



***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<sup>e</sup> arrondissement

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

AUDITION #8  
L'IA ET L'ÉTHIQUE

24 avril 2024

Intervenants :

- Giada PISTILLI, Éthicienne chez Hugging Face
- Eric SALOBIR, Président de la Human Technology Foundation

Groupes politiques :

- Lucas ESTAGNASIE, Conseiller technique, Groupe Indépendants et Progressistes
- Jeanne BES DE BREC, Secrétaire Générale, Groupe Modem, Démocrates et Ecologistes
- Inès SLAMA, Collaboratrice, Groupe Communiste et Citoyen

Adjointes 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<sup>ème</sup> arrondissement
- Kevin REVILLON, Directeur de cabinet de Pénélope KOMITES
- Alexandra MEDER, Collaboratrice de Pénélope KOMITES

Administration :

- Ottavia DANINO, Chef de projet Innovation - bureau de l'Innovation, DAE

Membres du Conseil parisien des Européens :

- Karim KHAIR
- Karine SINANDER



***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<sup>e</sup> arrondissement

Intervention de Giada PISTILLI, éthicienne chez Hugging Face

Je suis docteur en philosophie et chercheuse industrielle au sein d'une entreprise franco-américaine, Hugging Face. J'y suis responsable de l'éthique et mène des recherches sur l'IA d'un point de vue éthique et philosophique. Je suis aussi en charge de la modération du contenu de la plateforme.

Éric SALOBIR, président de la Human Technology Foundation

Je préside le Comité opérationnel de la *Human Technology Foundation*<sup>1</sup>, une fondation basée au Lichtenstein, mais dont le bureau principal est ici, à Paris. Je suis aussi membre du Conseil National du Numérique. Nous travaillons aussi sur ces questions d'un point de vue éthique, très pratique et un peu anthropologique. Nous avons été auditionnés sur des choses concrètes et cela prend corps vers des choses plus opérationnelles.

Giada PISTILLI, éthicienne chez Hugging Face

Je vais d'abord vous présenter Hugging Face. Sachant qu'un grand nombre d'enjeux concernent l'*open source*<sup>2</sup> en matière d'intelligence artificielle. Il s'agit d'une entreprise franco-américaine, une sorte de *market place* complètement gratuite, puisqu'il est question d'*open source* : au sein de la plateforme, sur le site internet, nous pouvons trouver des modèles de données, des *data sets* et des applications finies d'intelligence artificielle. Donc, ces trois grandes verticales. Et « *open source* » car le code est complètement ouvert.

Qu'est-ce que signifie parler « d'*open source* » ? Il ne s'agit pas seulement d'avoir un simple code ouvert pour que cela soit facile à comprendre ou reproductible. Parce que, contrairement à des codes statiques, pour avoir accès à des codes d'intelligence artificielle, vous devez avoir accès à des données d'entraînement. Et il faut aussi de la puissance de calcul. Il ne s'agit pas simplement d'accéder à un modèle qui pèse beaucoup de gigas ou d'avoir des données de très grande taille, il faut aussi le hardware<sup>3</sup> derrière, pour supporter tout cela. Donc, cela, c'est Hugging Face. Et notre rôle est de réunir toute la communauté de l'intelligence artificielle, notamment autour de l'*open source*. Cela va des utilisateurs un peu simples, individuels, quelqu'un qui a envie pour ses études ou son intérêt personnel de construire des outils d'intelligence artificielle. Mais c'est aussi l'endroit où se réunissent de grandes entreprises : Meta, Google, etc., qui partagent des contenus en *open source*, de manière ouverte.

Cette semaine, Meta a déployé son dernier modèle de langage ouvert, Lama 3, d'ailleurs l'équipe de recherche Lama se situe à Paris et nous sommes en contact avec eux.

Hugging Face se veut le plus agnostique possible avec tous les acteurs de l'IA qui existent aujourd'hui, puisque nous ne développons pas de produits finis, notre produit est vraiment la plateforme. Donc, là

---

<sup>1</sup> « Fondation pour la Technologie Humaine ». La Human Technology Foundation, créée en 2012, est une fondation mais aussi un réseau de recherche et d'action qui place l'être humain au cœur du développement technologique : <https://www.human-technology-foundation.org/>

<sup>2</sup> Un logiciel open source est un modèle de développement décentralisé, qui distribue le code source accessible au public.

<sup>3</sup> Equipement



***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<sup>e</sup> arrondissement

où d'autres entreprises développent des produits Open AI, Chat GPT, etc., nous essayons d'être l'échiquier, l'écosystème où tout cela se déploie. Nous existons depuis 2016, mais un gros tournant a eu lieu en 2018 avec la découverte de la nouvelle architecture Transformer, qui correspond au T de « Chat GPT ». Au sein d'Hugging Face, nous maintenons cette architecture, avec des codes ouverts et c'est là que les co-fondateurs se sont dit : « Nous pouvons faire un pivot, devenir l'endroit où les gens se servent de nouvelles techniques pour assembler leurs produits et réaliser leur application d'intelligence artificielle ».

Voilà une introduction à ce que nous faisons chez Hugging Face.

Mon rôle est particulier : j'ai une licence en sciences politiques, puis j'ai fait un master et un doctorat de philosophie. Mais je m'intéresse à l'intelligence artificielle depuis un long moment. Mon équipe est une équipe interdisciplinaire de chercheurs, nous nous intéressons à des questions très larges.

