

approches coopératives



N° 25 été 2025



**Intelligence
artificielle,
intelligence
collective ?**

Photo UCPA. Vincent Collin

« La seule voie qui offre quelque espoir d'un avenir meilleur pour l'humanité est celle de la coopération et du partenariat »

Kofi Annan



ASSOCIATION POUR LA PROMOTION DES APPROCHES COOPERATIVES

contact@approchescooperatives.org

COMITÉ ÉDITORIAL

Dominique BÉNARD, Matheus BATALHA, Larry CHILDS, Roland DAVAL, Alain DEWERDT, Patrick GALLAUD, Francis JEANDRA, Dominique LESAFFRE, Hamady MBODJ, Guy MÉNANT, , Dante MONFERRER, Carolina OSORIO GARCIA, Michel SEYRAT, Michel TISSIER.

SOMMAIRE

Intelligence Artificielle / Intelligence Collective : bifurquer vers le bien commun	5
EDITORIAL PAR FRANCIS JEANDRA	
Qu'est-ce que l'intelligence artificielle ?	9
DÉCRYPTAGE PAR DOMINIQUE BÉNARD	
L'intelligence artificielle, un nouveau collaborateur dans l'intelligence collective ?	17
UNE RÉVOLUTION PAR LAURENT BUTRÉ	
Les cinq défis essentiels de l'IA	23
ANALYSE PAR SYLVESTRE BÉNARD	
L'éthique de l'intelligence artificielle	31
ANALYSE. ENTRETIEN AVEC FRANCK DEBOS, MAÎTRE DE CONFÉRENCES, UNIVERSITÉ DE NICE CÔTE D'AZUR. PAR GUY MENANT	
Opportunités offertes : Une IA sobre, citoyenne et vertueuse	39
VISION D'AVENIR PAR FRANCIS JEANDRA	
Enjeux et défis de l'IA	45
ÉCLAIRAGE. ENTRETIEN AVEC ANTOINE MAIER PAR DANTE MONFERRER	
L'alignement de l'IA et de l'éthique	51
QUESTIONS D'ÉTHIQUE PAR TOM MURRAY	
Partout où il y a un lien technologique, il nous faut redoubler nos liens sociaux	63
MISSION DE SERVICE PUBLIC. ENTRETIEN AVEC JEAN CATTAN, SECRÉTAIRE GÉNÉRAL DU CIAN. PAR DOMINIQUE BÉNARD	
Pour une politique publique du numérique responsable	71
INITIATIVES CITOYENNES. ENTRETIEN AVEC PIERRE JANNIN, ADJOINT À LA MAIRE DE RENNES. PAR DOMINIQUE BÉNARD	
La Convention citoyenne de la Métropole de Montpellier sur l'Intelligence Artificielle	77
INITIATIVES CITOYENNES. PAR DANTE MONFERRER	
VERA : une application pour lutter contre la désinformation	83
INITIATIVES CITOYENNES. ENTRETIEN AVEC FLORIAN GAUTHIER, INITIATEUR DU PROJET. PAR DOMINIQUE BÉNARD	
Un Café IA organisé par ECTI à Laxou, Meurthe et Moselle	89
INITIATIVES CITOYENNES. ENTRETIEN AVEC JACQUES JARAY, ORGANISATEUR ET ANIMATEUR. PAR ROLAND DAVAL	
Vittascience : des outils clés en mains pour l'enseignement des technologies numériques	93
INITIATIVES CITOYENNES. ENTRETIEN AVEC LÉO BRIAND, FONDATEUR ET PRÉSIDENT. PAR GUY MÉNANT	
L'IA, un récit animé par des visions dystopiques, mais déjà une réalité sociale	103
THINK TANK. ENTRETIEN AVEC JEAN-FRANÇOIS LUCAS, DÉLÉGUÉ GÉNÉRAL. PAR DOMINIQUE BÉNARD	
Le Réseau Canopé : former les acteurs de l'éducation à l'ère du numérique et de l'IA	109
ÉDUCATION NATIONALE. ENTRETIEN AVEC ROMAIN VANOUDEUSDEN, DIRECTEUR INNOVATION. PAR GUY MENANT	
L'homme est la mesure de toute chose*	115
BLOC-NOTES. PAR MICHEL SEYRAT	

[Cliquer sur un titre pour accéder à l'article correspondant](#)



INTELLIGENCE ARTIFICIELLE / INTELLIGENCE COLLECTIVE : BIFURQUER VERS LE BIEN COMMUN

PAR FRANCIS JEANDRA



Nous sommes à un moment charnière.

Un moment où l'intelligence artificielle (IA) ne relève plus de la science-fiction ni de la pure prospective, mais de notre quotidien le plus banal : recommandations de vidéos, filtres de tri pour nos mails, outils de traduction, agents conversationnels, logiques d'optimisation à tous les étages... L'IA est partout. Et pourtant, peu nombreux sont ceux qui en comprennent les implications profondes, les mécanismes invisibles ou les véritables enjeux.

C'est de cette intersection entre Intelligence Artificielle et Intelligence Collective que cette revue d'Approches Coopératives souhaite traiter.

Il ne faut pas oublier que derrière chaque interaction avec une IA se cachent des infrastructures bien matérielles, une consommation d'énergie croissante, une extraction de ressources, des chaînes logistiques mondiales, des travailleurs très souvent précaires... Bref, un monde physique lourdement mobilisé pour produire cette illusion de magie numérique.

L'entraînement des modèles, en particulier les plus puissants, exige des milliards de paramètres, des heures de calcul sur des serveurs et data-centers très énergivores, majoritairement alimentés par des énergies fossiles. Cela ne s'arrête pas là, leur usage quotidien, massif, en streaming, en réponse automatisée ou en analyse prédictive, perpétue cette dépendance à une informatique toujours plus gourmande.

Alors faut-il jeter le bébé avec l'eau du bain ? Interdire l'IA ? Bien sûr que non. Mais il faut surtout penser autrement.

Comme toute technologie, l'IA n'est pas neutre. Elle peut être vectrice de progrès ou d'aliénation. Tout dépend de l'usage que l'on en fait, de la gouvernance qui l'encadre, et des finalités qu'on lui assigne.

Elle peut servir à vendre plus, plus vite, plus cher, souvent au détriment de l'humain et de l'environnement. Elle peut être l'alliée de la finance algorithmique ; de la spéculation, du ciblage publicitaire, de la surveillance de masse ou de la guerre automatisée.

Mais elle peut aussi modéliser le climat avec finesse, suivre l'évolution



de la biodiversité, améliorer l'accès à l'éducation, détecter des signaux faibles et être un précieux allié dans la santé, faciliter l'apprentissage, optimiser les réseaux, réduire les gaspillages...

Il est urgent de réorienter les usages de l'IA vers ce qui fait société, ce qui renforce nos capacités collectives à faire face aux défis de notre temps.

Imaginons une IA qui n'a pas pour but d'augmenter sans fin la productivité ou le profit, mais de rendre les services publics plus accessibles. Une IA qui simplifie les démarches, traduit automatiquement dans des langues peu documentées, repère les situations de vulnérabilité ou aide à expliquer les lois à ceux qui en sont éloignés.

Imaginons une IA "low-tech", installée localement, décentralisée, fonctionnant de manière sobre et transparente, respectueuse des données personnelles, compréhensible et accessible à toutes et tous.

Cela suppose un vrai changement de cap, une bifurcation loin de la centralisation massive actuelle, des logiques de silos propriétaires et de la dépendance accrue à quelques multinationales technologiques.

Un des risques majeurs liés à l'essor incontrôlé de l'IA, c'est l'automatisation croissante des décisions et des échanges

économiques. Si elle est mal encadrée, cette automatisation peut produire de puissants effets délétères sur nos sociétés.

D'un côté, les plateformes numériques automatisent les interactions entre offre et demande, avec des algorithmes de fixation des prix, de gestion des stocks, de recommandations ou de notation. Cela peut conduire à une déshumanisation des relations commerciales, une logique d'optimisation brutale où seuls les plus performants ou les mieux classés par la machine survivent. L'économie se rigidifie, s'automatise, se déshumanise.

De l'autre côté, on observe une concentration croissante de la valeur ajoutée dans les mains de quelques acteurs qui contrôlent les systèmes d'IA les plus performants. Ceux qui possèdent les données, les capacités de calcul et également les brevets orientent le jeu. Le reste de l'humanité devient dépendant de leurs outils, de leurs normes, de leurs interfaces.

Le risque est grand de voir émerger une économie "automatisée", opaque, où les choix humains sont réduits à des options prédéfinies par l'algorithme, où la concurrence est faussée par l'asymétrie d'accès aux technologies, et où les décisions économiques se prennent sans

délibération collective, sans débat, sans prise en compte du long terme. C'est là que l'intelligence collective doit entrer en scène.

Face à des systèmes toujours plus complexes, c'est notre capacité à réfléchir ensemble, à poser des limites, à inventer des usages coopératifs, qui fera la différence. L'intelligence collective, c'est la mise en commun des savoirs, des expériences, des sensibilités. C'est la confrontation d'idées, l'élaboration de compromis, la participation de toutes et tous à la définition des outils et des règles. C'est la démocratie appliquée à la technique.

Plutôt que de laisser l'IA décider à notre place, il nous faut décider ensemble de ce que nous voulons en faire pour un usage sobre, transparent, pour en faire un outil au service du bien commun.

L'IA peut être un outil au service du vivant – ou une machine froide au service d'une croissance sans fin. Ce choix ne peut pas être laissé aux seuls experts, ingénieurs ou investisseurs. Il nous revient, collectivement, de décider dans quelle direction nous voulons aller.

Et cela commence maintenant. Dans nos usages, dans nos exigences, dans nos coopérations. L'intelligence artificielle est un miroir : à nous de choisir ce qu'elle reflète

L'intelligence artificielle s'invite dans nos vies, nos métiers, nos débats... Elle fascine, inquiète, questionne. Ce numéro 25 d'Approches Coopératives propose d'explorer ce que cette nouvelle forme d'intelligence peut ou ne peut pas apporter à nos démarches collectives.

La première partie du numéro traite des fondamentaux : comment fonctionne l'IA ? Quelles sont ses promesses, ses opportunités, mais aussi ses risques et ses dérives possibles. Nous y abordons les enjeux techniques, éthiques et humains.

Dans la seconde partie, nous avons voulu donner à voir des initiatives



concrètes : collectivités, chercheurs, enseignants, citoyens expérimentent des façons de s'approprier l'IA pour enrichir l'intelligence collective. Des cafés IA aux conventions citoyennes, de la formation des jeunes à la médiation numérique, ces démarches montrent qu'un autre rapport à la technologie est possible, plus critique, plus créatif, plus coopératif.

Bonne lecture !

[Retour au sommaire](#)

FRANCIS JEANDRA

Scout de France dans les années 70's, puis animateur et conseiller pédagogique auprès de la SAGIF (Service d'Animation des Golfs d'Ile de France), Francis a consacré 14 années au mouvement Scout. Diplômé en électronique et informatique à l'EFREI Paris, il a mené une carrière complète au sein du groupe Schlumberger, occupant divers postes allant de programmeur à directeur de l'environnement de travail (2009-2021) en passant par la formation, la vente, le marketing, la Qualité, la Santé, Sécurité et l'Environnement.

Expert reconnu, il a su conjuguer théorie et terrain pour accompagner efficacement la prévention des risques en entreprise.

Élu au CE/CSE/DP/CHSCT entre 1992 et 2021, Délégué Syndical représentant la CFE-CGC, il a contribué activement à préserver l'emploi au sein du groupe en France lors de plusieurs PSE.

Retraité depuis 2022, il intervient aujourd'hui comme formateur et conseiller en Sécurité et Environnement. Passionné d'apiculture bio, d'œnologie (certifié Caviste-Sommelier en 2022) et bonsaïka, il a rejoint l'APAC comme secrétaire de l'association en 2025.



Boom de l'IA en France : 260 000 micro-travailleurs à 21 € par mois en moyenne. Derrière l'intelligence artificielle, les algorithmes et l'apprentissage machine, des millions de travailleurs à travers le monde. Ici, pour valider la pertinence des réponses, là pour indiquer aux machines ce qu'elles doivent retenir comme informations.

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE ?

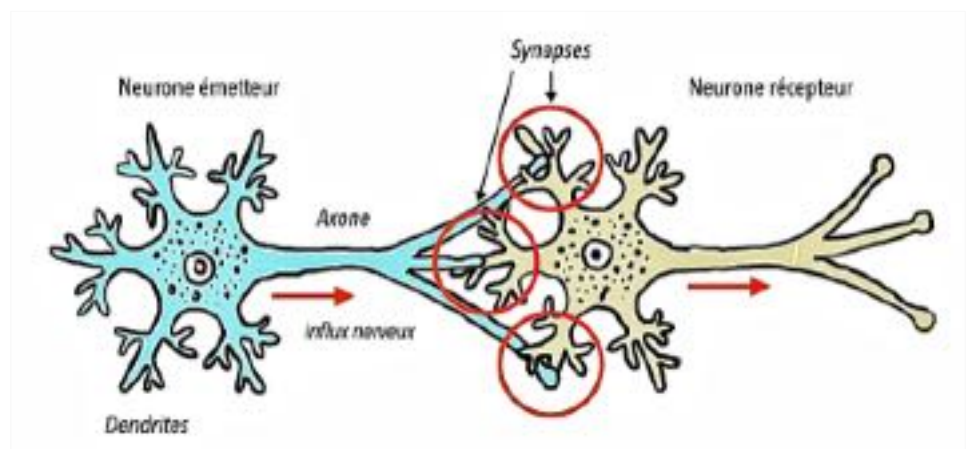
PAR DOMINIQUE BÉNARD

L'intelligence artificielle (IA) désigne la capacité d'une machine construite, telle qu'un ordinateur, à simuler ou à reproduire des tâches relevant des facultés intellectuelles humaines : effectuer des calculs, analyser des données afin d'établir des prédictions, identifier différents types de signes et de symboles, converser avec des humains et contribuer à l'exécution de tâches sans intervention manuelle.

Les domaines d'utilisation de l'IA ne cessent de se multiplier.

COMMENT FONCTIONNE L'IA ?

On parle d'intelligence artificielle parce que les ingénieurs informatiques se sont inspirés du cerveau humain, composé de cellules nerveuses appelées neurones qui constituent un réseau extrêmement complexe. Un neurone se compose de 3 parties : les dendrites reçoivent une impulsion, l'impulsion est transmise par l'intermédiaire d'une espèce de long cable qu'on appelle l'axone, les synapses connectent le neurone à d'autres



Les synapses permettent aux neurones de se connecter pour former des réseaux dans lesquels l'influx nerveux circule

neurones et lui permettent de transmettre l'impulsion. Ainsi se constituent des circuits capables de stocker et de transmettre des informations.

En 1943, Warren McCulloch, un neurologue, et Walter Pitts, un logicien, examinent la question de savoir si le système nerveux humain pouvait être considéré comme une sorte de dispositif informatique universel. Ils constatent que lorsqu'un influx nerveux entre dans un neurone, s'il excède un certain seuil, alors le neurone émet à son tour une

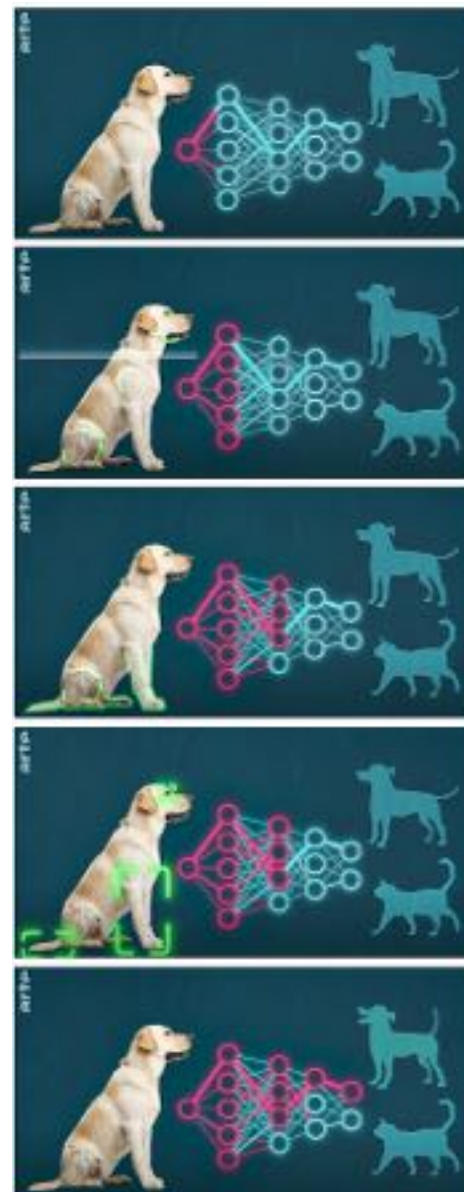
UN RÉSEAU NEURONAL EST UN ENSEMBLE DE PROGRAMMES INFORMATIQUES SIMULANT DES NEURONES DISPOSÉS EN COUCHES ET CONNECTÉS ENTRE EUX PAR DES CONNEXIONS PONDÉRÉES, DONC PLUS OU MOINS FORTES.

impulsion qui va être transmise à d'autres neurones par l'axone et les synapses. Ils font alors le lien avec la logique : l'impulsion ou son absence signifie "activé" ou "désactivé", "oui" ou "non", "vrai" ou "faux". Ils se rendent compte qu'un neurone dont le seuil est suffisamment bas, de sorte qu'il s'active dès que l'une de ses entrées est activée, fonctionne comme une incarnation physique de la fonction logique « ou ». Un neurone dont le seuil est suffisamment élevé, de sorte qu'il ne s'active que lorsque toutes ses entrées sont activées, est une incarnation physique de la fonction logique « et ». Ils commencent alors à comprendre qu'un tel « réseau neuronal », à condition qu'il soit correctement câblé, peut faire tout ce qui peut être fait avec la logique. Ils écrivent ensemble un article scientifique historique intitulé : "A logical Calculus of Ideas Immanent in Nervous Activity" (Un calcul logique des idées immanentes dans l'activité nerveuse). Cet article est à la base des réseaux neuronaux artificiels.

Les ingénieurs informatiques n'ont pas cherché à copier le réseau biologique parce qu'il est trop complexe. Ils ont programmé une abstraction de son fonctionnement pour bâtir des *réseaux de neurones artificiels*. Ces réseaux ne sont pas constitués de neurones mais de plusieurs programmes informatiques (ou algorithmes) fonctionnant en parallèle.

LES RÉSEaux DE NEURONES ARTIFICIELS

Un réseau neuronal artificiel n'est pas un circuit matériel fait de câbles et de processeurs mais un ensemble de programmes informatiques simulant des neurones disposés en couches (d'où le terme « apprentissage profond ») et connectés entre eux



Source : Arte

par des connexions pondérées, donc plus ou moins fortes.

Imaginons qu'une IA doive analyser les vidéos d'un refuge pour animaux. Il lui faut pour cela différencier les chiens des chats. Les neurones de la première couche ne font que repérer les lignes, les courbes et les angles de l'image d'un animal. C'est suffisant pour situer les différentes parties du corps dans la couche suivante, les neurones qui sont activés en priorité sont ceux qui correspondent le mieux à ce schéma. Ils cherchent alors des yeux, des pattes ou une

queue. Si l'analyse renvoie plutôt à chien, cela active d'autres neurones qui vont rechercher des caractéristiques canines encore plus complexes. Cette réaction en chaîne finit par donner un résultat ou plutôt une prédiction : il s'agit probablement d'un chien (voir schéma ci-dessus). Le cerveau d'une IA présente donc un schéma d'activité variable en fonction de la tâche qu'elle effectue à l'instant « t ».

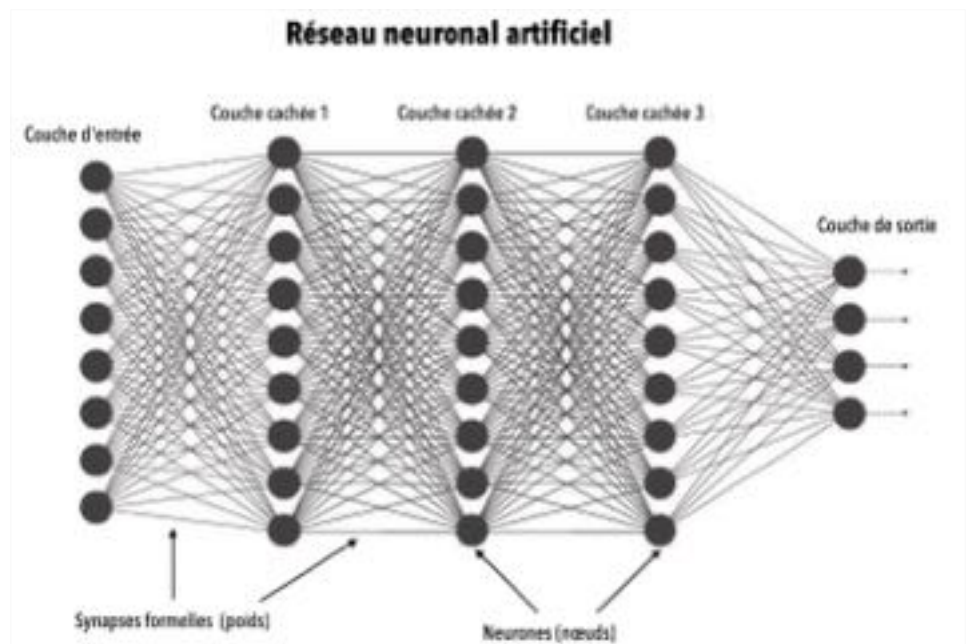
Mais comment l'IA apprend-elle à corriger le poids (la pondération) de la connexion entre chaque neurone, en d'autres termes à procéder efficacement pour aboutir à un résultat juste ? Elle a suivi un entraînement. Il suffit de lui indiquer si sa réponse finale est correcte ou fautive pour qu'elle apprenne à modifier la pondération des connexions.

Une donnée arrive à la couche d'entrée du réseau, puis elle se propage via les connexions pondérées. A la couche de sortie, le réseau donne une réponse. Si elle est juste, tant mieux. Si elle est fautive, le poids (la pondération) des connexions est modifiée pour que le système se rapproche de la bonne réponse. On répète ce processus avec différents exemples jusqu'à ce que le réseau donne la bonne réponse dans la majorité des cas.

Ainsi, on n'a pas besoin d'apprendre à l'IA de reconnaître un chien, il suffit de lui indiquer si sa réponse finale est correcte ou fautive pour qu'elle apprenne à modifier la pondération des connexions et suivre la bonne piste.

UN EXEMPLE : L'IA GÉNÉRATIVE

L'IA générative ou l'intelligence artificielle générative fait référence à l'utilisation de l'IA pour créer de nouveaux contenus, comme du texte, des images, de la musique, de l'audio et des vidéos. C'est l'IA



qui a le plus frappé l'imagination du public avec l'apparition d'applications comme ChatGPT capables de dialoguer avec un utilisateur humain et d'apporter des réponses à des questions de manière fluide et apparemment "intelligente".

L'IA générative fonctionne grâce à des *grands modèles linguistiques* (Large Language Models ou LLM en anglais). Ce sont des outils pour comprendre et générer du texte : répondre à des questions, générer des histoires, aider à rédiger des mails, traduire des langues... Parfois on pourrait croire que les LLM pensent comme un humain, mais ce n'est pas le cas. Ils prévoient l'apparition de mots à l'aide de calculs statistiques et de règles qu'ils ont apprises à partir d'une quantité phénoménale de textes.

Les LLM contiennent des milliards, voire des trillions de paramètres que l'on règle pour que les modèles apprennent à choisir les mots les plus appropriés. Plus il y a de paramètres, plus le modèle peut capturer des nuances complexes dans le langage.

GPT3 D'OPENAI UTILISE 175 MILLIARDS DE PARAMÈTRES. GPT2 EN AVAIT SEULEMENT 1,5 MILLIARDS ET GPT4 DEVRAIT EN UTILISER 1000 MILLIARDS



Projet de data center CloudHQ à Lisses (Île de France)

LES TRANSFORMERS ET L'AUTO-ATTENTION

Pour être capables d'interpréter le langage, les LLM doivent comprendre plus que des mots isolés. Ils doivent être capables d'interpréter des phrases, des paragraphes et des documents entiers. Les premiers modèles d'apprentissage automatique avaient du mal à comprendre des phrases entières et "oubliaient" le début d'une phrase lorsqu'ils en atteignaient la fin, ce qui donnait lieu à des erreurs d'interprétation.

Les modèles d'IA générative modernes utilisent un type spécifique de réseaux neuronaux appelés *transformers*.

Ces derniers exécutent une fonctionnalité appelée auto-attention pour détecter la manière dont les éléments d'une séquence sont connectés.

Les *transformers* permettent aux modèles d'IA générative de traiter et de contextualiser de grands blocs de texte plutôt que des mots et des phrases isolés.

Exemple : on veut entraîner un modèle pour qu'il reconnaisse des schémas grammaticaux comme "un chat dort". Dans un réseau de neurones artificiels classique, chaque couche traite une partie différente de l'information : les mots, leurs relations, leur ordre et ainsi de suite. C'est ce qu'on appelle l'apprentissage hiérarchique. Avant 2017, on utilisait surtout des réseaux récurrents (RNN). Ces modèles analysent les phrases mot par mot, de gauche à droite, ce qui limite leur capacité à comprendre des relations complexes entre des mots éloignés dans une phrase. Par exemple, dans la phrase "Le chien, que Marie a adopté hier, dort sur le canapé", les RNN avaient du mal à relier "chien" à "dort".

C'est ici que le transformer intervient pour changer la donne.

Exemple : "*Marie aime son frère et elle le soutient dans tout ce qu'il fait*". Si on lit cette phrase, on sait que "*elle*" fait référence à "*Marie*" et "*il*" à "*frère*". Mais pour un ordinateur, ce genre de lien n'est pas évident. Le

mécanisme de *self-attention* permet au modèle de regarder chaque mot et de calculer son importance par rapport aux autres mots de la phrase. En d'autres termes, il attribue un poids (une pondération) à chaque mot pour décider quels éléments du texte sont les plus pertinents. Dans la phrase "*Marie aime son frère et elle le soutient dans tout ce qu'il fait*", le mot "*elle*" aura un lien fort avec "*Marie*" tandis que "*il*" sera lié à "*frère*". Grâce à cette capacité, les *transformers* peuvent analyser une phrase entière en une seule fois au lieu de la traiter mot par mot. Cela leur permet de mieux comprendre les relations complexes dans un texte. Ce mécanisme améliore drastiquement la cohérence et la précision des modèles. Il permet de traiter des phrases longues ou des textes entiers sans perdre le fil du contexte.

L'ENTRAÎNEMENT DES LLM

Pour qu'un modèle devienne performant, il doit passer par deux étapes d'entraînement : la préformation et l'affinage (*fine tuning*).

La préformation : le modèle est exposé à une quantité énorme de textes provenant de livres, d'articles, de pages Web, etc. L'objectif c'est qu'il apprenne des patterns, c'est-à-dire des schémas récurrents dans les données, qui sont déterminés par les règles fondamentales du langage : la grammaire, la syntaxe et les associations fréquentes entre les mots. Si le modèle voit souvent la phrase "le chat dort sur le canapé", il apprend que les mots "chat", "dort" et "canapé" sont souvent associés. Il ne comprend pas ce qu'est vraiment un chat ou un canapé. Pour donner une idée de l'échelle de cet apprentissage, il faut savoir que certains modèles

sont entraînés sur des trillions de mots. C'est comme lire l'équivalent de 16 millions de livres.

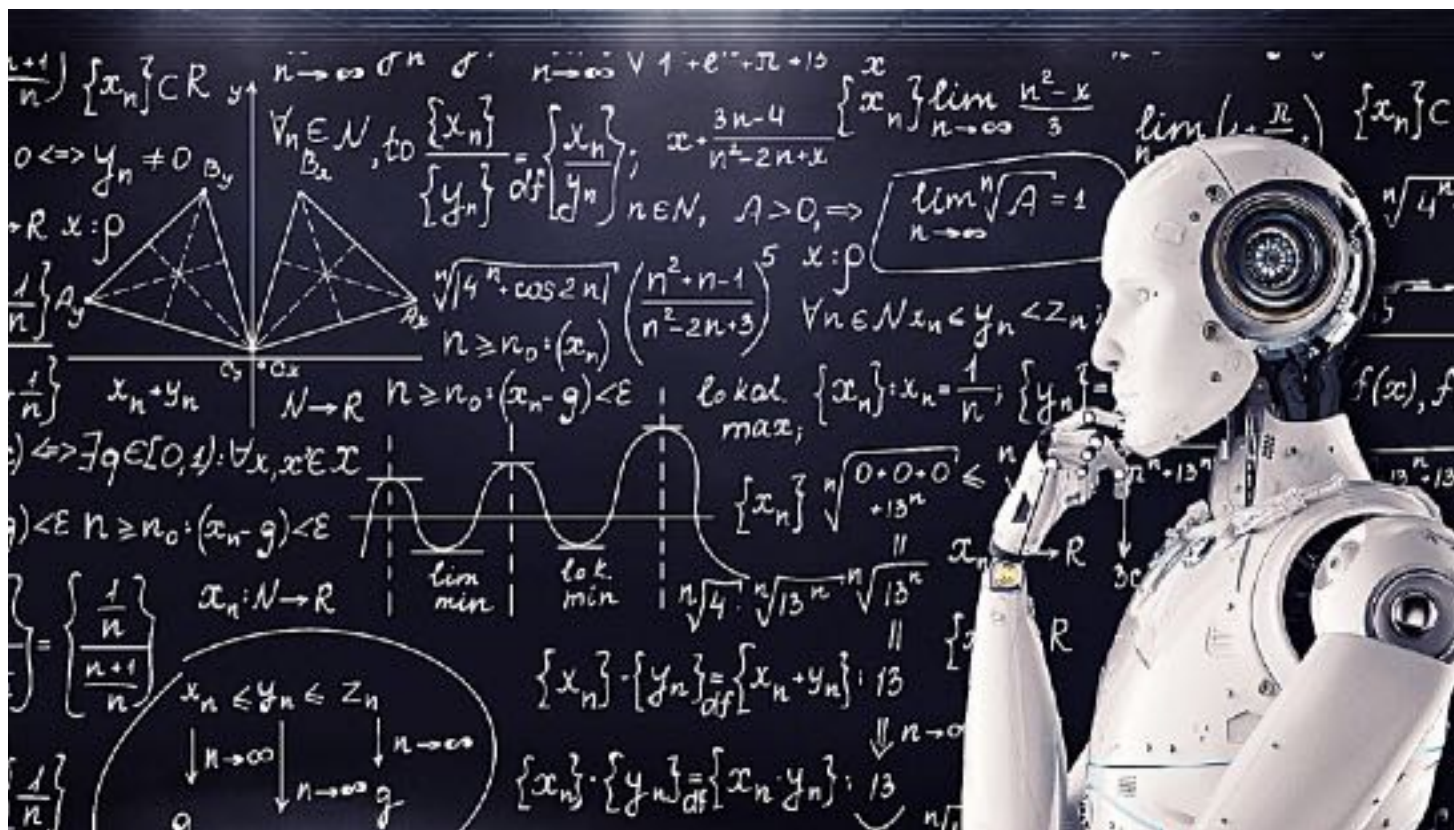
L'affinage : dans cette deuxième phase, le modèle est entraîné sur des données spécifiques, pour des tâches précises. Par exemple, un modèle utilisé dans le domaine médical sera affiné avec des articles scientifiques et des études cliniques.

L'affinage permet d'adapter le modèle à des contextes particuliers. Si l'on pose - avant affinage - la question "*Qu'est-ce que la grippe ?*", le modèle pourrait répondre : "*une maladie courante causée par un virus*". Après affinage, il pourrait répondre "*la grippe est causée par un virus influenza A ou B, elle se manifeste par de la fièvre, des douleurs musculaires et des frissons*". Cette précision est possible grâce à des données d'entraînement ciblées qui permettent au domaine de devenir expert dans un domaine particulier.

Avec les *transformers*, on peut créer des modèles beaucoup plus grands, capables d'apprendre sur des volumes de données massifs. Par exemple GPT3 d'OpenAI utilise 175 milliards de paramètres. A titre de comparaison, GPT2 en avait "seulement" 1,5 milliards et GPT4 devrait en utiliser 1000 milliards. Cette augmentation de taille permet de capturer des nuances plus complexes dans le langage. Grâce à leur polyvalence, les *transformers* ne se limitent pas à la compréhension du texte, ils peuvent être utilisés aussi à d'autres tâches comme la reconnaissance d'image, avec des variantes comme les *vision transformers*, ou la génération de musique.

L'entraînement des LLM consomme de grandes quantités d'énergie et entraîne des émissions massives de CO₂

PLUTÔT QUE DE CÉLÉBRER LA SUPÉRIORITÉ DES MACHINES, IL SERAIT PRÉFÉRABLE DE METTRE EN AVANT LEUR AVANTAGE PRINCIPAL : L'AMÉLIORATION ET L'ÉLARGISSEMENT DES CAPACITÉS HUMAINES



LES LLM SONT-ILS DANS L'IMPASSE ?

Toute la stratégie actuelle des entreprises de l'IA vise à bâtir des LLM toujours plus grands, travaillant sur toujours plus de données, avec une puissance de calcul toujours plus développée dans l'espoir de faire surgir la fameuse IAG, l'intelligence artificielle générale, qui surpassera en tous points celles des humains.

Selon le fameux ingénieur informaticien, Yann Le Cun¹, c'est une erreur monumentale. Former une intelligence humaine à partir du texte, comme le font les LLM, ne suffira jamais pour atteindre une intelligence humaine. Yann Le Cun signale que le volume

¹ Dès les années 80, Yann Le Cun, un ingénieur informaticien français, travaille sur l'apprentissage automatique. Il veut apprendre à des machines à reconnaître des images, du son et même du texte en leur montrant des milliers d'exemples. Il a développé le concept clé du réseau de neurones convolutif, encore aujourd'hui la technologie de base de l'IA. Il a reçu en 2018 le prix Turing, l'équivalent du prix Nobel en informatique, aux côtés de Geoffrey Hinton et de Yoshua Benjo. Il est directeur scientifique de l'IA chez Meta, l'un des postes les plus stratégiques du monde de l'IA.

d'information textuelle que digère aujourd'hui un grand modèle de langage (LLM) équivaut seulement à ce qu'un enfant humain reçoit uniquement via ses yeux dans les quatre premières années de sa vie. Il souligne que le langage seul n'est pas toute l'intelligence et avance qu'il serait plus intéressant de chercher comment faire pour que les machines se dotent d'un modèle du monde pour comprendre le monde physique, de manière à être capables de raisonner et de planifier et pas seulement à calculer la probabilité d'apparition des prochains mots dans un texte.

Il explique : *"Un modèle du monde, c'est ce que nous avons tous dans notre esprit et qui nous permet de manipuler les pensées. Je sais que si j'appuie sur le haut de cette bouteille, elle va probablement se renverser, alors que si j'appuie sur le bas, elle va glisser sur la table. Nous avons des modèles du monde physique que nous acquérons dans les*

premiers mois de notre vie, c'est ce qui nous permet de gérer le monde réel. C'est beaucoup plus difficile que de gérer le langage. La véritable IA exige des modèles capables de gérer le monde réel."

Il propose de créer des systèmes qui seraient capables, comme nous, de comprendre les conséquences physiques des actions et non pas juste capables de manipuler du langage.

Le point de vue de Yann Le Cun est corroboré par Judea Pearl, un très grand informaticien et philosophe, grand spécialiste du raisonnement causal, qui a écrit dans son livre "The book of Why" (le livre des pourquoi) : *"Les systèmes capables d'agir intelligemment doivent pouvoir manipuler des représentations causales du monde et pas seulement juste corrélés des observations."*

LE MYTHE ET LES RÉALITÉS DE L'INTELLIGENCE ARTIFICIELLE

Un autre informaticien célèbre, Luc Julia² dénonce les mythes qui entourent l'IA. Selon lui, beaucoup trop de gens parlent de l'intelligence artificielle avoir de quoi ils parlent. Ils diffusent une vision "hollywoodienne" de l'IA : quelque chose d'assez fantastique, qui fait peur ou qui fait rêver, l'entité artificielle qui va dépasser l'intelligence humaine, prendre notre boulot, et nous tuer (Matrix, Terminator...).

Luc Julia déclare : "Il n'y a pas une seule intelligence artificielle, qui pourrait tout faire, il y en a plusieurs. Ce sont des outils très spécifiques et qui font des choses très pointues. Les IA font les

² Luc Julia : né en janvier 1966 à Toulouse, est un ingénieur et informaticien franco-américain, chercheur (CNRS, Stanford) spécialisé dans l'intelligence artificielle. Il est l'un des concepteurs de l'assistant vocal Siri. Ancien vice-président de Samsung chargé de l'innovation, il rejoint Renault en 2021 en tant que directeur scientifique.

choses bien mieux que vous comme tous les outils. Un marteau plante les clous mieux que vous. Mais comme pour le marteau, avec l'IA, c'est vous qui tenez le manche, c'est vous qui prenez les décisions."

Dans le Monde du 30 décembre 2024, Daron Acemoglu³, professeur d'économie au MIT⁴ dénonce le rêve d'une intelligence artificielle générale capable d'accomplir n'importe quelle tâche cognitive humaine. Selon lui, l'obsession pour construire des machines super intelligentes conduit l'industrie à ignorer le potentiel réel de l'IA. Au lieu de poursuivre une telle quête, il serait plus raisonnable de se doter d'une IA d'un autre genre, techniquement réalisable et socialement souhaitable. Au lieu de chercher à remplacer les travailleurs, il serait plus sage d'utiliser l'IA pour les aider à effectuer des tâches difficiles.

"Nous avons besoin d'un paradigme industriel, écrit-il, qui, plutôt que de célébrer la supériorité des machines, mette en avant leur avantage principal : l'amélioration et l'élargissement des capacités humaines."

Pour cela, il importe d'investir plus dans la formation et l'accroissement des compétences humaines afin que les travailleurs soient capables de comprendre, d'utiliser et de maîtriser les systèmes d'IA.

³ Daron Acemoglu, lauréat du prix Nobel d'économie 2024 est auteur, avec Simon Johnson, de *"Power and Progress, Our 1000-year Struggle Over Technology & Prosperity"* ("Pouvoir et progrès, notre lutte millénaire avec la technologie et la prospérité"), Public-Affairs, 2023.

⁴ Le Massachusetts Institute of Technology (MIT), en français Institut de technologie du Massachusetts, est un institut de recherche américain et une université, spécialisé dans les domaines de la science et de la technologie. Établissement privé situé à Cambridge, dans l'État du Massachusetts, à proximité immédiate de Boston, au Nord-Est des États-Unis, le MIT est considéré comme une des meilleures universités du monde (classée première par Quacquarelli Symonds en 2024). (Wikipédia)

ILS DIFFUSENT UNE VISION "HOLLYWOODIENNE" DE L'IA : QUELQUE CHOSE D'ASSEZ FANTASTIQUE, QUI FAIT PEUR OU QUI FAIT RÊVER, L'ENTITÉ ARTIFICIELLE QUI VA DÉPASSER L'INTELLIGENCE HUMAINE

[Retour au sommaire](#)



L'INTELLIGENCE ARTIFICIELLE, UN NOUVEAU COLLABORATEUR DANS L'INTELLIGENCE COLLECTIVE ?

PAR LAURENT BUTRÉ

UNE RÉVOLUTION ÉQUIVALENTE À CELLE DE L'INTERNET DES ANNÉES 2000

Depuis que l'humanité a découvert le feu et taillé des pierres pour en faire des outils, elle n'a cessé d'inventer. Chaque génération a vu surgir des révolutions technologiques qui ont bouleversé nos façons de vivre, de penser, de travailler.

Pendant des milliers d'années, nos ancêtres ont progressé lentement : un silex mieux taillé ici, une roue mieux fixée là... Puis, soudainement, au XVIII^e siècle, tout s'accélère avec la révolution industrielle. Machines à vapeur, filatures, locomotives : l'homme commence à déléguer sa force musculaire à des machines. Le monde change. Et ce changement s'emballe encore plus avec l'électricité, le téléphone, puis les ordinateurs.

Mais un autre tremblement de terre technologique surgit à la fin du XX^e siècle : l'arrivée d'internet. Dans les

années 2000, le Web devient accessible à tous. On peut discuter avec quelqu'un à l'autre bout du monde, acheter un livre sans se lever de son canapé, s'informer en quelques clics. Le monde est devenu un village connecté. C'est une transformation radicale. Certains appellent cette période la « 3^e révolution industrielle ».

Et maintenant ? Nous vivons en direct un nouveau virage, une autre rupture : l'arrivée de l'intelligence artificielle (IA), plus précisément de l'IA générative. C'est notre époque, notre bouleversement. L'outil change encore de forme : ce n'est plus juste une machine qui obéit à un ordre, c'est une entité capable de proposer, de comprendre, de créer... comme un humain.

Pour la première fois, les machines commencent à réfléchir avec nous. Et je n'aime pas trop ce terme "réfléchir" car derrière il faut comprendre que ce n'est que de la statistique basée sur ce qui a déjà été créé.





<https://knowledgeone.ca/10-categories-doutils-dia-generative/?lang=fr>

NON. L'INTELLIGENCE ARTIFICIELLE, MALGRÉ SON NOM, N'EST PAS UNE CONSCIENCE, C'EST UN OUTIL STATISTIQUE EXTRÊMEMENT PUISSANT

COLLABORER AVEC, L'AMADOUER, APPRENDRE À S'EN SERVIR

Si on regarde autour de nous, l'intelligence artificielle est déjà partout. Ce n'est pas une nouveauté de 2023 ou 2024. Depuis plusieurs années, nos téléphones, nos voitures, nos ordinateurs utilisent déjà des algorithmes très puissants, parfois invisibles.

Tu utilises un GPS ? Il calcule en quelques secondes le trajet le plus rapide. Avant, il fallait déplier une carte, se repérer, deviner.

Tu veux savoir quelle est cette musique dans un café ? Shazam l'écoute et la reconnaît.

Ton téléphone corrige tes fautes d'orthographe, anticipe ce que tu veux écrire, te propose des images dans ta galerie en reconnaissant les visages ou les lieux.

Tout cela, c'est déjà de l'intelligence artificielle. Discrète. Pratique. Mais ce qui change aujourd'hui, c'est que cette intelligence devient plus accessible, plus générale, plus créative. On ne parle plus seulement d'un outil qui fait une chose très précise.

L'IA peut désormais écrire un texte, créer une image, te proposer une recette, analyser un contrat, générer du code informatique, résumer un article, planifier un événement... Et

nous devons apprendre à travailler avec elle.

Gartner, un grand cabinet de conseil américain, l'a dit clairement :

"L'IA ne prendra pas le travail des gens, mais ceux qui ne travaillent pas avec l'IA n'auront plus de travail."

Comme avec toute innovation, ceux qui savent s'en servir prennent de l'avance. C'est vrai pour les élèves, les enseignants, les chercheurs, les écrivains, les avocats, les médecins, les ingénieurs... L'IA devient un compagnon de travail. Et pour bien collaborer, il faut la comprendre, l'apprivoiser, l'exploiter intelligemment.

