur tout ce qui est prédictif et sur la reconnaissance faciale. Vous avez mentionné le fait de contrer ces dérives et conserver ce qui nous intéresse (et nous sommes une collectivité, donc pas forcément les meilleurs décisionnaires en la matière). Vous avez défendu une approche partenariale public-privé et je voulais savoir pourquoi vous faites ces recommandations.

Réponse de Renaud ALLIOUX, CTO et co-fondateur, Preligens

En matière d'IA, nous n'avons pas vraiment de définition. Ce qui est généralement admis, bien que certains ne soient pas d'accord, en parlant d'IA est de parler de *machine learning* et de *system inductive*, de la généralisation à partir d'exemple. Le système va apprendre à partir d'exemples. Ce qui explique pourquoi l'IA générative ne sera jamais intelligente, puisqu'elle ne fait que reproduire ce qui



Pénélope KOMITES
Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

est le plus probable. Il s'agit juste d'une histoire de perroquet, qui répète et engrange tant de choses qu'il répète des choses justes. Elles ont vu tout internet, donc une bonne partie de la connaissance humaine, elles ont donc appris beaucoup de choses. Mais elles sont incapables de générer des choses originales ou en-dehors de la norme.

Questions et remarques de Jérôme GLEIZES Conseiller de Paris, Groupe Les Ecologistes

Et elles peuvent répéter des erreurs.

Réponse de Renaud ALLIOUX, CTO et co-fondateur, Preligens

Oui. Donc, en général, quand nous parlons d'IA, il s'agit d'IA inductive. Ce qui pose des problèmes parce que cela est difficilement auditable. Nous savons comment cela fonctionne, ce n'est pas une boîte noire, mais ce n'est pas quelque chose que nous désignons. Si je lui apprend à reconnaître un chat, peut-être qu'au passage, elle aura appris à reconnaître d'autres choses. Telle est la limite de ces technologies.

S'agissant des recommandations juridiques, je ne suis qu'un simple ingénieur, chef d'entreprise, donc je ne sais pas, j'ai beaucoup de craintes. Étant dans un secteur régalien, je ne suis pas contre la régulation, je ne pense pas que la régulation soit un frein à l'innovation. Le secteur de la défense est l'un des plus régulés. Nous sommes souverains, 100 % français, nous ne pouvons pas prendre n'importe quels clients, nous en discutons avec la France, la DGEN et le ministère de la Défense. Pour les produits, nous pensons être au top de la technologie, même vis-à-vis des Américains, nous n'avons pas à rougir, donc il ne faut pas craindre de faire des recommandations.

Il y a un point sur lequel l'IA m'inquiète beaucoup, plus que dans le crédit social chinois (je suis peut-être naïf mais je me dis que nous sommes encore en démocratie, avec des garde-fous, une séparation des pouvoirs qui me laisse penser qu'un crédit social tel que celui-ci ne pourrait pas passer) : la manipulation de l'information et l'addiction aux réseaux sociaux me fait extrêmement peur avec l'IA générative. Un réseau social comme TikTok ou Facebook pourrait générer du contenu complètement adapté à la façon dont un adolescent surfe sur internet, lui générer du contenu et, à la fin, lui faire acheter des chaussures, parce que le but est de faire de la publicité. Déjà, les jeunes sont très « addicts » à TikTok. Et le jour où une machine va pouvoir générer le contenu spécifique pour que vous soyez ancré derrière votre écran m'inquiète beaucoup.

Je crains davantage cette manipulation de l'information par les réseaux sociaux que de la dérive sécuritaire, même si elle peut apparaître. Mais la sécurité est comme la défense : il s'agit d'un outil à la main du politique. Si nous voulons une dérive sécuritaire, nous en aurons une. Les militaires ne décident pas : ils mettent des outils dans la main du président de la République, à la fin, c'est lui qui décide. En particulier en Europe, où nous avons une protection des consommateurs, l'IA Act n'est pas mal, c'est un bon début.

Donc, ma recommandation serait de continuer à travailler sur les libertés individuelles, avec l'IA Act qui interdit les notions de surveillance de masse via l'IA, et de sensibiliser les pouvoirs publics aux dangers quand nous parlons technologies.