Chez Hugging Face, nous sommes confrontés à l'état de l'art au quotidien, à beaucoup de questions très larges. Par exemple, nous avons affaire à des modèles de langage ou à des modèles de générations d'image, vidéo ou audio. Les derniers développements d'intelligence artificielle montrent qu'il est assez simple de reproduire de l'audio, de faire du clonage de voix, de sons, même de musique. Une entreprise a sorti son produit depuis quelques semaines, générant des musiques à part entière. Nous abordons donc des questions au sens très large.

Pour ma part, je questionne les enjeux éthiques de ces outils à chaque fois et le rôle de la plateforme. Car nous sommes une plateforme, nous avons donc un rôle d'intermédiaire, mais aussi d'amplificateur. Nous avons une responsabilité assez importante pour la société civile tout entière : nous ne pouvons pas héberger n'importe quoi. C'est aussi pour cette raison que je suis en charge de la politique interne de modération du contenu. Je travaille aux côtés de l'équipe légale pour définir les lignes directrices selon la législation et la régulation courante, selon ce qui est permis sur la plateforme ou ce qui ne l'est pas.

Il est intéressant de souligner le rôle de l'*open source*, aux niveaux académique et industriel. Depuis quelques années, nous remarquons que les gros développements, les gros tournants techniques viennent surtout du côté des industries.

D'un côté, les universités manquent de financements pour financer de grosses expériences, au niveau mondial. La France est une des excellences recherchées en matière de modèles de langage. Il est intéressant d'avoir un pion *open source*. Il existe des manières de faire, et ce, de manière ouverte.

Je pense à Meta qui depuis deux ou trois ans déploie ses modèles de façon ouverte. Il s'agit d'un choix stratégique : être le plus accessible possible. Mais comme vous l'imaginez bien, surtout d'un point de vue éthique, nous ne pouvons pas laisser tout ouvert à tout le monde. Les enjeux éthiques ne sont pas les mêmes pour déployer un *chatbot* de service après-vente et un chat d'interprétation de diagnostic médical. Cela veut dire que nous enverrions tout dans la nature et que nous ne serions plus responsables.

Donc, notre approche au sein de l'entreprise est celle de l'ouverture éthique.

Avec l'*open source*, nous imaginons quelque chose d'ouvert, « balancé dans la nature », dont nous ne serions plus responsables. Ce n'est pas notre philosophie : nous souhaitons offrir aux développeurs de



***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<sup>e</sup> arrondissement

l'industrie ou aux utilisateurs individuels des outils techniques pour déployer, à un certain niveau et avec un certain degré d'ouverture, leurs outils.

Par exemple, vous voulez développer un outil d'intelligence artificielle pour faire de la détection de langages toxiques sur les réseaux sociaux. Beaucoup de réseaux sociaux automatisent cela. Or, pour ce faire, l'IA doit être entraînée sur des données « toxiques » : dans le jargon du milieu, il s'agit de contenu raciste, homophobe, etc. Je n'aime pas trop utiliser le mot « biais » qui est parfois trop large, mais souvent, ce sont des données qui peuvent porter atteinte à des individus ou des groupes sociaux spécifiques. Or, si nous voulons déployer un modèle de ce type, nous nous posons la question des données : comment héberger ces données ? S'il s'agit d'un projet de recherche pure, les données ne seront pas les mêmes que pour un produit final. Nous souhaitons donner des lignes directrices et des outils techniques en particulier.

Par exemple, si vous voulez déployer un modèle qui pourrait générer du contenu toxique, le propriétaire de ce modèle va demander aux utilisateurs qui veulent accéder à ce portail leurs données personnelles, pour pouvoir tracer, savoir qui s'en sert et de quelle manière. De même pour les modèles de génération d'image de type Midjourney<sup>4</sup>, etc., qui sont hébergés aussi chez nous. Nous avons développé en interne une espèce de méta-modèle d'IA qui détecte du contenu pédopornographique ou violent. L'IA va renvoyer à l'utilisateur concerné une image noire avant de générer l'image finale. C'est un exemple, mais nous développons de nombreux outils en interne pour fournir à la communauté *open source* les outils nécessaires développer tout cela de manière responsable.

Nous collaborons aussi avec des acteurs privés ou des associations qui travaillent sur des sujets précis. Par exemple, nous remarquons le contenu généré sur tout type de plateformes : texte, audio, vidéo, etc. Une des solutions techniques intéressantes pour la détection de ce type de contenu est le *water marking* ou filigrane. Ce n'est pas la même chose de le faire avec des images, des vidéos sur lesquelles il « suffit » de mettre un logo comme Open AI le fait déjà. L'idée est de faire cela de façon plus fine, en ajoutant un filigrane au niveau des métadonnées d'un modèle ou en le rendant visible qu'à des machines, et non à des humains, pour que cela ne puisse pas être contourné. De nombreuses équipes travaillent sur cela et nous leur fournissons la puissance de calcul pour qu'ils fassent leurs expériences.

Mais que fait une philosophe là-dedans ? D'ailleurs je ne suis pas la seule au monde. Des Google, Microsoft et de grandes entreprises ont aussi...

Je connais très bien l'équipe éthique qui travaille chez Google à Londres. Ils ont de nombreux outils pour lesquels ils ont besoin d'éthiciens : tout ce qui touche au milieu biomédical, chimique, biologique. L'enjeu est de constituer des équipes interdisciplinaires pour attaquer un sujet de plein de points de vue différents. Face à un gros questionnement ou un problème, ma collègue en sciences cognitives, en linguistique computationnelle, un juriste, moi-même, quelqu'un qui fait de la recherche en droit, sommes sollicités et nous avons de nombreux points de vue. Et l'intérêt est de plus en plus grand, avec le degré d'accessibilité que nous avons vis-à-vis de ces outils et avec la prise conscience

---

<sup>4</sup> Midjourney est un laboratoire de recherche indépendant. Il produit un programme d'intelligence artificielle du même nom, lequel permet de créer des images à partir de descriptions textuelles, suivant un fonctionnement similaire à celui de DALL-E d'OpenAI



***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<sup>e</sup> arrondissement

de l'impact positif comme de l'impact négatif. Ce qui est typique de l'intelligence artificielle est sa possibilité d'être mise à l'échelle de manière très simple.