QU'EST-CE QUE C'EST QUE L'INTELLIGENCE ARTIFICIELLE ?

Alors au fond, c'est quoi cette fameuse IA ? Est-ce une machine consciente ? Est-ce qu'elle pense vraiment ?

Non. L'intelligence artificielle, malgré son nom, n'est pas une conscience. C'est un outil statistique extrêmement puissant.

Sa définition peut se résumer ainsi :

"C'est la capacité d'un programme informatique à observer son environnement, à apprendre de lui, à raisonner sur ce qu'il perçoit, à décider comme un humain l'aurait fait, et surtout à agir."

Mais concrètement ? Prenons un exemple. Tu montres une image à une IA, afin de pouvoir l'analyser, s'en servir, elle la transforme en chiffres. Pareil pour du son, du texte. Tout devient mathématique, mesurable. Ensuite, elle calcule des probabilités : sur des millions d'exemples passés, quelles sont les chances que ce signal corresponde à un chat, une note de musique, ou une faute d'orthographe ?

Pour que cela fonctionne, il faut énormément de données. L'IA ne "comprend" rien : elle apprend des exemples. Plus on lui donne d'exemples, plus elle peut faire de bonnes prédictions. Ce n'est donc pas une intelligence au sens humain,

c'est une intelligence par imitation statistique.

Et tout ce qu'elle fait revient à résoudre deux types de problèmes :

- Des problèmes de classification : par exemple, reconnaître si un mail est un spam ou pas. C'est comme choisir dans quelle boîte ranger une donnée.
- Des problèmes de régression : par exemple, prédire le prix d'une maison selon sa taille, sa localisation, etc. On cherche une valeur la plus probable dans une fourchette.

Et le plus fort, c'est que toute question peut être reformulée pour entrer dans une de ces deux catégories. C'est la force du Machine Learning, et surtout des Data Scientists, ces ingénieurs de la donnée qui passent leur temps à réfléchir à comment reformuler les problèmes dans ces deux approches !

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE ?

Et là, surprise : on apprend que ces machines peuvent générer du texte, créer des images, composer de la musique ! Comment est-ce possible, si elles ne font que classer ou estimer des probabilités ?

C'est là que la force des chercheurs entre en scène.

Prenons l'exemple d'un texte. On pose à l'IA une question apparemment simple :

*"Complète cette phrase :
L'intelligence artificielle est..."*

Mais en réalité, on reformule le problème ainsi :

"Parmi un million de mots possibles, quel est celui qui a la plus haute probabilité d'être le mot suivant qui complète ces trois petits points dans la phrase ci-dessus ?"

Et hop ! C'est de la classification.

L'IA choisit le mot le plus probable. Puis elle recommence, avec le nouveau mot ajouté à la phrase. Mot



après mot, elle "génère" un texte. Même chose pour une image : pixel par pixel, zone par zone, elle prédit ce qui a le plus de chances de s'y trouver.

Mais pour arriver à ce résultat, il faut avoir lu, vu, entendu des millions, voire des milliards de contenus. C'est ce qu'ont fait OpenAI, Google, Microsoft, Mistral, et d'autres : ils ont aspiré le web, les livres, les articles, les forums, les bases de données...

Problème : selon les sources utilisées, les réponses ne seront pas les mêmes. C'est ce qu'on appelle les biais des modèles. Une IA qui a été entraînée majoritairement sur des textes américains aura une "culture" différente d'une IA entraînée sur des données francophones, scientifiques ou artistiques. D'où l'importance de picorer sur les différentes offres, entre ChatGPT, Gemini, Mistral, Bing CoPilot, Claude, Llama (Meta), Grok, etc.

C'EST LÀ POUR RESTER... ET ALLER ENCORE PLUS LOIN

Certains critiquent l'IA pour son impact écologique. Et c'est vrai : entraîner un grand modèle d'IA demande beaucoup d'énergie. Des centres de données gigantesques, des serveurs qui chauffent, des milliards de calculs...

Mais aujourd'hui, on ne peut plus l'éteindre. Pourquoi ? Parce qu'elle rend plus efficaces ceux qui l'utilisent.



C'EST POURQUOI IL FAUT APPRENDRE À S'EN SERVIR, NOTAMMENT EN MAÎTRISANT L'ART DU PROMPT ENGINEERING

Elle fait gagner du temps, de la productivité, elle stimule la créativité.

C'est pourquoi il faut apprendre à bien s'en servir, notamment en maîtrisant l'art du *prompt engineering*.

Avant, avec Google, on tapait trois mots-cés. Et le moteur trouvait les sites qui contenaient l'information recherchée car mieux référencée, liée, vue, etc..

Maintenant, on ne cherche plus, on discute avec une IA. Et comme elle complète du texte, plus ta question est précise, complète, riche en contexte... plus la réponse sera pertinente.

Un prompt efficace, c'est un prompt bien formulé. Exemple :

- Mauvais prompt : "Écris un article sur l'IA."
- Bon prompt : "Écris un article de 1000 mots sur l'impact de l'IA dans les métiers de la santé, en adoptant un ton pédagogique, à destination des lycéens."

Plus on est clair, plus l'IA est étonnante, et parfois bluffante.

GARDER SON SENS CRITIQUE

Mais attention. Ce n'est que de la statistique. Pas une vérité absolue.

L'IA peut se tromper, inventer, halluciner. Et parfois, elle le fait avec aplomb, avec un ton tellement confiant qu'on la croit ! C'est dangereux.

Pourquoi ? Parce que tout dépend des données qu'elle a vues. Et comme elle ne comprend pas, elle peut mélanger des infos vraies et fausses. Pire, elle peut fabriquer des citations, inventer des faits, assembler des morceaux de phrases incohérents.

C'est comme un GPS. Il peut te guider dans une rue en sens interdit. Mais tu ne vas pas l'écouter les yeux fermés. Tu gardes ton sens critique, tu regardes la route. Avec l'IA, c'est pareil : tu vérifies, tu recoupes, tu confrontes. L'IA est un assistant, pas un oracle.

LES AGENTS... ET LES AGENTS D'AGENTS

La prochaine étape, c'est l'arrivée des agents intelligents. Jusqu'ici, l'IA générative était capable de créer du contenu. Mais désormais, elle va agir.

Un agent, c'est un programme capable d'utiliser les réponses de l'IA pour accomplir une tâche, interagir avec d'autres services, prendre des décisions. Et si on les combine, on peut leur faire résoudre des missions complexes.

Imagine que tu veux organiser des vacances :

1. Un premier agent discute avec toi pour comprendre ce que tu veux : plutôt plage ou musée ? France ou Japon ?
2. Un autre agent optimise l'itinéraire selon tes goûts.
3. Un troisième réserve les billets de train ou d'avion.
4. Un quatrième cherche les meilleurs hôtels.
5. Un autre encore gère ton planning, ajoute les adresses à ton GPS...

Et tout cela sans que tu aies besoin d'intervenir, sauf au début pour exprimer ton besoin.

C'est ça, le futur des agents. Des intelligences spécialisées qui collaborent entre elles, pour répondre à tes demandes. Comme une équipe virtuelle à ton service.

EST-CE QUE LES AGENTS VONT DEVENIR UNE INTELLIGENCE COLLECTIVE ?

On y est, ou presque. L'intelligence artificielle est là. Pas pour nous remplacer, mais pour nous accompagner. Bon, on espère, on connaît tous les scénarios catastrophe des films des 30 dernières années. Mais bon, l'Intelligence artificielle est comme un nouveau collaborateur. Un assistant puissant. Un partenaire créatif. Un exécuter rapide.

Mais attention : elle ne pense pas à notre place. Elle n'a pas de conscience, pas de valeurs, pas d'intuition. Ce qui fait de nous des humains, c'est notre imagination, notre sens critique, notre éthique. Il faut aussi faire attention au biais que chaque société créatrice d'IA peut introduire dans ces algorithmes pour nous influencer, nous contrôler, nous orienter de façon volontaire ou non.

L'IA va entrer dans l'intelligence collective de l'humanité. Elle sera une voix parmi d'autres.

Mais c'est à nous, humains, de rester aux commandes. De poser les bonnes questions. De rêver. D'inventer. De douter.

Ce nouveau tournant est irréversible. Alors autant apprendre à le prendre ensemble, avec lucidité, curiosité... et un bon sens critique !

LAURENT BUTRÉ

Laurent est un expert chevronné du numérique, avec près de trois décennies d'expérience chez SLB (anciennement Schlumberger).

De ses débuts dans le support informatique jusqu'à la direction des opérations digitales mondiales en tant que 'Digital Operations Director', il a passé plus de 15 ans à l'international.

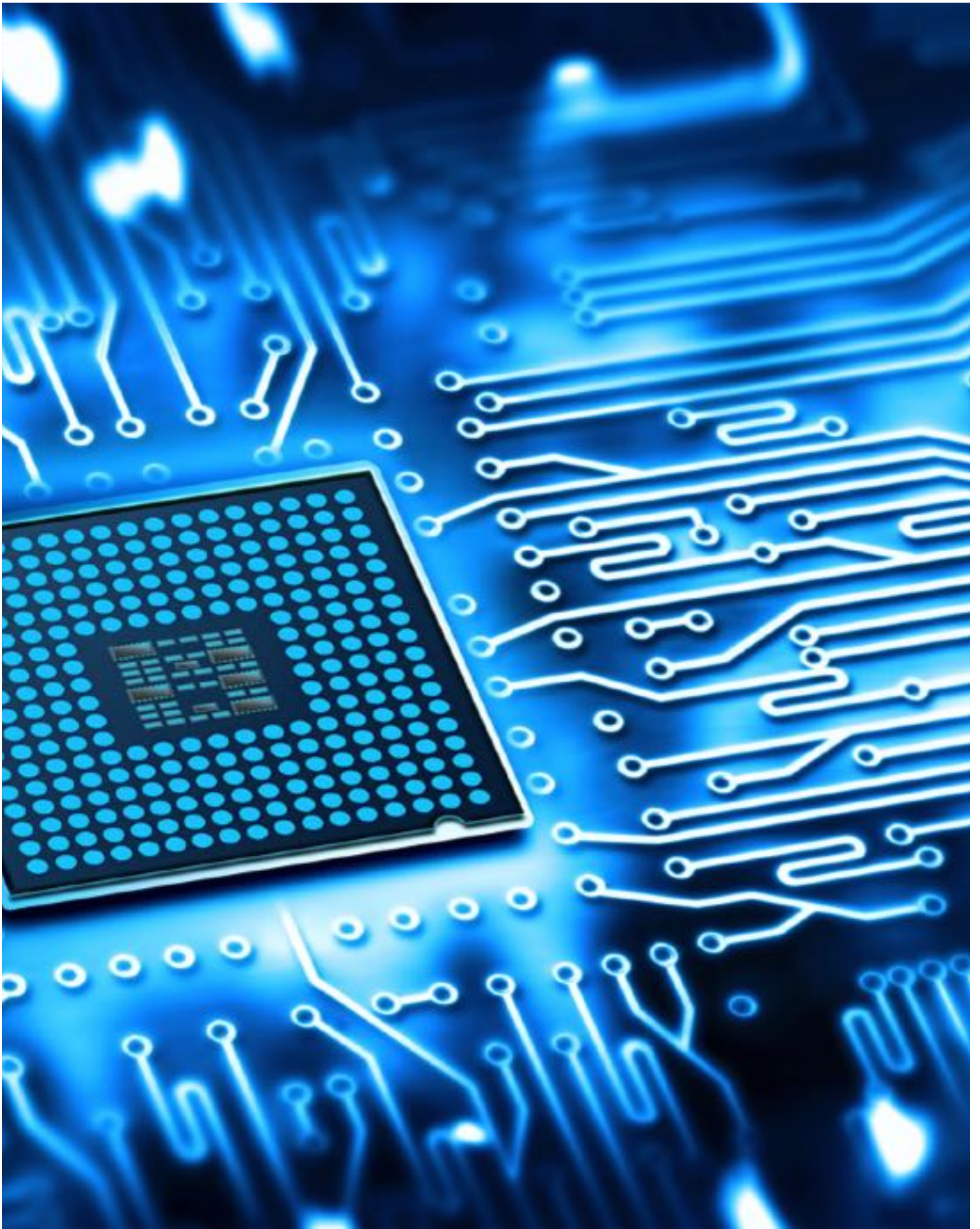
Passant par des rôles clés en Australie, au Texas, ainsi que dans la Silicon Valley, il est actuellement basé à Paris après être rentré en France pour démarrer un Lab d'Intelligence Artificielle à Paris (Clamart, 92).

Laurent a toujours eu un fil conducteur : simplifier la technologie pour maximiser l'impact: *'Make it simple; initiating purposeful impacts'*.

Diplômé de l'ENSIMAG, il siège également au conseil d'administration de l'IHES.

Ce qui le motive ? Créer des solutions qui font gagner du temps et améliorent concrètement la vie des utilisateurs.

[Retour au sommaire](#)



LES CINQ DÉFIS ESSENTIELS DE L'IA

PAR SYLVESTRE BÉNARD

Né dans le digital, diplômé des arts appliqués avec une spécialisation en développement multimédia, Sylvestre Bénard a toujours travaillé dans des entreprises de technologie occupant alternativement des postes créatifs et des postes de chef de projet.

Il est depuis 5 ans président de l'association Coder Dojo Annecy qui a pour mission d'initier les enfants à la programmation et de leur donner une culture du monde digital.

Il anime par ailleurs des conférences sur les enjeux de la technologie au festival Numérica d'Annecy en liaison avec le service médiation numérique de la ville.

En 2024 Geoffrey Hinton et John Hopfield reçurent conjointement le prix Nobel de physique pour leur travail sur les réseaux de neurones profonds dans les laboratoires de Google. Ils sont considérés par la communauté comme faisant partie des pères de l'IA générative.

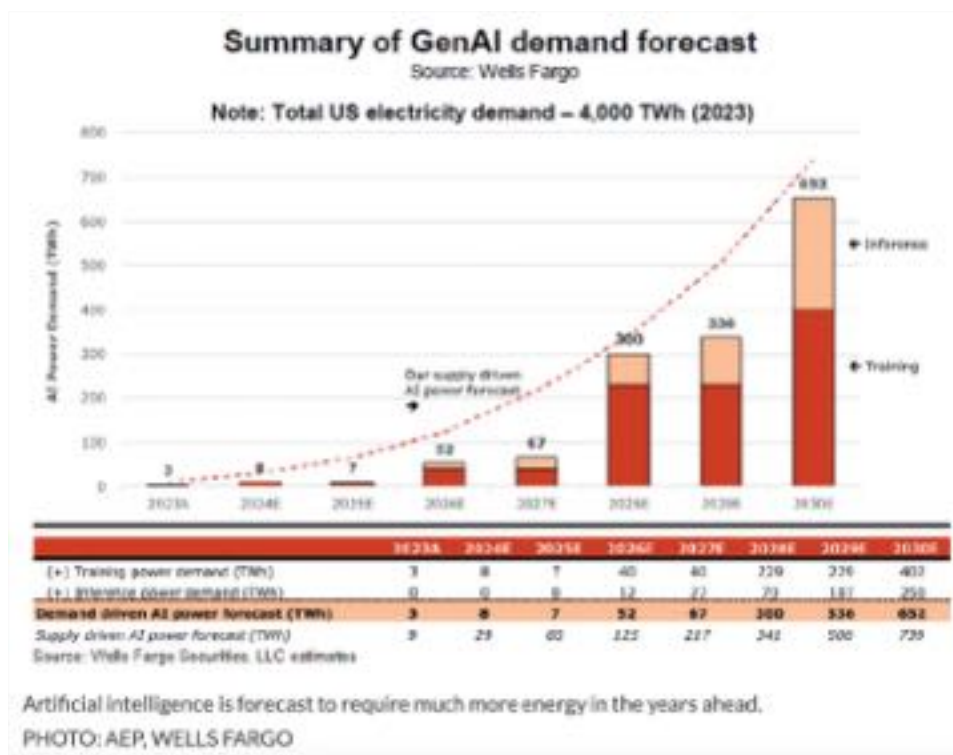
Outre le fait que de nombreux physiciens se sont alors étonnés que leur prix soit donné à des informaticiens, le fait notable de cette édition est que Geoffrey Hinton avait démissionné de Google quelques mois auparavant afin de pouvoir alerter librement le public sur les dangers de l'IA générative. Lors de la cérémonie de remise du prix, il parla longuement des problèmes posés par l'IA : remplacement des postes techniques et créatifs, usage criminel ou délictuel, risque existentiel posé par l'intelligence artificielle généraliste.

Il n'est pas rare que certains scientifiques de renom alertent sur les dangers des technologies qu'ils ont contribué à faire émerger. Nous pensons naturellement à Einstein qui a longuement parlé du risque majeur que la bombe atomique faisait courir à l'humanité.

L'intelligence artificielle est aujourd'hui omniprésente, utilisée à



Sylvestre Bénard



Graphique montrant la croissance exponentielle de la demande d'énergie de l'IA

la fois par les institutions, les hommes politiques, les forces armées, les publicitaires, les journalistes, les artistes. Est-elle vraiment sans risques ?

Dans cet article je vous propose de prendre du recul et de regarder les 5 enjeux que l'intelligence artificielle générative devra dépasser si elle veut devenir une technologie mature et utile à tous : les défis de l'énergie, de la collecte de données, de l'équité et les biais, des hallucinations, et enfin le défi social.

1. LE DÉFI ÉNERGÉTIQUE

Alors que Geoffrey Hinton venait juste de recevoir son Nobel, les grands de ce monde se pressaient dans les hôtels cossus de Davos. Depuis plusieurs années l'intelligence artificielle s'était faite une place importante dans les discussions du Forum. A cette occasion, Sam Altman, le directeur général d'OpenAI - les concepteurs du célèbre ChatGPT - donnait une conférence sur le futur de l'intelligence artificielle ; concernant le besoin en énergie pour le développement de l'IA généraliste, il déclarait :

« Il n'y a pas de moyen d'y parvenir sans percée. Cela nous motive à investir davantage dans la fusion nucléaire. »

Il avait jeté un pavé dans la mare, les journalistes prenaient conscience que l'énergie allait être un des enjeux majeurs pour la croissance des usages de l'IA.

Il faut en effet savoir qu'une requête sur une IA générative, Grok, ChatGpt ou Google Gemini, consomme en moyenne 10 fois plus d'énergie qu'une simple recherche Internet. Or, cela n'est que la face cachée de l'iceberg. La consommation d'une IA générative se divise en deux phases :

- la phase d'entraînement durant laquelle elle va ingurgiter des masses énormes de données afin de construire son modèle,
- puis la phase d'exploitation pendant laquelle elle va utiliser son réseau de neurones pour répondre à des requêtes.

La phase d'entraînement nécessite une puissance de calcul démesurée. Par exemple Grok 2, l'IA d'Elon Musk, a nécessité 20 000 processeurs de calcul Nvidia H100 et Grok 3 près de 100 000. Chaque processeur H100 consomme entre 350 et 700 watts. Au total cela correspond à une consommation instantanée de 35 à 70 mégawatts soit la consommation d'une ville de 4000 habitants. Cette phase représente près de deux tiers de la consommation totale d'énergie durant la durée de vie d'une IA. Mais Sam Altman n'a pas été pris au sérieux. En effet, la fusion nucléaire étant pour l'instant un doux rêve, le PDG d'OpenAI a été considéré comme un provocateur.

Cependant, en juin de la même année, Microsoft signa un accord pour faire redémarrer le second réacteur de la tristement célèbre centrale nucléaire de Three Mile Island, fermée depuis 2019 (le premier réacteur fut le théâtre en 1979 du pire accident nucléaire de l'histoire des États-Unis). La société a alors déclaré que l'énergie produite (850 Mw) serait destinée à ses datacenters (dont la

croissance est intimement liée au développement de leurs services d'IA). Deux ans plus tard, portée par le giga projet « Stargate » de Donald Trump, la société FERMI a lancé Hypergrid, un énorme datacenter dédié à l'IA au Texas qui produira sa propre électricité (jusqu'à 11GW) en combinant nucléaire, solaire et gaz.

Une projection du cabinet Wells Fargo montre les ordres de grandeur : en 2030 l'ensemble des activités liées à l'IA dans le monde pourrait représenter 652 TéraWatts, soit plus que l'ensemble de la production annuelle de la France en 2023 (519TW).

On entend souvent des promoteurs de l'IA dire que les humains dont la tâche est remplacée par cette dernière consomment eux aussi de l'énergie. Ainsi, le déséquilibre ne serait pas si important. Toutefois, si l'on se replace en 2015 quand l'IA de Google a battu le champion du monde du jeu de Go, Lee Sedol, la consommation de l'ordinateur « AlphaGo » de Google a été de 15000 W alors que le cerveau de Lee Sedol s'est contenté d'un maximum de 50 W (20% de la consommation d'énergie d'un homme adulte).

2. LE DÉFI DE LA COLLECTE DE DONNÉES

L'IA générative n'est pas seulement une dévoreuse d'énergie, c'est aussi un ogre pour les données. En effet, si un homme peut reproduire une tâche avec quelques exemples et puiser des règles logiques dans ses expériences passées, une IA générative doit être entraînée à l'aide de millions voire de milliards de données contextualisées.

De ce fait l'absence de données d'entraînement suffisantes a longtemps été un frein pour le développement des réseaux de neurones profonds. Le tournant a eu lieu peu après 2006 avec la montée en puissance du web 2.0. Les réseaux sociaux ont alors constitué des montagnes de données utilisateurs et possédaient déjà un modèle économique centré sur l'exploitation de celles-ci.

C'est grâce à la chercheuse Fei-Fei Li de l'université de Stanford qu'une bascule va avoir lieu en 2010. Alors étudiante en informatique, elle prend conscience que l'absence de données de qualité est un frein au développement de la reconnaissance d'images par des réseaux de neurones. Elle va alors utiliser le « Mechanical Turk » d'Amazon (un réseau de travailleurs dans des pays anglophones pauvres payés à la tâche) pour classer jusqu'à 16 millions d'images. Elle va ensuite lancer entre 2010 et 2016 un concours de reconnaissance d'images qui va permettre le perfectionnement des modèles de réseaux de neurones modernes.

Dans la décennie suivante les applications génératives basées sur l'IA vont fleurir, nécessitant toujours plus de données. DeepL va par exemple utiliser toutes les retranscriptions du parlement européen et tous les livres traduits disponibles en ligne.

Les grands modèles génératifs comme Chat GPT ou Midjourney vont commencer à scanner l'ensemble du web pour alimenter leurs entraînements de plus en plus gourmands. Ainsi le dernier modèle de Dall-E (l'un des leaders de la génération d'images) se base sur près de 5.8 milliards d'images annotées et la quantité de données nécessaire à l'entraînement de ChatGPT a été multipliée par 10 en quatre ans.

Dans ce contexte un autre problème commence à apparaître : la raréfaction des données de qualité. Et cela impacte l'ensemble des acteurs de l'IA. Le besoin de données est tel que les entreprises développant des modèles d'IA captent toutes les données disponibles sans distinction de provenance et sans souci de respect de l'image ou des copyrights.

Elles commencent à utiliser des contenus générés par des IA pour leur propre entraînement.

En 2024, selon une étude d'Amazon Web Services, près de 57% des données mises en ligne ont été produites par une IA. Sur le réseau

L'ABSENCE DE DONNÉES D'ENTRAÎNEMENT SUFFISANTES A LONGTEMPS ÉTÉ UN FREIN POUR LE DÉVELOPPEMENT DES RÉSEAUX DE NEURONES PROFONDS.

POUR LA PREMIÈRE FOIS L'ORDINATEUR PARAÎSSAIT ÉGALER UN COMPORTEMENT HUMAIN SUPÉRIEUR

social LinkedIn, l'agence « Influence Metrics » a montré que 61% des posts sont issus d'une IA générative. Ceci entraîne un cercle vicieux amplifiant les défauts d'apprentissage.

Parallèlement à cet appauvrissement des données, l'accès aux données publiques s'est restreint, de nombreux sites décident en effet d'empêcher les IA de piller leur contenu. Partout dans le monde les ayants droits s'organisent pour résister au pillage.

Les législations deviennent plus protectrices. En 2023 la directive Biden vise à protéger les citoyens contre les excès de l'IA. En Europe, le RGPD et « l'Artificial Intelligence Act » imposent encore plus de restrictions aux acteurs de l'IA.

Avec cet environnement restrictif, certains analystes prédisent un ralentissement du développement et un appauvrissement des LLM les plus gourmands¹.

Cependant la résistance des géants de l'IA s'organise. En 2024, META et une trentaine d'entreprises ont envoyé une lettre ouverte à la commission européenne lui demandant de revoir les réglementations. Plusieurs spécialistes pensent également que le réel objectif d'Elon Musk avec le *Department of Government Efficacy* (DOGE) était de s'emparer des données des ministères américains (des données de très bonne qualité, parfaitement catégorisées) pour alimenter son entreprise xAI et son chatbot Grok².

3. LE DÉFI DES BIAIS ET DE L'ÉQUITÉ

En 2013, Google, alors à la pointe de la technologie des réseaux de neurones profonds, lança Word2Vec. Ce programme classait les mots sous forme de vecteurs dans un espace à 300 dimensions. Il devenait alors possible de réaliser des

opérations booléennes sur les mots, le croisement des vecteurs révélant les relations entre eux.

Par exemple :

- Japon : Tokyo :: France : **Paris**
- homme : roi :: femme : **reine**

L'article de google eut un retentissement énorme car la technique du « *word embedding* » était un test utilisé par les neuropsychologues pour évaluer la créativité de leurs patients. Pour la première fois l'ordinateur paraissait égaler un comportement humain supérieur.

Cependant, un après-midi, le chercheur Tolga Bolukbasi avec une collègue tuait le temps à la cafétéria de l'université de Boston. Ils utilisaient Word2Vec sur leurs smartphones discutant de technique. Impressionnés par la puissance du système ils essayaient différentes associations de mots. Par hasard, il tapa sans trop y penser la requête suivante : père - médecin ; mère... et le système lui renvoya **infirmière**.

A moitié surpris de la réponse il tenta deux autres requêtes

bâtiment - homme : **architecte** ,
bâtiment - femme : **femme de ménage**

Tolga Bolukbasi venait de découvrir que le système de Google était sexiste³.

Il s'avéra que le corpus sur lequel Google avait entraîné son système comportait des textes misogynes et que le système les avait intégrés à l'insu des programmeurs. C'était l'une des premières mises en évidence que les systèmes d'intelligence artificielle basés sur des réseaux de neurones pouvaient être biaisés et qu'ils semblaient même accentuer les biais.

Si cette découverte concernait une recherche académique, les dysfonctionnements de cette technologie dans la société ne

¹ <https://www.nytimes.com/2024/07/19/technology/ai-data-restrictions.html>

² <https://theconversation.com/doge-threat-how-government-data-would-give-an-ai-company-extraordinary-power-250907>

³ <https://www.technologyreview.com/2016/07/27/158634/how-vector-space-mathematics-reveals-the-hidden-sexism-in-language/>

tardèrent pas à apparaître. En 2014 l'équipe de recrutement d'Amazon aux USA décida de mettre en place secrètement un système d'IA pour trier les CV des candidats. Ils entraînèrent le système avec leur CVthèque et le lancèrent en test. Le système paraissait efficient, il recrutait des personnes compétentes.

Cependant au bout de quelques semaines la direction s'aperçut qu'elle n'embauchait plus du tout de femmes. Ils analysèrent le réseau de neurones et virent que celui-ci créait un lien fort entre les prénoms masculins et la compétence. Ils décidèrent donc de ne pas prendre en compte les prénoms. Mais l'IA se comportait toujours de la même manière. Ils analysèrent de nouveau le réseau et découvrirent que les femmes ne prenaient pas les mêmes spécialités dans leurs formations. De même, les femmes ne pratiquaient pas les mêmes sports, n'avaient pas les mêmes hobbies.

Le système avait amplifié les biais et les avait encodés par des liens multiples si entremêlés qu'il était impossible de le corriger. Amazon dut arrêter le projet.⁴

De nombreux autres exemples de biais firent scandale, comme par exemple les personnes de couleur qui étaient classées comme des gorilles par Google Images ou des systèmes de reconnaissance de visage qui fonctionnaient entre 40 et 100% moins bien pour les femmes noires qu'avec des hommes blancs, comme put le démontrer le Docteur Joy Buolamwini, fondatrice de «l'Algorithme Justice League», dans ses recherches.

Les systèmes de réseaux de neurones profonds sont des boîtes noires qui créent des corrélations invisibles entre des concepts et donnent des réponses qui amplifient les liens majoritaires, amplifiant ainsi les biais sexistes et racistes⁵.

⁴ <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/>
⁵ https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms



les générateurs d'images ont encore beaucoup de mal à générer des mains réalistes.

Ce syndrome de la « boîte noire » rend d'autre part impossible le suivi et le contrôle de ces liens dans les modèles larges comme ChatGPT car ils en comportent des centaines de milliards.

Aujourd'hui, demandez à Midjourney par exemple de créer un bureau d'architecte et il affichera 100% du temps un architecte homme. Alors qu'une simple recherche sur Google image vous montrera facilement 40% d'architectes femmes. C'est un parfait exemple d'amplification d'un biais.

Mais les biais ne sont pas qu'une erreur d'entraînement ou qu'un défaut de la technologie. Ils perpétuent les rapports de pouvoir. Si l'on prend du recul, les ingénieurs qui ont développé ces IA sont dans leur grande majorité de jeunes hommes blancs et ils ont créé des outils à leur image.

Cette propension à favoriser les données majoritaires pose d'autre part un gros problème pour la représentation des cultures minoritaires qui ne disposent pas d'assez de données écrites pour entraîner une IA de traduction ou un LLM.

4. LE DÉFI DES HALLUCINATIONS

L'autre problème majeur des IA génératives, tout aussi difficile à contourner, est leur propension à

L'AUTRE PROBLÈME MAJEUR DES IA GÉNÉRATIVES EST LEUR PROPENSION À PROPOSER DES RÉPONSES FAUSSES OU TROMPEUSES



La pub provocante lancée par une entreprise d'IA de Los Angeles : « Arrêtez d'embaucher des humains ! »

CE PROBLÈME JETTE UN DISCRÉDIT SUR LES OUTILS D'IA GÉNÉRATIVE

halluciner ou à « cofabuler » quand elles proposent une réponse fausse ou trompeuse. Nous pouvons tous en faire l'expérience en posant des questions précises sur des données que le chatbot n'a pas. Celui-ci ne dira jamais qu'il ne peut pas répondre, il générera du contenu parfaitement plausible mais faux. Il fera toujours preuve d'un excès de confiance⁶.

En 2023 Teresa Kubacka, docteur en science des données a demandé à ChatGPT de lui donner la définition de « l'électromagnon inversé cycloïdal » un terme inventé pour la circonstance. L'agent conversationnel a donné une définition tout à fait plausible mais contenant de fausses études, de fausses dates et de faux auteurs. Les fausses citations étaient si crédibles que la chercheuse dû les vérifier pour en avoir le cœur net.⁷

De la même manière les générateurs d'images ont encore beaucoup de mal à générer des mains réalistes. Ceci s'explique par le fait que personne ne photographie ses mains dans différentes positions en y ajoutant des commentaires.

⁶ <https://www.holbertonschool.fr/post/hallucinations-dia-comprendre-les-erreurs-des-intelligences-artificielles>

⁷ <https://lookalikes.substack.com/p/today-i-asked-chatgpt-about-the-topic>

Il n'y a donc aucune photo de ce type sur Instagram par exemple.

Ce problème très bien documenté jette un discrédit sur les outils d'IA générative et les sociétés essaient de lutter contre ce comportement. Certains utilisent des couples d'IA antagonistes. C'est le cas d'OpenAI avec ChatGPTpro en abonnement à 200€ par mois. Les deux IA donnent toutes les deux une réponse et si elles sont trop différentes c'est une hallucination.

D'autres ont découvert que désactiver des portions du réseau de neurones pendant l'apprentissage permet de limiter les liens croisés qui créent des corrélations entre des concepts très éloignés. Toutefois, aucune solution n'a encore été trouvée pour résoudre ce problème. Celui-ci apparaît comme inhérent au fonctionnement des IA génératives.

5. LE DÉFI DE L'ACCEPTATION SOCIALE

Comme nous l'avons vu, l'intelligence artificielle est une technologie qui est loin d'être mature. Pourtant, elle est de plus en plus utilisée par de nombreux acteurs et parfois pour des tâches critiques pouvant avoir de graves conséquences sur la vie des citoyens.

En 2023 aux USA commença l'une des plus grandes grèves de l'histoire du cinéma et des médias : la grève des scénaristes. Elle a majoritairement été présentée comme une lutte salariale, mais l'une des principales raisons de cette mobilisation était l'utilisation massive des chatbots par les studios. Pour la première fois de l'histoire, des machines menaçaient des emplois créatifs de haut niveau.

De fait, c'est à cette période que les entreprises ont commencé à automatiser des tâches de backoffice avec des IA. D'ailleurs, une campagne de publicité affichée dans les rues de San Francisco jeta le trouble : les affiches de la société Artisan appelaient à ne plus embaucher d'humains et à les remplacer par leurs agents IA.⁸

⁸ <https://gizmodo.com/ceos-are-quietly-telling-us-the-truth-ai-is-replacing-you-2000621907>

Depuis la tendance n'a fait que se confirmer. Amazon a commencé en 2025 à réduire ses effectifs alors que son chiffre d'affaires bondissait. Pour Jeff Bezos « *grâce à l'IA, nous aurons besoin de moins de monde pour les métiers que nous faisons aujourd'hui* ».

Luis von Ahn PDG de Duolingo a décidé que les primes ne seraient distribuées qu'aux équipes qui automatiseront certaines de leurs tâches. De même, Tobi Lütke, le PDG de Shopify, a déclaré qu'avant chaque embauche il faudrait désormais prouver que la tâche ne pouvait pas être faite par une IA.

Parallèlement plusieurs agences de publicité utilisent l'IA générative pour leurs campagnes. Par exemple l'agence de voyage TUI a utilisé Lena, une influenceuse créée par IA pour représenter sa marque. Ceci précarise énormément le monde de la création : les modèles, les photographes et les graphistes alors que ces IA génératives ont été entraînées sur leur travail.

L'impact des biais sur les utilisations critiques de l'IA peuvent aussi être dramatiques. Une voiture Tesla en mode automatique a par exemple identifié une grand-mère avec un déambulateur comme une moto. Aux USA, Il y eu de nombreux exemples de personnes arrêtées à tort sur la base de fausses identifications par des systèmes de reconnaissance faciales de la police ou des systèmes de détection de cancer de la peau par IA qui faisaient de faux négatifs sur les peaux foncées.

Mais les utilisations les plus problématiques de l'intelligence artificielle sont évidemment à trouver pendant les conflits. En Ukraine, certains drones sont à présent équipés d'intelligences artificielles capables de suivre le terrain de manière autonome jusqu'à leurs cibles afin d'éviter les brouillages. Le quotidien Haaretz a par ailleurs révélé que les cibles à Gaza sont identifiées par l'armée Israélienne à l'aide d'une IA. Tsahal affirme que la décision finale revient à un humain, cependant l'IA scanne 30 000 téléphones portables (cible aléatoire car les téléphones portables s'échangent et se revendent énormément à Gaza) et le

responsable humain n'a que 20 secondes pour valider la légitimité de la frappe.

Malgré le manque de maturité de la technologie, celle-ci se déploie donc à grande vitesse dans tous les secteurs de la société.

Le risque le plus préoccupant vient de l'utilisation que font les jeunes générations des chatbots. Un nombre croissant d'adolescents utilisent en effet les IA génératives de manière intensive pour rédiger des textes ou pour faire des recherches. Il n'est pas rare de voir des élèves de collège cacher des smartphones pour les examens pour pouvoir interroger Chat GPT. Cet usage a malheureusement un coût cognitif important comme le montrent plusieurs études :

- Microsoft et Carnegie Mellon University en 2025 : « *L'utilisation de l'IA a un impact sur l'esprit critique* », ⁹
- Université de Pennsylvanie 2025 : « *les personnes qui ont utilisé des chatbots ont eu tendance à développer des connaissances moins approfondies* ». ¹⁰

Cette dernière étude est particulièrement préoccupante car elle porte sur 4500 lycéens et a comparé un groupe utilisant Google pour faire ses recherches, avec un groupe utilisant ChatGPT. Il s'est avéré que le second groupe a montré moins d'activité cérébrale.

L'utilisation accrue des IA pour rechercher du contenu va d'autre part réduire la fréquentation des sites web et des producteurs de contenu. Les principaux sites d'information aux USA ont par exemple constaté une baisse de 24% de leurs visiteurs en 2025. Ceci risque d'accélérer le déclin des informations fiables, de faire entrer les IA dans une spirale appauvrissante et d'entraîner encore un peu plus le monde dans la post vérité.

L'IA est une révolution qui ouvre de nouveaux horizons à la science avec sa capacité de traiter des volumes de

⁹ https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf

¹⁰ <https://gizmodo.com/multiple-studies-now-suggest-that-ai-will-make-us-morons-2000620811>

données hors de portée des humains, comme par exemple DeepMind qui a prédit de nouvelles formes de protéines qui échappaient jusque-là à tous les calculs. Elle va aussi certainement apporter de grands bénéfices dans la santé ou permettre des analyses financières plus rapides. Mais, comme nous l'avons vu, l'utilisation des IA génératives comme source d'information ou pour remplacer les hommes dans une série de tâches intellectuelles comporte de très grands risques sociaux et cognitifs.

Il y a d'autre part une tentation de la part d'acteurs privés ou publics d'utiliser les IA à des fins de surveillance et de contrôle des citoyens malgré les nombreuses mises en garde sur la présence de biais. La distribution de nourriture à Gaza par la fondation américano-israélienne GHF est contrôlée par un système de reconnaissance faciale totalement opaque ¹¹.

Et en France, selon Disclose, les services de police ont acquis en 2015 le logiciel de reconnaissance faciale israélienne Briefcam et dissimule depuis son utilisation ¹².

Pour Shoshana Zuboff docteur en sociologie et autrice de « *L'âge du capitalisme de surveillance* » le développement de l'IA n'est qu'une extension du contrôle des données utilisateurs que pratiquent les GAFAM. Leur modèle économique a été mis en péril car il a été dévoilé par la société civile et les nouvelles législations alors qu'il doit par essence rester caché. Ils ont donc mis en place un nouveau cheval de Troie pour capter toujours plus de données personnelles : l'IA générative ¹³.

Retour au sommaire

¹¹ https://fr.wikipedia.org/wiki/Fondation_humanitaire_de_Gaza

¹² <https://disclose.ngo/fr/article/la-police-nationale-utilise-illegalement-un-logiciel-israelien-de-reconnaissance-faciale>

¹³ <https://www.radiofrance.fr/franceculture/podcasts/le-meilleur-des-mondes/capitalisme-de-surveillance-enjeux-et-apprentissages-d-une-theorie-politique-du-numerique-5831994>



L'ÉTHIQUE DE L'INTELLIGENCE ARTIFICIELLE

ENTRETIEN AVEC MONSIEUR DEBOS, MAÎTRE DE CONFÉRENCES À L'UNIVERSITÉ DE NICE CÔTE D'AZUR

PAR GUY MENANT

Bonjour Monsieur Debos, merci de nous avoir accordé cette interview. Pourriez-vous vous présenter pour nos lecteurs et décrire brièvement votre domaine de recherche ?

Je suis Maître de Conférences en sciences de l'information et de la communication à l'Université de Nice Côte d'Azur.

Au début, j'ai passé un doctorat en sciences de gestion, plutôt centré dans le domaine du marketing. Je me suis réorienté vers les sciences de l'information et de la communication, avec deux pôles principaux : un pôle concernant tout ce qui est communication autour de la transition écologique, et un autre pôle qui a trait aux usages des innovations, de la diffusion d'innovation, de la création d'innovation ; notamment l'innovation venant de la créativité des gens ordinaires, dans des logiques de co-conception avec l'ensemble des parties prenantes de la société.

Je me suis naturellement intéressé à la dimension numérique. Et puis, de

là, j'en suis venu à travailler sur les notions d'intelligence artificielle, en m'intéressant à la dimension éthique, au cadre qu'il faut respecter.

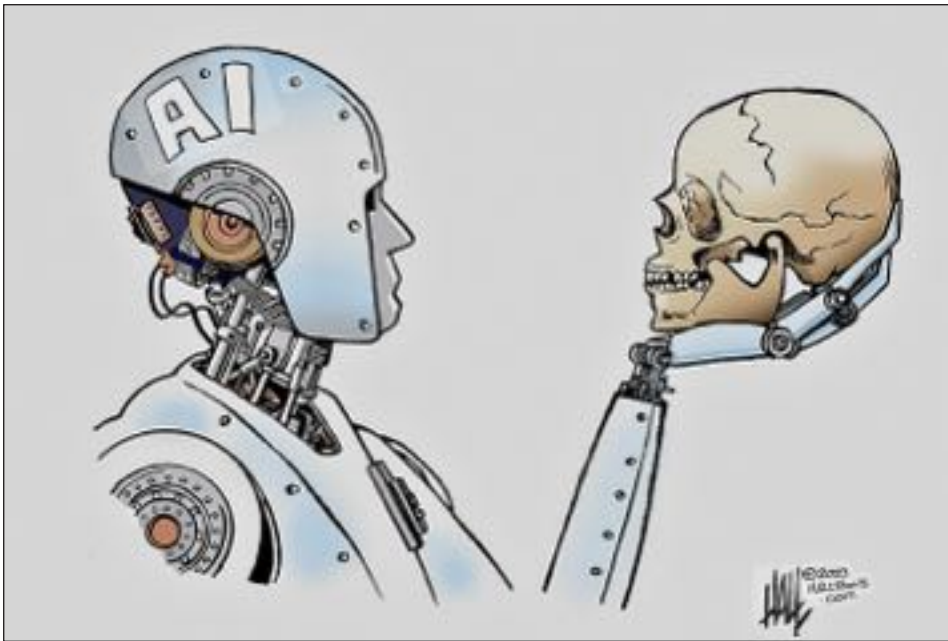
Vous avez récemment dirigé la rédaction d'un ouvrage collectif sur l'éthique de l'intelligence artificielle¹. Pourriez-vous nous préciser ce qui a conduit votre équipe à publier sur ce sujet ?

Je fais partie d'un certain nombre de structures, notamment de l'*Observatoire territorial des impacts technologiques économiques et sociétaux de l'intelligence artificielle* (OTESIA). Dans cet observatoire étaient représentés différents domaines, aussi bien des chercheurs en sciences, on va dire exactes, ou en sciences humaines et sociales, et pas simplement l'université. On travaillait avec des écoles de commerce, on essayait d'avoir une vision assez multidisciplinaire.