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

À propos de la reconnaissance de comportements, j'ai un ami qui travaille là-dessus. Aujourd'hui, leurs détecteurs cherchent à éviter des bagarres, ce qui n'est pas mal. La même technologie pourrait être développée pour des gens qui ne traversent pas bien au passage piéton ou qui ont des comportements antisociaux. La technologie est juste là pour faire ce que nous lui demandons de faire. Concernant l'erreur : c'est un sujet de l'IA. Il existe même un théorème mathématique derrière : l'IA ne peut pas avoir 100 % de bonnes réponses. Obligatoirement, vous allez vous tromper à un moment. Avec également la question d'où et comment est utilisée l'IA.

Mettrions-nous une IA dans un missile nucléaire ? Non, parce qu'il ne faut pas se tromper. Et le missile nucléaire ne marche pas avec un système inductif. En mettant de l'IA pour reconnaître des données sur des images satellites, il est possible de se tromper versus ne pas le faire. Faut-il mieux envoyer une alerte fausse ou pas d'alerte à l'humain, pour qu'il regarde une image satellite ? L'IA n'est pas adaptée à tous les cas d'usage. Elle est beaucoup utilisée dans le secteur médical, avec beaucoup de succès, mais à un moment, il ne faut pas qu'elle remplace le radiologue. Il faut qu'elle soit une aide à la décision, mais non pas qu'elle oriente la décision.

Ensuite, un sujet plus opérationnel est : comment valider les performances ? Ce sont des sujets que nous travaillons avec la Direction Générale de l'Armement. Nous leur envoyons un tank, pour être sûr qu'il réponde aux besoins. En matière d'IA, c'est plus compliqué parce qu'il faut être conscient du risque d'erreurs et faire avec. Ne pas dire : « Mon IA ne fera jamais d'erreurs », mais plutôt : « En cas d'erreur, que se passe-t-il ? ». Le transport aérien fonctionne de la même manière : deux pilotes sont là pour minimiser l'impact. Le raisonnement est le suivant : « Attendons-nous à des erreurs donc comment optimiser la chaîne de décisions pour minimiser les impacts ? ». Tout cela doit être testé pour mesurer la performance et la part d'erreurs. C'est tout un sujet, qui n'est pas clos et qui est même un sujet de recherche pour certains.

Nous avons toujours été vigilants dans les applications. De même que pour l'avènement des JO, le terrorisme, etc. Dans le domaine de la défense, nous avons un réchauffement des conflits : nous voulons mettre de l'IA partout, mais attention : c'est là qu'arrivent les bêtises. Pour la surveillance des JO, nous allons faire sauter une sécurité éthique, juridique et je pense que ce n'est pas toujours bien de le faire. Parce que nous mettons le pied dans quelque chose, puis, nous ne revenons pas dessus. Dernier point sur les partenariats public-privé et ses dangers : je pense qu'il est très important, dans les sujets régaliens ou les sujets de santé, d'avoir une co-construction forte de ces solutions, puisque lorsque nous allons voir un militaire, un adjoint ou un policier, nous allons lui dire : « Que veux-tu en IA ? » et globalement, il s'en moque.

Il veut une solution à ses problèmes. Pour cela, il faut construire des compétences en interne dans les différentes organisations. Côté défense, l'agence de l'IA a été créée récemment par Sébastien LECORNU⁵ et je pense que c'est plutôt une bonne idée parce que cela permet d'avoir des acteurs qualifiés en face, avec lesquels nous allons pouvoir nous comprendre, construire des solutions, des régulations et faire des recommandations aux politiques qui eux, vont faire des recommandations à leur autorité de tutelle avec le savoir-faire industriel, technique, doctrinal, etc.