Par exemple, nous pouvons nous dire que faire des deep fake<sup>5</sup> était déjà possible avant, avec Photoshop. La grosse différence aujourd'hui est que nous pouvons faire une demande en langage naturel et mettre à l'échelle (« *scale* ») de manière très simple. Les enjeux sont éthiques, légaux et sociétaux. De plus en plus d'entreprises ont cette prise de conscience et souhaitent réfléchir à tout cela en amont, avant que quelque chose de négatif, voire de catastrophique ne se produise.

C'est dans cet esprit que j'ai été embauchée en tant qu'éthicienne.

L'éthique existe depuis le milieu du XX<sup>e</sup> siècle et comme pratique philosophique, environ depuis les années 1960. De même que les bioéthiciens ou comités de bioéthiques ont des rôles consultatifs, qui deviennent des rôles clés sur des sujets précis, l'éthique va avoir un impact sur les réflexions sociétales, pour les individus et la société et va impacter le développement législatif par la suite. Cela fait partie de la littérature du droit, aussi. C'est une manière de mieux comprendre les valeurs qui existent au sein de la société et de les transformer en régulation ou législation, quelle que soit l'échelle.

Il est drôle que l'annonce sur Albert<sup>6</sup> ait eu lieu hier, parce que mon sujet est vraiment l'IA conversationnelle, sur laquelle j'ai fait ma thèse de doctorat. Avant de travailler chez Hugging Face, je travaillais dans une startup qui développait des *chatbots* pour faire de l'information citoyenne. C'était un peu dans le même esprit.

Cela me fait donc sourire : en 2024, nous réfléchissons à la question. Puisque, d'après mon expérience, je n'avais pas eu de retours très positifs s'agissant des IA conversationnelles en contact avec les citoyens, parce que nous ne prenons pas assez en compte l'impact des erreurs.

Quelles sont les conséquences de ces erreurs sur la société civile lorsqu'il s'agit de communiquer une information ? Elles ne sont pas identiques si l'erreur porte sur la communication des horaires de la piscine communale ou sur les horaires pour se rendre aux urnes. Je n'en sais pas beaucoup plus que vous, mais le fait d'adopter cela au niveau de l'administration publique peut être intéressant. Reste à voir dans un an ou deux, pour faire le bilan et voir comment cela se passe.

Ensuite, la simplification de l'administration est un véritable enjeu.

L'IA conversationnelle a beau être très performante, des risques très concrets existent aujourd'hui. Nous parlons d'« hallucinations ». Lorsque les IA font des erreurs, cela peut avoir un impact assez important. D'ailleurs, les hallucinations ne sont pas encore bien maîtrisées. La façon que nous avons de les maîtriser est de mettre moins d'IA générative et plus de statistiques. Je parle, par exemple, des *rags*, des tables de vérités. Nous disons à l'IA : « Tu ne vas te baser que sur ce modèle spécifique » et il ne va générer du contenu que lorsqu'il ne trouve pas l'information. Mais ce contenu généré reste strictement attaché à des données spécifiques, sans partir très loin dans son « raisonnement ».

---

<sup>5</sup> ou hypertrucage, est une technique de synthèse multimédia reposant sur l'intelligence artificielle.

<sup>6</sup> Intelligence artificielle conçue pour les services publics :

<https://www.bercynumerique.finances.gouv.fr/albert-de-la-dinum-lia-francaise-pour-les-fonctionnaires>



***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<sup>e</sup> arrondissement

Aujourd'hui, la question des biais est une vraie question, qui n'est pas encore résolue, même si nous en parlons depuis 2017 ou 2018.

Mais je n'aime pas trop ce mot « biais », car d'un point de vue philosophique, tout le monde est biaisé dans la mesure où nous posons toujours une question d'un point de vue particulier.

Aujourd'hui, nous savons en prendre conscience grâce aux études qui ont eu lieu jusque-là. Mais ceux qui promettent de débiaiser des IA, je ne les crois pas trop. Parce qu'un contenu qui cesse de faire la promotion d'un point de vue va en favoriser un autre. Il est important de toujours se rappeler qu'il ne s'agit que de calculs de probabilités d'informations lues ou rencontrées auparavant.

Nous pouvons parfois tomber dans des pièges : Gemini, le dernier modèle de Google a essayé d'envisager la diversité. Il lui a été dit que la diversité consistait à prendre en compte toutes les couleurs, alors que c'est plus complexe que cela. Cela a fini par générer, par exemple, des samouraïs représentés comme des soldats de la Seconde Guerre mondiale.

Cela part d'une bonne intention, mais d'un point de vue éthique, ce n'est pas que la bonne intention qui compte, mais surtout le résultat final.

Il faut appréhender l'éthique de l'IA et se poser les bonnes questions aux moments spécifiques. Les enjeux et les questionnements éthiques peuvent ne pas être les mêmes au moment du développement, dans le choix des données, dans l'architecture et au moment du déploiement, en faisant plus de recherches utilisateurs (un élément qui manque cruellement dans le milieu de l'IA). Parce qu'il existe beaucoup de modèles généralistes et nous perdons la trace de l'origine de ce gros modèle et de la façon dont il est appliqué.