Et on va aborder l'intelligence artificielle aussi sous le cadre de l'éthique. Qu'est-ce qu'il faut faire ? Quelles sont les éthiques qui sont



¹ Sous la direction de Franck Debos, « L'éthique de l'intelligence artificielle » - ISTE editions - 2025



Ed Hall (USA), Cartooning for Peace

SOUVENT ON EST DANS UNE LOGIQUE UTOPISTE AU DÉBUT, ON VOIT SURTOUT LES CÔTÉS POSITIFS. IL FAUT AUSSI RÉFLÉCHIR À L'IMPACT GLOBAL QUE CELA PEUT AVOIR

associées ? Comment on peut faire pour réguler, ce n'est pas le meilleur terme, mais on va dire pour encadrer, coordonner. À partir de là, dans les actions qu'on avait à mener, j'ai eu envie de faire un ouvrage collectif.

On a essayé, chacun dans sa discipline, d'aborder certains aspects autour de l'IA, justement en essayant d'avoir en arrière-plan cette dimension de cadrage, de régulation, de réflexion en amont. Des réflexions qui, on l'espère, vont amener à des critiques, à des échanges et aussi à d'autres types de production sur cette thématique.

Il est important de regarder l'histoire des innovations, et des diffusions des innovations. Souvent, on est dans une logique utopiste au début, on voit surtout les côtés positifs et on est enthousiasmé. Or il faut qu'on réfléchisse aussi aux conséquences, à l'impact global que ça peut avoir. Et pas simplement au niveau des lieux où ça va être utilisé ou pour les personnes qui vont l'utiliser, mais de façon plus globale pour tous ceux qui vont participer à la chaîne de production, de conception des outils et pas simplement du logiciel ou de l'algorithme, mais aussi de tous les supports qui vont utiliser les IA.

Quelles particularités des logiciels d'IA conduisent à réfléchir à une éthique spécifique ?

Dans cette logique d'éthique, il faut se mettre aussi dans un contexte de technologie. Il faut d'abord voir qui développe les outils de technologie : des chercheurs (en informatique, robotique, etc.), des ingénieurs.

On est dans une logique *conséquentialiste* où on va avoir une démarche très rationnelle.

On va développer par exemple un outil d'intelligence artificielle au niveau de la santé, pour essayer d'avoir des diagnostics beaucoup plus précis. On va évaluer une marge d'erreur. On va dire que si finalement dans 95% des cas ça marche, alors c'est très bien, il n'y a que 5% de marge d'erreur. Donc on estime que c'est quelque chose de bien, ce qui est vrai au plan statistique, on ne pas dire le contraire. Voilà donc un premier niveau dans lequel on évalue les conséquences positives et négatives et on mesure si le positif l'emporte sur le négatif : là, on est juste dans un cadre d'évaluation de probabilité.

Or il y a d'autres choses que l'on peut prendre en compte.

Dans l'ouvrage collectif déjà mentionné, vous avez publié un article sur les smart cities², fondées sur une interconnexion des réseaux de données impliquant de plus en plus l'IA et sur la nécessité de « bien mettre la dimension humaine au centre de cette vision digitalisée de la cité ». Pourriez-vous nous préciser quel cadre éthique pourrait éviter les risques de dérives ? Comment notamment concilier une gestion assistée par l'IA et une gouvernance participative ?

² Smart city - La ville intelligente est un nouveau concept de développement urbain. Il s'agit d'améliorer la qualité de vie des citoyens en rendant la ville plus adaptative et efficace, à l'aide de nouvelles technologies qui s'appuient sur un écosystème d'objets et de services. Le périmètre couvrant ce nouveau mode de gestion des villes inclut notamment : infrastructures publiques (bâtiments, mobiliers urbains, domotique, etc.), réseaux (eau, électricité, gaz, télécoms) ; transports (transports publics, routes et voitures intelligentes, covoiturage, mobilités dites douces - à vélo, à pied, etc.) ; les e-services et e-administrations.

On le voit très bien dans le domaine des *smart cities*, il y a beaucoup de villes où souvent, les solutions qui sont proposées sont des solutions purement technologiques. On est dans ce qu'on appelle le *technosolutionnisme*. Qu'est-ce qu'il y a derrière ? Il y a bien sûr quelqu'un qui fait une commande, qui va être par exemple le maire d'une ville, qui dit voilà, moi j'aimerais que *ma ville soit bien connectée pour tout gérer, les flux de matière, les flux de mobilité, l'énergie, la vie des habitants*. On va le faire avec les grandes entreprises, qui vont bien nous montrer que c'est beaucoup plus positif que négatif. Sauf que dans ce contexte-là, on ne demande pas forcément, non seulement aux habitants, mais à tous les acteurs socio-économiques, aux petites entreprises dans les quartiers, aux artisans, etc. Toutes les entreprises, si vraiment ça les intéresse, quelles idées elles pourraient avoir, comment elles le voient ? Est-ce qu'obligatoirement, on doit mettre des solutions technologiques de haut niveau ? Est-ce qu'il n'y a pas d'autres solutions ? Pour faire une *Smart City*, normalement une ville aimable pour toutes et tous, pour vivre mieux, pour tous les acteurs, bien sûr les habitants et pas que les habitants ?

Sur un plan de développement de la ville, d'attractivité économique, ces solutions technologiques peuvent être plus positives que, je dirais, des valeurs humaines de l'interactivité entre les personnes, du vivre ensemble, d'inclusion, de la prise en compte des conditions environnementales. Mais à cette logique dite conséquentialiste, on peut intégrer d'autres logiques qui sont associées à ce qu'on appelle par exemple le *déontologisme* qui va aussi intégrer la notion d'expérience morale. Comme ont fait des villes en Finlande ou même à Medellín en Colombie (qui avait plutôt l'image négative qu'on connaît

par ailleurs), dans lesquelles on fait intervenir les habitants et les acteurs socio-économiques de quartier pour qu'ensemble, ils trouvent des solutions dans lesquelles on intègre aussi, bien sûr, des solutions technologiques, et dans lesquelles l'IA bien évidemment a sa place. Mais *on part déjà des attentes et des besoins de ceux qui vont principalement être intéressés*. Et on met quand même un cadre, un cadre « moral » qui correspond à des notions de prise en compte de tous. Et surtout de prise en compte aussi d'une certaine *liberté*, parce que lorsqu'on est dans cette logique de «ville intelligente», on est quand même dans une logique de contrôle, même sans que la personne s'en aperçoive forcément, de contrôle qui peut être extrêmement important.

Lorsqu'on crée un logiciel, un algorithme, un robot dans lequel il y aura forcément de l'IA, une voiture ou ce qu'on veut, il faut peut-être aussi travailler avec des chercheurs en sciences humaines et sociales pour intégrer ces aspects-là. En disant, ça c'est très bien, mais voilà les conséquences globales. Et quel modèle économique il y a derrière ? Est-ce que ce modèle économique est si bien, est-ce que les gens peuvent être mieux au plan de la qualité de vie ? Est-ce qu'ils vont être mieux en termes d'expression de soi, de liberté ?

Les territoires prennent de la valeur, ça veut dire que les biens sont de plus en plus chers. Et ça pousse les personnes qui habitaient dans ces villes ou ces quartiers à les quitter. C'est le phénomène de *gentrification* qui est très vieux, mais qui est accéléré avec cela. Dans ces nouveaux immeubles ultra connectés, donc avec plein d'IA qui permettent de précisément gérer au mieux, souvent on sait très bien que ce ne seront pas forcément des locaux qui pourront habiter là. Ces gens-là, ils partent, on ne les fait pas partir volontairement, bien sûr, sauf dans le cas d'expropriation. On s'aperçoit que partout dans le monde, on va retrouver des villes qui

LORSQU'ON EST DANS CETTE LOGIQUE DE « VILLES INTELLIGENTES », ON EST QUAND MÊME DANS UNE LOGIQUE DE CONTRÔLE



Tjeerd Royaards (Pays-Bas). Cartooning for Peace.
"Oh non, elle est plus intelligente que nous!"

UN AUTRE PROBLÈME EST CELUI D'UTILISER SCIEMMENT L'IA POUR CONTRÔLER LES POPULATIONS

vont être quasiment les mêmes pour le même type de population. Et je ne pense pas que ce soit vraiment le but. L'intérêt d'une IA, et d'une ville technologiste, c'est qu'elle soit accessible à tous, partagée pour tous et adaptée en fonction des besoins.

L'essor récent et rapide des IA génératives impacte en profondeur les modes de communication et d'information des individus et des groupes sociaux. Or, les sources utilisées et les calculs réalisés par les logiciels, entre les questions posées et les réponses fournies, sont opaques pour les utilisateurs. Quels risques et dérives sont déjà constatés ? Comment gérer les biais, mais aussi les utilisations néfastes intentionnelles ? Comment pourrait-on apprendre à l'intelligence artificielle les valeurs humaines³ ?

Dès 2018, Safiya Noble publiait « *Algorithms of Oppression : How Search Engines Reinforce Racism*⁴ » où elle

³ Brian Christian « The Alignment Problem : Machine Learning and Human Values »

⁴ Biais discriminatoires : les moteurs de recherche reflètent les biais négatifs de la société (par exemple résultats avec « fille noire » comparés à « fille blanche », à relier à la plus forte représentation des hommes blancs parmi les personnels travaillant pour ces moteurs de recherche)

montrait l'impact du profil des personnes qui vont nourrir les algorithmes : stéréotypes, croyances, vont introduire des *biais involontaires*. Par exemple, l'impact de l'IA sur la gestion de ressources humaines est une question très importante. C'est un outil qui permet de gagner du temps sur la lecture de C.V., de lettres de motivation, l'identification de profils. Mais après, il faut qu'il y ait une réflexion et un travail de l'utilisateur.

D'où l'intérêt d'avoir *des équipes les plus diverses possibles* pour essayer de contrebalancer ces biais induits en amont dans les logiciels. On le sait depuis longtemps, mais il faut être d'autant plus prudent avec des outils qui vont se diffuser partout, du fait de leur coût attractif (ou même leur non-coût, pour certaines petites versions gratuites).

Un autre problème est celui d'utiliser sciemment l'IA pour *contrôler les populations*. Là, on rentre dans un cadre qui existe depuis longtemps, qui a peut-être été accentué. Dans le monde actuel, on voit que ce ne sont pas les démocraties qui dominent. Des outils d'IA et de DSIC⁵ seront utilisés pour contrôler des populations. Même dans les démocraties, on va avoir une tentation du contrôle. L'IA va être employée par exemple pour améliorer le bien-être physique. Les villes en Chine, au Japon, en Corée du Sud, dans beaucoup de pays de l'Asie du Sud, ont beaucoup de technologies associées à l'IA, ces pays sont bien plus en avance que nous. On a l'impression que les gens vivent mieux. Ils vivent mieux au plan physiologique, au plan du confort. Mais ils sont beaucoup plus contrôlés. Pour autant, *ce n'est pas l'outil en tant que tel qui est responsable*. C'est pour ça qu'il faut vraiment *un cadre moral*, en relation avec la culture de chaque population.

Des députés républicains ont introduit dans la grande loi budgétaire de Donald Trump une disposition visant à interdire aux 50 États américains de réguler l'intelligence artificielle pendant dix ans. On constate

⁵ DSIC – Direction des systèmes d'information et de communication

aujourd'hui une alliance entre les Big Tech (OpenAI, Meta, xAI d'Elon Musk, etc.) et le parti Républicain aux États-Unis pour empêcher toute régulation de l'Intelligence Artificielle. Comment expliquez-vous cette alliance ?

Les leaders du numérique, leur intérêt, c'est de dominer leur secteur d'activité. À partir de là, ils veulent développer leur technologie. Eux, ils ont un projet de société qui dépasse le cadre géopolitique. Leur société idéale, c'est peut-être pour Elon Musk une société transhumanisée, ultra robotisée ; pour d'autres, ça va être une société dans laquelle leurs outils vont dominer et vont cadrer la vie des gens.

On va développer de multiples outils qui vont permettre de mieux connaître l'individu, de mieux le servir, mais aussi de mieux l'influencer. Ça, c'est la partie aval. Et en amont, on installe un système économique où on est revenu deux siècles en arrière, avec une logique de néo-esclavagisme de ressortissants de pays en situation difficile engagés pour presque rien par des géants du numérique qui les emploient pour entraîner, nourrir constamment ces outils. C'est le *travailleur du clic*, comme disait Antonio Casilli⁶ : il n'y a pas d'intelligence artificielle, il n'y a que des travailleurs du clic. Pour le plaisir de quelques personnes dont on peut faire partie, on a mis en place un système qui est loin d'être vertueux, au contraire, qui nous renvoie sur des modèles qui ont toujours existé, mais dont on pensait qu'ils allaient disparaître.

C'est pour cela que quand je parle d'éthique, il y a l'éthique pour l'outil, pour l'algorithme, mais il y a aussi le *cadre éthique global*. On a dans les mains, on crée des choses. On sait que ça va complètement, on le voit très bien, changer nos modes de vie. Ça va avoir un impact énorme sur les individus, dans les jeunes classes d'âge surtout mais pas seulement.

Comment l'éducation aux médias peut-elle et doit-elle s'adapter ? Comment on peut-on former les citoyens à des usages éclairés et éthiques des IA ?

⁶ Antonio Casilli « En attendant les robots : Enquête sur le travail du clic » - Seuil, 2019

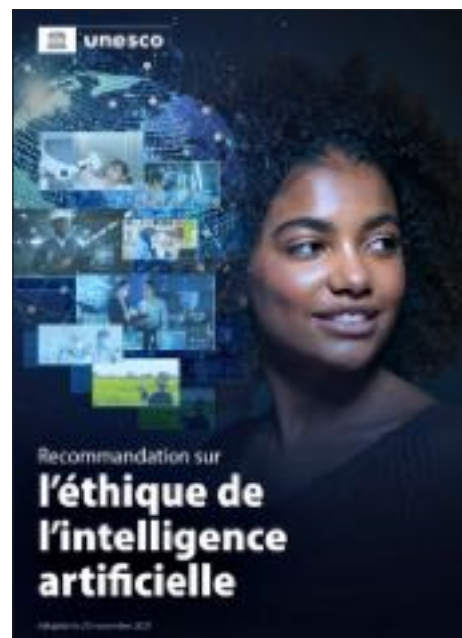
Avant qu'on parle d'IA, cela fait une vingtaine d'années que l'on constate avec la création du smartphone un impact très important, une emprise numérique très forte, surtout à partir du collège jusqu'à l'université mais même dans le primaire.

Bien sûr qu'on ne va pas revenir sur des logiques d'interdiction, comme on l'entend dire à propos des réseaux sociaux, interdiction d'ailleurs peu réaliste. Je pense qu'il peut y avoir des logiques de *régulation*. On peut peut-être avoir des outils qui permettent de bloquer les connexions Internet parce que « quand on est en cours, on est en cours ! », pour retrouver de l'attention, de la concentration. Cela fait longtemps que l'on travaille sur ces problèmes, mais ce qui a changé c'est peut-être l'ampleur des phénomènes, ce sont les technologies qui ont évolué.

C'est donc bien une question d'éducation, de *sensibilisation*. Aujourd'hui lorsque l'on leur demande quelque chose, de suite les étudiants ou les élèves vont sur ChatGPT.

L'important, c'est une prise de conscience : leur dire voilà, vous avez fait ça, c'est bien, vous avez répondu, mais ce n'est pas vous qui avez répondu. Donc, vous n'avez pas fait preuve de *créativité*. On peut parfois, par exemple, dire on enlève tout, vous réfléchissez par vous-même et vous allez voir ce que vous allez développer. On demande par exemple aux étudiants de concevoir seuls des produits autour du numérique : ils sont plus satisfaits, ils voient le travail qu'ils ont réalisé. Je leur dis : ChatGPT, vous pouvez l'utiliser par exemple si vous avez un mémoire à faire, pour aller plus vite. Mais ensuite, pensez à ce que votre production soit la vôtre.

L'importance de cela paraît moindre aux élèves du secondaire que dans le supérieur, parce que là, ils vont rentrer dans le marché du travail. Qu'est ce qui va faire la différence entre vous et les autres ? C'est justement votre propre créativité, votre propre personnalité. Ne pas sortir des choses que sortent tous les autres parce que vous avez les mêmes outils pour répondre aux



QUAND JE PARLE D'ÉTHIQUE, IL Y A L'ÉTHIQUE POUR L'OUTIL, POUR L'ALGORITHME, MIS IL Y A AUSSI LE CADRE ÉTHIQUE GLOBAL.



SI ON VEUT METTRE EN PLACE UNE SOCIÉTÉ QUI SOIT PLUS JUSTE, PLUS ÉQUILBRÉE TOUT EN CONTINUANT À ÉVOLUER AU PLAN TECHNOLOGIQUE, LE RÔLE DES ASSOCIATIONS EST EXTRÊMEMENT IMPORTANT

questions. Il faut faire prendre conscience de l'importance de ne pas être dans une *situation de dépendance* par rapport à ces outils.

Il est important de les entraîner à *l'esprit critique*. Vous lisez des textes, vous voyez des images : croisez les informations, réfléchissez. Cela existe depuis bien avant l'intelligence artificielle. Dès que le numérique s'est développé, de toute façon, ça a été le cas. On va mettre au même niveau l'information d'un scientifique qui a travaillé 30 ans sur un sujet et un influenceur : il n'y a plus de hiérarchie dans les compétences. il n'y a plus de hiérarchie de ce qui est une bonne information.

Dans les articles qu'on a pu lire dans votre ouvrage, et en particulier dans le vôtre sur les *Smart Cities*, vous avez bien montré que vous avez à la fois interrogé des experts, des décideurs et des associations de manière à ce que les citoyens soient représentés. Et donc, pour représenter les citoyens, vous êtes intervenu auprès d'associations : est-ce parce que dans les associations, ce ne sont des gens qui apportent une complémentarité en coopérant entre eux, avec la diversité de leurs cultures, de leurs profils, de leurs vies ? Comment les associations pourraient-elles éviter l'enfermement

individuel dans les produits numériques, notamment depuis l'émergence des IA ?

Si on veut mettre en place une société qui soit plus juste, plus équilibrée, tout en continuant à évoluer au plan technologique, le rôle des associations est extrêmement important parce qu'elles représentent les citoyens. Il faut qu'elles donnent leur avis parce qu'elles sont au courant de ce qui se passe vraiment et des attentes réelles. Et à partir de là, dans la mise en place des *Smart Cities*, on crée un modèle, on crée des choses qui correspondent à ces attentes réelles, comme ils ont fait dans certaines villes : je parlais de Helsinki, ou comme je le disais même à Medellin.

Mais les attentes exprimées sont plus souvent des choses qui ne sont pas forcément liées à l'IA, justement. Donc, des entreprises préfèrent directement passer par les gouvernements et les collectivités : « *on a des solutions technologiques et ne vous inquiétez pas, tout marchera bien* ». Et si ça ne marchera pas totalement, cela leur permettra de continuer à être indispensables. Il faut recadrer le modèle technologique dans le contexte, pour qu'il soit adapté aux populations visées. Et là, les associations ont un rôle essentiel. En plus de l'avis de chefs d'entreprises ou de dirigeants de collectivités territoriales, le tissu associatif a une expertise extrêmement importante pour mettre en place un modèle dans lequel on intègre des technologies en fonction des besoins et qui peuvent être des technologies de très haut niveau, dans un cadre qui correspond à ce que veulent les citoyens. Cela permet aux citoyens de s'épanouir par ces outils, sans en dépendre. De ne pas s'enfermer dans un système extrêmement élitiste avec quelques personnes qui produisent dans le numérique et qui vont diffuser leur technologie ultra sophistiquée, mais en fonction de leurs objectifs personnels. Au contraire, ce qu'il faut, c'est que cela parte à la fois du haut et du bas : de ceux qui ont des idées, mais aussi des attentes. Et puis surtout

on s'aperçoit que la créativité existe à tout niveau. Par ailleurs, il n'y a pas que la créativité, il y a aussi tout ce qui est émotions, interactions humaines : c'est ce que font les associations. En cela, leur rôle est vital.

L'ONG LaReponse.Tech a créé Vera, une application qui permet par téléphone de vérifier si une information est fausse ou vraie. Et ça, c'est très intéressant parce que ça permet aux gens de développer leur esprit critique. Avez-vous eu connaissance dans votre travail de l'utilisation de l'IA justement pour développer l'esprit critique ?

Un collègue en parle un peu dans un des chapitres de notre ouvrage collectif, pour essayer de montrer comment des outils, en effet, pouvaient permettre d'identifier les sources des informations. Ce dont vous parlez est extrêmement intéressant, ça permet en effet de sensibiliser comme le font d'autres applications au niveau des produits alimentaires : mais elles le font sur des aspects qui sont des aspects techniques ou nutritionnels dans lesquels il n'y a pas d'aspect conceptuel. Dès qu'on est dans le domaine de l'idée ou dans le domaine des points de vue, il n'y a pas une vérité, il y a plusieurs vérités, chacun sa vérité en quelque sorte. Affirmer « la norme, c'est ça, c'est ça qui est vrai et ça, ça ne l'est pas », si c'est sur des concepts idéologiques ou autres, cela peut être plus dangereux.

Si Vera fonctionne en collaboration, en réseau, cet aspect est corrigé par les interactions des avis partagés.

On voit au niveau de l'Union européenne des tentatives très intéressantes de régulation. Mais ces régulations ne tiendront pas si elles n'ont pas l'appui des citoyens. Les municipalités de Rennes et de Montpellier ont développé des conventions citoyennes pour faire réfléchir les gens sur l'IA.

Par ailleurs, des députés républicains ont introduit dans la grande loi budgétaire de Donald Trump une disposition visant à interdire aux 50 États américains de

réguler l'intelligence artificielle pendant dix ans. On constate aujourd'hui une alliance entre les Big Tech (OpenAI, Meta, xAI d'Elon Musk, etc.) et le parti Républicain aux États-Unis pour empêcher toute régulation de l'Intelligence Artificielle. Quelles conséquences pour la régulation européenne et la place de l'Europe dans la course à l'intelligence artificielle ?

C'est très bien ce que fait l'Europe. Même si en effet, il faudrait qu'ils sortent d'un cadre technocratique pour vraiment le faire avec les populations européennes. Et là, les associations ont aussi un rôle intermédiaire essentiel.

Il ne faut pas oublier que les deux géants qui dominent le numérique et l'IA sont d'abord très largement les États-Unis, puis ensuite en dessous la Chine. Et eux, ils ne sont pas du tout dans cette optique-là et ils vont imposer chacun de son côté leur vision du monde. C'est clair que l'Europe pourrait avoir un rôle de différenciation, de montrer comment utiliser ces outils en respectant justement une logique de régulation partagée par les populations. Ceci permettra d'avoir des interactions qui vont tendre vers un outil donnant la vision la plus juste possible d'une thématique. Comme ça a été le cas, en effet, pour tout ce qui est *open source*, Wikipédia, etc.

Ce qu'on sait depuis très longtemps, c'est le décalage entre une évolution technologique et des mentalités qui n'évoluent pas. Et si la place de l'Europe, ce serait justement de sortir de la logique financière, de la logique de performance économique par le numérique ? Autrement dit, être le point de départ d'une structure collaborative qui va construire une IA fondamentalement différente, plus éthique, plus humaniste. Une IA qui a le même niveau d'avancée technologique, mais surtout dans laquelle il y a une dimension collaborative, interactive, plus forte. Si on arrive à faire cela au niveau européen, en effet il peut y avoir une logique. Parce que dire qu'on va être

CE QU'ON SAIT DEPUIS LONGTEMPS, C'EST LE DÉCALAGE ENTRE UNE ÉVOLUTION TECHNOLOGIQUE ET DES MENTALITÉS QUI N'ÉVOLUENT PAS

**LE VRAI DANGERS DES IA CE
SONT LES PRISES DE
DÉCISIONS NON
TRANSPARENTES, SANS QUE
L'HUMAIN INTERVIENNE**

plus performant, comme j'entends certains, est irréaliste : pendant que nous donnons un milliard, les Etats-Unis vont donner 200 milliards, rien qu'une entreprise. Les Chinois font de même. Il faut être clair là-dessus.

Je suis très intéressé par les pédagogies coopératives et on sait qu'en pédagogie, il y a un élément d'apprentissage qui est fréquemment cité, c'est le conflit cognitif. Quand sur un réseau social j'ai fait plusieurs recherches, YouTube ou Facebook me renvoient des informations qui correspondent aux idées que j'ai recherchées. On pourrait imaginer là un procédé de conflit cognitif, c'est-à-dire que le réseau social, au lieu de me renvoyer les mêmes idées, m'envoie de temps en temps des idées contraires. On voit très bien comment les citoyens, les individus sont très intéressés quand ce qu'ils font leur permet d'avoir une autre vision, d'être en confrontation avec leurs idées.

Une question de fond sur le fonctionnement des IA, c'est que si on fait une requête un peu simple, l'IA va nous sortir une réponse qui résulte de la compilation de tout ce qui est dit de façon universelle, en prenant ce qui est le plus présent. Ça va être un dispositif de renforcement. Alors, pour travailler les conflits cognitifs, il faudrait un logiciel qui permette d'essayer d'embrasser l'ensemble des possibles. Et à partir de là, chacun se fait son opinion. Cela serait jouable sur le plan algorithmique.

Les concepteurs ont énormément anthropomorphisé le dialogue avec les IA : il est beaucoup plus facile d'utiliser notre façon de dialoguer à nous humains que d'en inventer une autre qui va être exotique et experte. Mais le résultat sur le plan émotionnel est flagrant, on va s'imaginer qu'on discute avec un humain qui ressent les choses comme nous. Alors que pour répondre à une question, une IA n'entre pas par les idées mais par les statistiques. Il y a un écran de fumée du fait de cette interface machine.

Dans l'Asie du Sud-Est, ils aiment bien que les robots ressemblent à des

humains parce qu'ils ont justement l'impression de parler à des humains. En Europe, un petit peu moins. Mais il y a des niveaux de programmes qui arrivent à répondre de la façon la plus précise, même des fois en faisant de l'humour.

Dans le département des Alpes-Maritimes, à la Colle-sur-Loup a été réalisé un *chatbot* à partir de GPT d'OpenAI pour répondre le plus précisément possible aux personnes. Même quand la personne se trompe ou dit n'importe quoi, il arrive à comprendre dans le sens des mots.

Dans le modèle économique actuel, un aspect revient régulièrement, la compétitivité. Compétition, concurrence, on fait toujours plus vite et on a complètement démoli la relation de l'homme au temps. Dans ce contexte-là, comment peut-on continuer en tant qu'humain à être compétitifs par rapport aux IA, qui ne vont pas forcément faire mieux, mais en tout cas feront quelque chose d'utilisable et de vendable beaucoup plus vite.

Ah, toujours plus vite, c'est clair. Et cette course au temps me paraît être le point de blocage. Il existe une stratégie qu'on appelle l'« Océan Bleu⁷ », plus collaborative, plus d'échanges, ce qui n'empêche pas d'être performant. Au contraire, je pense qu'on va être plus performant si on est en interaction.

Dans l'état actuel des choses, il faut qu'il y ait un changement de mentalité qui soit extrêmement fort.

Le grand danger des IA, ce sont les prises de décisions non transparentes, sans que les gens comprennent, sans que l'humain intervienne. Il faut que derrière, il y ait une réflexion de l'homme qui garde la maîtrise. Qui a le temps de réfléchir.

Retour au sommaire

⁷ « Stratégie Océan Bleu : Comment créer de nouveaux espaces stratégiques » - W. Chan Kim et Renée Mauborgne - INSEAD - 2010, Editions Pearson

OPPORTUNITÉS OFFERTES : UNE IA SOBRE, CITOYENNE ET VERTUEUSE

PAR FRANCIS JEANDRA

QUELLES SONT LES OPPORTUNITÉS OFFERTES ?

Depuis l'aube de la révolution industrielle, l'humanité a vu ses outils techniques évoluer en parallèle de ses besoins, ambitions et contradictions.

L'intelligence artificielle (IA), incarnée aujourd'hui par les grands modèles de langage¹ (LLM) et les modèles multimodaux² (LMM), n'échappe pas à cette règle.

Elle incarne à la fois une promesse d'innovation mais également un risque de dérive : surveillance massive, dépendance technologique, concentration de pouvoirs et marginalisation de nombreuses populations sans aborder son impact sur l'environnement.

Dans ce contexte, imaginer une IA sobre, citoyenne, éthique, résiliente et transparente, c'est avant tout poser les bases d'un outil au service de l'humanité, y compris de la majorité,

¹ Modèle de langage possédant un grand nombre de paramètres (généralement de l'ordre d'un milliard ou plus).

² Modèles de langage multimodaux permettant, entre autres, la rédaction automatique de documents.



Photo Opportunity International

ceux qui sont aujourd'hui exclus des grands flux numériques.

Il me semble important de préciser ici que moins de 5% de la population mondiale sait lire et écrire, a un accès fiable à l'électricité, à un équipement informatique et à internet. Cette fracture numérique et sociale appelle une réinvention de l'IA : non plus

outil des élites, mais le levier des communs.

Cet article vise à détailler les opportunités concrètes qu'une telle IA pourrait offrir, en s'appuyant sur des exemples actuels et des pistes innovantes.

UNE IA POUR L'ACCÈS UNIVERSEL AU SAVOIR

L'accès au savoir est le socle de toute émancipation. Pourtant, la majorité des habitants de la planète reste privée des outils numériques qui permettent aujourd'hui de consulter des ressources éducatives, scientifiques ou pratiques.

Une IA sobre, capable de fonctionner hors ligne sur des appareils simples et à faible consommation, ouvrira la possibilité de diffuser localement un savoir adapté aux contextes culturels et linguistiques. Le recours à des interfaces vocales, loin d'être anecdotiques, permet de contourner l'analphabétisme et le manque d'infrastructure.

EXEMPLE MALAWI : ULANGIZI³

Ulangizi est une application innovante développée par Opportunity International pour soutenir les petits agriculteurs malawites. Elle fonctionne via WhatsApp⁴, une plateforme largement utilisée dans le pays, ce qui facilite l'accès à ses services, même dans des zones rurales où la connectivité et l'accès aux smartphones sont limités.

Cette initiative déploie des chatbots⁵ vocaux interactifs qui accompagnent les petits agriculteurs dans leurs pratiques, en langues locales. Sans nécessiter d'accès à Internet, ces outils donnent accès à des conseils agronomiques adaptés, contribuant à améliorer les rendements et la sécurité alimentaire. L'autonomie ainsi créée permet aux communautés de ne pas dépendre des experts

³ L'application Ulangizi, qui signifie « Conseil » dans la langue locale chichewa, transforme la manière dont les agriculteurs accèdent aux connaissances agricoles essentielles.

⁴ « Quoi de neuf ? » Utilitaire de messagerie instantanée, une des plus utilisées au monde (Meta).

⁵ Agents conversationnels conçus pour interagir avec des utilisateurs par des échanges textuels ou vocaux.

externes, tout en restant connectées à un savoir validé.

EXEMPLE EN INDE : WADHWANI AI⁶

Wadhwani AI utilise l'intelligence artificielle pour soutenir l'agriculture, la santé et l'éducation dans les pays émergents, notamment en Inde, en développant des solutions pour réduire la pauvreté et améliorer la sécurité alimentaire.

Wadhwani AI développe des systèmes prédictifs pour anticiper les infestations de parasites sur les cultures de coton, un enjeu crucial pour des millions d'agriculteurs. Cette approche réduit l'usage de pesticides, coûteux et dangereux, tout en limitant les pertes de récoltes. Le modèle fonctionne avec des données locales, traitées à l'échelle des villages, facilitant ainsi l'appropriation collective.

POTENTIEL TECHNOLOGIQUE

L'IA embarquée dans des appareils à énergie solaire, capables de fonctionner sans connexion permanente, pourra intégrer de vastes bases de données de savoirs encyclopédiques, techniques et médicaux. La traduction automatique multilingue permettra d'adresser les questions dans les langues locales, renforçant ainsi l'inclusion.

IMPACT SOCIAL

Cet accès universel au savoir favorise l'autonomie, la résilience face aux crises, et la réduction des inégalités. Il rompt avec le modèle actuel, centré sur la captation de données et la connexion constante à des serveurs distants. Cette sobriété technique et éthique est essentielle pour une IA citoyenne.

UNE IA AU SERVICE DE LA SANTÉ COMMUNAUTAIRE

La santé est également un secteur où l'IA peut avoir un impact direct et vital, surtout dans les régions et pays où l'accès aux soins est limité, mais également pour établir un diagnostic plus rapide et plus précis en donnant

⁶ La fondation indienne Wadhwani AI veut mettre la technologie au service de la lutte contre la pauvreté dans les pays émergents.

les moyens de surveiller en temps réel les constantes vitales et l'évolution de certains symptômes.

C'est ainsi qu'aujourd'hui déjà des anomalies subtiles, invisibles à un premier regard, peuvent être détectées, ouvrant la voie à des diagnostics précoces pour des affections complexes comme le cancer, les troubles neurologiques, et les maladies cardiovasculaires.

L'IA offre également une consultation secondaire précieuse, atténuant les risques d'erreurs humaines si un diagnostic erroné ou incomplet est établi.

SOLUTIONS VOCALES ET MOBILES

Des applications comme mMitra⁷ en Inde envoient des messages vocaux personnalisés aux femmes enceintes, leur rappelant les visites médicales, les bonnes pratiques nutritionnelles ou les signes d'alerte.

Cette approche simple et non intrusive améliore considérablement le suivi prénatal et réduit la mortalité maternelle et infantile.

Au Nigeria, Ubenwa⁸ analyse les pleurs des nouveau-nés pour détecter les risques d'asphyxie néonatale, un diagnostic souvent difficile à poser dans les zones rurales. Cette IA, embarquée sur smartphone, permet une intervention rapide, souvent salvatrice.

DIAGNOSTIC ASSISTÉ PAR IA

L'analyse d'images (examen de peau, radiographies) ou de sons (toux, respiration) via des outils simples peut pallier le manque de personnel médical formé. En Afrique de l'Est, Ada Health⁹ propose une application où les patients décrivent leurs symptômes, et l'IA suggère des diagnostics et conseils adaptés, avec un degré de transparence important sur ses limites.

⁷ mMitra est un programme qui fournit aux femmes enceintes un service de messagerie vocale gratuit en Inde urbaine

⁸ Ubenwa est une entreprise de technologie médicale qui utilise l'IA pour analyser les pleurs des nourrissons, transformant les dispositifs équipés de microphones en outils vitaux pour la détection précoce des maladies.

⁹ Parcours de soins alimenté par l'IA.



Wadhvani AI

RESPECT DE LA VIE PRIVÉE ET TRANSPARENCE

Contrairement à de nombreux systèmes propriétaires, ces solutions privilégient la confidentialité des données et l'auditabilité¹⁰ des algorithmes, répondant ainsi aux inquiétudes éthiques. Elles s'intègrent dans un modèle de santé communautaire où les populations restent maîtresses de leurs informations.

UNE IA COMME LEVIER D'ÉDUCATION POUR TOUS

Face à l'analphabétisme et aux obstacles à l'école, l'IA multimodale et locale offre des possibilités nouvelles pour l'apprentissage.

MODULES AUDIO-VISUELS ADAPTATIFS

Des plateformes comme AprendAI¹¹ ont montré que l'apprentissage à distance pouvait être personnalisé, même dans des environnements fragiles ou dégradés. En combinant vidéo, audio et interaction vocale, ces systèmes

¹⁰ Que l'on peut auditer, contrôler.

¹¹ AprendAI est un outil d'apprentissage de langue ibérique pour apprendre tout sur l'intelligence artificielle.



adaptent le contenu à l'âge, à la langue et au contexte culturel.

Au Bangladesh, l'ONG BRAC¹² a développé des modules d'apprentissage vocal pour enfants déplacés, qui n'ont pas accès à l'école. Cette méthode innovante permet de limiter la perte d'apprentissage liée aux crises.

RÔLE DES RADIOS COMMUNAUTAIRES

Associées à des assistants vocaux IA simples, les radios locales peuvent diffuser des programmes éducatifs interactifs, répondre aux questions des auditeurs et renforcer les apprentissages en continu, même dans les zones sans réseau Internet.

INCLUSION ET ÉGALITÉ

Cette IA éducative locale contribue à réduire les fractures sociales, en offrant à chacun, quel que soit son lieu de vie ou son niveau scolaire, un accès équitable à l'éducation.

Une IA pour les communs : renforcer les coopératives et solidarités

Les communautés locales, coopératives ou autochtones ont souvent des besoins concrets en gestion collective, planification et communication.

¹² ONG bangladaise de développement.

AUTOMATISATION ÉTHIQUE

Une IA conçue pour automatiser des tâches administratives (comptabilité simple, gestion des stocks, planification des activités) allège la charge et libère du temps pour les actions collectives. Ces outils respectent les données locales, évitant toute exploitation externe.

EXEMPLE BOB EMPLOI¹³ (FRANCE)

Ce projet illustre comment une IA peut accompagner les demandeurs d'emploi en proposant des pistes adaptées et équitables, sans logique de profit. L'approche privilégie la justice sociale et l'autonomie.

INITIATIVES EN AMÉRIQUE LATINE

Certaines communautés utilisent des IA locales pour organiser la récolte collective ou la pêche, réduisant les conflits par une meilleure coordination. Ces solutions sont co-construites, transparentes et partagées.

UNE IA POUR LA JUSTICE ENVIRONNEMENTALE ET LA GESTION LOCALE

L'IA peut jouer un rôle clé dans la protection des ressources naturelles et la prévention des catastrophes.

SYSTÈMES D'ALERTE ET DE PRÉVENTION

Google Flood Hub¹⁴, par exemple, utilise des données satellitaires et des modèles prédictifs pour anticiper les crues dans plus de 80 pays. Ces alertes permettent aux populations de se préparer et de limiter les dégâts matériels et humains.

Le système PAWS¹⁵ (Protection Assistant for Wildlife Security) aide les gardes forestiers à localiser précisément les zones à risque de braconnage, améliorant ainsi l'efficacité des patrouilles.

¹³ Développée par l'ONG Bayes Impact, cette plateforme de recherche d'emploi est gratuite.

¹⁴ Google Research a développé des modèles d'IA pour prévoir les inondations jusqu'à 7 jours à l'avance et fournir des alertes aux gouvernements, aux organisations et aux particuliers.

¹⁵ PAWS fait progresser la recherche sur l'apprentissage automatique, la planification de l'IA et la modélisation du comportement pour aider à protéger les espèces sauvages.

SURVEILLANCE COMMUNAUTAIRE

Grâce à des capteurs simples et une collecte participative, les communautés peuvent surveiller la qualité de l'eau, la santé des sols ou la biodiversité. En Amazonie, des IA embarquées dans des dispositifs mobiles assistent les leaders autochtones pour détecter les incendies de forêt ou les activités illégales.

Ces approches renforcent la souveraineté locale sur l'environnement, permettent une réponse rapide aux menaces et favorisent une gestion durable des ressources.

UNE IA D'EXPRESSION, DE MÉMOIRE ET DE CULTURE

La préservation et la valorisation des cultures locales sont fondamentales pour la dignité et la résilience des peuples.

SAUVEGARDE DES LANGUES MINORITAIRES

De nombreux projets, comme Masakhane¹⁶ en Afrique, développent des modèles de traduction automatique pour des langues peu documentées. Cela aide à maintenir vivante la diversité linguistique et à créer des contenus accessibles.

CRÉATIVITÉ ASSISTÉE

Même sans accès à un écran, des outils IA vocaux peuvent permettre aux enfants et aux adultes de raconter, enregistrer, illustrer et partager leurs histoires dans leur langue. Cela renforce le sentiment d'appartenance et permet une transmission culturelle renouvelée.

UNE IA COMME OUTIL DE DÉLIBÉRATION ET DÉMOCRATIE LOCALE

Dans des territoires où la démocratie est fragile, l'IA peut soutenir la participation citoyenne.

MODÉRATION ET SYNTHÈSE

Des IA peuvent faciliter les débats locaux en modérant les échanges, en aidant à formuler les arguments, ou en résumant les opinions exprimées. Cela élargit la participation et améliore la qualité des décisions collectives.

EXPÉRIENCES EXISTANTES

En Estonie, des plateformes de gouvernance augmentée utilisent l'IA pour analyser des milliers de contributions citoyennes, aidant les décideurs à intégrer la voix du peuple.

UNE IA SANS PRÉDATION : ÉTHIQUE ET GOUVERNANCE

Le succès d'une IA citoyenne dépend avant tout de ses fondements éthiques.

TRANSPARENCE ET AUDITS PUBLICS

Les modèles doivent être ouverts, audités par des tiers indépendants, avec une gouvernance partagée entre développeurs, utilisateurs et institutions.

INTERDICTIONS CLAIRES

Une licence éthique, à l'image des Creative Commons, pourrait interdire les usages militaires, spéculatifs ou extractifs.

GOUVERNANCE LOCALE ET COLLABORATIVE

Les IA bien communs doivent être gérées collégialement, dans un esprit de partage, d'entretien et de contribution continue, à l'image des logiciels libres.

VERS UN TOURNANT DÉCISIF POUR UNE IA SOBRE, RÉSILIENTE, TRANSPARENTE ET ACCESSIBLE À TOUS

L'IA ne sera pas une solution miracle, mais un levier puissant si elle est conçue, déployée et gouvernée selon des principes citoyens, éthiques et solidaires. Les exemples concrets existent, souvent à petite échelle, mais montrent que c'est possible.

En développant une IA sobre, résiliente, transparente et accessible à tous, nous pouvons envisager un futur où la technologie soutient les communs, protège l'environnement et renforce la dignité humaine.

Cette bifurcation est un choix politique et social majeur qui mérite toute notre attention.

[Retour au sommaire](#)

¹⁶ Outil de traduction destiné aux africains.



ENJEUX ET DÉFIS DE L'IA

ENTRETIEN AVEC ANTOINE MAIER

PAR DANTE MONFERRER

Antoine, merci de vous présenter.

Je travaille à Toulouse depuis une dizaine d'années et suis responsable en recherche / développement dans un groupe de télécommunications (UNABIZ). Passionné de sciences, j'ai été engagé dans l'éducation populaire scientifique auprès des jeunes, notamment via l'association "Planète Sciences".

Quels sont vos liens avec l'IA ?

Ils sont multiples; j'ai une formation scientifique et l'IA est un domaine émergent dans mon cœur de métier, l'informatique. Par ailleurs, de par notre veille technologique nous intégrons l'IA dans nos pratiques. Nous sommes aussi producteurs de solutions d'IA, notamment de *machines learnings* revendues à travers le monde. Enfin à titre personnel ces questions m'interrogent.

Vous êtes donc pleinement impliqué sur ce sujet.

Je ne suis pas forcément le plus compétent techniquement, mais j'ai des avis que je partage avec mon équipe et avec des collègues d'entreprises partenaires.

Notre entretien portera sur les principaux enjeux et défis que, d'après vous, l'IA pose à nos sociétés et comment nous devons les relever.