⁵ Agence ministérielle pour une IA de Défense (AMIAD)



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

Pour finir, sans parler de bord politique, sur ces sujets-là qui sont importants et graves, nous ne pouvons pas dire : « Nous ne mettons pas d'IA dans la défense ». De même pour la sécurité, nous en parlons moins, mais avec les catastrophes naturelles, le réchauffement climatique, etc., ce serait une erreur stratégique de se passer de l'apport de l'IA. Mais ce n'est pas à Google, Facebook, Mistral ou PreliGens de vous dire quoi faire. Nous sommes là pour vous fournir des capacités, pour vous aider à comprendre ce qui est derrière. D'où le partenariat public-privé, parce que des décisions doivent être prises et plus que de laisser le marché *d'open source* décider, il faut voir qui pense quoi.

J'ai beaucoup de respect pour Yann LECUN en tant que chercheur, il est extrêmement brillant. Mais il travaille chez Facebook. Donc derrière son avis, il y a l'avis de Facebook. Ce qui n'est pas anormal. Derrière mon avis, s'entend celui de PreliGens, c'est-à-dire que j'y crois, ce n'est pas machiavélique. Mistral est derrière l'avis de Cédric Ô. Derrière ces avis, il existe des intérêts privés et à la fin, l'intérêt général doit primer. Ce n'est pas à moi de décider.

Réponse d' Olivier BABEAU, Président, Institut Sapiens

Avant, « L'IA » était un abus de langage, mais aussi un élément marketing qui permettait de vendre plus cher un programme. Nous disions aussi qu'il fallait taxer les robots, mais le mixeur ou le grille-pain sont-ils des robots ? C'est la question de la complexité. À partir du moment où la machine n'est plus un programme mais qu'elle va elle-même se programmer, qu'elle va faire évoluer elle-même les curseurs sur des milliards de couches pour arriver à la décision, là nous pouvons parler un peu plus de l'intelligence artificielle au sens de traitement de l'information. Mais qui gardera les gardiens, qui contrôlera les contrôleurs ? Pouvons-nous décider ce que nous décidons ? Si vous avez une machine qui fait un traitement extrêmement complexe, quelle est votre capacité à dire si la décision est bonne ou mauvaise ? À la fin, vous êtes beaucoup entre les mains de la machine, même si vous êtes décisionnaire.

Nous avons deux risques : le faux positif et le faux négatif. Le faux positif : vous avez pensé que c'était une bagarre et vous vous êtes trompé. Tant mieux. Le faux négatif est plus gênant : vous avez raté le coche. C'est le trou dans la raquette. Déjà, la fiabilité vis-à-vis de l'IA est bien supérieure à celle que nous avons vis-à-vis de l'humain. Des personnes qui se tuent en voiture, nous en voyons tous les jours. En revanche, le jour où cela sera la faute de l'IA nous dirons : « Non, ce n'est pas possible ! » La fiabilité de l'IA est bien supérieure à celle de l'humain. L'exemple d'Airbus est intéressant. Aujourd'hui, j'entends qu'il pourrait décoller sans personne. Le pilote est là pour rassurer, faire l'annonce au début et agir en cas de problème. Mais concernant le Rio-Paris, en l'occurrence, le problème a été créé par l'humain. Si ça avait été l'IA, cela se serait très bien passé.

Donc, à terme, nous aurons de moins en moins de personnes dans le *cockpit*, voire plus du tout. Et ce sera peut-être une bonne nouvelle. Le problème est plus culturel. Dans 10 ans, aller voir un médecin qui ne sera pas assisté par l'IA sera peut-être aussi absurde que d'aller voir un astrologue, cela aura aussi peu de sens et il faudra que ce soit aussi peu remboursé, d'ailleurs. La question est celle de la complémentarité avec la machine.

S'agissant des hallucinations, c'est tout à fait vrai. Des amis notaires et avocats disent : « L'IA générative c'est très bien, mais elle invente des arrêts qui ont l'air plus vrais que vrais ». Il s'agit quasiment d'une œuvre d'art à ce niveau. Il faut avoir la couche de séniorité pour porter un regard critique sur l'IA. Mais plus ce sera complexe, plus nous aurons besoin d'IA pour analyser les rapports de l'IA.