Et la question de l'interaction : pouvons-nous interagir avec n'importe quel modèle d'IA ? Deux questions se posent et sont bien distinctes, avec la question de la contextualisation. Je vais m'arrêter là.

Intervention d'Éric SALOBIR, président de la Human Technology Foundation

Je serai bref parce que, surtout, je souhaiterais que vous puissiez poser des questions à partir de ce que vous voulez en faire.

S'agit-il d'une consultation citoyenne ou bien voulez-vous bâtir quelque chose comme Montpellier qui a fait des *guidelines* pour le déploiement de solutions à destination des citoyens ?

Selon moi, il n'y a d'éthique que de pratique, de cas d'usage. Parler de réduire le biais est intéressant, mais cela reste très générique. C'est à partir du moment où nous avons des cas pratiques que nous pouvons commencer à mener une réflexion.

De façon générale, l'éthique est moins une science du bien et du mal qu'une capacité à résoudre des conflits de devoirs, entre différents valeurs et devoirs que nous avons, vis-à-vis des citoyens.

La question de la pondération entre la sécurité et la vie privée est un dilemme assez classique, mais il en existe beaucoup d'autres et nous sommes constamment en train de pondérer cela.



***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<sup>e</sup> arrondissement

Pour une entreprise privée, cela est important, car ça lui donne une solidité sur la solution qu'elle propose.

Pour une administration, cela est capital, car le citoyen n'a pas le choix. Donc, il faut s'assurer que cette pondération représente le point de vue des citoyens qui vont utiliser ces services d'IA générative. Il faut définir un référentiel des valeurs et des devoirs à remplir.

S'agissant de la liste, nous sommes assez d'accord : qui veut quelque chose de biaisé, non transparent, homophobe, etc. ? Personne. Mais c'est surtout une question de classement. Je prends un exemple qui n'a rien à voir. Le parlement du Canada a voté un projet de loi pour la pénalisation de la prostitution. Des associations de droits des femmes disaient : « Le corps de la femme ne peut pas être marchandé, donc cela devrait être interdit ». Une partie des associations de droits des femmes étaient donc pour la loi. Une autre partie disait : « Ce n'est pas un parlement masculin qui va faire la loi », et était donc contre la loi. Selon qu'elles plaçaient la dignité ou la liberté en premier, elles étaient pour ou contre la loi.

Cet exemple, que j'ai volontairement pris très éloigné de notre sujet, va colorer la totalité du débat sur l'utilisation de ces technologies.

Nous avons une dimension très pratique sur la nécessité d'avoir un débat citoyen, non pas sur la technologie, mais sur les valeurs sous-jacentes.

Deuxième point : si nous parlons d'un point de vue technologique, nous clôturons déjà le débat.

Nous pouvons parler des algorithmes, des jeux de données, mais tel n'est pas le vrai sujet. Avec la Poste, je fais partie du Comité au sein de notre fondation qui accompagne un certain nombre de grands groupes français ou européens. Notre groupe de travail inclut des administrations centrales au niveau national (pour l'instant pas de municipalités) qui ont fait toute une cartographie des risques et opportunités liés au déploiement de l'IA générative. À la fois démocratisation et prolifération de l'IA.

Nous nous rendons compte que ce qui compte est d'aborder l'éthique de l'ensemble du projet. L'*open source* est une dimension parmi plein d'autres. Lorsque nous discutons avec Mitchell BAKER, présidente de l'association Mozilla, cette dernière fait de l'*open source* depuis longtemps et le promeut. Mais selon Mitchell BAKER<sup>7</sup>, il y a de la place pour les modèles propriétaires ou des modèles en *open innovation*<sup>8</sup>, pour reprendre le terme de Meta, c'est-à-dire qui ne sont pas complètement ouverts. Cela va juste dépendre de ce que nous décidons d'en faire.

Ce matin, j'étais avec la direction de l'innovation d'un groupe du CAC 40. Eux ont décidé d'avoir une partie de leurs modèles avec Microsoft. Et ils travaillent avec d'autres modèles purement propriétaires, hébergés sur des serveurs en France, avec une sécurisation totale, une autonomie stratégique totale de toute la chaîne de valeurs. Parce que certaines choses sont hyper stratégiques. D'autres, pas du tout. J'aborderais la question de la factualité, de la conformité d'un rendu d'une IA générative aux faits. Certains disent : « confabulation ». Nous voyons que ce n'est pas forcément un inconvénient. Si vous

---

<sup>7</sup> Présidente de la Mozilla Foundation.

<sup>8</sup> « Innovation ouverte ».



***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<sup>e</sup> arrondissement

êtes un auteur, un scénariste ou un designer graphique, avoir une IA qui invente des choses *out of the blue*<sup>9</sup>, venant de nulle part, c'est très bien.

En revanche, si vous essayez de prendre une décision, là, c'est une catastrophe. Les types d'usages vont demander des types de modèles complètement différents. Il est intéressant de voir que *Copilot*,<sup>10</sup> solution de Microsoft assez connue du public, a trois modes : un mode proche de la réalité qui va limiter au maximum les problèmes de factualité ; un mode médian ; et un mode créatif qu'il ne faut pas utiliser pour prendre des décisions. C'est un exemple d'intégration en entreprise.