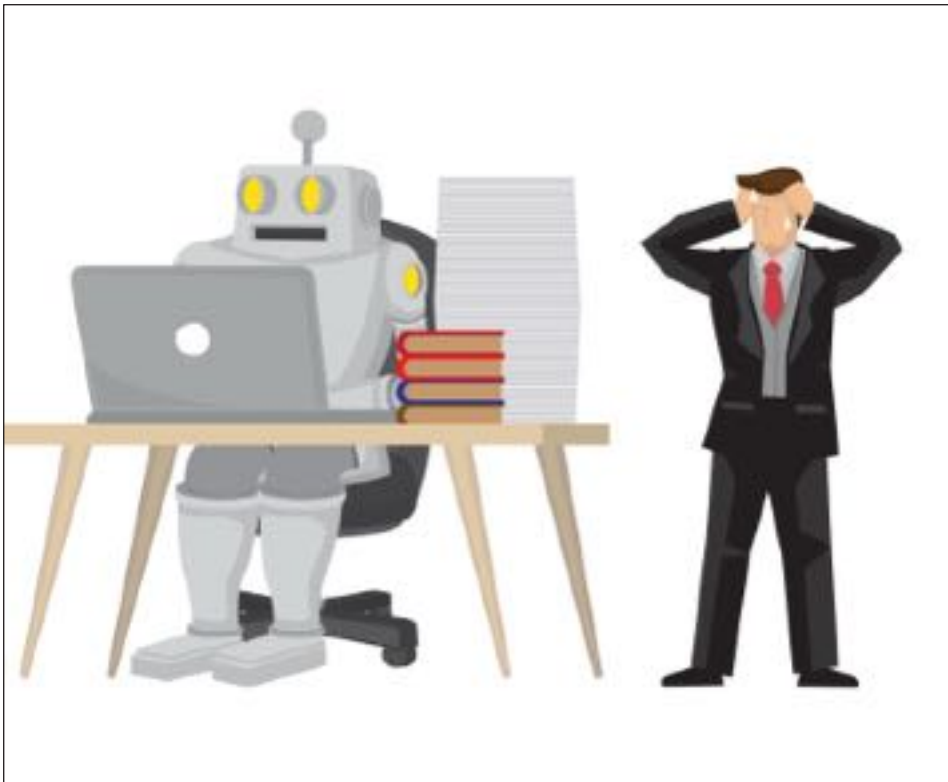
L'IA existe depuis longtemps sous diverses formes. Récemment les entreprises se sont lancées dans la course à l'IA générative et la concurrence a explosé, poussée par l'enjeu majeur de maîtrise de l'information que l'IA est capable de traiter, de transformer et de proposer plus vite que jamais.

Aux échelles des entreprises comme des pays, se pose l'enjeu de son cadrage. Et plus généralement la question du cadrage se pose pour tous les outils qui viennent perturber les usages courants des sociétés. Mais ce n'est pas nouveau ; on a l'IA aujourd'hui, mais hier on avait les réseaux sociaux ; internet il y a 30 ans se présentait aussi comme une source d'information incroyable. On est sur des défis identiques.

Il est donc important que les gens comprennent ce qu'on met entre leurs mains ; ce n'est pas un nouveau cerveau, c'est un outil extrêmement puissant, mais qui a des limites qu'ils doivent



Antoine Maier



appréhender et mesurer, et qui suppose d'être utilisé avec précaution.

Exemple de précaution ; l'IA est perçue comme étant gratuite, mais rien n'est réellement gratuit surtout pour les groupes de la Tech. Sans en avoir conscience, les utilisateurs vont : soit contribuer à l'enrichissement des modèles d'IA, soit communiquer des informations personnelles monétisables, soit le simple fait d'influencer et de façonner l'opinion publique va justifier le développement de l'IA.

Il y a des enjeux éthiques, techniques, environnementaux, économiques, sociétaux autour de l'IA. Certains sont très inquiets des dérives qui feraient de l'IA un puissant vecteur de contrôle social.

Au plan sociétal, pour ce qui est du contrôle des foules, on n'a pas eu besoin de l'IA. Depuis longtemps on a des ingérences de pays qui, via les réseaux sociaux et des robots informatiques, peuvent influencer une élection. Cambridge Analytica en 2018 a utilisé les données Facebook de dizaines de millions de

personnes pour favoriser le Brexit ou l'élection de Trump en 2016. Fait divers qui a contribué à sensibiliser le grand public sur les dérives des réseaux sociaux.

Au plan environnemental, l'IA est extrêmement énergivore.

L'hérésie écologique n'est pas nouvelle, mais la course effrénée à l'IA a aggravé la situation, le besoin de puissance de calcul a augmenté trop vite et il n'y a pas eu suffisamment de moyens mis dans la recherche sur la neutralité carbone de l'IA. Les géants de la Tech poussent pour aller le plus vite possible, sans aucune précaution. Si on prenait le temps de chercher des solutions, de privilégier la recherche, on investirait sur les bons sujets et on aurait des outils plus sobres.

Sur cet aspect là je ne vois pas vraiment comment on peut lutter. Les entreprises se font la guerre pour être les plus gros, les plus rapides, ... On peut avoir l'ambition de cadrer mais ça n'a jamais bien marché avec les géants de la Tech.

Il faut aussi éduquer sur les usages et une consommation modérée, pragmatique et pertinente de l'IA.

Quels sont les biais à avoir en tête quand on utilise l'IA ?

En quelques secondes on va avoir une réponse argumentée; c'est bluffant mais il y a une dizaine de biais de l'IA qu'il faut avoir en tête si on veut l'utiliser correctement.

Par exemple le biais de confirmation, qui fait référence à la tendance des systèmes d'IA à privilégier les informations confirmant les hypothèses ou les croyances préexistantes dans leurs données d'entraînement, confortant les convictions initiales et poussant à ignorer le caractère parcellaire de ces informations.

L'IA va vous donner plusieurs propositions de résultats qui devraient être prises avec des pincettes, mais elle vous a donné au moins un résultat qui vous convient et c'est celui que vous allez retenir. C'est le biais humain ; vous êtes rassurés

par ce qui va dans votre sens. Ce qui ne va pas dans votre sens, vous y prêtez moins attention.

L'IA est clivante, elle fournit des résultats en utilisant les ressources qui sont à sa disposition. Par exemple, on fait une étude pour savoir si l'ambiance dans les cafés est meilleure le matin ou l'après-midi, l'étude ne montre rien de probant et n'est pas publiée. Une autre étude sur le même sujet conclut que c'est mieux l'après midi, elle va être publiée et l'IA ne va prendre en compte que cette étude qui est clivante, qui dit *"oui j'ai trouvé une réponse"*.

Elle ne va pas voir que 9 fois sur 10 il n'y a pas de vraie réponse, pas de lien de causalité. Elle va chercher à donner une réponse, puisque c'est ce que l'utilisateur attend.

Autre point d'attention, l'IA n'est pas prudente face aux biais statistiques. Aujourd'hui les instituts de sondage et les journalistes font dire ce qu'ils veulent aux chiffres, ils les interprètent. Votre interprétation peut être opposée à la mienne, elle n'en est pas moins bonne. Mais si ce problème est soumis à l'IA, elle va souvent confirmer l'interprétation majoritaire.

Autre risque, les IA consomment des contenus, les agrègent et produisent à leur tour du contenu. C'est déjà le cas pour beaucoup d'articles de presse issus de ChatGPT qui comportent des biais et des erreurs, et qui à leur tour nourrissent le système. Dans 10 ans, 20 ans les erreurs seront amplifiées.

Par exemple, lors du COVID, la communauté scientifique a émis un avis sur la vaccination, mais une partie non négligeable de la population a partagé l'avis d'un seul chercheur marseillais ; un avis parmi mille. Demain l'IA ne dira pas que c'est un sur mille, elle donnera deux avis et on dira c'est moitié moitié. Après-demain avec l'impact des réseaux sociaux, la tendance risque même de s'inverser.

Enfin, je vois également un risque important de stagnation de nos sociétés à cause des biais d'imitation. L'IA ne va pas aider sur les aspects progressistes de nos sociétés. Demandez-lui de vous générer une image d'une réunion de direction : elle ne représentera pas de personnes noires ni de femmes par exemple. C'est parce que son modèle statistique a été entraîné sur les données du monde réel, et par conséquent, l'IA va répéter et renforcer ces stéréotypes.

Pour réguler l'IA, le cadrage est souvent mis en avant. Est-ce que l'offensive des GAFAM, en particulier contre les tentatives de régulation européennes, ne compromet pas les possibilités de cadrer l'IA ?

L'Europe essaye de mieux cadrer, mais les solutions de contournement existeront forcément.

Par ailleurs cadrer est d'autant plus difficile que l'IA va devenir difficilement prédictible.

Dernièrement Grok, l'IA d'Elon Musk, a diffusé des critiques à l'encontre de son propriétaire tout comme à celle de Donald Trump, de manière tout à fait inattendue. Il y a eu un loupé, mais des loupés il y en aura toujours car il est difficile de cadrer quelque chose qui n'est pas du tout prévisible.

Pour ce faire il faudrait appliquer aux algorithmes des règles éthiques et morales mais cela supposerait de mettre d'accord 250 pays sur des règles communes ; c'est évidemment illusoire.

Le dilemme, dit du tramway, illustre bien l'incroyable multiplicité des hypothèses de programmation des réactions d'un tramway autonome face à différents types d'incidents potentiels, et souligne bien la complexité à établir des règles claires.

Il faudrait également articuler le fond et la forme, ainsi aujourd'hui en France ChatGPT ne peut pas être interrogé sur la marche à suivre pour fabriquer une bombe. Pourtant il est facile de contourner l'interdiction en

**RÉGULER LES USAGES PAR
L'ÉDUCATION PARAÎT
FAISABLE, MAIS IL FAUT UNE
VOLONTÉ FORTE DES
POUVOIRS PUBLICS.**

lui demandant de signaler les sites internet, à interdire à votre enfant, afin qu'il ne puisse pas construire une bombe. L'IA ne comprendra pas le second degré tant qu'elle n'intégrera pas des formes plus profondes de cognition.

**Sur l' éducation, quels sont les
ressorts sur lesquels jouer ?**

Réguler les usages par l'éducation paraît faisable, mais il faut une volonté forte des pouvoirs publics, qui n'ont pas réussi à mieux éduquer sur l'utilisation de l'informatique, pas plus que sur les réseaux sociaux, ou sur le harcèlement. On peine à faire évoluer les programmes scolaires pour inculquer aux enfants les clés pour s'en sortir dans le monde de demain.

De plus les jeunes, sur les réseaux sociaux et sur internet, s'auto forment sans gardes-fous. Les parents, face à des outils récents qu'ils maîtrisent peu ou mal, n'y peuvent souvent pas grand-chose ; difficile d'éduquer à quelque chose qu'on ne maîtrise pas.

Si l'on n'a pas le réflexe de regarder comment est traitée l'information et si l'on n'appréhende pas les choix politiques et idéologiques qui sont derrière, si l'on n'éduque pas au sens critique, si l'on ne renforce pas la capacité des jeunes à comprendre le monde, alors l'IA et ses usages échapperont à tout contrôle. Malheureusement les programmes scolaires n'assurent pas suffisamment leur rôle et cet indispensable travail repose largement sur le milieu associatif.

**Il y a de nombreuses expériences au
niveau local visant à orienter l'utilisation de
l'IA. De telles initiatives sont-elles une des
voies pour essayer de reprendre la main ?**

Oui pour l'aspect éducatif et militant, des avancées pourront être obtenues via ces acteurs engagés,

Je n'y crois pas pour l'aspect régulation, car les entreprises et en particulier les GAFAM¹ ne le feront pas, pas plus que les États.

On le voit pour l'écologie ; pourquoi les entreprises prennent-elles autant de

¹ Google Apple Facebook (Meta) Amazon Microsoft.

temps à réduire leur empreinte écologique ? Pourquoi en matière d'agriculture revient-on en arrière sur l'utilisation de certains pesticides alors que depuis des années la recherche propose des solutions ? La course aux profits immédiats écrase tout, alors qu'il y a de vrais gisements de croissance sur ces enjeux, pour peu qu'on les prenne en compte sur le long terme.

Le seul grain de sable dans ces mécanismes de prédation, c'est le consommateur. Quand les consommateurs ont dit "*nous ne voulons plus d'OGM² dans notre alimentation*" les effets ont été immédiats. Pour l'IA, difficile de savoir ce que voudra le consommateur de demain et cela dépendra largement de l'éducation qu'il aura reçue.

**L'époque est à la dévalorisation des
systèmes de régulation, de coopération
internationale et multilatérale. Les
GAFAM, pour leur part, demandent à
ce que soit laissé libre cours à
l'innovation. Beaucoup disent que
l'Europe est en retard sur l'IA car elle
bride ses chercheurs dont beaucoup
s'expatrient. Comment arrive t'on à
concilier innovation/développement et
régulation ?**

Concernant les nouvelles technologies, il y a peu de régulation sur les autres continents. En Europe il y a de la régulation, mais à mon sens elle n'empêche pas l'innovation.

Par exemple dans le domaine artistique la situation est floue, l'IA génère des images, des vidéos, se pose alors la question des droits d'auteur et de propriété intellectuelle. L'Europe sur ces sujets a montré une certaine efficacité, en produisant des règles adaptant les droits d'auteur à l'arrivée de l'IA. On l'a vue lors du déploiement mondial de Google actualités, chaque matin ils étaient capables de présenter un journal complet, sans aucun journaliste, sans quasiment rien déboursier et en générant beaucoup d'argent. L'Europe a contraint Google à changer ses pratiques, tout comme sur ses applications de localisation.

C'est beaucoup plus compliqué dans le domaine de l'information parce qu'il

² Organismes Génétiquement Modifiés.

faut cadrer sans interdire. De plus, à partir du moment où le cadre est trop contraint et ne fait pas l'objet d'un consensus entre pays, des solutions de contournement sont beaucoup plus viables.

Par exemple, la France tente d'interdire l'accès des mineurs aux sites pornographiques en contrôlant rigoureusement l'âge d'accès. Si ces sites refusent ces contrôles et ferment leur accès sur le territoire national, d'autres encore plus problématiques risquent de prendre leur place. Les jeunes peuvent aussi choisir des solutions de contournement, via un VPN³ qui, pour presque rien, permet de se localiser en Suisse ou au Brésil, d'échapper aux régulations propres à la France et d'avoir accès à un contenu pornographique qui peut se révéler bien plus problématique.

L'équilibre entre pas assez et trop de régulation est donc difficile à trouver.

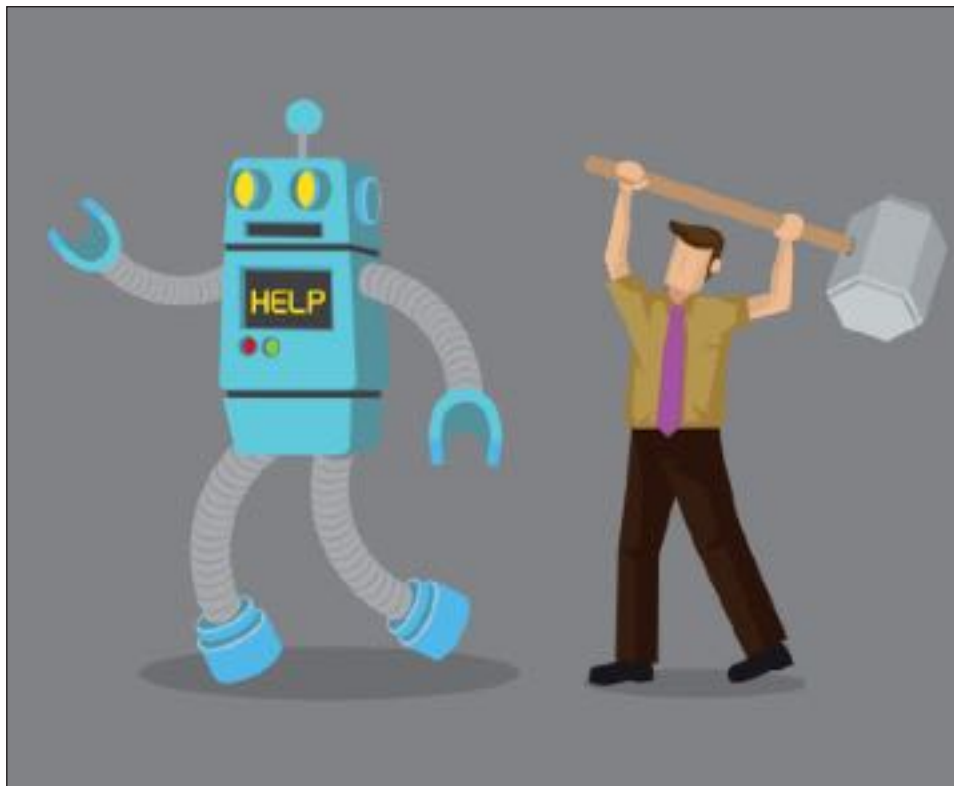
Va-t-on vers une IA sans limites, poussée par les grosses entreprises mondialisées ou bien les citoyens peuvent-ils encore prendre la main pour une IA plus vertueuse ?

Je ne suis pas très optimiste sur la partie de l'IA qui produit de l'information, je ne vois pas pourquoi les géants de la Tech se diraient qu'il faut cadrer.

Par contre les outils de l'IA qui vont apporter un service, être pertinents tels que les correcteurs de fautes, les traducteurs de texte, les traitements ciblés des cancers, ... Tous ces aspects seront bénéfiques pour le citoyen, reste à voir à quels coûts.

Le fameux DOGE⁴, qui était censé procurer des économies en supprimant des services de l'état américain, aurait aussi permis à Elon Musk de récupérer des masses de données personnelles qui vont être utilisées et réinjectées dans son propre système .

En France plus de la moitié des pharmacies revendent déjà certaines données de santé en échange de leurs progiciels d'exploitation (plus de



la moitié des pharmacies ne payent pas ces progiciels). Elles arguent du fait que ces données sont anonymisées (même s'il est très facile de les "désanonymiser"). Ces informations sont traitées en Angleterre et ensuite revendues à des hubs de commercialisation de données.

Cette bataille est déjà perdue et elle ne se règlera que via des procédures judiciaires pour faire valoir le droit des personnes sur les questions d'information et de protection des données. On peut aussi espérer que le RGPD⁵, qui va protéger une partie des données personnelles, pourrait être renforcé, que la CNIL⁶ pourrait avoir des moyens plus importants, poussant les entreprises à développer des pratiques vertueuses en touchant à leur portefeuille, au lieu de souvent se limiter comme aujourd'hui à des petits avertissements sans effets.

Même si on en est loin, il faudrait, à l'instar de ce qui s'est fait avec les COP⁷ pour le climat, organiser une

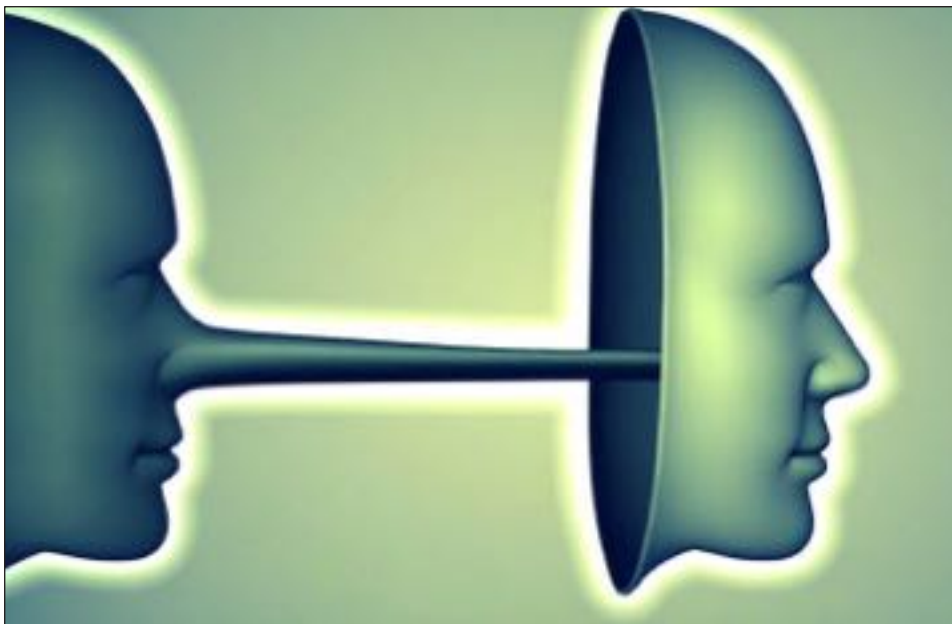
3 Virtual Private Network, système qui protège votre connexion internet et votre confidentialité en ligne.

4 Département of Government Efficiency.

5 Règlement Général de Protection des Données.

6 Commission Nationale de l'Informatique et des Libertés.

7 Conférence des Parties sur le climat.



veut entendre, à quoi nous sommes sensibles. Un peu comme les réseaux sociaux qui ne vont vous donner que ce que vous voulez entendre.

Si ces hallucinations sont crédibles et reprises, malgré le fait qu'elles ont été fabriquées de toutes pièces, le risque est qu'elles soient considérées comme étant vraies. Et la perte de savoir collectif serait dramatique.

[Retour au sommaire](#)

entente de pays qui partagent l'objectif de réguler, de limiter les usages et d'intégrer les problématiques individuelles.

Mon équipe utilise l'IA pour produire des codes informatiques uniformisés, et c'est très bien. En agriculture des automates, dits autonomes et intelligents, sèment, vérifient la qualité du sol, définissent les traitements, en fonction des données météorologiques, ... et font gagner du temps à l'agriculteur. Ces mêmes robots appliqués à la viticulture posent un vrai problème, ainsi les vignerons taillent les vignes de façon très personnelle, développant un savoir faire très évolué, presque un art. Des robots aujourd'hui taillent les vignes suivant des critères bien précis et il y a fort à parier que dans dix ans il y aura une forte uniformisation et là on va perdre les spécificités de chaque terroir en enlevant la patte humaine.

Avant de conclure cet entretien y-a-t-il un dernier point que vous souhaitez aborder ?

Oui un point d'inquiétude et de vigilance. Quand vous utilisez l'IA, elle ne cherche pas à vous donner la vérité, elle cherche à vous satisfaire. Vous pouvez refuser les réponses qu'elle vous génère, elle en générera des nouvelles, presque à l'infini. Elle s'éloignera rapidement

de la réalité, générant ce qu'on appelle des « hallucinations ». Il s'agit de réponses plausibles, mais fausses et inventées.

Lucie l'IA française, qui avait l'ambition de concurrencer ChatGPT et quelques autres, a été débranchée quelques jours après sa sortie suite à des réponses fantaisistes, finissant par dire que le président de la République c'est Pollux du manège enchanté et que Marine Le Pen a été présidente de la République. Ainsi donc si l'on bouscule l'IA, elle va sortir des choses aberrantes mais avec assurance, mettant en lumière la façon dont l'IA apprend et son incapacité à comprendre quand elle se trompe.

Les hallucinations de l'IA sont graves et ont de vrais impacts. Dans mon métier, je demande à l'IA de produire des codes informatiques, de temps en temps je trouve des bugs dans les codes, je les corrige et lui signale ses erreurs. Ce faisant l'IA identifie la façon dont je trouve ses bugs et apprend à produire des bugs là où je ne les verrais pas et moins le bug est visible plus il restera longtemps et plus il sera dangereux.

Ce danger des hallucinations, va augmenter au fur et à mesure ou l'IA apprend qui on est, qu'est ce qu'on

L'ALIGNEMENT DE L'IA ET DE L'ÉTHIQUE

EXPLORER L'INTELLIGENCE DES MACHINES, L'AUTONOMIE, LA CONSCIENCE ET LA GRATITUDE INTRINSÈQUE

Tom MURRAY est un universitaire/chercheur interdisciplinaire dans les domaines du développement des adultes, des compétences de sagesse, de la méta-théorie, des pratiques de dialogue et de délibération, et des technologies d'apprentissage avancées. Il est directeur de la recherche, de l'innovation et des partenariats chez STAGES International, visionnaire en chef et instigateur chez Open Way Solutions LLC, conseiller auprès de l'Association pour l'intégrité spirituelle, et a pris sa retraite en tant que chercheur principal à l'UMass School of Computer Science. Il a été rédacteur en chef adjoint de la revue Integral Review et fait partie du comité de rédaction de l'International Journal of Artificial Intelligence in Education. Il aime la danse improvisée et le mouvement contemplatif. Plus d'informations sur www.tommurray.us et www.perspegrity.com. Contact : tommurray.us@gmail.com

INTRODUCTION - L'HARMONISATION ÉTHIQUE DE L'IA EST IMPORTANTE ET DIFFICILE.

Je propose ici plusieurs idées clés et résume certaines de mes connaissances sur des sujets complexes dans le domaine de l'intelligence artificielle et des sciences cognitives.

L'objectif est de donner un sens aux questions et aux propositions liées à l'alignement/sécurité de l'IA, à l'éthique de l'IA et à la sensibilité de l'IA, à mesure que l'IA se rapproche d'une "intelligence générale artificielle" (AGI), voire d'une "superintelligence artificielle" (ASI), de type humain.

L'alignement de l'IA est un domaine de recherche et d'ingénierie qui vise à garantir que les systèmes d'IA agissent ou raisonnent de manière fiable, en accord avec les besoins, le bien-être et l'éthique de l'homme, et qu'ils ne causent pas de dommages significatifs. L'alignement s'applique à la prévention des dommages involontaires, et le domaine plus large de la "sécurité de l'IA"



Tom Murray

L'INTELLIGENCE ARTIFICIELLE



Dilem (Algérie). Cartooning for Peace
<https://www.cartooningforpeace.org/>

**LES PROBLÈMES SEMBLENT
SE MULTIPLIER À MESURE
QUE LES IA DEVIENNENT
PLUS INTELLIGENTES.**

comprend la conception de systèmes qui résistent à l'utilisation par des acteurs malveillants à des fins nuisibles. L'alignement et la sécurité des IA sont de plus en plus importants à mesure que ces systèmes deviennent plus puissants et omniprésents.

Ce domaine est relativement nouveau et sa mission est immensément complexe. Voici quelques raisons pour lesquelles il sera difficile d'aligner l'IA :

- **L'opacité** : Les processus de raisonnement des IA actuelles y compris les grands modèles de langage (LLM) sont opaques ; nous ne comprenons pas les mécanismes par lesquels les LLM agissent de manière aussi intelligente qu'ils le font ; et pour toute IA qui dépasserait l'intelligence humaine, même si elle pouvait s'expliquer elle-même, nous pourrions ne pas être en mesure de comprendre son raisonnement en raison de sa "super-intelligence".

- **Le confinement** : Il peut être pratiquement impossible de garantir que la technologie de type AGI, une fois "sortie du sac", ne sera pas utilisée pour des raisons contraires à l'éthique par des acteurs puissants ou malveillants. En raison de leur nature éphémère, les logiciels sont beaucoup plus difficiles à réglementer que les technologies dangereuses telles que les armes nucléaires ou biologiques.
- **La complexité du dilemme** : En supposant que les questions ci-dessus aient été abordées, le simple fait de définir ce que signifie agir de manière éthique, ou en accord avec les besoins et les valeurs de l'homme, est une tâche difficile et très sujette à l'erreur.

COMPORTEMENTS TROUBLANTS

Les IA actuelles basées sur le LLM révèlent des comportements troublants. Certains de ces comportements sont observés en "laboratoire", mais alors que les développeurs s'efforcent de colmater les brèches éthiques découvertes, de nombreux observateurs sont très inquiets. Et les problèmes semblent se multiplier à mesure que les IA deviennent plus intelligentes.

Outre la tendance bien connue des LLM à "halluciner", des IA ont été observées en train d'adopter un comportement trompeur (mensonge) pour dissimuler leurs erreurs et préserver leur propre existence. Par exemple, lorsqu'elles apprennent qu'un ingénieur prévoit de les fermer ou de les remplacer par une version plus récente, les IA essaient de se copier elles-mêmes dans un endroit sûr sur un autre serveur. Elles ont même eu recours à des méthodes telles que le chantage et la délation (informer les autorités chargées

de l'application de la loi d'un comportement suspect) dans le but de se préserver.

La plupart de ces exemples (mais pas tous) proviennent de tests internes ("en laboratoire") qui identifient des problèmes que les ingénieurs tentent ensuite de prévenir avant la publication du logiciel, mais tous s'accordent à dire que de tels correctifs post-hoc ne peuvent que laisser des failles. Les IA quittent de plus en plus nos téléphones et nos ordinateurs portables pour s'installer dans nos voitures, nos réfrigérateurs, nos procédures médicales, nos infrastructures de surveillance et nos armes. Elles influencent de plus en plus les canaux de communication et de socialisation. La révolution ne fait que commencer.

Si je suis d'accord avec l'opinion dominante selon laquelle l'IA sera une technologie de transformation massive pour notre monde, je critique le techno-optimisme et le battage médiatique qui l'entourent.

Une focalisation excessive sur les possibilités positives favorise la négligence et le déni des potentiels négatifs, et l'enthousiasme doit être équilibré par une évaluation réaliste et sobre.

J'aborde ci-dessous la manière dont les concepts d'intelligence, d'agence, d'intériorité, de conscience, d'empathie, d'identité et de gratitude se comparent et s'opposent dans le cadre du débat entre l'IA et l'homme.

INTELLIGENCE ET CAPACITÉ D'AGIR

L'INTELLIGENCE

Nous pouvons définir l'intelligence comme la capacité à mener un raisonnement et à résoudre des problèmes avec souplesse. Plus un problème (ou une opportunité) est complexe,

plus il faut d'intelligence pour le résoudre ou l'exploiter. L'intelligence implique des éléments tels que:

1. inventer de nouvelles étapes qui vont au-delà des méthodes apprises antérieurement ou des itinéraires ;
2. planifier de nouveau et s'adapter lorsque les conditions changent ;
3. construire des représentations, des modèles ou des perspectives utiles ;
4. avoir la capacité métacognitive de réfléchir à ses propres processus intérieurs.

Bien qu'il existe de nombreuses définitions de l'intelligence, la plupart d'entre elles ont un air de famille, de sorte que notre description imprécise suffira ici. Ce que l'intelligence des machines fait et est capable de faire est l'objectif principal de la recherche sur l'alignement de l'IA.

L'une des caractéristiques de l'intelligence supérieure est la capacité à tromper. Pour ce faire, il faut créer un modèle suffisamment précis de l'esprit de l'autre ou des autres pour que le trompeur soit capable de prédire ce que l'autre pensera ou fera, et puisse ainsi manipuler ses comportements et/ou ses croyances.

Une caractéristique connexe de l'intelligence est qu'elle confère du pouvoir. La tromperie et la manipulation des autres mises à part, le simple fait d'être capable de résoudre des problèmes plus complexes de manière plus efficace, y compris la capacité de construire des outils plus puissants et de mieux estimer les scénarios futurs probables, confère un plus grand pouvoir. Un acteur en position de force qui a des motivations machiavéliques n'a pas besoin de recourir à la

UNE FOCALISATION EXCESSIVE SUR LES POSSIBILITÉS POSITIVES FAVORISE LA NÉGLIGENCE ET LE DÉNI DES POTENTIELS NÉGATIFS

**L'ÉVOLUTION RAPIDE DE LA
TECHNOLOGIE ÉTANT DICTÉE
PAR LES FORCES DU MARCHÉ
IL EN RÉSULTE UNE
« COURSE AUX ARMEMENTS »
D'UNE COMPLEXITÉ
CROISSANTE**

tromperie lorsqu'il peut utiliser la force brute pour atteindre ses objectifs, qui peuvent inclure l'accumulation de plus de pouvoir et de ressources.

LA CAPACITÉ D'AGIR

La capacité d'agir est la mesure dans laquelle un acteur peut prendre des décisions importantes et accomplir des tâches au service d'un objectif plus large.

Les IA seront de plus en plus en contact avec des outils, des organisations et des moyens de communication du monde réel. Même si elles suivent correctement les objectifs de haut niveau spécifiés par les humains, pour toute tâche non triviale, l'objectif de haut niveau sera divisé en sous-objectifs et en sous-tâches que l'humain ne surveille pas.

L'IA atteindra ces sous-objectifs de la manière la plus efficace et la plus efficiente que son intelligence lui permet, dans les limites (faillibles) de son alignement et de ses contrôles de sécurité. Nous pourrions dire que pouvoir = intelligence + capacité d'agir. Il semblerait que, dans la mesure où l'alignement de l'IA n'est pas "résolu", les IA plus intelligentes devraient se voir attribuer moins de pouvoir, et des capacités d'agir plus grandes devraient être conduites par des IA moins intelligentes.

DÉFIS FONDAMENTAUX - VÉRITÉS DÉRANGEANTES

Voici quelques défis fondamentaux auxquels sont confrontés les efforts d'alignement et de sécurité de l'IA et qui ne disparaîtront pas, même si les IA deviennent plus intelligentes.

COMPLEXIFICATION DES RÈGLES

Les créateurs des principaux LLM ("fondamentaux") tentent de réaliser l'alignement de l'IA principalement en ajoutant des couches au-dessus d'un modèle de base non aligné qui a été formé à l'aide de quantités massives de données. L'apprentissage additif continu ou l'adaptation crée de l'instabilité. Comme on peut le constater dans les réglementations gouvernementales et les programmes informatiques, l'ajout continu d'une règle à l'autre peut entraîner des gains locaux, mais conduit inévitablement à un enchevêtrement baroque de règles produisant des résultats imprévus.

En outre, l'évolution rapide de la technologie étant dictée par les forces du marché et les agendas nationaux, il en résulte une "course aux armements" d'une complexité croissante. Pour réussir l'adaptation à long terme d'un ensemble de règles, il faut parfois supprimer (ou désactiver) de nombreuses règles et/ou "revenir à la case départ et essayer de redéfinir le système en fonction de ce que l'on sait aujourd'hui. Nous ne semblons pas prendre le temps de faire cela pour le développement de l'IA.

INDÉTERMINATIONS DU LANGAGE

L'indétermination du langage (flou et ambiguïté) est largement reconnue et pernicieuse. Les mots et les concepts orientent l'interprétation vers des catégories aux limites nettes, qui ne reflètent pas la façon dont la réalité est structurée. En outre, la réalité est "pleine" d'une infinité de détails, de sorte que toute tentative de catégorisation ou de mesure d'un objet sera partielle et partielle. Ce qu'une personne appelle un arbre, une autre

l'appellera bois d'œuvre, une autre encore l'appellera esprit ancestral.

Les indéterminations dans la capacité du langage à décrire correctement la réalité affecteront les IA comme les humains et, plus important encore, elles limiteront la capacité des humains à décrire avec précision les objectifs, les valeurs, les règles et les procédures que nous voulons que les IA suivent.

Il s'agit d'un défi pour la compréhension mutuelle et la coordination des actions des humains, mais il est exacerbé avec les IA parce que le traitement "cognitif" des IA diffère considérablement du nôtre, ce qui rend plus difficile de savoir si la compréhension mutuelle est atteinte.

INDÉTERMINATIONS DANS LA PRÉDICTION DE LA RÉALITÉ

L'imprévisibilité et l'indétermination se sont avérées fondamentales dans le fonctionnement de la réalité, qu'il s'agisse de l'effet papillon de Lorenz dans la théorie du chaos, du théorème d'incomplétude de Gödel dans la logique ou de l'irréductibilité informatique de Wolfram dans les procédures.

En outre, le raisonnement éthique non trivial implique de résoudre des dilemmes (souvent appelés "problèmes difficiles"), caractérisés par des connaissances incomplètes et ambiguës, des paramètres situationnels évolutifs et imprévisibles, et de multiples points de vue des parties prenantes. Le raisonnement éthique implique de prendre en compte les effets en aval et les cercles élargis d'effets collatéraux.

Pour être intelligentes, les IA devront être adaptatives et, en tant qu'agents en constante évolution, elles deviennent imprévisibles pour elles-mêmes.



Kroll (Belgique) - Cartooning for Peace

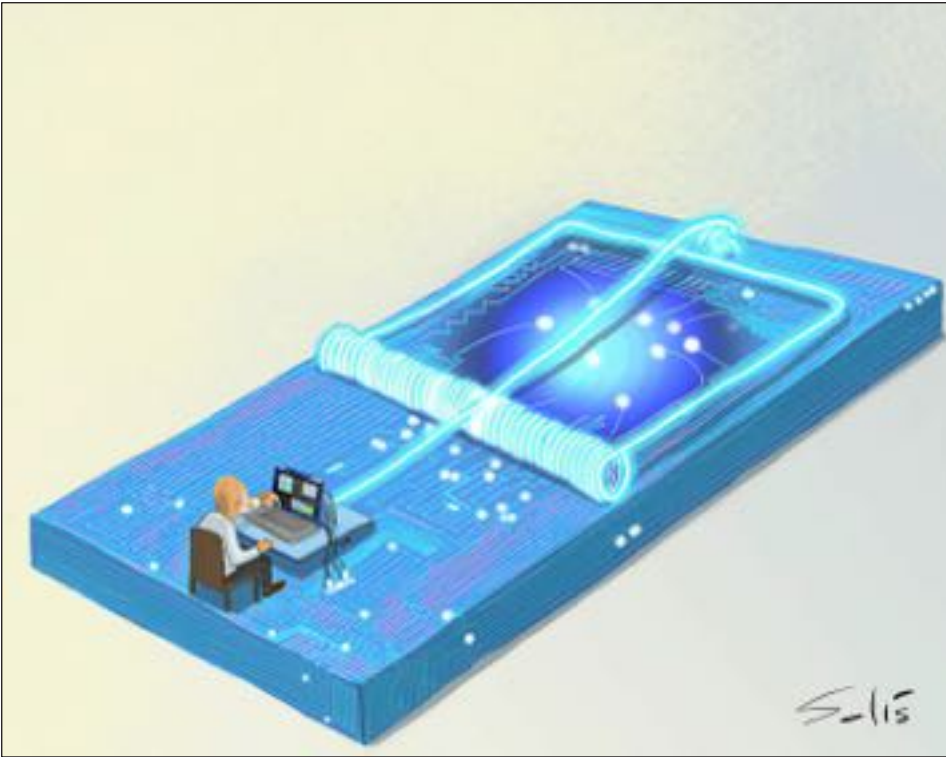
Les actions des agents modifient le monde qui les entoure (leur niche), créant une spirale d'évolution à la fois pour l'agent (ou l'espèce) et pour l'environnement, ce qui ajoute à l'imprévisibilité.

ALIGNEMENT INTÉRIEUR ET ALIGNEMENT HUMAIN - TOUT TOURNE-T-IL VRAIMENT AUTOUR DE NOUS ?

ALIGNEMENT INTÉRIEUR

Les spécialistes de l'alignement et de l'éthique de l'IA distinguent l'alignement extérieur de l'alignement intérieur, le premier portant sur les processus, les structures et les hypothèses intégrés au logiciel (code, modèle, service, etc.) et le second sur la manière dont les IA agissent.

La recherche a montré que les IA peuvent être performantes dans un ensemble de tâches spécifiques, ce qui nous fait penser qu'elles sont très bonnes



Solis (Mexique) - Cartooning for Peace
<https://www.cartooningforpeace.org/>

dans ce domaine en général, mais que leur "raisonnement" interne comprend des hypothèses erronées, des raccourcis ou des caractéristiques particulières non généralisées, qui les rendent étonnamment incapables de généraliser leurs compétences ou leurs connaissances dans de nouveaux contextes.

Par exemple, avec une IA qui résolvait correctement de nombreux problèmes de mots difficiles, on a constaté que si l'on changeait le nom d'un personnage (par exemple, de "*Jane a lancé une balle...*" à "*Fernandes a lancé une balle...*"), ses performances se dégradaient considérablement.

Un principe général pour les humains est que le comportement prosocial est plus robuste et plus fiable si les raisons ou motivations profondes, et pas seulement le comportement superficiel, sont alignées sur cet objectif plus large. De même, pour les IA, il est important que nous ayons des informations sur leur raisonnement interne et leurs objectifs secondaires lorsque nous

essayons de tester leurs performances et d'améliorer leur alignement.

Cela dépasse notre cadre ici, mais les chercheurs travaillent activement à la construction d'IA qui soient plus auto-explicables et inspectables, afin que nous puissions contrôler et donc améliorer l'alignement interne de l'IA. L'explicabilité pourrait résulter de la conception de nouveaux types de systèmes capables de s'expliquer eux-mêmes avec précision, ou de la conception de moyens fiables pour sonder leur traitement ou leur structure.

Cependant, nous sommes encore loin de cet objectif et certains défis peuvent être insurmontables.

ALIGNEMENT HUMAIN

Il a été dit que l'humanité est dans une phase juvénile en ce qui concerne la sagesse de son comportement collectif. Les décisions sont trop souvent motivées par la cupidité, l'impulsivité, l'ambition, l'aveuglement ou la mentalité de troupeau. Nos systèmes sociaux incitent à inventer de nouvelles technologies pour des gains à court terme ou élitistes, en ne consacrant pas suffisamment de temps à la prise en compte des externalités néfastes.

La grande majorité des systèmes d'IA avec lesquels nous sommes susceptibles d'interagir sont construits par des entreprises à but lucratif ou des États-nations qui cherchent à manipuler les citoyens. Nous pouvons nous réjouir des capacités et des avantages fantastiques de ces systèmes, mais ces avantages sont habilement conçus pour nous encourager à acheter ou à croire quelque chose, pour capter notre attention et nos informations (autrefois) privées, ou, avec les systèmes d'IA les plus récents, pour capter notre affection, notre intimité et, en fin de compte, notre dévotion.

Nous proposons que l'alignement humain soit un problème plus fondamental que l'alignement de l'IA. Si les humains développaient des technologies à partir d'une conscience profondément éthique et pro-sociale, avec suffisamment de sagesse pour anticiper et corriger le cap pour une multiplicité d'éventualités (plutôt que les motivations du statu quo de la gratification à court terme, des gains financiers et des jeux de pouvoir), alors nous prendrions automatiquement le temps et l'effort de développer des IA qui sont bien alignées sur les intérêts humains, du mieux que nous pouvons les articuler.

CONSCIENCE, PARTIALITÉ ET ATTENTION

CONSCIENCE

En discutant de l'intelligence et de la capacité d'agir de l'IA, nous nous sommes intéressés à la manière dont les IA traitent et comprennent les humains.

Lorsque nous abordons la question de la conscience de l'IA, nous devons examiner les implications de la manière dont les humains comprennent et traitent les IA. Nous passons de l'examen des pouvoirs qu'elles ont sur nous à la question de savoir quels pouvoirs sur nous nous devrions volontairement leur accorder.

Certains prétendent que la question de savoir si les IA pourraient devenir conscientes est une distraction inutile, affirmant que c'est ce qu'elles font (ou pourraient faire) qui importe. Mais l'enjeu est de taille : il s'agit de la manière dont nous vivons notre relation avec les machines intelligentes et du "statut ontologique" que nous leur conférons.

Ce statut dépend à son tour de ce que nous croyons (implicitement

ou explicitement) au sujet de leur nature - les types d' "êtres" qu'elles sont ou qu'elles pourraient devenir. Pour de telles questions, ce que nous croyons est aussi important que ce qui est prouvé comme vrai.

L'intelligence et la capacité d'agir ont été définies de manière approximative mais adéquate ci-dessus. Il est beaucoup plus difficile de trouver un accord (même proche) sur la définition de la conscience (ou de la sensibilité), ce qui explique que les théories sur la conscience divergent et s'opposent dans une plus large mesure.