Pénélope KOMITES
Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

S'agissant de l'apport et de l'anonymisation, l'IA dans la ville va de plus en plus avec les *smart grips* pour permettre de gérer l'énergie. Je discutais avec Veolia : nous avons de grands espoirs. Quasiment 50 % de l'eau est perdue, mais nous avons grand espoir d'optimiser la gestion de la ressource. En ce qui concerne le flux de personnes, nous pouvons très bien anonymiser et bénéficier de l'apport de l'IA. Il faut donc que les modèles soient fixés en ne prenant que l'information nécessaire, pour avoir une gestion des flux, de la sécurité, tout en respectant la vie privée. Cela ne me fait pas très peur.

Maintenant je relève un risque : vous connaissez le principe du « *Pay as you drive* ». Des capteurs dans la voiture mesurent la façon dont vous conduisez. Si vous êtes un bon conducteur, vous payez moins d'assurance. C'est très bien, mais intrusif. Demain, cela peut être « *Pay as you live* ». Cela peut être une simple montre, qui peut enregistrer une chute, par exemple. C'est cela le problème de ces technologies. Elles présentent des apports extraordinaires. Vous avez vu Neuralink. Les aveugles voient, les sourds entendent, c'est miraculeux, au sens propre. Mais la machine pourrait dire : « Tu en es à ton deuxième verre de vin, arrêtes, sinon ton assurance va augmenter. » Si votre risque est trop bien connu, les gens à risque, malades, ne pourront plus s'assurer et les gens qui n'auront pas de risques ne paieront pas parce qu'ils n'auront pas envie de s'assurer.

Nous devons nous imposer des voiles d'ignorance, comme c'est le cas aujourd'hui dans l'assurance. Et demain, il faudra s'imposer un voile d'ignorance dans les comportements, alors que nous savons que nous pouvons connaître plus de choses, que ces choses ont un intérêt, par exemple, en termes de santé. Nous utilisons très mal nos données de santé, alors que nous sommes assis sur une mine d'or par rapport à la Sécurité Sociale. En comprenant des *patterns*, demain, nous pourrions faire de la prévention de façon extraordinaire. Mais à quel moment cela devient-il un contrôle des comportements. À quel moment allez-vous rendre les gens responsables de leurs actions ? Si vous fumez, non seulement vous payez plus d'impôts en achetant des cigarettes, mais vous pouvez imaginer que demain vous payeriez des impôts pour la maladie que vous êtes en train de vous créer. Où mettons-nous le curseur ? Idem pour l'obésité.

Réponse de Renaud ALLIOUX, CTO et co-fondateur, Preligens

Par nature, la cybersécurité reste discrète, nous n'en parlons pas beaucoup. Mais c'est un sujet d'importance où l'IA aide beaucoup. Son sujet est de détecter des intrusions. Il existe de nombreuses façons de le faire et l'IA peut aider en détectant des comportements anormaux. Elle va regarder comment vous utilisez votre ordinateur, en observant les flux réseaux, et signaler s'il se passe un comportement anormal. L'IA est très puissante pour détecter que ce n'est pas normal. Nous en revenons à des sujets de liberté et de vie privée, parce que pour faire cela, l'outil doit tout regarder dans votre PC. Et plus il regarde de choses, plus il est performant pour détecter des comportements anormaux. Qui le fait, la société souveraine ?

En France nous avons de petites pépites, le problème est qu'elles sont en train de se faire racheter par des fonds américains, par Microsoft qui a tendance à racheter tout le monde. Mais c'est une application extrêmement souveraine parce que tout commence à être interconnecté. Aujourd'hui, il est possible d'*hacker* une centrale nucléaire à distance, dans les armées tout est connecté. Dans les



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

politiques publiques, c'est pareil, les caméras de surveillance vont remonter les feux de signalisation qui sont déjà connectés. Il s'agit d'un sujet très complexe, et nos adversaires Chinois et Russes regardent beaucoup cela. Des attaques se produisent très régulièrement.