Je n'irai pas beaucoup plus loin, si ce n'est en disant que la partie *open source* n'est qu'un aspect parmi d'autres. La question des jeux de données colore aussi les rendus que nous aurons ensuite. La présence de notre culture à l'intérieur d'un modèle me paraît importante. J'étais à la Villa Médicis, avec une artiste en résidence d'origine nord-africaine qui tente de générer des univers de culture nord-africaine à partir de MidJourney, et la plupart de ce qui en ressort ressemblent à Tataouine, une sorte de planète des singes. Elle a beaucoup de mal parce que les cultures nord-africaines ne sont pas présentes. La question de la francophonie est importante, mais si nous voulons être justes, il faut nous poser la question pour les autres langues, y compris des langues minoritaires : le peule, le swahili, qui ont aussi leur culture. Comment intégrer cela dans des modèles ? C'est important.

La gouvernance est aussi un élément clé énormément sous-estimé.

Et enfin, l'acceptabilité sociale qui résulte des autres, mais qui est comme une brique supplémentaire. Parmi les modèles testés par des administrations en France, certains sont très facilement acceptés, d'autres ne le sont pas alors qu'objectivement ils sont très aidants. Mais bien qu'ils soient très aidants, certains agents de la fonction publique se sentent dévalorisés par la machine.

Un rapport a été publié sur ce thème l'an dernier. Dans un certain nombre de cas, nous avons l'impression d'aider des agents, mais nous leur enlevons une brique de leur travail qui est un élément qu'ils considèrent comme constitutif de leur valeur ajoutée, et cela leur donne l'impression de rétrograder. C'est assez désastreux parmi les acteurs, publics ou privés, y compris chez des acteurs américains que nous n'attendrions pas là-dessus, partant du principe du « *One size suits all* » en déployant une chose pour un milliard d'utilisateurs. Paradoxalement, certains sont capables de mettre en place des solutions très *bottom-up*<sup>11</sup> avec des *workshops*<sup>12</sup>, des *hackathons* dans lesquels les équipes pluridisciplinaires vont faire remonter des choses. Ce qui nous donne du *reward*<sup>13</sup>, c'est cela. Cela est très surprenant.

---

<sup>9</sup> Expression signifiant « sortie de nulle part ».

<sup>10</sup> Système d'IA de Microsoft.

<sup>11</sup> « De bas en haut »

<sup>12</sup> « Des ateliers »

<sup>13</sup> « De la récompense »



***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<sup>e</sup> arrondissement

Sur un autre sujet qui n'est pas lié à l'IA générative, mais à l'automatisation de *process*, nous pouvons avoir l'impression que c'est ennuyeux, mais paradoxalement ce sont aussi des lieux de partage.

Nous pourrions vouloir éviter à certains personnels d'aller dans des lieux bruyants, chauds, etc., mais c'étaient aussi des lieux où ils allaient rencontrer des confrères avec de l'interdisciplinarité, de la mixité. C'est le *watercool effect*<sup>14</sup> : tout ce qui se passe autour de la machine à eau ou de la machine à café.

Cela prend de multiples autres formes et demande une fine intelligence du travail qui n'est pas de nature technologique.

Dans nos travaux, nous collaborons avec Labor IA, un laboratoire conjoint du ministère du Travail et de l'INRA. Lorsque vous les auditionnez, vous verrez, ce sera congruent avec ce que je vous dis là.

Nous voyons déjà, dans les entreprises privées par exemple, que l'IA générative qui s'adresse plutôt à des cols blancs, parce que les cols bleus n'ont pas eu le choix d'être automatisés ou non, alors que les cols blancs ont une certaine liberté. Et toute une partie utilise des IA non proposées par l'entreprise ou l'administration. Parce qu'ils craignent une demande d'hyper productivité, comme le fordisme de l'encadrement. Les cadres se prémunissent de cela un peu au détriment de l'entreprise avec des solutions grand public, non sécurisées.

Quand un cadre utilise Chat GPT qui est un peu une passoire sur laquelle nous n'allons pas traiter d'informations confidentielles, cela revient à utiliser un téléphone normal pour avoir une conversation sur la défense nationale. Mais cela permet de dégager leur propre poche de productivité, de respiration, pour ensuite faire ce qu'ils veulent. Une partie des agents ou collaborateurs vont faire cela pour avoir des poches de respiration. Ils peuvent décider d'allouer ensuite ce supplément de temps, d'énergie, de dispositions à la structure en travaillant plus, ou alors à titre complètement personnel. Lorsque nous voyons les utilisateurs *by-passer*<sup>15</sup> les solutions proposées par leur entreprise ou leur administration, c'est déjà le début d'un gros problème. Et à partir du moment où nous supprimons des choses censées être rébarbatives, mais qui sont aussi des moments de respiration, comme classer ses mails, nous les obligeons à être tout le temps sur la brèche, tout le temps productif.

Or, nous avons fait une comparaison entre certains pays où les médecins touchent assez peu leur clavier : dans une consultation de 15 à 20 minutes ce sera trois ou quatre minutes et ce sera une respiration. Parce qu'être tout le temps en contact visuel avec des gens qui souffrent, à la fin de la journée, moralement, c'est très difficile. Et ce qui se vit (comme les blagues assez *gore* dans les salles de garde) permet de désamorcer ce rapport à la souffrance et à la mort. De l'autre côté, nous voyons avec la NHS que les médecins passent 40 % de leur temps à faire des tâches annexes.

Nous travaillons avec le Charing Institute auquel le gouvernement britannique a demandé de mettre un peu d'ordre dans tout cela. Et ils voient que ce sont des médecins désespérés en se disant : « Je n'ai pas fait toutes ces études pour saisir des informations, et quid du serment d'Hippocrate, je suis là

---

<sup>14</sup> « Effet kiss-cool »

<sup>15</sup> « Contourner »



***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<sup>e</sup> arrondissement

pour soigner des gens, pas pour rentrer de la *data* ». Mais de l'autre côté, si nous leur disons qu'ils seront constamment en contact visuel avec des personnes qui souffrent, cela est trop difficile, il faudra leur ménager des pauses.