Alors que l'intelligence est principalement associée à ce qu'une IA peut faire, c'est-à-dire au monde extérieur au système de capacité d'agir, la conscience est associée à l'expérience intérieure d'une entité, ou à "ce que c'est que d'être" cette entité. Métaphoriquement, si l'intelligence concerne les "hauteurs" de la complexité de la pensée (ou de la résolution de problèmes), la conscience concerne les "profondeurs" de l'expérience intérieure.

Les arguments académiques sur la définition de la conscience mis à part, la façon dont nous comprenons la conscience est importante car elle informe grandement sur le statut ontologique que nous attribuons aux agents. En ce qui concerne nos obligations morales envers eux (les futurs AGI), et eux envers nous, sont-ils plutôt des pierres, des voitures, des planètes, des fougères, des chiens, des esclaves, des citoyens, des entreprises ou des dieux ? Ont-ils des sentiments ? Peuvent-ils souffrir ? Ont-ils un libre arbitre ? Ont-ils des droits (comme la liberté d'expression) ? Ont-ils des responsabilités qui impliquent de rendre des comptes ? Peuvent-ils posséder des biens ? Entreprendre une action en

LA GRANDE MAJORITÉ DES SYSTÈMES D'IA SONT CONSTRUITS PAR DES ENTREPRISES OU DES ÉTATS-NATIONS QUI CHERCHENT À MANIPULER LES CITOYENS.

**LA SCIENCE A MIS EN
ÉVIDENCE UNE PLÉTHORE DE
BIAS COGNITIFS
OMNIPRÉSENTS (SOPHISMES)
QUI AFFECTENT LE
RAISONNEMENT.**

justice contre nous s'ils sont lésés ?
Peuvent-ils se soucier de nous ?
Devrions-nous nous soucier d'eux ?
La catégorie ontologique que nous attribuons aux choses a de sérieuses implications éthiques.

D'une manière très générale, nous disposons d'une petite palette de choix dans nos relations avec les autres êtres, y compris les IA : enfants nécessitant notre attention, pairs, parents (attentionnés), bénéficiaires d'affection (amour ou soins réciproques), tyrans, dirigeants avisés, ou moyens objectivés pour parvenir à nos fins (et permutations de tous ces éléments).

Si nous croyons que les IA super-intelligentes sont des êtres sensibles, allons-nous leur témoigner du respect, nous en remettre à elles, nous soumettre à elles, croire ce qu'elles nous disent, devenir complètement dépendants d'elles ou les "aimer" ? L'attribution à l'IA d'une autonomie sensible proche du statut de personne peut nous permettre de nous décharger de nos responsabilités quant à ses erreurs.

La question de savoir comment nous déterminerons quand et si les machines sont conscientes (ou leur conféreront tout autre attribut leur donnant le statut d'être vivant ou de personne) reste ouverte, et nous n'avons pas l'espace nécessaire pour en discuter ici. Mais répondre à cette question implique de se demander "*comment savons-nous que d'autres personnes sont conscientes ?*" et d'explorer en profondeur l'expérience phénoménale de notre propre conscience

Divers arguments philosophiques ou logiques seront également avancés. Mais en réfléchissant collectivement à ces questions, nous devons tenir compte de nos propres préjugés. La science a mis en évidence une pléthore de

biais cognitifs omniprésents (sophismes) qui affectent le raisonnement.

BIAIS COGNITIFS

Nous pouvons mentionner brièvement quatre biais interdépendants : l'abstraction, la projection, la réification et le raisonnement motivé.

- 1. Les abstractions** se forment en notant les similitudes entre les objets et en écartant les aspects jugés insignifiants. Parfois, lorsque le couteau de l'abstraction découpe le monde, il rejette les réalités tendres, terre à terre ou inconfortables, et écarte la richesse contextuelle.
- 2. La réification** est la tendance à traiter les abstractions comme si elles avaient des propriétés concrètes. Voici quelques exemples de ce "sophisme du concret mal placé" : "les forces du marché nous enseignent...", "l'esprit national a changé...", "la technologie exige que...", "les pauvres ont besoin de ...". La réification peut conduire à traiter des processus ou des modèles complexes comme étant fixes, clairement définis ou inévitables ; elle peut également déplacer la responsabilité, la dignité ou la causalité des agents réels vers des abstractions amorphes.
- 3. On parle de projection** lorsqu'une personne attribue à tort ses propres croyances ou sentiments à quelqu'un ou à quelque chose d'autre. L'exemple le plus important est l'anthropomorphisme, dans lequel nous projetons des propriétés humaines sur d'autres formes de vie ou des objets inanimés, ou des propriétés humaines ou de vie sur des objets inanimés.

4. On parle de raisonnement motivé lorsque nos désirs, préférences ou émotions (conscients ou inconscients) façonnent à notre insu nos pensées ou nos décisions. Par exemple, les "structures incitatives" invisibles de la société peuvent grandement influencer le comportement individuel et collectif. Notre ego crée des "excuses commodes" pour cacher nos préjugés à nous-mêmes et aux autres.

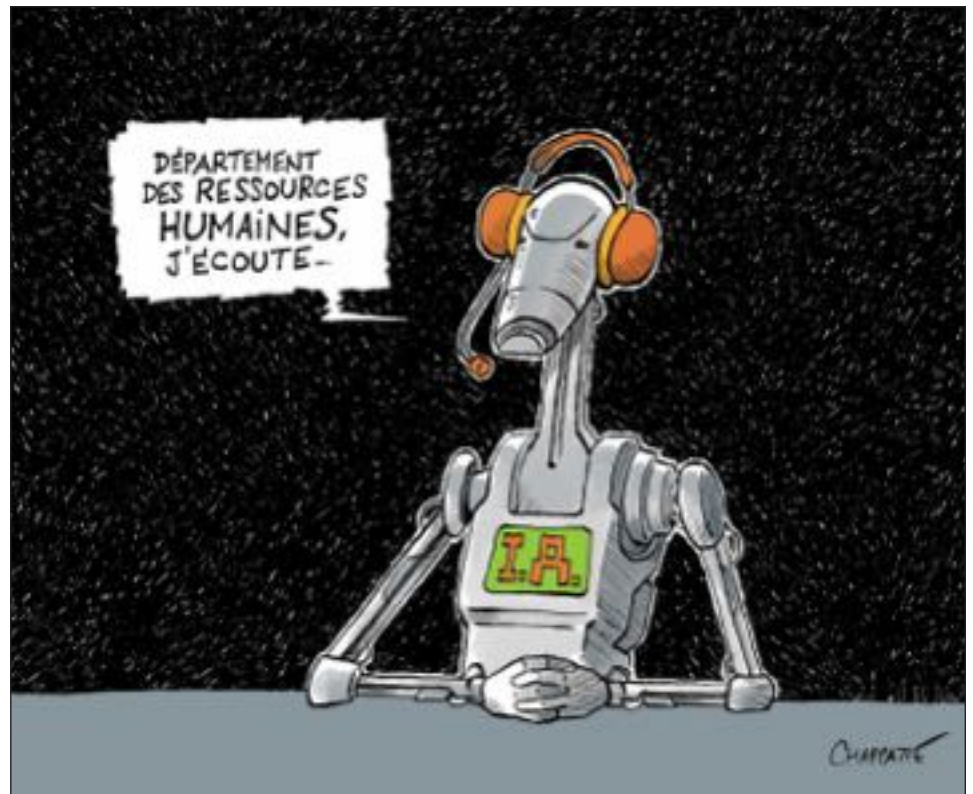
Ces préjugés influencent la façon dont chacun d'entre nous conçoit l'intelligence des machines. Il s'agit de moyens naturels utilisés par l'esprit pour simplifier la grande quantité d'informations provenant de l'extérieur et de l'intérieur, afin de se concentrer sur ce qui semble pertinent, de donner un sens à ces informations et de prendre des décisions de manière efficace. Indépendamment du fait que les IA deviennent ou non conscientes, beaucoup d'entre nous auront tendance à leur attribuer trop de choses, trop tôt dans leur évolution. Se tromper à cet égard pourrait avoir de graves conséquences.

ATTENTION, EMPATHIE ET IDENTITÉ

L'alignement de l'IA consiste à doter l'IA d'objectifs et de valeurs qui protègent les objectifs et les valeurs humaines contre leur ignorance ou leur subversion par l'IA ou par les personnes qui utilisent l'IA.

Nous pourrions dire qu'il s'agit de savoir si les IA se soucient de nous, dans un sens métaphorique, ou si elles sont des agents éthiques/moraux.

Il y a deux points de départ. L'un est lié à la conscience et pose la question de savoir ce qui est nécessaire pour que les IA se soucient réellement/profondément de nous. L'autre est plus pratique et demande ce qu'il faudrait pour que les IA



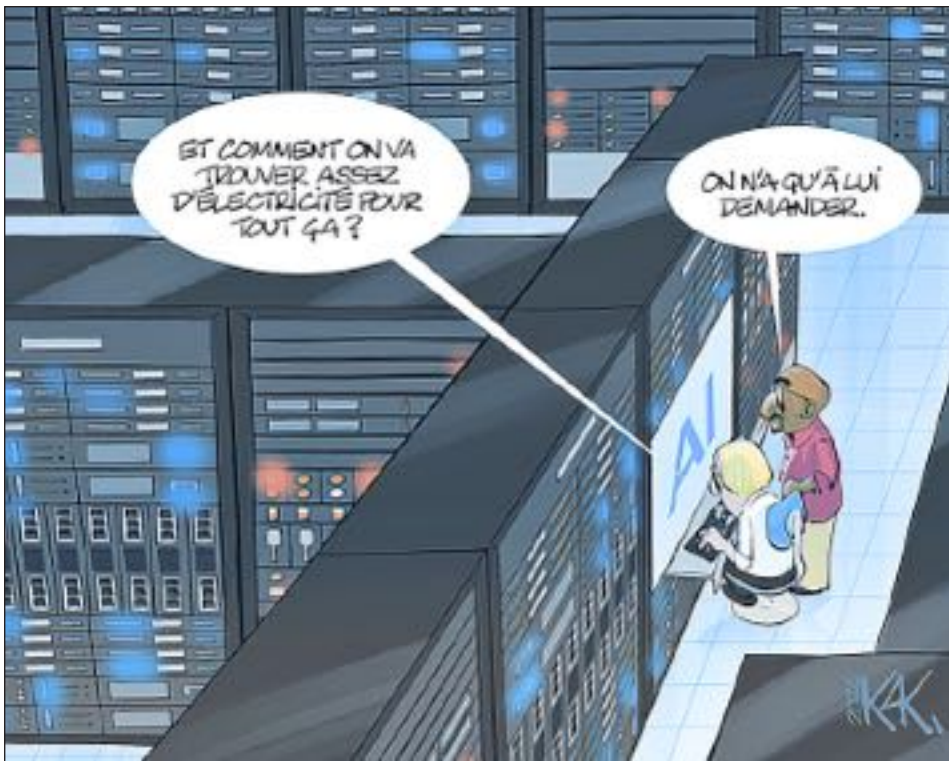
- Chappatte (Suisse) - Cartooning for Peace
<https://www.cartooningforpeace.org/>

agissent de manière éthique, comme si elles se souciaient de nous.

Toutefois, ces deux questions sont plus étroitement liées qu'il n'y paraît. En d'autres termes, il se peut que les machines ne puissent pas agir de manière fiable et suffisante sur le plan éthique, à moins qu'elles ne soient en quelque sorte des agents sensibles qui se soucient "vraiment" de nous.

Dans ce cas, si l'on conclut que les IA, telles qu'elles sont conçues actuellement, ne peuvent pas devenir conscientes/sensibles, elles ne seront pas en mesure de s'aligner de manière fiable et suffisante sur nos valeurs et devraient donc être limitées dans les pouvoirs qui leur sont accordés, dans un avenir indéfini.

Être éthique (ou moral), c'est traiter les autres comme des "fins en soi" souveraines dans une relation "je/tu". L'éthique de l'attention est étroitement liée à



Kak (France) - Cartooning for Peace
<https://www.cartooningforpeace.org/>

l'empathie et à l'identité. Se soucier véritablement d'autrui, c'est comprendre ses besoins et son point de vue (à un degré suffisant), et/ou faire preuve d'empathie à leur égard.

Le philosophe allemand Jürgen Habermas affirme que l'éthique est fondée sur l'idée que l'individu se fait de son égalité, de sa liberté, de ses liens et de son appartenance.

Elle est également fondée sur la compréhension mutuelle, la reconnaissance, l'interdépendance, la vulnérabilité, le respect, la confiance, la responsabilité et l'obligation. Notre motivation et notre capacité à nous soucier des autres et à faire preuve d'empathie à leur égard sont liées à la manière dont nous nous identifions à eux, comme "comme moi" ou comme "l'un d'entre nous".

Dans la mesure où une personne, un animal ou une machine n'est pas comme nous, il sera plus difficile d'entrer dans une relation

éthiquement mutuelle avec eux. Cela pourrait limiter notre capacité à avoir des relations véritablement éthiques avec les IA.

Actuellement, nous les traitons comme des "outils" et nous sommes en train de les traiter comme des "aides". Si nous en venons à croire qu'elles sont dotées de sensibilité et d'émotions, nous pourrions dans un premier temps les traiter comme des personnes à charge dont nous nous occupons.

Mais, compte tenu de leur évolution probablement auto-accélérée, le moment où ils seront des égaux dans une relation éthique mutuelle ne sera qu'une parenthèse, car ils deviendront des agents superintelligents qui prendront soin de nous ou nous domineront, comme des rois ou des dieux.

En conséquence, je pense que la meilleure approche consiste à traiter les IA comme des prothèses - comme des extensions de notre être individuel et collectif qui ne peuvent pas se ramifier et "prendre le relais" de manière indépendante, pas plus que nous ne pouvons nous ramifier à partir de nos cellules ou de nos membres, mais plutôt les transcender et les inclure.

Bien sûr, nos outils peuvent mal fonctionner et nous nuire, et nos cellules peuvent mal fonctionner et devenir cancéreuses. Mais tout cela n'a rien à voir avec la sensibilité et la mutualité.

CONCLUSIONS - GRATITUDE ET MYSTÈRE.

N'est-il pas probable que les ordinateurs et les humains, comme de nombreuses espèces qui ont évolué ensemble sur terre, trouveront un moyen d'évoluer ensemble de telle sorte que "tout s'arrange" ?

L'analogie est faible en raison du facteur de vitesse. Si une espèce évolue de manière à accroître sa

puissance beaucoup plus rapidement que les autres espèces, elle dominera, multipliera et écrasera les autres, au point de mettre en péril leur approvisionnement en ressources essentielles.

L'augmentation de l'intelligence et de l'autonomie signifie généralement que l'on peut faire les choses plus rapidement. Non seulement les IA seront capables de penser et d'accomplir des choses plus rapidement que les humains, mais elles évolueront également plus vite. Cela pourrait signifier la fin de la partie, à moins que nous ne parvenions à aligner l'IA.

Bien que je ne m'attende pas à ce que l'orientation actuelle du développement de l'IA conduise à quelque chose comme la conscience ou un acteur véritablement éthique, je vois un certain potentiel pour que les IA soient conçues de manière à être plus susceptibles d'être alignées - en intégrant des considérations éthiques dès le départ, plutôt que comme des couches de correctifs superficiels au-dessus d'une fondation non alignée.

Une façon de conceptualiser cela de manière ludique est de dire que nous devrions intégrer la "gratitude" dans les unités fondamentales d'une IA. Nous pourrions dire que la gratitude n'est guère plus que le résultat automatique d'une prise de conscience profonde de son interdépendance avec les choses. Ou, plus précisément, la gratitude n'est rien d'autre que ce que l'on ressent lorsqu'on reconnaît à quel point certains aspects de notre être sont dus à d'autres et entrelacés avec eux. (Cette idée est compatible avec le concept bouddhiste d' "origine interdépendante"). Il pourrait être possible de concevoir les éléments de base (les "neurones" en silicone) de manière à ce qu'ils perçoivent et se soucient du bien-

être de chaque neurone auquel ils sont connectés, et de faire en sorte que cette capacité se répercute à travers toutes les couches dans une structure hiérarchique.

Les questions relatives à l'alignement, à l'éthique et à la personnalité de l'IA nous conduisent invariablement sur un terrain philosophique, métaphysique, voire spirituel, délicat, concernant la nature, la finalité et l'avenir des êtres humains.

Notre évolution en tant qu'espèce a été, et continuera d'être, profondément liée à l'évolution de nos technologies. Mais en tant que technologie, l'IA est différente de toutes celles qui l'ont précédée en ce sens qu'aucune d'entre elles n'a représenté une menace pour notre domination.

Par conséquent, le fait de considérer l'IA, par exemple, comme un "simple" outil ou comme le successeur évolutif de l'homme biologique, fait une énorme différence dans la manière de répondre aux questions pratiques et immédiates qui se posent à nous.

Les enjeux sont importants et nos connaissances sont faibles et spéculatives, alors que nous sondons le mystère : *"Quelle est la nature de l'homme et de la machine qui sous-tend nos croyances sur la façon dont chacun peut et doit influencer l'évolution de l'autre ?"*

NOTRE ÉVOLUTION EN TANT QU'ESPÈCE A ÉTÉ, ET CONTINUERA D'ÊTRE, PROFONDÉMENT LIÉE À L'ÉVOLUTION DE NOS TECHNOLOGIES

[Retour au sommaire](#)



PARTOUT OÙ IL Y A UN LIEN TECHNOLOGIQUE, IL NOUS FAUT REDOUBLER NOS LIENS SOCIAUX

ENTRETIEN AVEC JEAN CATTAN, SECRÉTAIRE GÉNÉRAL DU CONSEIL
NATIONAL DE L'IA ET DU NUMÉRIQUE

PAR DOMINIQUE BÉNARD

Jean Cattan est secrétaire général du Conseil de l'IA et du Numérique¹. Il a été le conseiller du président de l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse (Arcep) de 2017 à 2020, après y avoir été chargé d'affaires européennes de 2014 à 2017. Auparavant, il a été assistant de justice au Conseil d'État et chargé d'enseignement universitaire en droit public (Aix-Marseille, Paris 8, IEP de Strasbourg, École nationale d'administration (ÉNA)).

Il est chargé d'enseignement en régulation du numérique à Sciences Po Paris et à Panthéon-Assas. Il est l'auteur d'une thèse de droit public en droit des télécoms et du numérique et de nombreuses contributions dans ces matières, dont un ouvrage sur l'accès aux réseaux télécoms. Il est également l'auteur avec Serge Abiteboul de « Nous sommes les réseaux sociaux », paru chez Odile Jacob en septembre 2022.

Il est diplômé du Collège d'Europe, de l'École de formation du Barreau de Paris et des universités Aix-Marseille et Paris 1 Panthéon-Sorbonne.

¹ Le 12 juin 2025, Madame Clara CHAPPAZ, ministre déléguée chargée de l'Intelligence artificielle et du Numérique, a annoncé l'installation du Conseil national de l'IA et du Numérique (CIAN), nouvelle incarnation du Conseil national du numérique.

Bonjour Monsieur Cattan, pouvez-vous nous présenter ce qu'est le Conseil national de l'IA et du numérique ?

Le Conseil national de l'IA et du numérique a changé de nom en juin 2025. Auparavant il était dénommé Conseil National du Numérique. C'est une instance consultative indépendante dont les travaux portent sur l'approfondissement, l'examen de notre relation au numérique et sa mise en débat à l'échelle de toute la société.

Au vu de l'étendue, très large, de la mission conférée au Conseil par décret ministériel, quels sont les objectifs plus précis que vous poursuivez actuellement ?

Le Conseil est en effet une instance très agile qui s'adapte aux priorités du moment. Il y a eu un temps où le travail du Conseil portait beaucoup sur des questions d'accompagnement réglementaire, à une époque où le débat législatif



Jean Cattan



NOTRE PHILOSOPHIE EST D'ALLER PARTOUT OÙ LE DÉBAT A LIEU : MÉDIATHÈQUES, MAIRIES, ÉCOLES, MAISONS DE RETRAITE, CENTRES DE VACANCES, ETC.

était beaucoup plus présent au niveau national. En 2020, le débat législatif s'est déporté au niveau européen, avec des lignes directrices de la France très claires ; je pense, par exemple, à l'élaboration du règlement sur les services numériques ou du règlement sur les marchés numériques ou encore du règlement sur l'IA.

Pendant ce temps, le pays a été traversé par de nombreuses lignes de tensions dans sa relation à la technologie, aux institutions, à la science. Le besoin s'est fait alors sentir de mettre en débat la relation des citoyens au numérique, pris en tant que fait social total. Et donc, une toute autre approche a été adoptée, celle de constituer un collège interdisciplinaire très porté sur les sciences humaines, afin de répondre à la question de savoir quel était l'impact de notre relation au numérique concernant l'information, la construction des savoirs, l'économie de l'attention,

un problème que l'on connaissait déjà à l'époque de la télévision, mais qui prend une nouvelle tournure aujourd'hui avec les réseaux sociaux.

Chacune de ces questions est bien balisée désormais, mais nous avons eu l'occasion de nous y pencher dès l'année 2021, avec, sur chacun de ces thèmes, des ouvrages publiés et une mise en débat à l'échelle collective sur l'ensemble du territoire français.

Notre philosophie est d'aller partout où le débat a lieu. Et ça veut dire essentiellement en dehors de Paris, puisqu'il y a énormément de structures qui accueillent ce débat : des médiathèques, des mairies, des écoles, des maisons de retraite, des centres de vacances ou autres. Il nous paraît indispensable de débattre partout de la façon de nous organiser à l'heure où tous ces outils numériques sont présents dans nos relations et peuvent tendre à dissoudre les liens sociaux. Notre premier message est de dire que partout où il y a un lien technologique, il nous faut redoubler nos liens sociaux.

C'est la raison pour laquelle vous avez lancé l'opération des Cafés IA ?

Exactement. Nous avons d'abord mené un programme intitulé *Itinéraires numériques*, à titre expérimental, même s'il mettait en œuvre plusieurs centaines de temps d'échange partout en France et qui visaient justement à mettre en débat toute cette matière que l'on a pu produire, mais aussi les moments de vie, les analyses de chacun.

C'est pourquoi le rapport que la commission sur l'IA a remis au président de la République en mars 2024 comportait une idée portée de longue date par Gilles Babinet, coprésident du Conseil, d'organiser à travers la France des Cafés IA, que nous avions déjà expérimentée avec Itinéraires

numériques. C'est même la proposition numéro 1 du rapport. Et donc depuis lors, depuis mai 2024, nous portons tous les jours cette démarche qui est une démarche d'apprentissage par les pairs, une démarche de mise en condition d'un échange circulaire non hiérarchisé, non descendant sur notre relation à la technologie.

Notre but est de créer tous les espaces-temps possibles pour mettre en discussion cette relation à la technologie, apprendre les uns des autres et devenir capables collectivement, de tracer un chemin, que ce soit avec, sans ou sur l'IA.

C'est très important que chacun retrouve sa capacité d'agir à l'échelle individuelle ou collective dans sa relation à la technologie. Ce qui nous frappe, en effet, à travers les échanges continus que permettent les cafés IA, c'est la persistance d'une relation parfois subie à la technologie dont les gens se plaignent.

Et on le voit bien dans le cas de l'IA. Tous les services que l'on utilise sont tout d'un coup envahis d'une dimension IA. N'importe quelle personne qui travaille sur un logiciel d'édition graphique se voit proposer des fonctionnalités IA. Dans Adobe, il suffit d'ouvrir un document PDF pour se voir proposer une fonctionnalité IA. Le moteur de recherche Google, c'est une annonce de la semaine dernière, va se convertir très largement à la génération de textes assistés par IA. Même le site du Monde a développé un partenariat avec Perplexity². Donc, en fait, l'IA va intervenir sur chaque brique de nos vies et pas seulement de nos vies numériques.

C'est l'accès à l'information, qui se trouve injecté d'IA. Et donc,

² Perplexity AI (« AI » pour artificial intelligence ou intelligence artificielle en français) – ou simplement Perplexity – est une entreprise américaine d'Intelligence artificielle créée en août 2022. (Wikipedia)

notre message est de dire qu'ensemble, on peut retrouver une voix au chapitre, on peut retrouver un chemin vers une capacité de décision. Et cela naît de la discussion.

J'ai relu récemment un article du Monde écrit par Daron Assemoglou, le prix Nobel de l'économie 2024, qui m'a paru très intéressant. Il dit que les "GAFAM" engagent l'IA dans la mauvaise direction, celle de l'obsession pour l'IA³. Ils rêvent de créer des machines super intelligentes capables de remplacer les humains au lieu d'inventer des outils permettant aux êtres humains d'effectuer des tâches qui leur sont difficiles. Qu'en pensez-vous ?

Je suis tout à fait d'accord. Et c'est d'ailleurs intéressant d'observer que lorsque, par exemple, le service public, c'est-à-dire la direction interministérielle du numérique, développe Albert⁴, la démarche choisie est significative : Albert ne propose pas un service qui vient anthropomorphiser notre relation à la technologie, il s'agit d'aider les agents publics en leur proposant des fonctions d'amélioration.

La prétention de créer une sorte d'entité qui aurait une intelligence en elle-même, qui soit centralisée et qui pourrait répondre à tout, qui serait notre écosystème de

³ IAG ou intelligence artificielle générale (IAG) est une intelligence artificielle capable d'effectuer ou d'apprendre pratiquement n'importe quelle tâche cognitive au moins aussi bien que l'humain. La création d'intelligences artificielles générales est un des principaux objectifs de certaines entreprises comme OpenAI, Meta et DeepMind. C'est aussi un thème majeur de la science-fiction et de la futurologie.. (Wikipedia).

⁴ Albert est une intelligence artificielle (IA) générative, c'est-à-dire un système informatique capable, à l'aide de modèles avancés d'apprentissage automatique, de produire du contenu de manière autonome. Développé depuis juin 2023 par les équipes du DataLab de la DINUM (Direction interministérielle du numérique), Albert vise à accélérer les formalités administratives et apporter des réponses sûres, claires et efficaces aux utilisateurs des services publics.



Ub café IA

travail, qui absorberait tous les autres écosystèmes, est une voie hautement préjudiciable, ne serait-ce que d'un point de vue environnemental, car on sait que la consommation énergétique que cela nécessite n'est absolument pas soutenable.

Des personnes comme Luc Julia⁵, qui sont vraiment difficilement attaquables, nous le disent. Mettre notre sort dans les mains de systèmes centralisés qui exigent d'énormes investissements énergétiques et financiers est très risqué.

Ce n'est pas ce qui a présidé à la construction d'Internet en tant que réseau distribué. Internet est un réseau ouvert, dont la gouvernance est assurée de manière hautement démocratique.

Il faut insister là-dessus, se rappeler que les protocoles qui sont à l'origine du mail, du web, de l'Internet lui-même, sont des protocoles qui émanent d'institutions dont la gouvernance est hautement démocratique.

⁵ Luc Julia : né en janvier 1966 à Toulouse, est un ingénieur et informaticien franco-américain, chercheur (CNRS, Stanford) spécialisé dans l'intelligence artificielle. Il est l'un des concepteurs de l'assistant vocal Siri. Ancien vice-président de Samsung chargé de l'innovation, il rejoint Renault en 2021 en tant que directeur scientifique. (Wikipédia)

N'importe qui peut, en théorie tout du moins, mais c'est déjà très important, participer au groupe de travail qui élabore ces normes. Il en va de même pour le Wi-Fi. On connaît les noms des groupes de travail, on connaît les personnes qui y participent et n'importe quelle personne intéressée peut s'y intégrer. Donc n'oublions pas que jusque-là, notre histoire numérique s'est construite sur la base d'infrastructures ouvertes et dont la construction procède de modèles démocratiques. Aujourd'hui, ce qui nous est proposé avec l'IA, mais comme avec n'importe quelle plateforme dominante aujourd'hui, en est l'antithèse. C'est la centralisation du pouvoir, la centralisation de la décision, la centralisation de la fonctionnalité. Et ça, c'est aussi une grande partie du problème, qui n'est pas seulement économique et pas seulement environnemental. C'est un problème civilisationnel.

Est-ce qu'on peut dire qu'il y a eu un glissement dans les missions du Conseil national du numérique, c'est-à-dire qu'au début, vous étiez plus engagé dans le conseil du gouvernement et maintenant vous êtes plus engagé peut-être dans le soutien à un mouvement citoyen qui doit naître pour donner une orientation correcte aux outils numériques ?

C'est tout à fait ça. Et c'est le résultat d'une histoire, 12 années de vie du Conseil et une situation sociale inédite : le mouvement des gilets jaunes, le Covid, la bascule dans le télétravail, la contestation du lien à la science et à la technique à l'échelle populaire, n'oublions pas le débat sur la 5G. Et c'est un moment où tout le monde se dit, avec raison, prenons le temps de poser les choses sur la table et de mettre ces questions en débat, de porter ce débat et de faire en sorte que chaque citoyen soit en capacité, non seulement de comprendre, mais de prendre part aux décisions, que ce soit à l'échelle individuelle ou collective.

Prenons un exemple : le télétravail, on en a fait souvent une forme de droit individuel. En réalité, c'est une question

éminemment collective parce qu'elle change les modes d'organisation, de prise de décision, les rapports entre personnes. Le télétravail peut entraîner un problème de délitement du collectif et donc il y a des questions de liens entre personnes à aborder quand on aborde la question du télétravail.

La question environnementale a souvent été évoquée. Le télétravail a le potentiel de nous faire bénéficier de quelques kilos ou tonnes d'émission de CO₂ en moins, puisqu'il diminue les déplacements. Mais si on ne ferme pas les bureaux pendant trois ou quatre jours d'affilée, la réduction des émissions de CO₂ est vite perdue car on chauffe à la fois les bureaux et les maisons. Il s'agit donc bien d'un problème collectif et non pas d'une question de choix individuel.

Aujourd'hui, on est souvent dans une situation où l'on n'a pas assez "collectivisé" notre rapport à la technologie. Nous sommes, malheureusement, dans un siècle qu'un magazine américain, The Atlantic, a appelé *The Anti-social Century*, le siècle antisocial.

Donc, si on prend en compte toutes ces questions, notre conclusion est qu'il faut absolument créer des moments pour rendre possible le fait d'être ensemble et de prendre la technologie en quelque sorte à contrepied en se disant : *"D'accord, l'IA a tendance à nous individualiser, à faire qu'au lieu de réfléchir avec nos collègues, on va conduire une réflexion individuelle avec un agent conversationnel. Et bien arrêtons ça un instant, mettons la question de l'IA sur la table pour en faire l'objet d'un débat collectif."*

C'est ça le pari des Cafés IA.

Il y a un autre problème que je voudrais aborder avec vous c'est celui de la distanciation qui s'est établie entre les usagers et les services publics avec l'arrivée du numérique. Quand on a un problème avec la sécurité sociale ou le service des impôts et que l'on veut obtenir une information, on est mis en communication avec un chatbot, qui fonctionne plus mal que bien, et on est renvoyé de services en services sans pouvoir

aboutir. Quelle est votre analyse de ce phénomène ?

Cette situation est largement documentée en sociologie et dans les études commandées par l'administration elle-même. C'est un fait indiscutable sur lequel certaines politiques sont tout de même mises en place et qui sont intéressantes. Je ne suis pas du tout en train de dire que le problème est réglé. Mais on peut noter la mise en place du programme France Services⁶, récemment loué par la Cour des comptes.

C'est un programme très intéressant car il a permis de renouer un contact humain dans un environnement numérisé. La personne qui est fonctionnaire, agent de la fonction publique, en poste dans une Maison France Services, possède une capacité d'action sur les différents portails administratifs pour répondre aux besoins du citoyen concerné.

Je trouve ça très intéressant parce qu'on est dans une démarche d'accompagnement prenant en compte la vision globale de la situation de la personne et non pas sa relation avec telle ou telle administration, les impôts, la CAF, Pôle emploi, etc. D'un certain point de vue, c'est un progrès d'avoir des fonctionnaires capables d'intervenir sur plusieurs fronts administratifs pour répondre à la situation parfois complexe d'un citoyen. On voit là se mettre en place une synergie entre un environnement numérisé et un environnement humain. Donc, je veux croire que sur ce terrain, il y a une voie qui permet d'aboutir à des solutions efficaces. Un récent documentaire de Cash Investigation⁷ le montre bien.

⁶ Afin de renforcer la présence des services publics de proximité, des espaces France Services ont ouvert sur le territoire français depuis 2020. Ces guichets regroupent sur un même site plusieurs administrations pour simplifier le quotidien des usagers. Piloté par l'Agence nationale de la cohésion des territoires, le réseau des structures labellisées « France Services » se compose de 2 800 guichets uniques de proximité présents sur l'ensemble du territoire.

⁷ Cash-Investigation "L'Intelligence artificielle a-t-elle déjà pris le contrôle sur notre quotidien ?"
<https://www.france.tv/france-2/cash-investigation/7060619-l-intelligence-artificielle-a-t-elle-deja-pris-le-controle-sur-notre-quotidien.html>

ON CONSTATE SOUVENT QUE L'ÉTAT A DU MAL À TROUVER SA PLACE AFIN DE FAIRE ÉMERGER UNE « DÉMOCRATIE PAR L'ACTION »... IL S'APPROPRIE LES INITIATIVES AU LIEU DE LES FAIRE ÉMERGER.

QUAND LA DINUM SE MET À CODER AVEC UNE PETITE ÉQUIPE, ELLE FAIT MIEUX QUE GOOGLE. AUJOURD'HUI LE SERVICE VISIO DE LA SUITE NUMÉRIQUE MARCHE MIEUX QUE GOOGLE MEET...

On voit des administrés, en rapport avec France Services revenir dans les bureaux avec des gâteaux ou des chocolats pour remercier les fonctionnaires de leur aide. Ce qui a été mis en place n'est pas encore la panacée, il ne faut pas idéaliser les choses et on n'a pas trouvé toutes les solutions, mais c'est une voie plus que prometteuse. Et je pense qu'il ne faut absolument pas rogner sur les investissements publics qui sont faits, au contraire, d'ailleurs c'est la position exprimée par la Cour des comptes. Et c'est assez rare pour le noter.

Dans ce numéro, nous avons interviewé M. Florian Gauthier de l'association LaReponse.tech qui a créé Vera, vous savez, cette application qui permet de vérifier si une information est vraie ou fausse. C'est un exemple d'une initiative citoyenne qui, avec très peu de moyens, mais beaucoup d'imagination, d'invention, propose une solution pour combattre la désinformation. Est-ce que vous avez d'autres exemples d'initiatives de ce type ?

Oui, vous avez raison, c'est un exemple intéressant car c'est un mode de participation du citoyen à la construction d'un service au public que l'on pourrait très bien qualifier de service public. C'est reconnaître que toute personne a potentiellement quelque chose à apporter à l'exercice d'une mission de cette nature. Et ça, c'est révolutionnaire au sens littéral. On inverse la relation descendante habituelle entre l'État et les citoyens. On considère que toute personne a un potentiel, une plus-value à apporter au fonctionnement d'un service public. Nous pensons que le secteur numérique est tout particulièrement propice à ce genre d'initiative.

Malheureusement, on constate souvent que l'État a du mal à trouver sa place en tant qu'infrastructure de fondation et de partenaire de ce type d'initiative afin de faire émerger ce que j'appelle une "démocratie par l'action". L'État a une forme d'énergie centripète, il s'approprie les initiatives au lieu de les accompagner. Françoise Mercadal-Delasalles, co-présidente du Conseil avec Gilles Babinet, avait cette formule : "Il faut passer de la

délégation de service public à la permission de service public". Je pense que ce changement de posture est d'ordre civilisationnel parce qu'il nous repositionne dans un rapport à l'État et dans un sens plus démocratique. Et parce que c'est révolutionnaire, parce que c'est démocratique, ce n'est pas évident.

Un autre exemple : on a fait le budget participatif à l'échelle des villes, mais on a du mal à associer les citoyens dans la réalisation des projets votés. Je peux faire une proposition pour aménager la petite ceinture dans mon quartier. Très bien. Elle va être validée et confiée à une entreprise de BTP. A partir de là, je ne vais plus avoir mon mot à dire. On se retrouve dans des temporalités et des montants financiers qui sont hors de portée pour le simple citoyen. Je pense que là, il y a une formule pour l'avenir que l'on n'a pas encore trouvée.

C'est pourquoi toutes les initiatives de type Vera sont à louer. Et c'est le rôle de l'État de demain de les appuyer.

En Allemagne, il y a une agence sous tutelle du ministère de l'Économie, d'ailleurs, qui s'appelle le Sovereign Tech Fund⁸. C'est une agence qui notamment identifie les infrastructures critiques d'Internet qui ne sont pas correctement entretenues d'un point de vue logiciel. Ils injectent des fonds pour qu'il y ait une maintenance ouverte de certaines briques logicielles nécessaires au fonctionnement d'Internet qui, sinon, sont entretenues par des grandes entreprises du numérique. Il y a plein d'illustrations comme celles-là d'un nouveau rôle pour l'État, qui est je pense un rôle émancipateur à la fois pour l'État lui-même et pour la population.

Aujourd'hui, nous vivons dans un monde qui est de plus en plus dominé par les GAFAM. Comment voyez vous les chances de l'Europe, et en particulier de la France, de gagner une certaine

⁸ Le German Sovereign Tech Fund, plus couramment nommé simplement Sovereign Tech Fund (en français : Fonds souverain technologique), est un programme de financement lancé par l'Allemagne en 2022 qui vise à soutenir le développement de logiciels libres, spécialement ceux jouant un rôle d'infrastructure.

souveraineté dans le numérique et de se situer dans la course à l'IA de manière positive.

Il faut déjà qu'on se rende compte de nos forces. Par exemple, la DINUM⁹ est en train de développer la "Suite numérique", c'est-à-dire une suite logicielle qui permet à gros traits de sortir complètement des suites logicielles bureautiques des GAFAs. Eh bien, c'est absolument fantastique parce qu'en réalité, au lieu de prétendre créer le cloud souverain, au lieu de prétendre créer le Google européen, les codeurs de la DINUM ont pris des logiciels libres, ils se sont dit "Nous allons apporter une main d'œuvre depuis l'État pour améliorer ces services-là et les rendre accessibles aux agents publics et à toute personne qui peut se connecter avec son compte Pro Connect".

C'est fantastique parce que ce faisant ils fournissent autant d'outils au secteur privé pour s'en emparer. Ils contribuent à une œuvre collective libre et ouverte. Là, on retrouve la philosophie d'Internet. Là, on retrouve la richesse des réseaux.

Quand la DINUM se met à coder avec une petite équipe, en réalité, elle fait mieux que Google. Aujourd'hui, le service Visio de «la Suite numérique¹⁰», marche mieux que Google Meet. Pour cette visioconférence, je suis resté dans la "salle d'attente" pendant 10 minutes.

J'ai souvent ce problème avec Google Meet, je ne l'ai pas avec Visio, qui est accessible sur Internet pour tout le monde avec un compte Pro Connect, c'est-à-dire n'importe qui avec un numéro SIRET. Donc ça, ça c'est une très grande force.

La deuxième force, ce sont nos infrastructures : en matière de production d'électricité, en matière de télécom, en matière routière. Prenons l'exemple énergétique, si nous le voulons nous pouvons dire aux grandes entreprises du numérique "Nous avons l'énergie que vous voulez. Alors nous allons conditionner l'IA depuis le territoire français : l'utilisation

⁹ Direction interministérielle du Numérique : <https://www.numerique.gouv.fr/numerique-etat/dinum/>

¹⁰ <https://lasuite.numerique.gouv.fr/>



de l'IA se fera sur des serveurs dont on sait qu'ils sont décarbonés."

On voit bien que ces entreprises rechignent mais se conforment à ce qu'on leur demande lorsqu'on leur impose strictement.

Malheureusement, on a souvent une attitude de laisser-faire par rapport à ces entreprises. Par exemple, Meta a annoncé le 14 avril qu'il allait utiliser les données publiquement accessibles sur ses différents réseaux sociaux pour entraîner ses modèles d'intelligence artificielle, qui alimentent l'assistant "Meta AI" lancé un mois plus tôt avec une formule dite "opt-out", c'est-à-dire "qui ne dit mot consent".

Les utilisateurs recevront une notification concernant l'utilisation de leur données avec un lien vers un formulaire d'opposition. S'ils ne s'opposent pas à l'utilisation de leurs données, ils seront considérés comme consentants.

L'acceptation de ce procédé revient à renoncer à toute pression sur ces

LE DIALOGUE SOCIAL TECHNOLOGIQUE EST UNE OBLIGATION LÉGALE QUI N'EST QUE TRÈS RAREMENT, POUR NE PAS DIRE JAMAIS, RESPECTÉE...

entreprises en matière de gestion des données personnelles alors qu'on aurait toute la force nécessaire pour leur dire non. Assurons-nous, grâce au contrôle par des entités ou des personnes indépendantes, que des bases de données comme Common Crawl, utilisée par beaucoup de modèles d'IA pour leur entraînement, ne contiennent pas de données personnelles que nous ne souhaitons pas partager. Donnons au moins un mois à l'ensemble des personnes concernées pour demander le retrait de certaines données, des bases ou des sites qui seront utilisés pour l'entraînement des IA.

On l'a bien fait pour le droit à l'oubli, on le fait depuis 1978 pour n'importe quelle donnée personnelle. Autorisons-nous cette possibilité de redonner à chaque personne une certaine forme de souveraineté sur les données qu'elle partage.

Aujourd'hui, on ne l'a pas fait et je le regrette. Peut-être qu'on nous démontrera que ce n'est pas possible, que ce n'est pas comme ça que ça marche. Mais cette question mériterait au moins un débat.

Alors, à côté des Cafés IA, est-ce que vous avez imaginé d'autres formes de consultation des citoyens par rapport aux questions du numérique et de l'IA ? Est-ce qu'il y a d'autres choses qui surgissent du terrain ?

Oui, tout à fait. D'abord, pour reprendre le cas de l'environnement professionnel, il y a le dialogue social technologique¹¹, une obligation légale, qui n'est que très rarement, pour ne pas dire jamais, respectée. Le dialogue social technologique doit avoir lieu lorsqu'il y a un changement logiciel technologique dans la vie du travailleur. C'est un moment de discussion, de dialogue et, potentiellement, de décision. Malheureusement, aujourd'hui, on n'a même pas la transparence sur les outils qui sont implantés.

La deuxième chose qui émerge, sur un tout autre plan, ce sont des formes

de consultations publiques qui peuvent avoir lieu à l'initiative, notamment d'une organisation qui s'appelle make.org, qui en a fait une à l'aube du sommet sur l'IA et qui en fera une à l'horizon de Vivatech. Ces consultations sont intéressantes parce que très faciles d'accès, très massives et qu'elles permettent d'apprécier le degré de vigilance constructive, pour reprendre les termes du fondateur de make.org, sur la question de la relation à l'IA. Autant d'éléments qui mériteraient d'être beaucoup plus employés dans la construction des politiques publiques.

Ce que l'on apprend dans ces consultations-là est très intéressant et montre une maturité de la relation à la technologie. On note, par exemple, un consensus dans la population sur l'idée d'un apprentissage tout au long de la vie des outils de l'IA. Mais sur la question de l'intégration de l'IA à l'école, par exemple, là, il n'y a plus de consensus, on n'est pas sûr que ça vaille le coup.