Cela fait partie des priorités de se renforcer là-dessus. Mais comment être souverain en étant un petit pays par rapport à d'autres ? Selon moi, il faut créer une industrie souveraine européenne pour cela, avec une question de force de frappe. Notre société représente 200 ou 250 personnes, ce qui est peu pour une société de défense par rapport à Thalès, Safran ou Airbus. Aux États-Unis, pour Lockheed Martin, nous sommes l'équivalent d'un étage. Les sociétés de défense américaines qui travaillent sur l'IA arrivent à lever des milliards de dollars ; quand en France, si vous levez de l'argent avec des capitaux privés, avec les fonds français vous êtes limité, le seul contre-exemple est Doctolib, au-delà il faut aller voir des fonds américains. Cela dans la cybersécurité ou la défense, c'est un sujet.

Questions et remarques de Pénélope KOMITES, Adjointe à la Maire de Paris

Nous avons vu beaucoup de *start-ups* dans le domaine médical, incubées dans de grandes écoles. Les levées de fonds se font en Chine, en Inde, aux États-Unis, etc. Les directeurs de ces grandes écoles me disaient que les grands comptes français n'ont pas encore saisi l'opportunité qu'ils avaient à investir et laissent ces pépites partir.

Réponse de Renaud ALLIOUX, CTO et co-fondateur, Preligens

C'est vrai, mais avec une question de profondeur financière. Avec le PIB européen, nous pouvons être compétitifs, mais avec le PIB français nous sommes trop petits pour mettre autant d'argent. Il s'agit d'un vrai sujet : les start-up sur les sujets souverains.

Réponse d'Olivier BABEAU, Président, Institut Sapiens

Avec un problème d'ordre de grandeur. Un spécialiste de LLM de 25 ans est aujourd'hui recruté à plus de deux millions de dollars. Le CNRS va le payer 2 000 € par mois. Voici une histoire vraie que m'a racontée une entreprise qui fabrique des avions. Elle était très contente de proposer à un jeune ingénieur frais émoulu d'école 80 000 € ; ce qui est quand même bien pour commencer. Mais Google lui a proposé 800 000.

Que voulez-vous ? Autrefois, vous pouviez dire : « Je crée une entreprise pour devenir Google ». En réalité, vous créez une entreprise pour être racheté par Google. Et souvenez-vous, What's App a été racheté 250 millions de dollars pour chaque employé. Avec tant d'argent, vous attendez que le marché fasse ce qu'il a à faire et quand cela arrive, vous avancez vos dizaines de milliards. Et nous n'avons pas le début du commencement de quelque chose.

Je vous dis deux mots sur la cybersécurité : demain, je serai à Lille dans un grand forum qui s'appelle InCyber. C'est un monde en soi : une banque est attaquée 1 000 fois par jour. Ce sont des châteaux assiégés, une sorte de course où les moyens des attaquants montent en même temps que ceux des défenseurs. La grande angoisse est l'informatique quantique. Le rapport entre l'informatique d'aujourd'hui et l'informatique quantique est d'un à 100 000 milliards. Le microprocesseur quantique fait en 3,20 secondes ce que les meilleurs ordinateurs font aujourd'hui en 10 000 ans. L'informatique



Pénélope KOMITES

Adjointe à la Maire de Paris,
chargée de l'Innovation, de l'Attractivité,
de la Prospective Paris 2030 et de la Résilience
Conseillère de Paris et du 12^e arrondissement

quantique permettra de craquer n'importe quel code, par exemple. Beaucoup de choses ne fonctionneront donc plus.

Questions et remarques de Jérôme GLEIZES Conseiller de Paris, Groupe Les Ecologistes

Ce n'est pas encore au point avec des problèmes techniques de température.

Réponse d' Olivier BABEAU, Président, Institut Sapiens

Bien sûr. Mais plus vous augmentez notre capacité informatique, plus vous augmentez votre capacité informatique d'attaque. La technologie des attaquants évolue au fur et à mesure que les technologies de défense évoluent. Il faut donc toujours avoir un peu d'avance.

Réponse de Renaud ALLIOUX, CTO et co-fondateur, Preligens

S'agissant de cybersécurité, le premier point d'entrée reste les humains. Et l'éducation, la formation, devraient avoir lieu au lycée pour expliquer comment cela fonctionne et ce qu'il ne faut pas faire.