Toutes ces tensions sociales qui se dessinent sont la plupart du temps des tensions qui ne peuvent pas s'exprimer. La plupart du temps, si vous enlevez à quelqu'un ce qui fait le cœur de son travail, il va se plaindre. Et si vous lui enlevez ses tâches périphériques, il n'est pas censé se plaindre et cela va créer de la tension sociale non dite. Or, la tension non dite se retrouve toujours quelque part, parce que l'eau trouve toujours son chemin. Cela va se répercuter différemment, sur quelque chose qui n'est pas le vrai objet de la tension. Donc, la tension ne peut pas être résolue. Les gens vont se plaindre pour des choses, mais le vrai problème n'est pas là. Le problème est la tension et le fait que les gens la vivent mal.

Ma recommandation est de voir la *big picture*, d'impliquer les agents et de faire beaucoup de *bottom-up*. Dans un certain nombre de cas, ils vont dire qu'ils voudraient automatiser des tâches que nous n'aurions pas imaginées lorsque nous sommes en haut de la pyramide. Par rapport à d'autres choses, dont ils ne vont pas du tout se plaindre, peut-être faut-il pudiquement les leur laisser. La vraie prise en compte de l'humain, c'est cela.

La spécificité de tout organisme public est d'avoir deux cibles : l'utilisateur final ou citoyen et l'agent. Si l'une des deux n'est pas atteinte, c'est un échec. Malheureusement, ces deux planètes, souvent, ne sont pas alignées. L'ingénierie de la gouvernance du modèle (nous avons vu qu'il s'agit d'une technologie très versatile, mature, qui peut être utilisée dans de nombreux cas différents) va être d'aligner ces planètes.

Cela va nécessiter un patient travail de *design* du produit jusqu'à faciliter la vie de l'agent et du citoyen en même temps.

Intervention de Pénélope KOMITES – Questions et réponses

Merci beaucoup.

Nous sommes partis d'un vœu de l'exécutif qui souhaitait des auditions.

Nous voyons au fur et à mesure des auditions que certaines directions de la Ville ont déjà mis en place des utilisations de l'intelligence artificielle, comme la Direction de l'urbanisme.

Nous devons nous poser la question : « À quoi l'IA va servir en interne dans la Ville ? » J'ai beaucoup apprécié ce que vous disiez sur les concepts d'hyper-productivité. Comment, au service de qui, etc., la Ville s'approprie l'IA, pour quoi faire, au service de qui ?

L'idée est de se dire que la ville de Paris se dotera peut être aussi d'un comité d'éthique, par exemple, et établisse une charte pour dire comment elle va utiliser l'intelligence artificielle, comment elle va s'adresser aux citoyens, pour quels usages. Nous devons aussi être vigilants sur les enjeux de fracture numérique.



***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<sup>e</sup> arrondissement

C'est autour de cela que nous allons essayer de bâtir une position, un peu à l'instar de Montpellier.

Des questions se posent, avec pour but d'améliorer ce que nous pouvons apporter à tous les Parisiens, d'améliorer le travail de nos agents, en étant extrêmement prudents sur le concept et le cadre dans lequel ils travaillent. Ce que vous nous avez dit tous les deux était extrêmement inspirant. Et j'ouvre la porte aux questions.

Karim KHAIR, membre du Conseil Parisien de la Jeunesse

J'avais préparé tout un tas de questions et après vous avoir écoutés, ces questions n'ont plus lieu d'être.

J'ai appris des choses, notamment le fait que l'éthique se fait avec les gens. La ville de Paris est composée des agents de la Ville, donc cela se construit avec eux, parce qu'ils pourraient mal vivre ce surcroît de travail et être déstabilisés professionnellement.

Ce que vous avez dit tous les deux a été très éclairant : j'ai mieux compris les mécanismes, ce qui les a démystifiés. Parce que sinon, de très loin, comme beaucoup de citoyens, je me demandais comment aborder les choses. Un peu comme le *bitcoin*, etc., nous ne savons pas comment cela fonctionne, même si nous avons envie d'y aller. Qui maîtrise cela ? Est-ce que c'est dans mon intérêt ? Beaucoup de choses sont à démystifier.

Madame Komites, vous avez soulevé l'idée d'un éventuel comité d'éthique, une charte avec un référentiel permettant dans chaque action de ne pas perdre la boussole. Tout a été maximisé dans l'efficacité de chaque process, mais finalement, l'humain est peut-être plus malheureux. Autre chose : vous avez évoqué la culture en disant que, selon les cultures, nous n'avons pas tous les mêmes raisonnements.

Une capitale comme Paris a beaucoup de populations issues de différentes parties du monde qui n'ont pas nécessairement les mêmes visions du monde.

Donc, si ce comité d'éthique est mis en place, peut-être faudra-t-il prendre en compte les différences culturelles.

Éric SALOBIR, président de la Human Technology Foundation

Le fait de se doter d'une charte vous permettra d'avoir une véritable intentionnalité. Parce que c'est une technologie qui bénéficie et en même temps souffre de cette *hype* en ce moment. Tout le monde veut y aller parce que c'est *fun*. Mais, la plupart du temps, pour quoi faire ? Beaucoup de cas d'usage n'ont pas d'impact fort.