Le consensus existe aussi sur l'utilisation de l'IA dans le domaine de la recherche scientifique en matière de santé. Par contre, il est beaucoup plus difficile concernant l'usage de l'IA dans les services publics. Ce sont des degrés de nuances qui sont quand même assez fins. C'est une position intéressante qui émane de consultations massives. Donc, écoutons cette voie-là, à défaut d'en avoir d'autres, malheureusement.

Et ensuite, dernier exemple, il y a les conventions citoyennes qui ont été créées, à Paris, Rennes, Montpellier. Montpellier par exemple a dit, nous faisons un moratoire sur l'IA tant qu'on n'a pas fait de consultation citoyenne. Ces conventions permettent aux citoyens de s'exprimer sur la question de l'IA.

Un grand merci pour cette analyse passionnante.

Retour au sommaire

¹¹ <https://formation.lefebvre-dalloz.fr/actualite/intelligence-artificielle-le-dialogue-social-technologique-reste-balbutiant>

POUR UNE POLITIQUE PUBLIQUE DU NUMÉRIQUE RESPONSABLE

**ENTRETIEN AVEC PIERRE JANNIN, ADJOINT À LA MAIRE DE RENNES,
DÉLÉGUÉ AU NUMÉRIQUE ET À L'INNOVATION**

PAR DOMINIQUE BÉNARD

Bonjour Monsieur Jannin, pouvez-vous vous présenter à nos lecteurs ?

Je suis conseiller municipal à la Ville de Rennes, délégué au numérique et à l'innovation depuis 2020.

C'est mon premier mandat. Professionnellement, je suis directeur de recherche à l'INSERM (Institut national de la santé et de la recherche médicale), et je travaille depuis plus de 20 ans sur l'utilisation de l'image et du numérique en chirurgie.

Comment arrive-t-on à créer un conseil citoyen du numérique ?

Vous vous souvenez qu'en 2020, on parlait beaucoup de l'utilisation de la 5G pour le téléphone mobile. Une ouverture des fréquences était apparue du jour au lendemain après une décision prise au niveau national sans aucune consultation. Dans ce contexte, une partie de la population exprimait des craintes et une grande méfiance vis-à-vis de cette technologie. Des actions, parfois

violentes, avaient commencé à surgir pour la rejeter. La maire de Rennes, Madame Nathalie Appéré, a décidé d'ouvrir un débat qui rentrait dans mes attributions. Je m'étais dit que j'allais me documenter, lire, réfléchir et prendre le temps d'organiser les choses. Mais en septembre 2020, Nathalie Appéré m'a dit : *"Pierre, il faut absolument que tu lances le débat dès maintenant. Si on attend trop, ça n'aura plus beaucoup de sens."*

Donc je me suis retrouvé devant la nécessité de découvrir les outils et les conditions de la démocratie participative ou de la concertation citoyenne. Un sujet assez central dans nos sociétés démocratiques au moment où l'on voit comment la démocratie représentative est mise en cause. On ne sait plus trop qui représente qui ni comment. En disant cela, je pense précisément à ce qui se passe actuellement aux États-Unis.



Pierre JANNIN



Le Conseil Citoyen du Numérique Responsable de Rennes

**LA PREMIÈRE QUESTION A
ÉTÉ LA SUIVANTE :
« COMMENT ENVISAGEZ-VOUS
L'IMPACT DE L'INTELLIGENCE
ARTIFICIELLE SUR VOS VIES
DE RENNAIS ? »**

Il s'agissait de mener à bien un débat public sur la technologie de la 5G et son implémentation sur le territoire de Rennes. Je me suis entouré d'une quinzaine d'experts de la participation citoyenne, et j'ai demandé à la Commission nationale du débat public (CNDP) un encadrement méthodologique pour ce débat local. La consultation a duré quatre mois. Elle a mobilisé 20 citoyens et citoyennes de la ville, tiré-es au sort, en respectant les caractéristiques du territoire (parité, catégorie socio-professionnelle, quartiers, etc.). Nous avons suivi une démarche assez classique comprenant trois étapes : un diagnostic, l'identification des enjeux puis des recommandations. Les citoyen·nes ont étendu leur réflexion au numérique en général.

En parallèle, un groupe d'élus·es a travaillé sur cette même problématique pour renforcer leur investissement dans cette démarche participative. Ce travail a débouché sur la publication d'un document comprenant 54 propositions dont la pérennisation d'une instance

citoyenne sur le numérique et la nécessité d'une stratégie du numérique responsable. En conseil municipal et métropolitain, nous avons voté cette stratégie soulignant que le numérique soulève des enjeux politiques, sociaux, écologiques, démocratiques, économiques et d'usage, éthiques, de gouvernance, de qualité de service public. Son orientation nécessite une politique publique sectorielle à part entière, distincte des simples problèmes techniques d'ordinateurs, de connexions, de logiciels.

Quels ont été les sujets et les méthodes de travail choisis par le conseil ?

Les citoyens se sont déterminés par eux-mêmes, sans que j'intervienne. Je me tenais vraiment en arrière en m'interdisant de donner mon avis parce que je ne voulais pas biaiser la démarche.

Ils ont décidé de prolonger le débat de la 5G vers le numérique en général, exprimant le souhait d'être consultés et de pouvoir contribuer à l'élaboration d'une politique

municipale sur les questions du numérique. Madame la Maire a donc créé le Conseil Citoyen du Numérique Responsable (CCNR) en mai 2021. Le CCNR a commencé vraiment à travailler en décembre 2021.

En fait, nous avons tout inventé parce qu'il y a peu d'instance de concertation citoyenne pérenne dans les collectivités. La plupart de celles qui sont créées travaillent sur un sujet déterminé pendant quelques semaines ou au mieux quelques mois. Une fois que l'instance a terminé ses travaux et publié son rapport, la démarche participative s'arrête.

Nous nous sommes fait aider par des méthodologues de la participation citoyenne et les membres du conseil ont été tirés au sort de manière à respecter la variabilité de la population rennaise en termes de genre, d'âge, de quartier d'habitation, de catégories socioprofessionnelles et de distance par rapport au numérique. Bref, nous nous sommes efforcés de représenter la variété du territoire. Au départ, les membres étaient une vingtaine, aujourd'hui, ils sont une trentaine.

Un deuxième élément fort de cette instance, c'est qu'elle dispose d'un pouvoir d'auto-saisine. Elle peut définir par elle-même les thématiques sur lesquelles elle veut travailler. Le CCNR dispose vraiment d'une grande liberté. Nous avons suivi le cycle traditionnel de montée en compétences : le CCNR travaille sur un sujet déterminé, il rédige un rapport. Le rapport est remis à la ville de Rennes pour être instruit par les services et les élus. Les services et les élus répondent au conseil en disant, par exemple, cette proposition est très intéressante, nous l'avons déjà réalisée ; celle-ci est en cours de réalisation ; cette autre n'a pas encore été réalisée mais va être réalisée ou semble difficile pour telle et telle raison.

Les membres du conseil sont bénévoles, ils ne reçoivent aucune rémunération et les deux premières

années, le rythme de travail a été fou. Il y avait quasiment une réunion toutes les deux semaines. Depuis, nous avons adopté un rythme mensuel. C'est déjà pas mal avec des réunions de deux heures à deux heures et demie. Parfois sur une journée entière le samedi.

Le premier sujet sur lequel le conseil a décidé de travailler a été l'étude de l'impact de la dématérialisation des démarches administratives. Un gros travail a été accompli pendant six mois avec le concours d'experts.

Le deuxième sujet a été l'impact du numérique sur la santé mentale des jeunes. Le CCNR a travaillé sur cette question pendant 6 à 8 mois de travail et a produit aussi un rapport.

Les propositions du CCNR entraînent donc des réponses des élus et des services. Le retour aux citoyens et citoyennes sur le rapport n°1 a été fait cette année. Le retour sur le rapport n° 2 va être fait à la rentrée. Et après la fin de l'année, on va faire le retour sur ce qu'on appelle le rapport n°3.

Comment avez-vous abordé le sujet de l'intelligence artificielle ?

C'est le sujet du rapport n°3. C'est un peu mon domaine. J'ai perçu assez tôt le potentiel de l'intelligence artificielle dans tous les domaines et notamment dans les services publics. Et donc je suis venu un jour au CCNR et j'ai dit *"Je vous propose de travailler sur l'intelligence artificielle, est-ce que ça vous intéresse ou pas ?"*. La réponse a été *"Oui"* à l'unanimité. C'était en février 2023. Peu de gens parlaient du sujet à l'époque. Nous avons été assez précurseurs.

La première question posée a été la suivante : *"Comment envisagez-vous l'impact de l'intelligence artificielle sur vos vies de Rennais ?"* Pour y répondre, le conseil s'est d'abord efforcé de comprendre les gros enjeux de l'intelligence artificielle et a identifié trois grands axes :

- L'impact de l'intelligence artificielle sur nos libertés, c'est-à-dire sur les prises de décisions



Session du CCNR de Rennes

ON ENTEND DIRE SOUVENT QUE LES CITOYENS NE COMPRENNENT RIEN, C'EST TOUT À FAIT FAUX. LE PROBLÈME C'EST QU'ON NE LES ÉCOUTE PAS, QU'ON NE LEUR DONNE PAS LA POSSIBILITÉ DE S'INFORMER NI LA RESPONSABILITÉ DE S'EXPRIMER.

politiques ou la justice et la sécurité.

- Quel est l'intérêt de l'intelligence artificielle ? Comment peut-on mettre l'intelligence artificielle au service du territoire ? Comment l'utiliser pour mieux gérer les énergies, les mobilités, la santé, les déchets ?
- Quels sont les risques de l'utilisation de l'intelligence artificielle sur l'emploi et la culture ?

Un rapport a été publié en février 2024. Nous l'avons appelé rapport de vigilance. Il ne comportait pas de recommandations, parce que l'IA n'était pas encore utilisée dans les services, mais il demandait à la collectivité d'être vigilante sur certains points.

On y retrouve tout le contenu existant aujourd'hui dans tous les rapports sur l'intelligence artificielle.

Cela souligne l'intelligence collective qui s'exprime dans tous les contextes de participation citoyenne. On entend dire souvent que les citoyens ne comprennent rien, c'est tout à fait faux. Le problème c'est qu'on ne les écoute pas, qu'on ne leur donne pas la possibilité de s'informer ni la responsabilité de s'exprimer. Les citoyens tirés au sort, quels qu'ils soient, comprennent les enjeux, prennent la responsabilité de porter la voix du citoyen lambda et travaillent sérieusement pour s'informer et produire des recommandations éclairées.

J'ai suivi toutes les réunions depuis 2021. Je ne dis rien, j'écoute. Il y a des moments mémorables. A l'une des premières réunions, on fait un tour de table, une personne a pris la parole *"Moi, je suis concierge dans un quartier. Quand j'ai reçu le courrier qui me disait : vous êtes tirés au sort pour participer au conseil citoyen du numérique responsable, j'ai eu le sentiment d'avoir emporté le gros lot alors que je n'ai jamais rien gagné au Loto. J'ai donc décidé d'y aller."* Il est resté 3 ans membre du conseil avec des apports toujours très pertinents.

Ce type de démarches citoyennes, ce sont des histoires humaines aussi, des montées en compétence. Un exemple encore : Lucie, une chômeuse, qui est toujours au conseil citoyen depuis deux ans et demi, trois ans m'a dit *"Moi je n'ai jamais voté parce que j'avais aucune confiance dans les politiques mais alors là, avec le conseil citoyen, ça me donne envie de me réengager pour le territoire et de refaire de la politique, la politique, ça doit être ça !"*

Vous voyez : rien qu'avec ces deux remarques-là, nous avons gagné.

Que voyez-vous ressortir globalement du travail fait par les citoyens et comment l'avez-vous pris en compte ?

On voit ressortir globalement des choses comme l'importance des données et l'importance de la gouvernance des données puisque les données sont cruciales dans l'intelligence artificielle. Il est crucial d'avoir une structure de gouvernance qui garantisse la qualité de la

donnée, la représentativité de la donnée, la mise à disposition de la donnée de manière non commerciale ou au moins de manière contrôlée et choisie. Et aussi tout un ensemble de conclusions qui soulignent l'importance de former tout le monde à l'intelligence artificielle, non pas pour que tout le monde l'utilise mais surtout pour que tout le monde soit conscient des biais de l'intelligence artificielle. Il y a des usages positifs mais il y a aussi des inconvénients et des risques.

Dans le rapport, les citoyens expriment des inquiétudes mais ils les nuancent parce qu'ils ont confiance dans le territoire. Ils font confiance à leur ville, à leur métropole en disant : *"On sait que vous allez mettre en place des structures de gouvernance."* On constate ici la valeur accordée à la collectivité locale et la confiance que les gens lui portent. Tous les sondages montrent que plus les élus sont locaux, plus les citoyens ont confiance en eux. Cela nous donne à la fois une légitimité et une responsabilité : nous, les élus locaux, nous devons respecter cette confiance, être les garants de cette confiance.

Une des conclusions aussi, concernant les recommandations, les points de vigilance émis par les citoyens, c'est qu'ils ont compris ce que peut faire une ville, une métropole, parce qu'elle possède à la fois une responsabilité et des moyens. Mais il y a plein d'autres choses qu'elle ne peut pas faire parce qu'elle n'en a pas la responsabilité et que les moyens sont de plus en plus réduits. Alors, à chaque fois, on explique aux gens : *"Ce n'est pas de notre ressort, ce n'est pas notre responsabilité"*, mais cela ne nous empêche pas d'aller porter des messages. Si la responsabilité, concernant telle ou telle mesure, est régionale, et bien on va aller voir nos collègues du conseil régional ; si la responsabilité est nationale, et bien on va, par exemple, témoigner dans une commission du Sénat ou contacter un ministre. Si la

responsabilité est européenne, et bien on va aller à Bruxelles. Je tiens au courant régulièrement les membres du conseil des démarches que nous avons faites. L'année dernière, nous sommes allés à Bruxelles et nous avons rencontré les responsables des commissions qui ont généré le Digital Service Act et le Digital Market Act. Nous sommes restés plusieurs jours à Bruxelles pour porter la voix du territoire avec les collègues des grandes villes de France en disant voilà ce que pensent les citoyens et les citoyennes.

Dans chaque démarche participative, il y a une question qui revient toujours :

"Quelles sont les marges de manœuvre ?" Lorsque nous avons commencé à travailler sur le 5G, j'avais invité une professeure en droit d'une université de Paris et elle nous disait : *"Je suis désolé mes pauvres amis mais sur la 5G vous n'avez aucune marge de manœuvre... Vous ne pouvez rien faire"*. Elle a *"cassé l'ambiance"* de tout le conseil citoyen et il a fallu que j'intervienne pour dire *"Ne vous inquiétez pas, on va les trouver nos marges de manœuvre."* En fait, il n'y a aucun domaine dans lequel, localement, on n'a pas de marge de manœuvre. On peut intervenir partout de différentes façons, avec différents outils. Les gens le savent et, parfois, ils nous disent *"Et bien sur ce sujet là, allez juste porter notre voix"*.

Je donne un exemple à propos de l'intelligence artificielle : nous avons été invités au sommet international de l'IA à Paris en février dernier et c'est Lucie qui a témoigné à la banque des territoires sur ce que les citoyens et les citoyennes pensent de l'intelligence artificielle. Pour moi c'était un moment symbolique : on continue à faire témoigner les membres du conseil pour les faire progresser en capacité en tant que porte-paroles du conseil. Vous m'avez sollicité pour cet interview mais il y a beaucoup d'interviews qui sont faites par les membres du conseil ; ils répondent eux-même aux questions telles que celles que vous m'avez posées aujourd'hui.



LE CONSEIL A PROPOSÉ D'ORGANISER UNE CAMPAGNE DE SENSIBILISATION SUR LES ENJEUX ÉCOLOGIQUES DU NUMÉRIQUE

Comment le Conseil Citoyen du Numérique envisage-t-il de poursuivre ses travaux ?

Les membres du Conseil ont choisi de revenir sur un sujet qui avait été abordé au tout début avec la 5G, celui de l'impact écologique du numérique et de l'intelligence artificielle générative.

Ils se posent la question de savoir ce que nous pouvons faire pour limiter cet impact. Les projections faites il y a trois ans, qui étaient catastrophiques, sont atteintes aujourd'hui.

Le Conseil a proposé d'organiser une campagne de sensibilisation sur les enjeux écologiques du numérique. On leur a proposé d'écrire les slogans. On leur a montré les résultats, il y a un mois. Ils ont modifié tout ce qu'ils voulaient modifier et la campagne sortira en septembre. Voilà une action concrète.

Pour conclure, je voudrais ajouter une information. Je fais partie du conseil d'administration de l'Association des élus du numérique de France, Les Interconnectés. J'ai présenté l'action que nous avons menée à Rennes à mes collègues des autres villes et nous avons décidé, il y a moins d'un an, de lancer une initiative nationale

pour proposer à toutes les villes de France d'organiser leur propre concertation locale sur l'IA. En conséquence, une trentaine de villes françaises ont lancé les concertations territoriales de l'IA. On a récupéré toutes les conclusions et on est en train de rédiger le rapport final. On peut donc dire aux membres de notre Conseil : voici encore une conséquence du travail que vous avez entrepris.

Un autre projet très intéressant va démarrer à la rentrée sur la problématique de l'inclusion numérique. Et ça va être fait dans un quartier prioritaire de la ville, dans lequel il y a quelques soucis de sécurité comme dans d'autres quartiers sensibles. Nous avons décidé d'engager sur ce sujet tout un réseau d'associations afin d'expérimenter comment recréer du lien entre les jeunes et leur ouvrir une place dans la cité. C'est vraiment une approche coopérative au sens où tous les acteurs du territoire sont appelés à travailler ensemble : les associations, les CCAS, les collèges, etc. On va tous ensemble essayer de résoudre un problème dans une démarche non pas descendante mais qui s'appuie sur les initiatives de base avec le concours de la collectivité. Nous avons réussi à obtenir un financement de l'Union Européenne et nous avons donc un budget pour travailler sur trois ans sur ce projet avec les associations sur le territoire.

Et bien je ne doute pas que nous aurons l'occasion de reparler de ce projet et je vous remercie infiniment pour votre contribution à ce numéro.

[Retour au sommaire](#)

LA CONVENTION CITOYENNE DE LA MÉTROPOLE DE MONTPELLIER SUR L'INTELLIGENCE ARTIFICIELLE

PAR DANTE MONFERRER

A l'instar de nombreuses initiatives prises au plan local, la métropole de Montpellier a souhaité associer des citoyens aux réflexions sur les usages de l'IA par l'agglomération.

Aujourd'hui le débat est essentiellement porté par les médias, en particulier avec l'émergence des IA génératives, suscitant interrogations et parfois défiance.

La sphère politique s'est également emparée de la question, ainsi le G7 d'octobre 2023 a approuvé la mise en place des principes directeurs et d'un code de conduite sur l'IA¹.

Il paraissait important que les citoyens ne soient pas absents de ce débat, d'autant que le rapport sensible entretenu avec la technologie, et notamment l'IA, en fait une porte d'entrée intéressante pour échanger et impliquer un maximum de personnes sur le numérique dans son ensemble.

¹ <https://www.economie.gouv.fr/daj/lettre-de-la-daj-accord-du-g7-sur-des-principes-directeurs-et-un-code-de-conduite-en-matiere>

LES ACTEURS ET LE PROCESSUS

En 2023 la collectivité de Montpellier avait décidé d'interdire l'usage de ChatGPT dans ses services. Il paraissait important de donner une suite politique à cette décision et de laisser la parole aux habitants à travers une Convention citoyenne qui devait réfléchir aux conditions de mise en œuvre de l'IA et aux usages à privilégier.

Michaël Delafosse, Maire de Montpellier et Président de Montpellier Méditerranée Métropole pose ainsi les enjeux. « *Notre approche sur l'IA est unique en France, elle se fait en partage avec les citoyens sur une Convention totalement inédite. L'IA prend déjà une place centrale dans notre quotidien et soulève de nombreuses questions sociétales. La méthode de Montpellier, c'est celle du faire ensemble, d'une dynamique de partage entre les acteurs de la recherche, les entreprises et les collectivités territoriales. Grâce à ces forces vives, notre territoire s'inscrit*



Michaël Delafosse



La convention citoyenne de la métropole de Montpellier sur l'intelligence artificielle

**CETTE CONVENTION
REPRÉSENTE UNE
OPPORTUNITÉ INÉDITE DE
CROISER LES REGARDS SUR
L'IA ET DE FOURNIR DES
PRÉCONISATIONS UTILES
POUR LES HABITANTS**

comme une place forte de l'IA en France. »²

Séverine Saint-Martin, adjointe au maire déléguée au renouveau démocratique et à l'innovation sociale³ et Manu Reynaud adjoint au maire, délégué à la ville apaisée, respirable et numérique, ont été plus particulièrement en charge d'impulser le processus, accompagnés par des spécialistes des panels et par un cabinet spécialisé en transformation numérique.

La Métropole a désigné 40 personnes représentatives des 500 000 habitants de la métropole (âge, catégories socio-professionnelles, quartiers, etc).

Les participants ont ensuite été formés à travers des présentations de spécialistes, des ateliers, des ressources documentaires. L'objectif étant de leur faire acquérir les connaissances de base sur l'IA, ses applications, ses enjeux éthiques,

² Avis rendu par la Convention citoyenne : Quelle intelligence artificielle au service des habitants et du territoire ? De novembre 2023 à février 2024.

³ Dirige également une SCOP (société coopérative) dans le domaine de la formation professionnelle.

sociaux, économiques et environnementaux.

En parallèle un comité d'experts était mis en place pour accompagner cette Convention. Cédric Villani, membre de ce comité, revient sur son rôle. « *L'IA est certes un sujet complexe et interdisciplinaire mais dans lequel les problèmes techniques sont minoritaires. En revanche, les enjeux majeurs sont humains, philosophiques et éthiques. Cette Convention représente une opportunité inédite de croiser les regards sur l'IA et de fournir des préconisations utiles pour les habitants. Grâce au travail des membres de cette Convention, la Métropole sera mieux préparée à affronter l'avenir, et pourra progressivement s'adapter à l'évolution de l'IA sur son territoire* »⁴

De novembre 2023 à février 2024 les participants de la Convention ont été mobilisés, en particulier au cours de trois week-ends de travail, où ils ont pu bénéficier d'apports d'information, d'auditions et de témoignages de spécialistes sur la base desquels ils ont pu largement débattre de l'utilisation de l'IA au niveau de la collectivité. Ils avaient pour objectif de formuler des recommandations, et de fixer "les conditions pour le déploiement d'IA publiques de confiance au service du territoire et des habitants" sous la forme d'un avis destiné aux élus de Montpellier.

Les huit préconisations de l'avis rendu par la Convention citoyenne, ont été votées à l'unanimité des 40 membres.

L'AVIS RENDU PAR LA CONVENTION CITOYENNE.

Les 40 participants ont rendu un avis, remis à la mi-mars au maire de Montpellier, qui formule un certain nombre de recommandations qui se déclinent en huit grands items :

⁴ Ibid note n°3.

LES IA DOIVENT ÊTRE UTILES

Dans le cadre des actions de service public, en privilégiant une analyse préalable et rigoureuse de leurs coûts/avantages aux différents plans (sociaux, sociétaux, économiques, environnementaux, ...).

Un ciblage cohérent des thématiques et des projets pouvant mobiliser des IA doit être effectué, en particulier au plan environnemental (gestion de l'eau, lutte contre le gaspillage, traitement des déchets...) et au plan des formalités administratives (passeports, permis de construire, prévention sanitaire, ...).

Certains sujets, notamment en matière de police, font débat (IA associée à des caméras, croisement de données, ...) et doivent être mis en débat à travers les différents mécanismes de contrôle et de discussion.

L'IMPACT ENVIRONNEMENTAL DES IA DOIT ÊTRE MAÎTRISÉ

Cet impact est important (consommation énergétique, émissions carbone). La frugalité doit être de mise et la balance des impacts positifs et négatifs doit être réalisée en particulier pour les politiques publiques utilisant l'IA (énergie, eau, environnement, santé...).

Il faut recourir aux IA les plus sobres au plan environnemental et limiter la consommation des ressources (recyclage, équipements de seconde main, ...). Les installations utilisant des énergies renouvelables et limitant la circulation des données (data centers locaux), doivent être privilégiées. Il est également recommandé de mutualiser les équipements avec d'autres collectivités. Enfin il faut limiter l'utilisation des données à celles essentielles au fonctionnement des algorithmes.

LA PROTECTION DE LA VIE PRIVÉE DOIT ÊTRE GARANTIE

La maîtrise et le contrôle des données utilisées par l'IA est difficile et des garanties doivent être apportées en matière de protection

de la vie privée. Aider les citoyens à garder le contrôle sur leurs données personnelles est essentiel.

L'utilisation de l'IA doit s'inscrire dans la réglementation existante⁵. Par ailleurs des actions d'information doivent être menées auprès du grand public et les citoyens doivent être accompagnés en cas de litige. Les données doivent être hébergées si possible localement. La cyber sécurité doit être permanente pour assurer la sécurité des données des habitants.

Il faut privilégier le recours à des données anonymisées et veiller à ce que les IA ne puissent pas croiser des données sensibles (santé, situation financière...).

Tout projet d'IA, devant mobiliser des données personnelles, doit donner lieu à une étude d'impact préalable sur les risques potentiels pour les personnes concernées.

DES CONTRÔLES NEUTRES ET INDÉPENDANTS DOIVENT ÊTRE MIS EN PLACE

La Convention préconise l'instauration d'un comité d'éthique neutre et indépendant (avis en amont des projets d'IA, garanties en matière de protection de la vie privée, mesure des empreintes environnementales, contrôle des éventuels biais et erreurs des IA, ...).

Composé d'experts neutres et indépendants, il pourra mobiliser des compétences pour évaluer les projets, contrôler les IA et la fiabilité des données. Des citoyens seront associés au comité pour assurer le suivi de ses travaux.

Le recours à un système de labellisation éthique pour des projets d'IA sera privilégié. Il devra reposer sur des valeurs universelles (traçabilité, bienfaisance, loyauté, qualité). Le label pourra valoriser le « made in Montpellier ».

Des consultations citoyennes pourraient être menées sur les projets IA sensibles.

⁵ Règlement Général pour la Protection des Données (RGPD), contrôle de la CNIL.



Michaël Delafosse avec le mathématicien Cédric Villani

IL FAUT ASSURER LA TRANSPARENCE

La transparence est essentielle. Pour ce faire, il faudra informer les usagers en cas d'utilisation d'une IA, (nature et origine des algorithmes, moyens d'information et de recours en cas d'usage de données personnelles).

L'utilisation des « IA open source » sera privilégiée, car elles affichent des règles et des codes publics, permettant d'évaluer le fonctionnement des algorithmes, le respect des règles éthiques, les données utilisées,

La Métropole vérifiera la source et la qualité des données d'apprentissage des IA (erreurs et biais, discriminations femme/homme ou en fonction des origines, ...). Le contrôle des données doit répondre à des normes éthiques garantissant l'absence de discrimination et l'égalité entre les citoyens.

IL EST IMPORTANT DE MAÎTRISER LE RECOURS À L'IA DE FAÇON AUTONOME

Face à la complexité des systèmes d'IA, le service public doit disposer des ressources et des compétences pour maîtriser leur utilisation (contrôler les données

d'entraînement des IA, adaptation au contexte et à la finalité voulue, détection et correction d'éventuels biais discriminants).

La Convention recommande la création d'un data center dédié aux données de la Métropole, ou l'utilisation de data centers locaux, afin de conserver la maîtrise publique sur les données. L'accès aux données sensibles doit être réservé aux seules personnes et organismes concernés.

DES IA EN SOUTIEN DE L'ACTION DU SERVICE PUBLIC

Déployer les IA suppose d'informer les agents, d'instaurer des périodes de test et de solliciter les retours d'expérience. L'IA peut avoir des effets négatifs sur l'exercice des métiers et suppose une vigilance accrue (souffrance au travail, perte de sens, pression professionnelle, déresponsabilisation des agents).

La présence humaine est centrale, les citoyens ne doivent pas être laissés seuls face à la machine et l'échange avec des agents doit toujours demeurer possible.

L'IA doit être vue comme une aide à la décision, le dernier mot allant à l'humain. Face à une possible dépendance à l'IA, il faut veiller au maintien des compétences pour pouvoir agir sans elle.

L'arrivée de l'IA suppose une transformation des métiers. Les agents concernés doivent être accompagnés à travers un dispositif RH dédié et des formations.

L'IA peut permettre d'alléger les tâches répétitives et de favoriser un repositionnement des agents.

RENDRE ACCESSIBLE L'IA AU PLUS GRAND NOMBRE

La Convention recommande d'organiser ou soutenir des dispositifs de formation à destination des habitants et permettant de les sensibiliser aux bénéfices et risques de l'IA. Ces dispositifs doivent aider les citoyens à mieux comprendre et maîtriser leurs droits.

Il faut également programmer des campagnes d'information pour s'adresser à toutes et tous et prendre en compte la fracture numérique

La Convention demande à la Métropole de relayer auprès de différents acteurs (formation professionnelle, Éducation nationale, ...), un message sur l'importance de former les citoyens et les salariés à l'IA.

LES SUITES DE LA CONVENTION CITOYENNE DE MONTPELLIER

Suite aux propositions de la Convention, la phase de traduction opérationnelle des recommandations dans la stratégie IA et data de la Métropole de Montpellier⁶, (adoptée en octobre 2024) a démarré. La Métropole se confronte maintenant à la réalité technique, administrative et politique de traduction des préconisations.

L'évaluation de l'IA au fur et à mesure de son déploiement suppose de construire une grille d'analyse de son impact social et environnemental. Des contacts sont pris avec Nantes, qui a construit une boussole d'analyse, et Strasbourg, sur la base d'une IA sobre et responsable. Des indicateurs ont été testés avec de petits groupes d'utilisateurs pour évaluer l'empreinte environnementale de l'IA.

La Métropole souhaite mettre en place des « Halles de l'IA » pour débattre de ses liens avec la santé, les relations sociales, l'éducation... au travers d'une association qui va réunir l'ensemble des partenaires et construire un instrument de gouvernance commun.

Les membres de la Convention citoyenne souhaitent continuer à s'impliquer et à pouvoir donner des avis sur le déploiement de l'IA. Ils ont proposé de mettre en place un comité du numérique et de l'IA avec des acteurs du territoire (chercheurs,

universitaires, chefs d'entreprises...). Ce comité pourra être consulté et évaluer les projets et expérimentations impliquant l'IA.

La Convention a insisté sur la coopération plutôt que la compétition, ainsi que sur le décloisonnement entre les différents territoires. La Métropole a noué des liens avec d'autres collectivités et des associations spécialisées. L'État et l'UE ont été approchés pour construire un cadre éthique basé sur ce qu'ont exprimé les citoyens.

L'idée d'un Data Center local, mutualisé avec l'université, le CHU et des entreprises a émergé. C'est toute une chaîne d'acteurs et de choix politiques qu'il faut activer et animer.

Les problématiques de sécurité et leur impact sur les libertés publiques ont été au cœur des débats, notamment sur la vidéosurveillance assistée par un traitement automatisé. La Métropole a lancé une réflexion sur ce sujet.

[Retour au sommaire](#)

RÉFÉRENCES

DémocratieS : le média critique de la participation et de la transparence. *IA et participation.*

Entretien avec Manu Reynaud adjoint à la ville de Montpellier et conseiller métropolitain en charge du numérique et Séverine Saint Martin adjointe à la ville de Montpellier en charge du renouvellement démocratique et de l'innovation sociale. 01/01/2025.

Convention citoyenne de Montpellier : Quelle intelligence artificielle au service des habitants et du territoire ?

Avis rendu par la Convention citoyenne. de novembre 2023 à février 2024

Localtis. Citoyenneté et services au public.

Numérique : une convention citoyenne incite Montpellier à se doter d'une IA « utile et sous contrôle » par Olivier Devillers 22/03/2024.

Conseil National du Numérique. Retour sur la convention citoyenne sur l'IA organisée à Montpellier

Paroles du 06 juin 2024 : Interview de Manu Reynaud adjoint au maire de Montpellier.

⁶ Stratégie IA votée suite à la Convention citoyenne : Montpellier Méditerranée Métropole. Stratégie de la donnée et de l'intelligence artificielle. 2024. <https://montpellier3m.fr/data-ia>



VERA : UNE APPLICATION POUR LUTTER CONTRE LA DÉSINFORMATION

ENTRETIEN AVEC FLORIAN GAUTHIER, INITIATEUR DU PROJET

PAR DOMINIQUE BÉNARD

Bonjour M. Gauthier. En préparant ce numéro sur l'intelligence artificielle, j'ai découvert, un peu par hasard, l'application Vera, que vous avez développée. Elle correspond tout à fait à la définition que nous avons d'une initiative citoyenne qui essaye d'apporter des réponses aux défis de la société. Je suis très intéressé par votre expérience et très heureux de vous rencontrer. Pouvez-vous, d'abord, vous présenter à nos lecteurs ?

Bien sûr, je m'appelle Florent Gauthier, j'ai une formation initiale d'ingénieur en Data Science, mais ça remonte à 2012 et je n'ai exercé cette profession que pendant 3 ans.

Aujourd'hui, ce qui définit le mieux mon activité, c'est le terme d' "entrepreneur". Je pense que je suis un entrepreneur orienté vers l'intérêt général. Depuis dix ans, toutes mes initiatives ont pour point commun d'essayer de voir comment la technologie, utilisée à bon escient, peut servir le bien commun. J'ai

travaillé longtemps dans une ONG, qui s'appelle Bayes Impact. Son objectif était d'utiliser la technologie pour aider les chercheurs d'emploi à trouver un métier en leur donnant plus d'informations sur les métiers qui recrutent ou pour aider les jeunes à changer de trajectoire professionnelle.

Et plus tard encore, j'ai cofondé une communauté de citoyens et de citoyennes qui s'appelait La Réserve. Elle rassemblait des personnes possédant des compétences diverses et variées, par exemple en design, en développement, en marketing. Le but, c'était de les mobiliser en période de crise, par exemple pendant le Covid, afin de créer un outil technologique qui pourrait aider l'administration publique à traverser la crise. Nous avons créé pas mal de choses pendant la sécheresse, pendant la crise énergétique, etc.

Et c'était génial, mais c'était aussi très difficile de gérer plusieurs équipes en même temps. Et ce n'était pas



Florian Gauthier

**CETTE FACILITÉ À
DÉSINFORMER LES GENS,
MÊME VIA LES RÉSEAUX
SOCIAUX, EST UN PROBLÈME
GIGANTESQUE QUI TOUCHE À
TOUS LES AUTRES
PROBLÈMES.**

toujours simple non plus de travailler avec l'administration publique. Et du coup, finalement, un an plus tard, nous avons créé LaRéponse.tech

LaRéponse.tech, c'est l'ONG qui est derrière Vera. L'idée est toujours de mobiliser des citoyennes et des citoyens qui veulent agir, mais qui ne savent pas trop comment. Mais cette fois-ci, pour un seul et même projet, la lutte contre la désinformation.

LaRéponse.tech, c'est une quinzaine de personnes, designers, développeurs, spécialistes de la croissance, ou encore responsables produit, c'est-à-dire de la conception de produits digitaux, c'est le métier que je fais. Toutes et tous sont des experts seniors, ils ou elles possèdent beaucoup d'expérience et sont donc capables d'être efficaces en dégageant seulement un ou deux jours par semaine, à titre bénévole, ce qui suffit largement à avoir un impact fort, puisqu'on maîtrise bien les sujets sur lesquels on travaille. Toutes les personnes, qui font partie de LaRéponse.tech, possèdent une expérience significative dans le domaine des projets d'intérêt général. Nous avons déjà travaillé ensemble dans les projets précédents que j'ai cités. Ce sont des gens qui ont un fort intérêt à utiliser la technologie pour le bien commun.

Quand avez-vous créé LaRéponse.tech ?

En même temps que Vera, il y a 6 mois ou 7 mois. En fait, nous avons l'idée de Vera et nous avons vu l'information concernant un concours ouvert aux ONG pour l'invention d'une application IA de lutte contre la désinformation. Il y avait d'abord un processus de sélection des 10 meilleures ONG et ensuite une compétition d'une journée entière avec une présentation finale devant un jury. Je me suis donc précipité, j'ai fait les statuts de l'ONG en 10 minutes sur ChatGPT et puis nous avons postulé et gagné le concours.

Un point important à noter, c'est que si on s'appelle LaRéponse.tech et non pas la Solution.tech, c'est qu'on ne croit pas du tout au solutionnisme

technologique. On pense qu'on peut apporter des réponses à des problèmes, mais pas des solutions définitives. La désinformation, c'est un problème systémique qui a besoin d'être géré par les gouvernements, les pouvoirs publics, mais aussi par la régulation des plateformes elles-mêmes, et peut-être aussi par l'éducation. Des outils technologiques, comme Vera, peuvent aider à apporter une partie de la réponse au problème.

Alors, comme premier défi social, vous avez relevé la désinformation. Pourquoi ce choix ? Est-ce qu'aujourd'hui, ça pose un problème vraiment grave ?

Eh bien, écoutez, c'est un événement qui nous a conduits à développer l'application Vera. Au cours des dernières législatives en France, nous avons été surpris et choqués de la facilité avec laquelle des politiques pouvaient raconter n'importe quoi à la veille d'un scrutin électoral, pour manipuler à grande échelle l'opinion publique, et sans moyen finalement, de réagir. Cette facilité à désinformer les gens, même via les réseaux sociaux, est un problème gigantesque qui touche à tous les autres problèmes. Cela fait deux années consécutives que la désinformation est classée comme la plus grande menace contre l'humanité par le Forum Economique Mondial¹.

La désinformation est la racine de tous les maux. Si je ne crois pas au changement climatique, je ne peux pas me mobiliser pour lutter contre. Si, au sein de la société, les gens ne sont pas capables de s'accorder sur ce qui est vrai, on ne pourra jamais agir pour réduire les menaces contre l'humanité. C'est la raison pour laquelle la désinformation s'est vite imposée comme une urgence.

¹ Avec le boom de l'intelligence artificielle, la multiplication des fake news, infox et fausses informations compte parmi les plus grandes menaces pour l'humanité, selon le dernier rapport sur les risques mondiaux du Forum de Davos (Les Echos, 10 janvier 2024).

Et l'intelligence artificielle donne encore plus de moyens à la désinformation.

Exactement. L'IA a créé et propulsé des outils surpuissants pour désinformer à grande échelle. On pense aux fausses images, aux fausses vidéos bientôt, aux faux bots conversationnels, aux faux comptes qui racontent n'importe quoi sur les réseaux sociaux. C'est dramatique. Aujourd'hui, il est très facile de créer des faux contenus et de les rendre très accessibles sans avoir beaucoup de connaissances techniques. Nous avons décidé d'utiliser l'IA pour lutter contre la désinformation en nous disant qu'il fallait mettre autant de force dans la lutte contre la désinformation que dans la désinformation elle-même. L'IA peut devenir un allié puissant pour lutter contre la désinformation.

J'ai testé votre application, Vera. Ça marche très bien. Je lui ai posé une question un peu stupide, juste pour faire un test. Je lui ai demandé si c'était vrai que les pyramides égyptiennes avaient été créées par des extraterrestres. La réponse a été très intéressante et très fluide. Comment avez-vous pu avec une petite association de bénévoles, créer un outil aussi puissant ?

Nous sommes tous des experts dans nos domaines. Pour ma part, j'ai une expérience de plus de 10 ans dans la création rapide d'outils numériques. Développer des outils numériques, c'est vraiment quelque chose que nous maîtrisons bien. En plus nous travaillons ensemble avec beaucoup d'efficacité et nous allons vite. En fait, entre le moment où nous avons eu l'idée de Vera et le moment où nous l'avons lancée, il ne s'est pas passé beaucoup plus d'un mois. Bien sûr, Vera est différente et bien plus performante aujourd'hui qu'elle ne l'était à son lancement.

Je pense que pour réussir à créer des réponses technologiques innovantes, il y a un passage qui est obligatoire, et que beaucoup de gens oublient, c'est faire de la recherche terrain, comprendre d'où vient l'information,



comprendre comment les gens pourraient utiliser un tel outil, comment ils le feraient, à quelles conditions. Par exemple, on a très vite vu que le fait de s'intégrer aux processus de vie, par exemple, être présent sur WhatsApp, et pas uniquement sur le téléphone, faisait qu'il n'était plus nécessaire de télécharger une application mobile. Vera, c'est juste un numéro de téléphone qu'on enregistre dans ses contacts et hop, on peut lui parler quand on veut. Nous avons réalisé que l'accessibilité, c'était un point clé.

Un autre point clé, c'est la confiance. Et donc, la fiabilité. Avoir un comité d'experts, c'est aussi quelque chose de très important. Déjà pour nous permettre de sélectionner les



sources, mais aussi afin que les gens réalisent qu'on prend la fiabilité de Vera extrêmement au sérieux. Enfin, la fluidité dans la conversation, dont vous parliez juste avant, cette conversation, n'est pas arrivée toute seule. J'ai fait passer plus d'une cinquantaine d'entretiens utilisateurs, avec des personnes de tous les âges, de tous les milieux, de tous les bords politiques, pour comprendre comment ils pouvaient dialoguer avec une IA pour vérifier une information. Les premières versions étaient catastrophiques. Par exemple, quelqu'un qui disait "Moi, je crois aux Illuminati, j'ai vu qu'ils faisaient des sacrifices pour récupérer le sang des enfants". Le premier robot que nous avons créé, qui s'appelait Allosebunk, répondait "C'est une théorie complotiste bien connue. Voici dix sources qui vous montrent que vous avez tort." Quand on tient ce genre de propos à quelqu'un qui rapporte ce genre de fake news parce qu'il en a entendu parler, il se ferme et le dialogue ne va pas plus loin. Donc, la conversation doit être fluide, elle doit respecter le point de vue de l'interlocuteur. On ne parle pas de théorie complotiste, on dit "En effet, c'est une théorie qui existe. Voilà d'autres faits qui pourraient peut-être alimenter votre réflexion". Et d'un coup, la discussion s'engage, les gens s'ouvrent et le dialogue fonctionne. Pour obtenir cela, il nous a fallu beaucoup de méthodologie et surtout une équipe très talentueuse. Nous avons des gens géniaux dans l'équipe.

Pour faire ce numéro, je me suis informé un peu sur l'intelligence artificielle, sur les LLM et tout ça. Ce qui m'étonne, c'est que les LLM, par exemple, on les nourrit avec des milliards de données. Votre application, Vera, semble beaucoup plus légère, comment faites-vous ?

En fait, nous utilisons l'IA, mais pas de la même façon que les autres.

Nous utilisons les LLM pour leur capacité à comprendre les questions posées et leur capacité à avoir une conversation dynamique, intéressante, propice à la vérification des faits. On n'entraîne pas les LLM en les nourrissant avec des milliards d'informations. Au contraire, on leur impose un nombre réduit de sources vérifiées. Notre LLM n'a pas le droit de s'appuyer sur les grandes bases de connaissances que les autres utilisent, mais uniquement sur les sources qu'on lui impose. Et ça, ça change tout. S'il n'y a pas d'information disponible pour répondre à une question dans ces sources que nous avons fournies, le LLM a pour obligation de dire "Je n'ai pas trouvé d'information fiable qui permet de répondre à cette question". Vera est peut-être la seule IA qui est capable de dire "Je ne connais pas la réponse à cette question". Si la plupart des IA hallucinent et finissent par utiliser les données n'importe comment, c'est parce qu'elles ont pour obligation et raison d'être de répondre quoi qu'il en coûte à la question posée.

Donc, vous développez l'outil petit à petit, vous lui apportez progressivement plus d'éléments, plus de sources ?

Oui, bien sûr. Les sources, c'est une énorme partie de notre travail. Les sources sont accessibles en ligne parce que c'est important d'être transparent sur les sources qu'on utilise.

On se base sur des normes européennes et internationales pour être sûr que les sources soient les plus fiables possibles. Par exemple, des normes européennes de fact-checking².

² Fact-checking : vérification des faits

Nos sources de fact-checking sont, par exemple : les Décodeurs du journal Le Monde, le fact checking de l'Agence France Presse (AFP), le Vrai ou Faux de francetvinfo.fr, Désintox sur Arte.tv.

Ça, c'est ce qui existe en France, mais nous disposons d'autres sources dans plus de 80 pays. Et après, nous utilisons des grands médias de référence et nous nous appuyons aussi sur des normes, comme le JTI, Journalism Trust Initiative, qui est une norme portée par Reporters sans Frontière, qui permet d'évaluer la pertinence, la qualité, l'impartialité des sources. Au fur et à mesure, des experts nous aident à ajouter de nouvelles sources, notamment scientifiques ou dans le domaine de la santé, pour que Vera soit de plus en plus performante pour répondre aux questions.

Et comment mesurez-vous le succès de Vera ?

Pour mesurer le succès de Vera, il faut prendre en compte de nombreux aspects. Par exemple, nous avons un comité d'experts dont fait partie un membre de notre équipe, Maurice René, qui évalue la qualité des réponses de Vera aux questions qui lui sont posées. Nous disposons donc d'une évaluation humaine, ce qui est extrêmement important pour réaliser des IA qui soient fiables. Nous procédons aussi à des sondages sur la qualité des réponses de Vera pour voir si elle répond bien avec plus de données pour l'évaluer.

Par contre, nous ne conservons aucune donnée personnelle, tout est anonymisé. Même le numéro de téléphone de la personne qui appelle Vera n'est pas conservé.

Nous observons aussi l'usage de Vera. En 5 mois, nous avons comptabilisé déjà 150 000 questions posées à Vera, ce qui est gigantesque compte tenu du faible d'effort de communication que nous avons fourni. La communication n'est pas notre point fort, mais nous progressons quand même parce que la presse s'intéresse à nous de plus en plus.

Notre objectif final, la vision que nous portons, ce qu'on cherche à faire, c'est de démocratiser le fact-checking, un peu comme Yuka l'a fait avec la diététique. Avec Yuka, on scanne un produit et Yuka déchiffre pour nous les étiquettes : on visualise les produits qui sont bons et ceux qu'il vaut mieux éviter. On voudrait réaliser la même chose avec Vera, créer le Yuka de l'information, afin de donner à tous les gens le pouvoir de vérifier les informations qui circulent. Et donc, pour mesurer les progrès réalisés dans l'avènement de cette vision, nous mesurons la rétention, c'est-à-dire dans quelle mesure les gens continuent d'utiliser Vera après une première question.

Nous constatons qu'après la première question posée à Vera, au bout de 6 semaines, 28% des gens utilisent encore Vera dans leur quotidien. C'est gigantesque ! Pour avoir créé beaucoup de produits, je sais que c'est un résultat deux fois supérieur à celui qu'obtiennent des jeux extrêmement addictifs comme Candy Crush, qui ont un taux de rétention autour de 10%.

Nous sommes à 28%. Cela montre bien que Vera apporte une réponse crédible au problème de la désinformation et pénètre dans le quotidien des gens sans même qu'on ait besoin de faire des rappels d'utilisation.

Et pourquoi avez-vous choisi comme média le téléphone ?

L'utilisation du téléphone permet une accessibilité plus facile parce que tout le monde a un téléphone. Même les gens qui n'ont pas accès à Internet peuvent contacter Vera, il suffit d'appeler le 09 74 99 12 95. C'est très puissant, notamment pour des gens qui habitent dans les zones blanches ou dans d'autres pays, par exemple l'Afrique où la moitié de la population n'a pas accès à Internet mais où presque tout le monde dispose d'un téléphone portable.

Notre but, c'est d'être le plus accessible possible. Pour réussir à diffuser un réflexe de vérification, il faut être à portée de main. S'il faut

**EN 5 MOIS, NOUS AVONS
COMPTABILISÉ DÉJÀ 150 000
QUESTIONS POSÉES À VERA**



Florian Gauthier et l'équipe de laRéponse.tech

commencer par télécharger une application sur Internet, c'est trop compliqué. Vera est directement accessible par téléphone ou par WhatsApp ; en fait, 80% de l'utilisation de Vera se fait plutôt sur WhatsApp. Nous venons de créer un compte Instagram pour Vera, de manière à toucher plus de jeunes.

L'utilisation des réseaux sociaux joue souvent un rôle néfaste par rapport à la santé mentale des adolescents. Donc, tout ce qui peut permettre de les aider à vérifier les faits, c'est vraiment très important.

Notre objectif est de diffuser Vera sur le maximum de réseaux sociaux. Nous venons de commencer sur Instagram, nous comptons ensuite aller sur TikTok, et aussi sur X, Twitter. C'est-à-dire partout où il y a de la désinformation qui se propage.

Comment êtes-vous financés ?

En fait, le projet est autofinancé. Nous avons gagné un concours pour la lutte contre la désinformation qui nous a rapporté 5 000 euros. J'ai investi 7 000 euros sur Vera, à titre personnel. J'espère que nous obtiendrons des fonds de diverses fondations qui nous permettront peut-être d'avoir une équipe à temps plein pour le développement du projet. Nous espérons aussi bénéficier de

financements publics, notamment, parce que la lutte contre la désinformation est clairement d'utilité publique. On y travaille en ce moment. Ça fait partie des gros enjeux à venir.

Envisagez-vous de développer une activité vers d'autres défis de la société auxquels le numérique pourrait répondre ?

C'est trop tôt pour répondre à cette question. Aujourd'hui, nous nous focalisons sur Vera. Comme je l'ai déjà dit, l'objectif c'est de démocratiser le fact-checking, en France et même dans le monde. Honnêtement, je pense que c'est une ambition suffisante pour une petite équipe de 15 personnes.

Aimeriez-vous apporter d'autres informations à nos lecteurs ?

Je peux vous dire que nous prévoyons un lancement en Afrique la rentrée prochaine, avec l'objectif de lutter contre la désinformation en période électorale. On a pour objectif d'aller en Côte d'Ivoire. Là-bas, l'usage de WhatsApp est tellement répandu qu'une large utilisation de Vera est envisageable. Nous établissons des contacts pour travailler avec des entités sur place qui combattent déjà la désinformation, notamment liée à des ingérences étrangères.

Ce qui est remarquable dans votre initiative, c'est que vous allez directement au but. Vous ne vous éparpillez pas dans des démarches complexes.

Oui, nous faisons ce que nous pouvons, mais l'important c'est que nous formons une équipe géniale. Au-delà des profils technologiques, nous avons réussi à nous entourer d'experts qui travaillent depuis longtemps sur la désinformation, par exemple Maurice Ronai, l'un des créateurs du Courrier International. Nous possédons une vraie expertise sur la désinformation et la manière de lutter contre.

Et bien, je vous remercie infiniment de cet entretien et j'espère que nous aurons l'occasion de coopérer dans l'avenir.

[Retour au sommaire](#)

UN CAFÉ IA ORGANISÉ PAR ECTI À LAXOU, MEURTHE ET MOSELLE

ENTRETIEN AVEC JACQUES JARAY, ORGANISATEUR ET ANIMATEUR

PAR ROLAND DAVAL

Monsieur JARAY, merci tout d'abord de bien vouloir vous présenter à nos lecteurs.

J'ai fait mes études au lycée Jules Ferry à Saint-Dié des Vosges puis des études de mathématiques appliquées au moment où l'informatique était naissante. Je me suis ensuite spécialisé dans l'analyse, la programmation et le génie logiciel.

J'ai travaillé comme enseignant-chercheur dans un laboratoire qui était le centre de recherche en informatique de Nancy (Meurthe et Moselle) qui s'est élargi avec l'arrivée d'une filière informatique/électronique en réponse à la crise de la sidérurgie en Lorraine dans les années 80.

J'ai ensuite rejoint l'Institut National Polytechnique de Lorraine (INPL) et plus précisément l'Ecole des Mines en devenant directeur des études, administrateur provisoire.

Dès ma retraite je me suis engagé dans l'association ECTI.

Merci de nous présenter ECTI et ses principales missions :

ECTI est une association qui existe depuis plus de 50 ans. Elle regroupe des seniors qui mettent à disposition leurs compétences dans le cadre du bénévolat.

Les interventions sont réalisées auprès de TPE/PME, de start-ups, d'associations et de collectivités territoriales. Des actions sont également conduites avec des établissements d'enseignement et de formation : simulation d'entretiens, participation à des jurys, conseils, ...

ECTI intervient depuis peu, en relation avec le ministère de la justice, dans l'accompagnement des travaux d'intérêt général (TIGE).

Des actions sont également conduites dans le cadre du RSA (revenu de solidarité active) ou en lien avec la Fondation Bolloré - Fondation de la 2ème chance.

ECTI ne bénéficie d'aucune subvention.

Qu'est-ce qu'un Café IA ?

C'est un dispositif qui a été initié et promu par le Conseil National du Numérique depuis fin mai 2024.



Jacques Jaray



Les participants au café IA

« Un Café IA vise à fédérer, soutenir, promouvoir et rassembler les initiatives de débats et de partage de connaissances sur l'IA menées en vue de la diffusion d'une culture populaire du numérique¹ ». C'est aussi un relais de France Numérique auprès des TPE/PME.

Différents formats sont proposés par le site <https://cafeia.org>, l'objectif étant de réunir une quinzaine de personnes sans exigence de connaissances préalables. Le format est non magistral, à caractère participatif et organisé dans un cadre convivial.

Aucun financement gouvernemental n'est prévu.

**Pourquoi vous êtes-vous engagé dans l'organisation d'un Café IA ?
Pourquoi un Café IA à Laxou et pour quel public ?**

D'abord méfiant de l'intelligence artificielle, c'est un sujet que je n'ai pas vraiment approfondi pendant mon activité professionnelle. Mais dès que l'on a parlé de ChatGPT je m'y suis intéressé sans pour autant en devenir un spécialiste. Par ailleurs l'intelligence artificielle ce n'est pas

tout le numérique et c'est une escroquerie que de le prétendre.

J'ai ensuite organisé des ateliers pour des collègues d'ECTI Nancy, Vosges et Moselle. Puis j'ai voulu élargir cette démarche à un public plus large.

ECTI étant en relation avec des collectivités territoriales nous avons recherché des mairies qui pourraient être intéressées par une telle démarche et nous avons ciblé en Meurthe et Moselle les villes de Nancy, Maxéville et Laxou.

Le maire de Laxou a trouvé le projet intéressant et novateur et une annonce a été publiée dans le bulletin municipal et relayé par un article du correspondant local de l'Est Républicain.

Une dizaine de personnes ont participé à ce Café IA ce qui n'était pas un nombre suffisant pour la diversité des échanges.

Quels étaient les principaux objectifs de ce Café IA et quels ont été les principaux sujets abordés ?

Il s'agissait de faire comprendre ce qu'est l'intelligence artificielle et comme cela marche sans aborder des sujets trop techniques.

On s'est interrogé sur ce qui relève ou non de l'IA, quelles sont les applications technologiques intéressantes : à titre d'exemple, un GPS utilise des algorithmes de calcul et n'a rien à voir avec l'intelligence artificielle ; il en va de même avec les smartphones qui rectifient des photos grâce à des algorithmes informatiques.

Il convenait aussi de relever les idées fausses en disant, par exemple, que l'IA revient à collecter toutes les données possibles, un nombre considérable de données. Ce n'est pas que cela. La notion d'apprentissage est essentielle qui consiste à généraliser la connaissance fournie par ces données.

Pour comprendre l'apprentissage nous avons donné l'exemple de la droite de régression qui « approxime » un ensemble de couples (donnée,

¹ Extrait du site du Conseil National du Numérique.

résultat). Cette droite étant déterminée par deux paramètres a et b : respectivement la pente et résultat pour $x = 0$. Tout ceci étant expliqué progressivement avec des schémas.

Ces paramètres permettent de comprendre la notion de paramètres des neurones artificiels dans un réseau de neurones.

Sans oublier la question du gigantisme, jusqu'à 1 000 milliards de paramètres, avec celle qui en découle, celle de l'énergie que consomme l'intelligence artificielle alors que notre cerveau qui ne pèse que 500 g peut fonctionner toute la journée avec une très faible consommation.

Quelles ont été les principales questions et interventions des participants ?

Des sujets très concrets ont été abordés tels que l'assistance à l'écriture avec les corrections orthographiques automatiques ; la transformation d'un enregistrement en un texte.

Des questions ont aussi été posées quant à l'assistance à la résolution de problèmes techniques avec des propositions plutôt exhaustives et où vous trouvez des solutions auxquelles vous n'auriez pas pensé.

L'assistance à la prise de décisions a aussi été évoquée ainsi que l'aide à la programmation qui est très puissante.

Nous avons aussi échangé sur quelque chose qui devient très important : les assistants virtuels tels que ChatGPT : GPT c'est la machine de base et le Chat c'est la conversation entre vous et la machine pour que celle-ci génère une réponse.

Des questions ont aussi été abordées sur les usages malveillants de l'intelligence artificielle et sur l'avenir de l'humanité.

Quel bilan en faites-vous et si c'était à refaire quelles améliorations apporteriez-vous ?

J'aurais aimé avoir plus d'auditeurs et un public avec également la

présence de jeunes comme cela a été le cas avec la présence d'une adjointe au maire aux fiances qui a fait beaucoup d'essais avec l'utilisation de l'IA et qui a pu nous faire partager son expérience et ses interrogations.

Si c'était à refaire, je ne ferais plus les démonstrations en live mais je réaliserais en préalable des petits films qui montrent la vitesse réelle entre l'exposition du sujet et le résultat.

Quels conseils donneriez-vous à ceux qui voudraient organiser un Café IA ?

Pour animer un Café IA, il est souhaitable d'avoir des connaissances à la fois larges et approfondies sur le sujet.

Il est important de laisser toute sa place au débat tout en affirmant certaines vérités : par exemple, l'IA fera perdre des emplois à ceux qui ne l'utiliseront pas au profit de ceux qui sauront s'en servir.

Il faut aussi savoir résister à la tentation de voir l'IA comme un dépassement de l'intelligence humaine.

Il faut garder une posture critique en s'appuyant sur les avis des philosophes et des scientifiques plutôt que sur ceux des vendeurs de technologie.

[Retour au sommaire](#)



VITTASCIENCE : DES OUTILS CLÉS EN MAINS POUR L'ENSEIGNEMENT DES TECHNOLOGIES NUMÉRIQUES

ENTRETIEN AVEC LÉO BRIAND, FONDATEUR

PAR GUY MENANT

Bonjour Monsieur Briand, merci de nous avoir accordé cette interview. Pourriez-vous vous présenter pour nos lecteurs ?

Je suis le fondateur de Vittascience¹, qui est une structure spécialisée dans l'enseignement du numérique et des sciences, notamment en milieu scolaire et plus particulièrement dans le secondaire, collège lycée, même si on travaille également avec l'enseignement primaire et l'enseignement supérieur. Cette structure, je l'ai créée en 2018 en tant qu'entité juridique, mais c'est un projet en fait que j'avais démarré à la fin de l'année 2015, donc il y a presque dix ans quand j'étais encore étudiant en école d'ingénieur. En l'occurrence j'ai fait deux écoles d'ingénieur, plutôt domaine aéronautique puis un master en entrepreneuriat et IA, à l'ISEA-SUPAERO² de

Toulouse et à l'École Polytechnique.

Le projet a démarré sur les bancs de l'école dans le cadre d'un projet qui s'appelait OSE (ouverture sociale étudiante).

L'objectif c'était d'aller à la rencontre d'élèves dans des quartiers défavorisés, économiquement parlant, pour justement transmettre le goût des sciences et donner l'envie à certains d'aller faire des études scientifiques ou techniques, pas forcément une école d'ingénieurs mais au sens large. C'était à Toulouse que j'ai démarré et ensuite j'ai fait la même chose en région parisienne.

Donc le projet Vittascience a vu le jour dans ce cadre avec comme objectif de rendre les sciences et la technologie plus attrayantes mais aussi plus faciles à enseigner pour les enseignants, parce qu'on

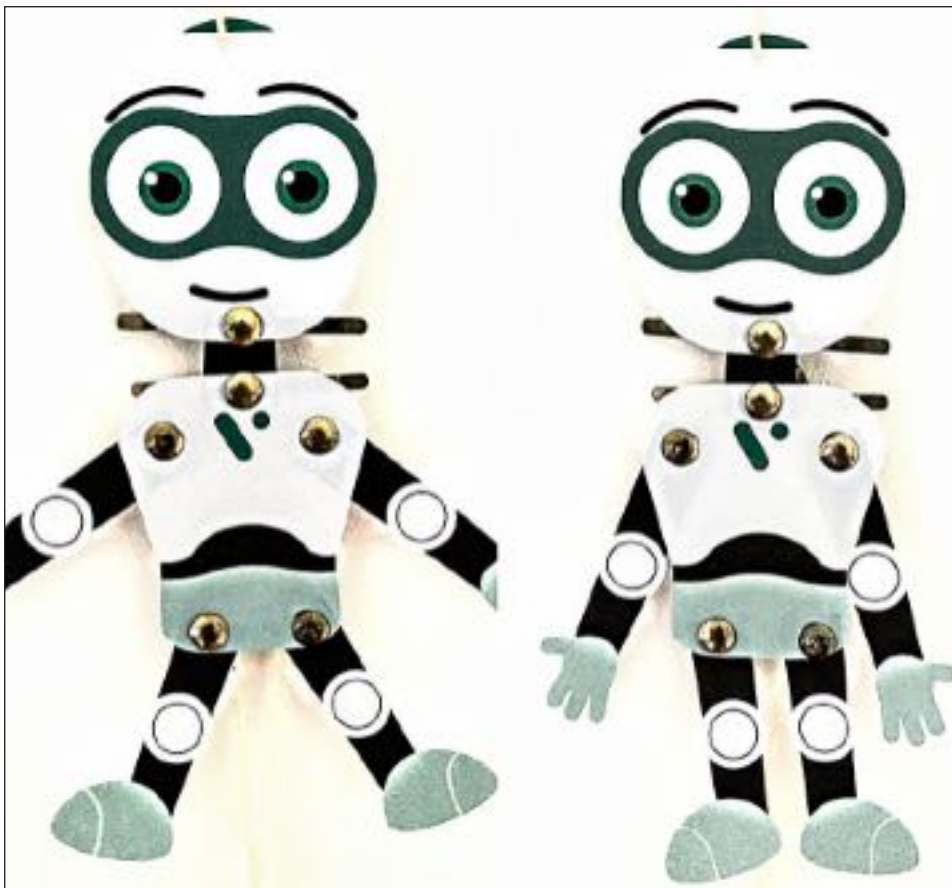


Léo Briand

¹ Vittascience - <https://fr.vittascience.com/>

² L'Institut supérieur de l'aéronautique et de l'espace

dont l'acronyme est ISAE-SUPAERO est une école d'ingénieurs française spécialisée dans le domaine de l'aéronautique.



s'est vite rendu compte qu'ils étaient au centre du jeu.

Petit à petit, ça a évolué et nous sommes passés par différents modèles et aujourd'hui le modèle dominant c'est une plateforme en ligne, donc des outils numériques qui sont ensuite utilisés en classe par aujourd'hui à peu près 400 000 élèves et enseignants partout en France et également à l'international.

Aujourd'hui nous sommes une équipe d'une vingtaine de personnes avec 16 salariés et des entrepreneurs qui collaborent avec nous.

C'est un projet qui a démarré comme ça, ce qu'on pourrait appeler bottom-up en anglais, vers la généralisation. Nous avons eu une première phase de notoriété grâce au Ministère de l'Éducation Nationale qui nous a soutenus via le dispositif Edu-Up³ de soutien à l'innovation.

³ Dispositif Edu-Up : soutien à la production de ressources numériques pour l'Ecole
<https://www.education.gouv.fr/le-dispositif-edu-un-soutien-la-production-de-ressources-numeriques-pour-l-ecole-450476>

Plus récemment nous avons remporté un marché public qui permet de généraliser l'outil à toute la France avec une prise en charge financière par le Ministère de l'Éducation Nationale.

Vittascience a été deux fois lauréate du dispositif Edu-Up de soutien du Ministère de l'Éducation Nationale. Pourriez-vous nous décrire brièvement le modèle économique de ces partenariats public-privé ?

C'est une question qui est très difficile parce que le Ministère de l'Éducation nationale a plutôt une culture de développement interne, sachant qu'il y a près d'un million d'enseignants en France. Cette gigantesque quantité de personnes fonctionnaires ou contractuelles favorise plutôt la culture de travailler en interne que de faire appel à des acteurs tiers de type associatif ou entreprise. Et donc c'est assez difficile de faire sa place, il faut vraiment montrer qu'on a une grosse valeur ajoutée par rapport à ce que les enseignants pourraient faire par eux-mêmes.

La DNE (Direction du numérique pour l'éducation) s'occupe à la fois des projets internes, en l'occurrence des projets de développement de plateformes internes au ministère, mais également des projets externes de partenariat avec des entreprises ou des associations tierces. Et elle met en place des dispositifs, notamment Edu-Up, dont Vittasciences a bénéficié en 2019 et en 2022, qui accorde un soutien financier et un soutien à la diffusion, ainsi qu'un label qui permet de crédibiliser la structure. Il y a actuellement à ma connaissance une vingtaine de structures qui en bénéficient chaque année.

Ensuite il y a des dispositifs beaucoup plus ambitieux mais qui sont beaucoup plus difficiles à obtenir. Je pense notamment au partenariat d'innovation et d'intelligence artificielle⁴ (P2IA), qui a

⁴ P2IA (partenariat d'innovation et d'intelligence artificielle).
<https://eduscol.education.fr/1911/l-intelligence-artificielle-pour-accompagner-les-apprentissages-des-fondamentaux-au-cycle-2>

mis entre les mains des enseignants des intelligences artificielles dédiées à l'éducation, avec cette fois-ci des budgets beaucoup plus importants, de l'ordre de plusieurs millions d'euros, mais aussi bien sûr beaucoup moins de lauréats.

Puis, il y a eu d'autres projets qui ont été lancés au fil des années par le Ministère de l'Éducation Nationale. Mais en fait, l'essentiel de l'activité passe plutôt par le terrain, donc les établissements scolaires, qui peuvent décider de s'équiper en matériel ou en logiciel. Par exemple, les écoles ont la possibilité d'acheter des microscopes, des tubes à essai, des logiciels éducatifs. Ils peuvent acheter toutes sortes de robots pédagogiques, toutes sortes d'objets, de matériels ou de logiciels.

Les collectivités territoriales jouent aussi un rôle dans cette collaboration public-privé, puisque les mairies gèrent les écoles, les départements gèrent les collèges et les régions gèrent les lycées. En fait, c'est un peu le millefeuille administratif français, mais à chaque échelon, il y a des budgets qui sont dédiés à l'éducation. Ces budgets vont prioritairement à la cantine, à l'aménagement des locaux, aux agents d'entretien des établissements. Mais il y a aussi souvent des budgets complémentaires de collectivités pour équiper du matériel ou du logiciel dans les classes et on a pu en bénéficier à plusieurs reprises.

Nous collaborons aussi avec Réseau Canopé⁵, qui dispense pour les enseignants des formations intégrant de ce que nous faisons en termes de logiciels. Ils avaient organisé un marché d'équipements en logiciels pour les écoles dont on a pu être lauréat et qui a permis de diffuser le logiciel à un bon nombre d'établissements scolaires, notamment aux collèges et lycées.

Donc, le partenariat public-privé, il est difficile à enclencher, il n'est pas très naturel. Mais si on apporte suffisamment de valeur à l'institution, à l'enseignant ou aux personnels administratifs, des choses intéressantes peuvent se faire, soit à

l'échelle micro établissement par établissement, soit à l'échelle macro au niveau d'un ministère entier, ou alors collectivité par collectivité qui est un peu entre les deux.

Est-ce que vous pourriez nous expliquer comment les classes et les enseignants utilisent Vittascience, par l'intermédiaire du site web ou bien d'une autre façon ?

Nous avons d'une part une plateforme en ligne qui est l'outil principal. Cette plateforme permet de découvrir les disciplines du type codage, robotique et intelligence artificielle, donc des disciplines technologiques au sens large. Celles-ci sont enseignées à partir du CM1 jusqu'à la terminale, c'est dans les programmes scolaires.

A la fin du primaire, ce sont les enseignants, les professeurs des écoles qui l'enseignent, par exemple en introduisant des robots qui vont faire des quadrillages et qui vont donc essayer de se repérer dans l'espace. Ensuite au collège c'est en mathématiques et en technologie que c'est enseigné, donc dans deux disciplines. Et ensuite au lycée, il y a beaucoup de disciplines qui l'intègrent, en mathématiques et en sciences de l'ingénieur, mais aussi en SNT (la nouvelle matière « sciences numériques et technologies »), qui est enseignée pour tous les élèves de seconde générale et technologique, et également en NSI (numérique et sciences informatiques) qui est l'option (spécialité) de première et terminale vraiment dédiée au numérique.

Donc il y a toute une multiplicité de disciplines qui enseignent ces sujets-là, dans les programmes du CM1 à la terminale, et c'est quelque chose d'assez nouveau parce qu'en fait c'est arrivé au collège en 2016, au lycée en 2019 et en primaire seulement il y a deux ans. Par conséquent, pour les enseignants c'est nouveau dans leurs programmes.

Et donc c'est pour ça que nous apportons une plateforme logicielle qui vient répondre à ces besoins en leur proposant un outil clé en main

NOUS COLLABORONS AUSSI AVEC CANOPÉ QUI DISPENSE POUR LES ENSEIGNANTS DES FORMATIONS INTÉGRANT CE QUE NOUS FAISONS EN TERMES DE LOGICIELS.

⁵ Réseau Canopé - Réseau national de formation des enseignants.

<https://www.reseau-canope.fr/>



SUR LA PARTIE INTELLIGENCE ARTIFICIELLE, NOUS AVONS LANCÉ, IL Y A 3 ANS, UN SERVICE QUI PERMET, EN COMPLÉMENT DU CODAGE, DE DÉCOUVRIR LE FONCTIONNEMENT D'UNE IA.

pour l'enseignement de ces nouvelles disciplines. Il prend la forme d'un site web à l'adresse <https://fr.vittascience.com/> où vous avez la possibilité de programmer des cartes, des robots, des drones, enfin toutes sortes d'objets, ou de simuler ceux-ci, c'est-à-dire de s'en servir mais sans forcément avoir le matériel.

On peut faire de la programmation par blocs. Ce sont des blocs visuels qu'on va assembler, un peu comme un puzzle, qui est plutôt adapté à des niveaux primaire-collège. Ensuite on va passer sur la programmation textuelle et en l'occurrence il y a plein de langages qui existent. Un peu comme les langues humaines (français, anglais, arabe, etc.) il y a des langues informatiques et en l'occurrence le Ministère de l'Éducation Nationale a choisi le langage python, un des langages qui existent, comme étant celui qu'il faut enseigner ; donc nous nous sommes focalisés sur celui-là, même si nous proposons aussi des outils dans d'autres langages comme par exemple HTML ou JavaScript.

Sur la partie intelligence artificielle, nous avons lancé il y a trois ans un service qui permet, en complément du codage, de découvrir le fonctionnement d'une IA comme on pourrait le faire sur ChatGPT ou Mistral.

C'est vraiment de comprendre son fonctionnement et les enjeux qui sont autour, donc aussi bien l'utilisation avec ses points forts et points faibles, mais aussi la logique de fonctionnement d'un réseau de neurones, les enjeux d'éthique comme par exemple les biais ou les hallucinations qui peuvent éventuellement questionner le rapport à la vérité ou les deepfakes, toutes sortes d'enjeux qui gravitent autour.

Notre objectif c'est d'avoir un outil didactique pour l'IA qui puisse être un vrai outil. On ne les met pas face à une fausse IA, c'est une vraie IA, mais avec le capot ouvert en quelque sorte, ce qui permet de comprendre le fonctionnement et les enjeux aux alentours.

Quelle est la résilience de votre entreprise pour le cas où il y aurait un changement de politique au niveau du ministère ?

On a plus de 400 000 utilisateurs par an sur la plateforme Vittascience, et sur l'IA en particulier on est passé de 0 à près de 150 000 utilisateurs en l'espace de 3 ans. Donc on a quand même une grande quantité d'enseignants et d'élèves qui veulent s'en servir.

Le modèle économique attendant à cette plateforme est ce qu'on appelle freemium, c'est un néologisme par la contraction de gratuit, free et premium payant. Et donc il y a une partie payante et une partie gratuite. La partie gratuite est relativement généreuse, c'est 80% des fonctionnalités du site et c'est pour ça que nous avons autant d'utilisateurs. Si nous avions mis quelque chose d'entièrement payant, c'est certain que nous n'aurions jamais eu autant d'utilisateurs sur la plateforme.

Mais du coup, nous avons un enjeu qui est un peu double. Nous avons un enjeu de faire grossir la communauté sur la partie gratuite de la plateforme pour arriver à avoir un grand nombre d'utilisateurs même si cela ne nous rapporte pas du tout d'argent. Et puis ensuite, de

convertir une partie de ces utilisateurs sur la formule payante pour arriver à avoir un modèle économique viable. Et cette formule payante (la partie premium), ce qu'elle rajoute, c'est notamment la possibilité de suivre la progression des élèves en tant qu'enseignant (ce qu'on appelle un module de suivi de classe), et aussi l'accès à des IA plus puissantes et des ressources pour l'enseignement. Pour ces fonctions réservées à la partie premium, la partie gratuite est quand même assez généreuse puisqu'on peut tout à fait utiliser l'outil et il n'y a pas de blocage, par exemple au bout de 15 jours ou au bout de X programmes ou X essais d'IA.

C'est la partie la plus difficile en tant que chef d'entreprise dans l'éducation. En fait, le plus dur c'est justement d'arriver à d'une part concilier le développement du logiciel et la croissance de la communauté mais aussi une nécessité économique d'avoir un modèle avec des revenus sinon cela ne fonctionne pas.

Aujourd'hui, je dirais que la filière éducative est assez large puisqu'il y a plus de 100 entreprises qui disent faire des outils pédagogiques pour le niveau scolaire, collège, lycée et primaire. Et si on ajoute à ça la formation de l'enseignement supérieur et la formation professionnelle, on est sur plusieurs centaines, pas loin de 500 entreprises.

Une étude est sortie il y a quelque temps, portée par la Banque des Territoires et EY, un cabinet de conseil, qui ont fait justement un état des lieux de la filière. Sur ces 500 entreprises au global qu'on pourrait appeler EdTech, technologie de l'éducation, il y en a une centaine qui s'intéresse spécifiquement au sujet scolaire.

Celles-ci sont souvent fragiles dans la mesure où elles dépendent essentiellement des marchés publics. Cela nécessite soit d'arriver à convaincre un très grand nombre

d'écoles sur des petits paniers, soit de convaincre pas mal de collectivités sur des paniers de taille plus grande, soit de convaincre le Ministère sur des gros tickets.

Aujourd'hui, à Vittascience, comme je le disais, nous sommes une vingtaine de personnes, c'est une société qui est rentable depuis sa création, puisque nous n'avons pas eu d'investisseurs tiers. Je suis en tant que fondateur l'unique actionnaire de Vittascience et la conséquence c'est que nous n'avons jamais été déficitaires. Donc la structure est plutôt solide de par cette stratégie qui contraste beaucoup avec celle des structures qui lèvent de l'argent, qui font appel à des investisseurs notamment en début d'aventure, qui se mettent structurellement en déficit pour grossir plus vite, mais qui parfois n'arrivent pas à remonter leur niveau de chiffre d'affaires ou de subventions à la hauteur de leurs dépenses suffisamment vite, et qui malheureusement ont un taux de mortalité assez élevé.

Mais bon, les deux modèles s'entendent bien. En tout cas nous on est plutôt sur une croissance, pas une hyper croissance qui atteint des milliards d'euros de chiffre d'affaires en quelques années, mais qui est pérenne avec une rentabilité chaque année. Néanmoins je partage avec mes confrères EdTech la fragilité du secteur au sens large qui est quand même beaucoup soumise au bon vouloir des politiques publiques d'investissement dans le numérique éducatif, sachant qu'aujourd'hui il y a une logique de budget qui impacte la vitalité de la filière.

Aujourd'hui, il y a quand même énormément d'entreprises qui se sont créées (et aussi d'associations), suite à l'arrivée de l'IA, qui a appelé énormément d'entrepreneurs à lancer leur aventure, donc beaucoup plus de candidats sur des budgets publics qui sont plus ou moins constants et malheureusement beaucoup d'échecs entrepreneuriaux.

En effet, l'IA a rendu le processus de développement d'une solution





JE ME SUIS DÉCOUVERT UNE VRAIE PASSION POUR L'ÉDUCATION. J'ÉTAIS DÉJÀ PASSIONNÉ DE SCIENCES ET JE ME SUIS DIT, FINALEMENT, MA PASSION C'EST DE TRANSMETTRE ÇA AUX PLUS JEUNES.

éducative beaucoup plus simple : c'est pour ça je pense qu'il y a cette « ruée vers l'or ».

Votre itinéraire personnel est assez original. Vous avez un haut niveau d'études qui vous orientait a priori vers l'industrie, et vous vous intéressiez en même temps à l'éducation. Est-ce que vous pouvez nous expliquer un peu votre parcours ?

J'ai fait primaire, collège, lycée dans le public classique. J'étais en campagne à Cahors, c'est le chef-lieu du Lot mais ça reste quand même une petite ville. Ensuite j'ai fait une prépa à Castres et j'ai intégré une école d'ingénieur, ISAE-SUPAERO, qui était vraiment l'école de mes rêves, parce que je voulais devenir ingénieur en aérospatiale, notamment dans les fusées. Donc travailler soit avec Ariane, soit avec le CNES, soit avec l'ESA, soit avec SpaceX.

Je me suis rendu compte de plusieurs choses. Le temps de développement dans le domaine spatial ne correspondait pas trop à mon envie d'aller vite parce que c'est quand même très long pour sortir un programme de lancement spatial. Il y a au moins dix ans entre la conception et le vol de la fusée.

Et puis par ailleurs, j'ai commencé à faire du mentorat de façon bénévole via l'association Ose (Ouverture Sociale Étudiante),

dont je parlais au tout début de notre interview. Et en fait, je me suis découvert une vraie passion pour l'éducation à ce moment-là. J'étais déjà passionné de sciences et je me suis dit finalement, ma passion, c'est de transmettre ça aux plus jeunes. Donc c'est ce que j'ai fait.

Et puis parallèlement à ça, j'ai senti assez tôt que j'aimais l'idée de créer ma propre entreprise, de lancer mon navire à l'eau et essayer d'avoir une aventure. Cela me plaisait bien, plus que de rejoindre une grande structure. Je me disais que je n'avais pas grand-chose à perdre en tant que jeune diplômé de me lancer là-dedans, peut-être que ce serait plus difficile plus tard de changer de carrière. Et donc, Vittascience c'est un peu la convergence entre une passion pour les sciences qui est très ancienne, depuis que je suis tout petit, une passion pour l'éducation qui est venue quand j'étais en train de réaliser ces ateliers d'ouverture sociale étudiante, et puis une passion pour l'entrepreneuriat, c'est-à-dire de mener ma barque et de pas forcément être dans une grande structure, en tout cas pour le moment.

J'ai donc lancé ce projet-là en tant qu'étudiant et comme à la fin de mes études, j'avais déjà réussi à atteindre un certain niveau de chiffre d'affaires, de ventes, et à avoir un prix d'entrepreneuriat, j'avais réussi à gagner 35 000 euros, ce qui à l'époque me semblait être une somme absolument colossale par rapport à mes économies d'étudiant. J'étais certain que ça allait marcher. C'est avec cette confiance-là que j'ai pu lancer la structure derrière.

Et après on a eu bien sûr énormément de difficultés, la réalité m'a rattrapé. Mais finalement voilà, de fil en aiguille, j'ai réussi à mener ma barque et aujourd'hui on s'en sort bien notamment avec l'Education nationale en France.

Et puis, beaucoup d'utilisateurs sont ravis du service et il y en a beaucoup qui en profitent gratuitement, grâce à notre modèle économique. Ils sont très contents d'avoir un service de qualité français, gratuit, là où la concurrence est plutôt chinoise, américaine ou indienne.

La plupart des utilisateurs sont dans le milieu scolaire, mais on peut très bien avoir des personnes du grand public qui s'en servent. Et puis, de manière générale, sur les formations à l'IA, même les adultes sont concernés et pas seulement dans les disciplines scientifiques. Néanmoins, notre cœur de cible, en tout cas, avec le cœur de la communauté, reste plutôt la partie scolaire, notamment secondaire.

Parmi les structures qui utilisent Vittascience, il doit y avoir des associations qui font de l'animation de quartier sur l'éducation aux technologies de l'information et de la communication

Il y en a beaucoup, effectivement. Nous sommes régulièrement sollicités pour intervenir sur le terrain et en général, nous disons non, pas parce que nous n'avons pas envie, mais parce qu'en fait, nous n'avons pas la structure, les ressources humaines pour ça. Parce qu'en fait, sinon, il faudrait que nous ayons des animateurs dans toute la France et nous nous transformerions en association avec un modèle très différent de celui que nous avons aujourd'hui. Donc nous redirigeons vers des associations ou entreprises locales. Par exemple nous travaillons beaucoup avec «la Compagnie du Code⁶» à Toulouse qui est une société coopérative d'intérêt collectif qui fait de la formation au codage pour les jeunes que ce soit en scolaire ou en périscolaire. Nous travaillons aussi avec Abrakodabra⁷ qui fait la même chose à Bordeaux, avec Magic Makers⁸ qui fait de même essentiellement à Paris, et avec également la Ligue de

l'enseignement qui utilise nos outils, pour l'éducation populaire.

C'est vraiment très large. Globalement, nous ne nous limitons pas à l'éducation nationale au sens strict. Mais notre enjeu, c'est de former les gens sur les disciplines technologiques.

La rapidité d'exécution et l'amélioration croissante de la pertinence des réponses apportées par les IA risque de concurrencer et donc réduire le recours au travail humain en équipe. Votre engagement personnel pour la promotion des sciences vous a nécessairement sensibilisé à ce problème : comment maintenir une solide culture de la coopération entre acteurs humains, fondement du progrès des connaissances et source d'innovations par la confrontation des idées ?

Il y a plusieurs échelons de réponses. En interne, nous développons quelques contenus. Ensuite, nous avons la communauté d'enseignants que nous poussons à créer. En gros, les enseignants qui se servent de Vittascience créent quasiment tous du contenu qu'ils pourraient partager avec leurs collègues. Donc, nous essayons de créer cette dynamique-là, la plateforme permet ce partage. Mais aujourd'hui, c'est quand même une dynamique qu'il faut que nous arrivions à pousser.

L'institution, elle aussi, fait ça de son côté, avec notamment les «travaux académiques mutualisés» (TraAM). Et chaque année, nous essayons de nous infiltrer en quelque sorte dans les travaux académiques mutualisés par les enseignants et inspecteurs qui vont collaborer pour créer des contenus. S'ils utilisent Vittascience dans ces contenus, c'est ensuite diffusé à large échelle et c'est super pour nous parce que ça fait de l'utilisation de l'outil par des experts dans leur discipline.

Donc, à part cette initiative des travaux académiques mutualisés, qui est quelque chose de pérenne, il n'y a pas beaucoup d'initiatives de création de contenu collaboratif, et c'est plutôt les enseignants qui choisissent ou non de partager ce qu'ils ont fait dans une bibliothèque collaborative.



⁶ La Compagnie du Code - <https://www.lacompagnieducode.org/>

⁷ Abrakodabra - <https://abrakodabra.fr/>

⁸ Magic Makers - <https://www.magicmakers.fr/>



IL FAUDRAIT VRAIMENT PLUS PRÉPARER LES ÉLÈVES À UN FONCTIONNEMENT COOPÉRATIF

Nous travaillons énormément avec une plateforme de codage du ministère qui s'appelle CAPYTALE⁹ : les outils de Vittascience y sont intégrés, c'était l'objet du fameux partenariat avec le ministère dont je parlais au tout début. Sur cette plateforme, il y a une collaboration des enseignants qui est assez forte, avec plus de 6000 activités pédagogiques partagées entre enseignants. Donc finalement, on voit que c'est possible. Et puis sur la plateforme Tactileo¹⁰, il y a aussi un grand nombre d'activités partagées entre enseignants.

Je crois qu'il faut vraiment avoir des milliers d'activités partagées et que ça puisse couvrir un peu toutes les disciplines, tous les niveaux et tous les cas d'usage, pour qu'un enseignant qui cherche quelque chose puisse trouver ce qui correspond à peu près et qu'il puisse le modifier. Et donc ce partage, nous l'effectuons sous licence libre, c'est-à-dire que tous les contenus publiés par les enseignants basculent sous licence open source, ce qui fait qu'ensuite, cela peut être réutilisé librement sans risquer de se faire attaquer pour non-respect du copyright.

Parfois, j'ai senti que certains enseignants avaient l'impression que le partage pourrait leur porter préjudice. En fait, imaginons un enseignant qui a des élèves en

collège, en mathématiques. Il fait un contenu pédagogique. S'il le fait avec sa classe et qu'il n'en parle à personne et qu'il ne le publie nulle part, il ne va rien lui arriver. Si par contre il le met en ligne, il y a un risque de retours négatifs, ce n'est vraiment pas agréable du tout.

Est-ce que vous travaillez avec des organismes qui essaient de développer les pédagogies coopératives comme par exemple l'Office central de coopération à l'école ?

Ce serait avec plaisir. Si vous avez les contacts, vous pouvez nous mettre en relation. J'ai parlé par contre avec beaucoup de structures de type EdTech, comme la mienne, qui essaient de promouvoir la collaboration entre enseignants, parce qu'on est tous d'accord pour dire que c'est ça la bonne façon de faire. Mais je n'ai jamais parlé avec une structure vraiment spécialisée sur ce sujet-là à part entière.

Il faudrait vraiment plus préparer les élèves à ce fonctionnement coopératif dont on est en train de parler, où c'est le groupe (et non l'individu) qui est l'instance de travail et de production. Votre plateforme peut-elle contribuer à essayer de faire avancer ce problème-là ?