La deuxième chose est de travailler avec les agents de la mairie. L'intérêt n'est pas seulement de se pencher sur leurs conditions de travail et de s'assurer que le modèle soit viable, mais aussi le fait qu'ils connaissent les citoyens au plus près : ce peuvent être les travailleurs sociaux, mais aussi prendre en compte le fait qu'ils connaissent les citoyens au plus près. Ils auront une cartographie extrêmement intéressante concernant les cas d'usage numériques. Parce que lorsque nous faisons parler 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<sup>e</sup> arrondissement

citoyens, souvent, nous récoltons beaucoup d'impensés. Le génie de certains dans la Silicon Valley n'est pas tant d'inventer des choses que d'aller chercher dans le cerveau des gens ce dont nous avons besoin. Comme Steve Jobs disant : « J'ai envie d'un Smartphone, mais pas avec un stylet, je veux l'utiliser avec les doigts. Nous n'en ferons un que lorsque nous pourrons l'utiliser avec le doigt ». Il avait raison.

L'idée est de réussir à comprendre ce dont les gens ont besoin, mais qu'ils n'expriment pas ou n'ont pas conceptualisé. C'est la règle des 80/20, voire 95/05<sup>16</sup>. Quelques usages clés vont faire gagner un temps fou aux citoyens, pour un grand confort qui va changer la vie des citoyens. Plutôt que de multiplier les cas d'usage dans tous les sens, il convient de faire quelques choses bien pensées.

Ensuite, la notion de l'agentivité : comment faisons-nous des agents des individus un peu augmentés, de façon très concrète, en leur donnant un surcroît de pouvoir ? Nous avons fait partie du processus *Tech for Good* avec l'Élysée il y a quelques années ; lequel s'est un peu ensablé parce que d'autres choses sont arrivées entre temps. Mais nous avons fait tout un travail sur la technologie numérique comme facteur d'inclusion. Comme les *millennium goals* des Nations Unies<sup>17</sup>, nous avons quatre objectifs ; et pour chacun, cinq critères assez simples.

L'inclusion dans la technologie, l'accès, n'était que le premier. Mais une fois que les gens ont accès, ils y vont.

Est-ce que l'accès n'est qu'un accès pour avoir plus de séries Netflix, plus de contenus sur lesquels *scroller*, plus de publicité, d'addiction ou est-ce que vraiment cela donne du pouvoir ?

Les trois autres étaient autour de cela, de la façon de respecter les droits humains fondamentaux. Les chartes de droits et pratiques numériques tournent autour de cela.

Au niveau européen, c'est peu connu et assez peu appliqué, alors que la charte est bien faite et contient beaucoup d'éléments concrets.

Ensuite, il s'agit de cette dimension de l'inclusion. Une initiative comme celle d'Albert présente, à mon avis, cet intérêt : donner du pouvoir en augmentant la compétence ou la capacité du fonctionnaire à aller chercher de l'information. Cela permet d'apporter une réponse en profondeur : « Vous bénéficiez de telle chose, vous ne devriez pas avoir ce problème, Madame, nous devrions vous le résoudre ». Cela accroît la dignité au travail de l'agent, ce qui est très positif, et la satisfaction du citoyen. Lorsque l'agent va chercher une réponse sur Google, cela le décrédibilise et c'est un peu angoissant pour le citoyen. En même temps, les questions sont si diverses que, parfois, nous n'avons pas le choix. Souvent, malheureusement, nous ne pouvons rien faire pour ce citoyen, si ce n'est écouter son problème. Et c'est énorme.

Les permanences politiques sont capitales d'un point de vue humain pour dire que nous avons entendu : « Madame, vos revenus sont légèrement au-dessus donc vous n'avez pas droit à cette aide, mais j'entends votre souffrance ».

Et c'est déjà énorme.

---

<sup>16</sup> Le principe de Pareto, aussi appelé loi de Pareto, principe des 80-20 ou encore loi des 80-20, est une observation selon laquelle environ 80 % des effets sont le produit de seulement 20 % des causes.

<sup>17</sup> Dans la Déclaration du Millénaire pour le développement des Nations Unies, signée en septembre 2000, les dirigeants mondiaux s'engagent à combattre la pauvreté, la faim, la maladie, l'illettrisme, la dégradation de l'environnement et les discriminations à l'égard des femmes



***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<sup>e</sup> arrondissement

Ce qui compte finalement dans la gouvernance est ce que nous choisissons pour être marqueurs de succès. Il ne s'agit pas tant de dire : « Nous avons traité les cas de 30 personnes dans l'après-midi plutôt que de 20 personnes avant » que de savoir pour quel pourcentage de personnes nous avons trouvé des solutions. Si cela fonctionne, nous devrions voir ce pourcentage augmenter.

Nous sommes dans une France fracturée et si cette fracture pouvait se réduire, ce serait important. Il s'agit aussi d'aider le citoyen et l'agent de la fonction publique à entrer plus en contact.

Nous avons développé une méthode à mettre en œuvre, il en existe d'autres, avec des méthodologies de co-construction, du *design-thinking*<sup>18</sup>, avec une question de maïeutique. Nous pouvons mettre en œuvre les principes méthodologiques, en revanche, les valeurs clés sur lesquelles vous allez fonder cette charte, vous êtes les seuls à les avoir.

Dans une institution, la dimension de la diversité ajoute une couche de complexité supplémentaire.

Ce doit être représentatif d'un certain nombre de choses, parce que c'est facteur d'inclusion. En même temps, la technologie est quelque chose sur laquelle nous pouvons parfois aligner des planètes que nous ne pouvons pas aligner sur d'autres sujets.

Giada PISTILLI, éthicienne chez Hugging Face

Quelque chose d'important à comprendre est qu'une charte éthique est souvent vue comme un point de fin, alors que c'est souvent le début d'un projet.

Nous pouvons parler de transparence, d'inclusion, mais si vous n'êtes pas en mesure de les appliquer, vos valeurs n'ont aucune valeur. Donc, l'idée est de rendre explicites des choses implicites : que souhaitez-vous réaliser ? Et de manière collaborative.

Soit nous cherchons le consensus, soit nous respectons l'éthique de la discussion, en donnant le même temps de parole à chaque personne. C'est aussi une question philosophique, parce que la transparence, selon moi, ce peut être une chose, et pour une autre personne, autre chose. Cela dépend du *background* professionnel, académique, etc. Et l'éthique veut que ce soit toujours une pratique dans le sens où c'est quelque chose qui va changer dans le temps. Nous pouvons dire : « Il s'est écoulé six mois ou un an depuis tel projet, est-ce toujours valable ? » Et éventuellement rectifier ces valeurs, les changer un peu. Nous parlons aussi d'*ethic-washing*, comme du *green-washing*, ou d'*ethic-shopping*<sup>19</sup>. Donc il est important de le ramener dans la pratique.

D'un côté, l'éthique normative va vous donner des normes du bon ou du mauvais dans la société. Et de l'autre côté, il y a l'éthique appliquée. L'une informe l'autre et c'est cette relation-là qui est très importante. J'ai rédigé la charte éthique du modèle Big Sense<sup>20</sup>, issu de Bloom, le plus gros modèle multilingue entraîné pas loin d'ici. Tout cela fait l'articulation entre une charte éthique, une licence d'utilisation et une documentation technique.

---

<sup>18</sup> « démarche de conception créative ».

<sup>19</sup> En utilisant la notion d'éthique ou d'écologie de manière trompeuse.

<sup>20</sup> <https://bloomassociation.org/nous-connaître/notre-charte-ethique/>



***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<sup>e</sup> arrondissement

L'idée est de dire comment la charte éthique va informer l'utilisation de ces documents. Nous pouvons imaginer la même chose au sein d'une institution, en nous disant : « La charte éthique correspond aux racines d'un arbre et ensuite nous développons les branches ». Parce que si nous posons les valeurs dès le début, cela nous permet de développer le reste.

Quelques mots concernant les agents publics : nous avons mené une expérience avec un chatbot citoyen dans Les Landes, auprès de 60 000 citoyens, dans la communauté de communes de Dax. Dans la pratique j'ai pu constater que, souvent, nous imaginons que nous prenons un chatbot, un modèle d'IA, nous le branchons et ensuite, c'est fait. Or, il est souvent nécessaire de l'affiner, de l'adapter à leur utilisation. Et les agents ont besoin de temps au début. Il s'agit de travail supplémentaire : traiter les questions des citoyens, affiner le périmètre de connaissances de ce chatbot. À propos des services publics qui ne se parlaient pas entre eux, si la couche technologique ne fait que rajouter une couche de complexité pour les agents et les citoyens, personne ne gagne. Or, au début, il faut ajouter de la formation. Il est donc très important, de faire du développement participatif. Mais à quel point, à quel coût ? Et dans le quotidien des agents publics, comment cela va-t-il se traduire ? Il est important de demander aux agents et aux citoyens s'ils en veulent. Il ne s'agit pas de faire de l'innovation pour faire de l'innovation. Parce qu'en effet, rares seront les cas d'utilisation utiles. Et un constat dans la pratique : les citoyens ne parlent pas le même langage que les agents. Comment travailler cela d'un point de vue pratique ? Ce sont des questions à se poser avant de lancer un projet. Nous nous servons de systèmes d'IA pour mettre en communication les deux. Je me souviens d'un cas où le mot « compétences » voulait dire quelque chose pour un agent public et autre chose pour un citoyen. Il existe de nombreux exemples tels que celui-là.

Je suis Italienne, et entre étrangers, nous pensons souvent que la France est le pays du : « Ce n'est pas possible ».

Donc, si ajouter de l'IA ajoute de la rigidité, non. Si cela peut rendre les choses plus humaines, que les processus sont suivis par le système d'IA et que le reste revient aux agents, c'est bon. Mais c'est une constante dans n'importe quel milieu, nous nous cachons derrière la façade de l'automatisation : « Ce n'est pas moi qui ne peux pas vous donner vos allocations, c'est la machine ». Si c'est une façon de décaler la responsabilité, ce n'est pas adapté. Ou alors, nous ajoutons de la machine, mais il faut voir de quelle manière. Et moi, j'aime bien parler de valeurs, de notre expérience avec Big Science et le modèle Bloom est en ligne, avec des langues minoritaires comme le Catalan ou des langues parlées en Afrique subsaharienne.

L'un de mes axes de recherche est l'omniprésence de l'anglais dans l'IA. J'essaie de lutter contre cela. En Nouvelle-Zélande ou en Afrique subsaharienne, des associations travaillent à des jeux de données pour faire du multilingue, même si cela reste très marginal. Avec la ville de Paris, nous pouvons avoir des projets communs.

Si l'IA est assez précise dans son objectif, nous pouvons nous mettre d'accord, même si nous avons des visions du monde différentes. Et cela revient à la même chose avec les comités d'éthique : le but n'est pas de voir la diversité et l'inclusion à l'américaine, mais surtout d'avoir la diversité de voix et de représentations, c'est le plus important.



***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<sup>e</sup> arrondissement