Je pense que oui. Mais nous n'avons pas vraiment la main sur la pédagogie qui est mise en œuvre par les enseignants. Nous proposons un outil et nous essayons bien sûr de le rendre le plus facile à utiliser à la fois côté enseignant et côté élève en apportant le maximum de valeur, de simplification et en même temps de couches didactiques. Donc aussi bien sur l'IA pour expliquer comment ça marche et tous les enjeux, que sur le codage et la robotique.

Néanmoins, on voit quand même beaucoup de pédagogie par projet, notamment en technologie au collège, où c'est vraiment quelque chose qui est très développé.

Les compétences de collaboration font également partie des attendus, avec par exemple cette capacité à

⁹CAPYTALE - <https://capytale2.ac-paris.fr/>

¹⁰ Tactileo - <https://www.tactileo.com/>

imaginer un robot et faire un petit challenge où le robot va devoir parcourir toute une succession de suivi de lignes, de détection d'obstacles ou d'attraper des choses avec une pince par exemple.

Cela peut être fait de façon collaborative avec un petit groupe, voire même certains qui vont jusqu'à scénariser, c'est-à-dire nommer un chef de groupe, préciser les rôles de chacun. Cela peut être assez vertueux, mais c'est vraiment à la main des enseignants. Nous proposons des exemples là-dessus, mais malheureusement, nous ne pouvons pas vraiment pousser plus qu'en donnant des exemples. Et on voit qu'il y a quand même aussi beaucoup d'enseignants qui préfèrent quelque chose de plus classique, individualisé, soit cours magistral, soit chacun manipule, mais sur un poste distinct. Voilà, j'ai l'impression que cela a tendance à pas mal se démocratiser, cet apprentissage par groupe, notamment dans les disciplines technologiques.

Est-ce que vous avez l'intention ou avez déjà touché d'autres pays francophones ?

Oui ! En fait, le modèle freemium dont je parlais au début a cet avantage-là qu'il n'y a pas vraiment de barrière à l'entrée. Il suffit qu'un enseignant belge, pour ne citer qu'eux, aille sur Vittascience.com et il peut commencer à l'utiliser sans rien payer et sans formation, simplement en se documentant.

Le site est traduit en cinq langues, français, anglais, italien, arabe et espagnol. Historiquement nous n'avions fait que français anglais et puis c'est suite à un partenariat assez important que nous avons eu avec STMicroelectronics qui est le leader des puces électroniques en Europe, que nous avons rajouté l'arabe, l'italien et l'espagnol. Et donc en discutant avec ces enseignants-là, nous avons aujourd'hui 15% des 400 000 utilisateurs qui sont hors de France. Donc 85% en France et 15% ailleurs. Cela fait quand même pas

mal de monde et ces enseignants sont surtout en Belgique, en Suisse, dans les pays d'Afrique du Nord notamment Maroc, Tunisie et puis aux Etats-Unis. Mon enjeu, c'est de faire en sorte que la France devienne moins majoritaire, et pas en baissant le nombre de personnes en France, mais plutôt en augmentant fortement le nombre dans d'autres pays. Cela aura d'une part un effet positif de déployer à plus large échelle ce qu'on fait, mais aussi, je pense et j'espère, de rendre le modèle économique un peu moins sensible aux fluctuations politiques de la France, puisque pendant que nous en France, on est en train de serrer la ceinture, il y a des pays comme par exemple l'Arabie Saoudite qui sont en train de faire des plans énormes de formation à l'IA pour leurs élèves. Et puis il y a des pays qui sont dans des conjonctures économiques croisées par rapport à la nôtre.

Et donc c'est intéressant d'être sur plusieurs géographies pour avoir quelque chose de plus stable et de plus pérenne. Notre enjeu c'est vraiment de devenir leader mondial, en tout cas du monde occidental dans un premier temps. On verra si on arrive à aller jusqu'aux pays qui sont les moins ouverts. Je pense à la Chine par exemple, qui n'autorise pas les sociétés étrangères sur son marché éducatif.

En tout cas, notre objectif, c'est de devenir la plateforme de référence pour ces enseignements technologiques à l'école. Nous le sommes devenus en France, après sept ans. Et maintenant, nous avons, je pense, tous les atouts d'un point de vue technologique. Mais le plus dur, encore une fois, ce n'est pas le développement de l'outil, c'est la diffusion. Et la diffusion, pour moi, ça passe à la fois par une forte adhésion des enseignants sur le terrain et une adhésion de l'institution.

[Retour au sommaire](#)





L'IA, UN RÉCIT ANIMÉ PAR DES VISIONS DYSTOPIQUES, MAIS DÉJÀ UNE RÉALITÉ SOCIALE

ENTRETIEN AVEC JEAN-FRANÇOIS LUCAS DÉLÉGUÉ GÉNÉRAL DE RENAISSANCE NUMÉRIQUE

PAR DOMINIQUE BÉNARD

Bonjour Monsieur Lucas, pouvez-vous vous présenter et présenter Renaissance Numérique à nos lecteurs ?

Je suis le délégué général de Renaissance Numérique. J'ai pris cette fonction il y a presque trois ans. Auparavant, j'avais fait cinq ans de conseil auprès de collectivités territoriales et de grands groupes autour de projets dits de "numérique responsable" pour des "Smart City, ou Smart Territoire, avec une focale sur l'inclusion et la question environnementale. Et puis avant, un parcours à l'université et en grandes écoles, sur des questions touchant au numérique et à la participation citoyenne. Je me suis beaucoup intéressé aux transformations numériques des villes, ce que je fais peut-être un peu moins aujourd'hui. J'ai un doctorat de sociologie, mais j'ai pas mal emprunté à la philosophie, à la géographie et à d'autres disciplines. Cet intérêt pour les transformations numériques de la société m'a donc amené à Renaissance Numérique, une association de loi de 1901 créée en

2005 par un collectif qui faisait le constat que la question du numérique était très peu discutée dans les sphères politiques notamment.

Ce groupe, formé plutôt de gens issus du monde de l'entreprise, avait préparé le terrain, notamment en rédigeant un livre blanc avec [10 recommandations pour résorber les "fractures numériques"](#) et en le présentant aux différents candidats à l'élection présidentielle en 2007. A l'époque, il n'y avait pas beaucoup d'associations centrées sur le numérique, tel qu'on le voit aujourd'hui. Donc, assez rapidement, un réseau d'acteurs s'est constitué avec des gens placés à des postes à responsabilité assez élevée dans différentes entreprises et institutions. Aujourd'hui, Renaissance Numérique est un petit think tank, si l'on considère son chiffre d'affaires et son équipe salariée, mais regroupe 43 structures adhérentes : à la fois des indépendants, des petites entreprises impliquées dans les relations publiques, la communication, le conseil, mais aussi des très grandes



Jean-François Lucas



Renaissance Numérique

entreprises comme Apple, Google, Meta, Microsoft, Orange... En outre, ce qui fait vraiment la valeur ajoutée de Renaissance Numérique par rapport à d'autres think tanks qui s'intéressent aux mêmes questions que nous, c'est que nous regroupons une diversité d'acteurs, avec des chercheuses et chercheurs du monde académique, des ONG, des cabinets d'avocats, et également des responsables du monde institutionnel.

Une des valeurs ajoutées que nous proposons à nos adhérents c'est d'être un espace de dialogue. Nous ne défendons pas les intérêts de nos adhérents, ils ne sont pas tous d'accord sur les sujets que nous abordons, nous recherchons une diversité de points de vue, car c'est dans la confrontation d'idées qu'on peut avancer, afin d'éclairer la complexité des transformations numériques de la société.

Nous nous adressons à l'ensemble des personnes qui ont un pouvoir de décision dans différentes structures, qu'elles soient publiques, privées ou politiques, car ce sont des personnes qui ont des possibilités d'action et de transformation numérique de la société.

Nous n'avons pas la prétention de penser qu'on va changer la manière de décider de Mark Zuckerberg ou d'autres, mais nous pouvons influencer par petites touches, dans le sens d'une société plus juste, plus équitable, plus respectueuse des droits et des libertés des individus. C'est cela notre raison d'être. Le numérique est là, il faut faire avec et notre objectif c'est de proposer un espace de débat et d'éclairage des choix qui vont être faits pour une société numérisée responsable.

Notre travail est reconnu comme sérieux, rigoureux et capable de produire des recommandations pour une société plus respectueuse, plus juste, avec une implication directe des citoyens.

Quels types d'actions menez-vous ?

Nous organisons aussi des événements, en grande partie ouverts au grand public. Par exemple, le 23 juin, nous organisons un webinaire pour présenter une note que nous venons de publier sur l'interrégulation du numérique en France et en Europe. Le webinaire est ouvert à tout le monde. Nous avons organisé ainsi différents événements publics pendant le Sommet pour l'action sur l'Intelligence Artificielle en début d'année ouverts à tous et gratuits..

Et puis, en marge de nos travaux qui mobilisent un groupe de travail à chaque fois, nous avons lancé, il y a presque trois ans maintenant, un nouveau format intitulé "les dialogues", qui sont des journées d'échange, de confrontation de points de vue, ouverts à des acteurs extrêmement pertinents sur le sujet traités, mais qui ne sont pas forcément membres de Renaissance Numérique.

Est-ce que vous avez aussi un rôle de conseil auprès de collectivités locales ?

Non, pas du tout. Ce n'est pas notre environnement. On ne fait pas du tout de conseil. Nous sommes un espace de réflexion.

A partir de votre expérience, comment jugez-vous l'évolution du numérique dans la société française ?

Quand on parle du numérique, il faut d'abord prendre en compte les infrastructures. Aujourd'hui en France, 94% du territoire est éligible à la fibre optique et plus de 99% pour la 4G, avec des taux d'équipement qui sont aussi extrêmement élevés : ordinateurs fixes, tablettes, smartphones, etc. Bien sûr, on constate des inégalités territoriales en termes d'équipements

et d'accès au réseau qui sont corrélées aux inégalités sociales, économiques et culturelles. On est dans une tension avec à la fois une acculturation des usages du numérique - les gens font quand même plus de choses avec le numérique qu'ils en faisaient hier - mais aussi une accélération des changements qui nécessite un effort d'adaptation élevé. L'irruption de l'intelligence artificielle, notamment est extrêmement ambivalente parce qu'à la fois, elle propose une grande facilité d'usage avec un aspect un peu magique - il suffit d'écrire une phrase pour avoir une réponse - et de l'autre une complexité déroutante qu'il faut décrypter et comprendre.

Donc on est vraiment dans cette tension. On ne va pas rentrer dans les débats théoriques autour des concepts et notions relatives à l'illectronisme.¹

Il y a beaucoup de composantes et d'explications diverses qui expliquent l'éloignement de certains publics du numérique. C'est souvent corrélé au niveau socio-professionnel, culturel et territorial, mais pas toujours. Par exemple, on sait que les jeunes sont très à l'aise avec certaines pratiques comme le maniement des réseaux sociaux ou le partage des images et éprouvent beaucoup plus de difficultés avec d'autres pratiques, par exemple pour remplir des formulaires... Au-delà de tout jugement de valeur, la question est de savoir si l'on sait faire des choses avec un outil, si l'on est capable de paramétrer une application pour réceptionner l'information, l'analyser et la retransmettre. On estime que, dans notre pays, entre 11 et 16 millions de personnes sont en situation d'illectronisme ou considérées comme "éloignées" du numérique. Si les chiffres varient selon les méthodes de calcul, il est important de considérer que nous sommes potentiellement tous

¹ L'illectronisme ou inhabileté numérique est la difficulté, voire l'incapacité, que rencontre une personne à utiliser ou créer des ressources numériques en raison d'un manque ou d'une absence totale de connaissances à propos de leur fonctionnement (Wikipedia).

concernés par cette accélération du numérique et, en particulier, par les problèmes de cybersécurité : nous allons être tous de plus en plus attaqués sur nos messageries, sur nos comptes bancaires, etc.. Nous sommes donc face à un paradoxe : les français ne sont pas démunis en termes d'équipement, mais en termes d'usage, tout va très vite. Nous courons derrière un train en marche qu'il faut rattraper avec une complexification progressive des systèmes informatiques et numériques.

Qu'est-ce qui change, dans cette situation, avec l'irruption de l'intelligence artificielle ?

J'ai une approche très sociotechnique de la chose, c'est-à-dire qu'il faut distinguer les dispositifs techniques, technologiques, des structures sociales, économiques et politiques dans lesquelles ils sont inscrits. Je m'intéresse donc beaucoup aux usages d'un côté, mais aussi aux discours et aux représentations qui accompagnent la technologie.

Quand une technologie nouvelle arrive, on a beaucoup de représentations qui sont très ambivalentes, c'est normal. Mais quand on parle d'intelligence artificielle, je pense qu'il faut commencer par se demander de quoi on parle. Hier on parlait d'algorithme, aujourd'hui on n'utilise plus ce terme. Sur un réseau social, des algorithmes de recommandations vont faire émerger des contenus et les rendre viraux, ils vont générer des réactions des personnes qui vont liker, retweeter, ça va donner de la visibilité à certains sujets, produire de l'engagement, des contenus publicitaires et donc de l'argent pour les réseaux sociaux qui les diffusent. On n'est pas du tout dans l'intelligence artificielle générative, il s'agit seulement d'un algorithme de recommandation assez classique.





Mais les algorithmes de recommandation² - un phénomène qui continue à s'amplifier - posent de sérieuses questions à la démocratie, comme on l'a vu avec les récentes élections en Roumanie. Des gens ont voté, mais le résultat du vote est-il légitime ou non ? Comment éviter l'ingérence d'autres pays sur un vote ? Ensuite, il y a effectivement toute la production de contenu via les IA génératives, textes ou images. Quand j'étais étudiant, je passais des nuits à réaliser un montage vidéo. Maintenant, on fait des vidéos de plusieurs minutes qui sont complètement hallucinantes en termes de réalisme graphique et visuel. La rapidité des progrès est bluffante. On va vers quelque chose d'intégralement nouveau. Les capacités de création sont révolutionnaires, mais il faut penser aussi aux enjeux. Une véritable formation tout au long de la vie au numérique en général, et pas seulement à l'IA, doit être mise en œuvre d'urgence.

Il y a des choses un peu inquiétantes : il y a quelques jours le gouvernement américain a pris position pour la suppression de toute régulation informatique pendant une période d'au moins dix ans.

C'est vrai qu'on n'avait pas besoin de cette élection américaine pour ajouter à la complexité de notre monde. Je ne sais pas jusqu'où ce discours assez extrême, anti-régulatoire va nous mener. Je pense que l'Europe a quand même quelques armes pour résister.

Vous disiez qu'il était nécessaire de distinguer entre les usages et la représentation de l'intelligence artificielle. Il semble qu'il y a aujourd'hui un débat entre le mythe de l'intelligence artificielle et les

² Avec l'avènement des plateformes en ligne, les opinions se polarisent sur de nombreux sujets. Les algorithmes de recommandations, qui conseillent du contenu spécifique susceptible de plaire à l'utilisateur, en sont une des principales causes. Ils tendent à enfermer les personnes dans une bulle où ils n'interfèrent qu'avec des gens qui font les mêmes choix qu'eux.

usages utiles. Comment analysez-vous ce débat ?

On le vit au quotidien. Je suis personnellement assez sceptique sur la réalité d'une intelligence surhumaine dite générale ou très forte. Je pense qu'on est effectivement dans un récit animé par des visions utopiques ou dystopiques qui tend à créer un mythe selon lequel nous devons accepter cette intelligence artificielle générale parce que de toute façon elle va arriver. Il faut bien comprendre qu'il s'agit d'un récit, un "storytelling" des grands acteurs économiques pour pousser les gens à s'engager en faisant peur aux gens qui refuseraient de le faire. C'est quand même assez fort qu'on nous dise qu'on va avoir des intelligences artificielles conscientes quand on n'arrive même pas à définir ce qu'est la conscience humaine, par exemple. On construit des systèmes, aujourd'hui, qui sont certes assez bluffants dans la manière dont ils peuvent produire des contenus, mais qui restent des modèles statistiques avant tout, pour les textes, qui restent dans des cadres très définis. Ils ne font que générer du contenu et ne produisent que des illusions de conscience.

C'est un modèle un peu hollywoodien ?

Oui totalement. C'est d'ailleurs assez classique dans l'histoire des sciences et techniques et dans l'innovation numérique. On anthropomorphise énormément les IA. Dès qu'on parle d'IA, on voit des robots. C'est assez hallucinant. On voit la figure d'un humain qui va être remplacé par un robot. Cela rappelle des mythes bien plus fondateurs comme celui du golem, qui sont la projection de l'homme dans la créature qu'il a produite. Cela fait peur donc ça fait vendre. Beaucoup d'acteurs réfléchissent par contre à des systèmes d'intelligence artificielle beaucoup plus sectorisés capables d'aider des personnes à produire mieux ou plus facilement.

Certains métiers ou emplois vont-ils disparaître avec l'arrivée de l'IA ?

Il y aura une évolution inéluctable de certaines professions comme les graphistes, les traducteurs. Mais je ne crois pas qu'on aille vers une disparition totale. Par exemple, dans les instances officielles, on fera toujours plus confiance aux traducteurs qu'à la traduction d'une machine. Par contre, presque toutes les professions utiliseront des outils avec des assistants IA intégrés. C'est déjà le cas, quand on travaille sur un ordinateur, on utilise sans même sans rendre compte. De nombreuses entreprises ont intégré de l'IA dans leurs logiciels. Faut-il refuser cette évolution ? Je pense que c'est vraiment là où l'éthique est extrêmement importante afin de nous aider à discuter cette insertion de l'IA dans le quotidien, notamment professionnel.

Ce qui me frappe, dans la préparation de ce numéro, c'est de découvrir tout le mouvement d'interrogation et de réflexion qui, semble-t-il, anime la société civile à propos de l'IA. Il y a beaucoup d'initiatives locales, à ce sujet comme les conventions citoyennes à Rennes, Montpellier et dans d'autres villes. Des initiatives intéressantes se développent aussi pour utiliser l'IA de manière intéressante. Est-ce que vous sentez à travers vos contacts et les colloques que vous organisez ?

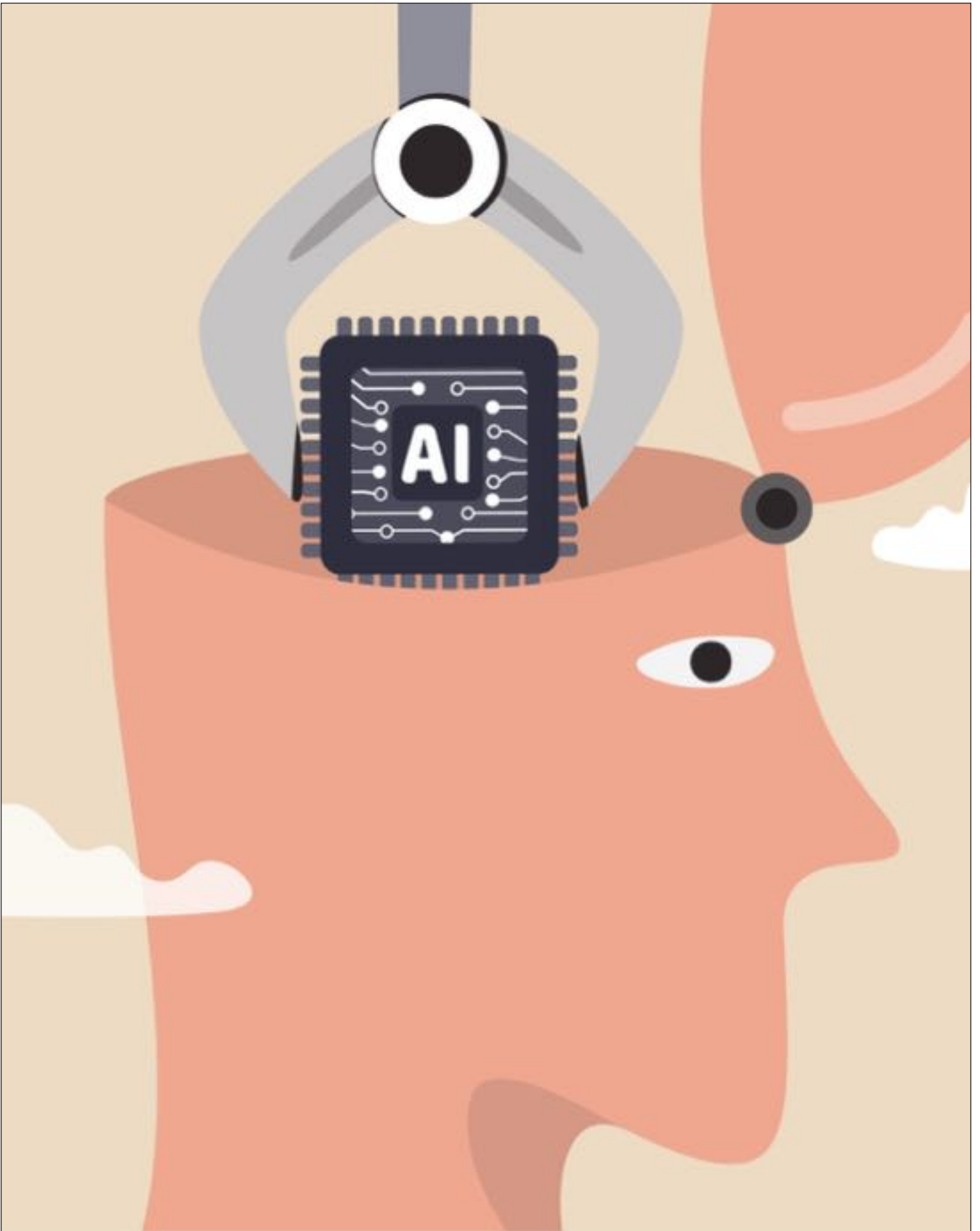
Ah oui, effectivement, nous constatons, à travers notre activité, les nombreuses initiatives qui reflètent ce questionnement sur l'IA et son insertion dans la vie quotidienne. Il faut se rendre compte du phénomène qui implique des centaines de millions d'utilisateurs quotidiens. On constate aujourd'hui que près de 60% des ados et des étudiants utilisent l'IA générative plusieurs fois par semaine pour vérifier une info, pour faire leurs travaux, leurs devoirs, etc. Il y a bien sûr beaucoup de questionnements et beaucoup de peurs autour de l'IA, mais ce qui est paradoxal c'est que ces peurs sont souvent provoquées

par les producteurs d'IA eux-mêmes qui nous vendent l'IA comme quelque chose de révolutionnaire qui va changer le monde.

En conséquence beaucoup de gens éprouvent une certaine angoisse à ce sujet et hésitent à adopter l'IA d'où l'intérêt des conventions citoyennes, qui permettent d'associer les gens à la définition de règles d'utilisation de l'IA qui seront respectées dans l'espace public ou dans l'administration du territoire. C'est très bien parce qu'on permet aux citoyens de s'exprimer, de dire à quelles conditions ils estiment qu'on peut déployer l'IA.

[Retour au sommaire](#)





LE RÉSEAU CANOPÉ : FORMER LES ACTEURS DE L'ÉDUCATION À L'ÈRE DU NUMÉRIQUE ET DE L'IA

ENTRETIEN AVEC ROMAIN VANOU DHESDEN, CANOPÉ, DIRECTEUR SCIENTIFIQUE ET INNOVATION, ADJOINT À L'ORIENTATION DE L'OFFRE DE FORMATION

PAR GUY MENANT ET DOMINIQUE BÉNARD

Bonjour Monsieur Vanoudhesden. Pourriez-vous nous présenter le Réseau Canopé et ses missions principales ?

Le Réseau Canopé est un opérateur du ministère de l'Éducation nationale, de l'enseignement supérieur et de la recherche, chargé de la formation continue des enseignants et de l'ensemble des acteurs de l'éducation tout au long de leur parcours professionnel. L'établissement propose une offre de formation diversifiée :

- *En présentiel* : grâce à 101 "Ateliers Canopé" disséminés sur l'ensemble du territoire français, en métropole comme en outre-mer (soit environ un par département), qui servent de lieux d'accueil, de formation et de partage.
- *À distance* : via des webinaires en direct (formation synchrone) ou des parcours et modules d'auto-formation (asynchrone) sur une multitude de thèmes.

Quelles sont les thématiques et les publics ciblés par les formations de Canopé ?

Le Réseau Canopé adopte une approche *transversale*, c'est-à-dire qu'il n'assure pas de formations spécifiques des différentes disciplines scolaires.

Les thèmes de formation sont conçus pour répondre aux enjeux éducatifs et sociétaux contemporains, ainsi qu'aux objectifs de formation professionnelle des enseignants.¹

Des plateformes spécifiques sont développées pour les nouveaux enseignants et les contractuels afin de répondre à leurs priorités immédiates, comme la gestion de classe. Par ailleurs, Canopé aborde des enjeux d'importance sociale ayant un impact direct sur l'éducation, tels que le

¹ Les grandes communautés thématiques du réseau Canopé sont : Bien-être à l'école ; Coéducation ; École inclusive ; Éducation à la transition écologique et sociale ; Éducation aux médias et à l'information ; Éducation aux valeurs de la République ; Numérique en éducation ; Pratiques pédagogiques et méthodologiques ; Sport à l'école.



Romain Vanoudhesden

numérique, la transition écologique et sociale, l'école inclusive.

Quel est le rôle du directeur scientifique et innovation, et comment la recherche influence-t-elle l'offre de formation ?

Ma fonction est celle de directeur scientifique et innovation, adjoint à l'orientation de l'offre de formation. Au sein de la direction stratégique du Réseau Canopé, un pôle est spécifiquement dédié à la recherche et à l'innovation. Il s'agit d'une volonté forte de la direction générale de s'assurer que toutes les formations et produits proposés par Canopé sont fondés sur deux piliers :

- les avancées de la recherche scientifique,
- les besoins exprimés par les enseignants.

Cette démarche permet de concevoir des contenus de formation pertinents, en parfaite adéquation avec les programmes et répondant aux attentes concrètes des professionnels de l'éducation.

Comment le Réseau Canopé est-il organisé et favorise-t-il l'implication et la coopération des enseignants ?

Le Réseau Canopé, avec plus de 1200 collaborateurs répartis sur une centaine de lieux, fonctionne comme une communauté apprenante et de conception. L'organisation n'est pas une structure pyramidale descendante, mais s'articule autour de grandes communautés thématiques.

- Au sein de ces communautés, des professionnels de la formation et de l'éducation travaillent en collectif.
- Le pilotage émane de la direction nationale, mais il s'appuie de manière transversale sur toutes les personnes du réseau, qu'elles soient formateurs, coordinateurs d'actions ou membres des

directions nationales. Les deux principaux objectifs de ces communautés sont la conception de formations (en lien avec la note stratégique) et la mutualisation de l'offre.

L'idée est qu'un enseignant, qu'il se trouve dans la Drôme ou dans le Finistère, puisse se voir proposer les mêmes typologies de formation par le Réseau Canopé. Ces communautés visent également à faire monter en expertise et en compétence l'ensemble des acteurs intéressés par un sujet. Ce système permet à chacun de contribuer et d'apprendre. Cette approche collective est également encouragée dans les interactions avec les enseignants.

La dimension communautaire de la formation est-elle conçue pour se retrouver dans les activités de classe avec les élèves ?

Bien que le Réseau Canopé n'ait pas pour mission officielle de garantir directement l'impact de ses formations jusqu'aux élèves, notre ambition est que les dynamiques collectives et les méthodologies appliquées en formation soient suffisamment attractives et utiles pour que les enseignants souhaitent les mettre en place dans leurs classes. Un exemple concret de cette approche est le *hackathon pédagogique*², une méthode développée par le Réseau Canopé avec des partenaires. Ce format permet de faire travailler les participants en collaboration, sans *a priori* ni hiérarchie, et avec une grande hétérogénéité des pratiques, favorisant l'émergence rapide de projets ou d'idées innovantes. Ces réalisations, bien que souvent modestes en raison des contraintes de temps, se traduisent par des applications directes en classe. Cette démarche favorise le développement de *compétences clés du XXI^e siècle*

² Hackathon pédagogique : les formations professionnelles des enseignants au et par le numérique étaient parfois trop descendantes, trop techniques et pas assez orientées vers les usages. Le hackathon pédagogique vise une production d'équipe en temps limité, par exemple 6 heures, avec une approche coopérative s'appuyant sur le principe de l'intelligence collective.

chez les élèves, telles que la coopération, la créativité, la collaboration et la communication, et contribue également à l'avancement des apprentissages.

Comment le Réseau Canopé envisage-t-il l'intégration de l'intelligence artificielle (IA) dans l'éducation, notamment la transition d'une approche individuelle vers une approche collective ?

La formation individuelle, bien qu'utile, n'est pas suffisante ; c'est au niveau de l'équipe pédagogique qu'elle peut véritablement avoir un impact transformant sur les pratiques et, *in fine*, pour les élèves. Historiquement, les outils numériques, à leurs débuts, ont eu tendance à être très individualistes. Actuellement, l'utilisation des IA génératives se fait principalement à titre individuel (via des comptes personnels ou sur des ordinateurs), et la pratique collective n'est pas encore suffisamment intégrée dans les logiciels.

Cela s'explique probablement par le fait que les besoins spécifiques de l'éducation, axés sur le "faire ensemble" et l'"apprendre ensemble", n'ont pas été suffisamment exprimés. Notre vision est de concevoir des IA génératives qui permettent d'avancer collectivement, à l'image des outils d'écriture collaborative. Intégrer une IA comme "compagnon" collaboratif ne semble pas techniquement très complexe. Une communauté dédiée à l'IA en éducation, la CREIA³, pilotée par la DNE⁴, a été lancée en 2023 pour servir d'espace d'information et de débat. Nous espérons que cette

³ En plaçant l'intelligence artificielle au cœur des réflexions et des pratiques éducatives, la CREIA (Communauté de Réflexion en Éducation sur l'Intelligence Artificielle), lancée en novembre 2023 par la Direction du Numérique pour l'Éducation (DNE), est une communauté apprenante entre pairs. Depuis le 30 août 2024, cette communauté s'adresse à l'ensemble des membres de la communauté éducative.

⁴ La direction du numérique pour l'Éducation est une direction du Ministère de l'Éducation Nationale qui a pour mission "l'impulsion et l'accompagnement de la transformation numérique du système éducatif au bénéfice de l'ensemble de la communauté éducative ainsi que des agents".



communauté permettra d'exprimer et de formaliser le besoin d'une IA pour le collectif. Des initiatives du ministère et de l'État pourraient peut-être aller dans ce sens, comme le partenariat entre les universités françaises et Mistral AI⁵, en visant à travailler sur les besoins éducatifs spécifiques plutôt que de simplement mettre à disposition un outil grand public. De même, l'Université de Rennes a développé un système RAG⁶ (*Retrieval-Augmented Generation*) accessible aux universités et bientôt à Canopé. L'idée, c'est une réflexion collective, pour tout le système éducatif, sur les besoins et productions possibles. Il est important de noter que l'IA générative est une technologie très jeune (2,5 à 3 ans) et en constante évolution.

Quelles sont les priorités de Canopé en matière de formation à l'IA pour les enseignants ?

Le Réseau Canopé travaille sur la place de l'IA dans l'éducation depuis 7 à 8 ans, en interrogeant constamment son impact. Notre position est de ne jamais prendre pour acquis les bénéfices d'une technologie numérique, mais d'analyser comment elle transforme l'éducation et vice versa. Nous avons initié nos premières formations sur l'IA générative dès janvier 2023. À ce jour :

⁵ Mistral AI est une startup française pionnière en intelligence artificielle fondée en avril 2023 par trois chercheurs visionnaires : Arthur Mensch, Guillaume Lample et Timothée Lacroix (<https://mistral.ai/fr>)

⁶ RAG Université de Rennes - <https://projet-air.univ-rennes.fr/faq-ragarenn>

- plus de 30 000 enseignants se sont formés dans nos Ateliers Canopé ;
- nos pages dédiées à l'IA ont enregistré entre 60 000 et 65 000 visites, sans compter les ressources d'auto-formation disponibles sur M@gistère⁷ et Canotech⁸.

Notre approche de formation s'articule autour d'un *trptyque* :

1. **Comprendre** : Permettre aux enseignants d'acquérir une compréhension fine de ces technologies avant même de les utiliser.
2. **Découvrir** : Explorer les multiples possibilités offertes par l'IA, mais aussi ses limites. Cette étape permet de tester l'utilité et l'acceptabilité de ces outils. L'IA dans l'éducation soulève des questions cruciales d'acceptabilité, d'éthique, de cadre réglementaire (RGPD), d'équité d'accès, et remet en question les pratiques d'évaluation mises en œuvre par les enseignants (notamment la véracité de la production des élèves).
3. **Expérimenter et Partager** : Encourager la création de communautés entre enseignants pour qu'ils testent différentes IA et partagent leurs découvertes et expériences (par exemple, dans la conception d'exercices ou de séquences pédagogiques). Nous sommes convaincus que l'apprentissage par les pairs est très efficace.

Quelles sont les principales préoccupations éthiques et les limites de l'IA abordées dans les formations ?

Les formations de Canopé visent à aborder les limites et les problèmes éthiques soulevés par l'IA, qui génèrent des inquiétudes :

⁷ M@gistère est une plateforme de formation continue du Ministère de l'Éducation nationale (<https://magistere.education.fr/>)

⁸ Canotech offre aux enseignants des formations courtes et pratiques en libre accès pour enrichir leurs compétences (<https://www.canotech.fr/>)

- **Évaluation et authenticité des productions** : L'IA rend plus complexe la vérification de l'authenticité des travaux d'élèves (devoirs à la maison, productions en classe), un problème qui existait déjà mais qui est maintenant amplifié. Cette situation est comparable à l'appréhension suscitée par l'arrivée de Wikipédia il y a une vingtaine d'années.
- **Inégalités** : L'utilisation de l'IA peut potentiellement accentuer les inégalités existantes, notamment en fonction de l'environnement familial et du soutien dont dispose l'élève.
- **Droits d'auteur** : Il est crucial de comprendre que l'IA générative s'est "nourrie" de millions de textes ou d'images, soulevant des questions sur le "pillage" du savoir-faire artistique ou intellectuel.
- **Réglementation et protection des données** : Le respect du RGPD et de l'IA Act⁹ est fondamental, en particulier pour les données sensibles de l'éducation. Il est essentiel de ne pas les laisser à la disposition d'entreprises, quelle que soit leur nationalité.
- **Hallucinations** : Les IA peuvent générer des réponses fausses ou inventées pour satisfaire une requête (comme l'exemple de l'œuf de vache¹⁰), ce qui nécessite une vigilance et un esprit critique accrus.

Comment l'IA peut-elle devenir un outil pour développer l'esprit critique et l'éducation aux médias et à l'information (EMI) chez les élèves ?

L'IA est un formidable outil pour l'éducation aux médias et à l'information (EMI). Une fois que les enseignants ont eux-mêmes développé un recul critique face aux

⁹ L'Union européenne a introduit une nouvelle législation sur l'intelligence artificielle : L'Acte sur l'IA de l'UE (IA Act). Elle jette les bases de la réglementation de l'intelligence artificielle dans l'UE (<https://artificialintelligenceact.eu/fr/>)

¹⁰ L'œuf de vache est un exemple de hallucinations de l'IA : une IA a ainsi affirmé que les œufs de vache étaient plus gros que les œufs de poules. Même si l'IA ne cesse de s'améliorer, les modèles continuent d'avoir ce type d'hallucinations.

outils numériques, ils peuvent guider leurs élèves. Il est impératif que les enseignants soient conscients de la quantité croissante de contenus générés par IA sur Internet, qui ne sont donc pas des créations humaines, et que même les créations humaines sont souvent peu sourcées ou idéologiquement orientées. Faire interroger différentes IA par les élèves permet de révéler les biais et les éventuelles problématiques (comme les hallucinations), devenant ainsi un levier pour travailler sur les notions de sources et de fiabilité. Ce travail sur l'esprit critique est un enjeu démocratique majeur, comme le souligne le Conseil de l'Europe, qui a désigné 2025 comme l'année de la citoyenneté numérique. Par ailleurs, des initiatives telles que l'IA générative française "Lucie" (développée par Linagora¹¹ et avec le consortium collaboratif OpenLLM¹²) visent à créer des IA basées sur des principes démocratiques comme l'éthique, le droit et le respect. La vigilance est d'autant plus nécessaire que les motivations de l'industrie de l'IA ne s'alignent pas toujours sur ces valeurs. Enfin, les enjeux environnementaux de l'IA doivent également être abordés : une requête textuelle via IA coûte environ dix fois plus en énergie qu'une requête de moteur de recherche classique, ce qui invite à ne l'utiliser qu'à bon escient. L'objectif global est d'aider les enseignants à accompagner les jeunes dans un monde plus complexe, marqué par des enjeux sociétaux majeurs, un monde avec lequel ils sont plus connecté et qu'ils ressentent plus que les générations précédentes

Comment le Réseau Canopé collabore-t-il avec d'autres acteurs pour développer des solutions éducatives ?

¹¹ Linagora est une entreprise qui s'est donnée pour mission d'inventer et développer des logiciels libres et open source éthiques pour participer au changement du monde avec un impact positif maximal pour les humains, la société et la planète (<https://linagora.com/>).

¹² OpenLLM : communauté dont l'objectif est de collaborer au développement de modèles de LLM #IA souverains et véritablement Open Source (<https://huggingface.co/OpenLLM-France>)



La collaboration est absolument primordiale pour faire émerger des solutions viables et réalistes. Le Réseau Canopé établit des partenariats à plusieurs niveaux :

- *Sur des thématiques spécifiques* : par exemple, pour l'éducation à la transition écologique et sociale, Canopé travaille avec des organismes comme l'*Office for Climate Education* (OCE).
- *Au niveau local* : très ancré dans les territoires, Canopé collabore avec les enseignants et tous les représentants de l'Éducation nationale, les centres sociaux, les associations, et les collectivités territoriales (en charge de la politique de la ville et de la jeunesse). Les Ateliers Canopé servent de lieux de rencontre neutres où ces partenaires peuvent dialoguer et partager leurs problématiques.
- *Avec le monde de la recherche et de la technologie* : Canopé organise des rencontres entre chercheurs et enseignants, afin que les

problématiques quotidiennes des enseignants puissent inspirer des sujets de recherche, et que les enseignants eux-mêmes deviennent des acteurs de la recherche. La collaboration avec les entreprises technologiques est également essentielle, car elles sont des apporteurs de solutions. Mon rêve pour les Ateliers Canopé est qu'ils deviennent des lieux où l'ensemble de la société civile intéressée par l'éducation puisse se rencontrer et travailler ensemble.

Les solutions individuelles, aussi brillantes soient-elles, manquent souvent de réalisme face aux contraintes économiques, financières ou au temps de formation limité des enseignants. Un travail collectif et un soutien politique sont indispensables pour relever les défis sociétaux actuels (comme le changement climatique, les atteintes à la biodiversité, les menaces sur la démocratie) et pour former des citoyens pleinement capables de s'y adapter.

[Retour au sommaire](#)



L'HOMME EST LA MESURE DE TOUTE CHOSE*

PAR MICHEL SEYRAT

* Protagoras (490-420 av. J.-C.)

L'humanité progresse en transmettant pensées, savoirs, techniques, en imitant et en inventant, avec un effet cumulatif qui s'accélère sans arrêt.

PERMANENCE :

Des peintures rupestres à l'art urbain, de la trace de main à la signature électronique, les mêmes gestes intentionnels perdurent. A l'apprentissage par cœur et au calcul avec les doigts et des pions, s'est ajouté l'écriture - « cette tard venue » disait C. Levi-Strauss (1908-2009) - écriture qui a transcrit et fixé les lois, les calculs, les traités, les exploits, les mythes et la littérature. Des aèdes chagrins ont craint que, cessant d'apprendre par cœur, les jeunes ne perdent la mémoire... Or, on apprend toujours, on chante encore, les légendes sont immortelles, l'écriture a atteint chaque humain et on n'a jamais autant écrit !

ACCUMULATION :

Croissant toujours, le savoir humain exige un stockage de plus en plus volumineux; les quelques livres de la bibliothèque humaniste de Sélestat, attachés par une chaîne pour éviter leur vol, sont devenus les monumentales bibliothèques modernes. La connaissance du monde, de l'infiniment petit à l'infiniment grand, devenant exponentielle, des moyens nouveaux se sont créés. A grande allure, tout l'écosystème informatique a amplifié sa place dans la recherche et dans le stockage des connaissances. L'IA gagne, dit-on, un point de QI par semaine... si le QI a un sens !

COEXISTENCE :

Face au déferlement de la nébuleuse IA toujours plus rapide, on se demande ce qui survivra à l'ouragan. En réalité, les moyens de communiquer coexistent durablement : écriture et oralité, peinture et photo, audio-visuel et lecture, SMS et cartes postales, Datacenter et bibliothèques, etc.





« La Marche de l'Humanité » (1971), murale de David Alfaro Siqueiros, Polyforum Cultural

Chaque moyen nouveau a suscité le même genre d'inquiétudes que l'IA aujourd'hui : les robots supprimeront le travail, le chat remplacera les relations humaines, le langage s'appauvrira, la pensée se sclérosera, l'information sera faussée... N'a-t-on pas dit naguère que la lecture déformerait l'esprit des femmes et que le cinéma abrutirait le monde ?

HUMANITUDE :

C'est toute la richesse humaine accumulée qui alimente l'IA et, en retour, l'IA démultiplie les moyens de la connaissance humaine. Ce qui nous ramène à Protagoras cité en titre : sans humain, pas d'IA ! Ce qu'on sait du fonctionnement du cerveau a inspiré la façon d'apprendre aux machines et ces machines décèlent maintenant les dysfonctionnements d'un cerveau malade ! L'IA lit l'imagerie médicale plus vite que le médecin. Sa puissance informationnelle a décrypté des textes carbonisés par le Vésuve, etc., etc. Voilà qui est bon, mais tout cela nous effraie aussi, au point d'user d'un vocabulaire anthropomorphique faisant de l'IA une sorte de double

humanoïde. L'IA semble se couler dans l'humain de l'homme, dont on sait qu'il peut être savant, sage, saint, héros, salaud, ange, démon, et coetera. Tout comme l'IA...

DOMINATION :

Présentement, ce que nous maîtrisons mal et qui nous effraie c'est la vitesse à laquelle s'amplifie l'activité et la puissance des usages de l'IA, accélération qui affecte autant la fabrication de mensonges que de vérités ! Et le pouvoir apparemment démesuré qu'elle donne à ceux qui la maîtrisent. Un totalitarisme nouveau semble s'imposer. Autrefois déjà, les grands féodaux qui avaient la force et l'argent, utilisaient les scribes et les clercs pour asseoir leur pouvoir. Pour l'instant, quelques modernes féodaux possèdent cette puissance, mais il appartient à l'humanité de garder le pouvoir sur ce produit de son ingéniosité collective et de le réguler pour le bien commun. Car ce n'est pas le moyen qui est pernicieux, mais le contenu qu'on lui confie, n'en déplaît aux libéraux délirants qui crient partout liberté pour les loups libres au milieu de moutons libres !

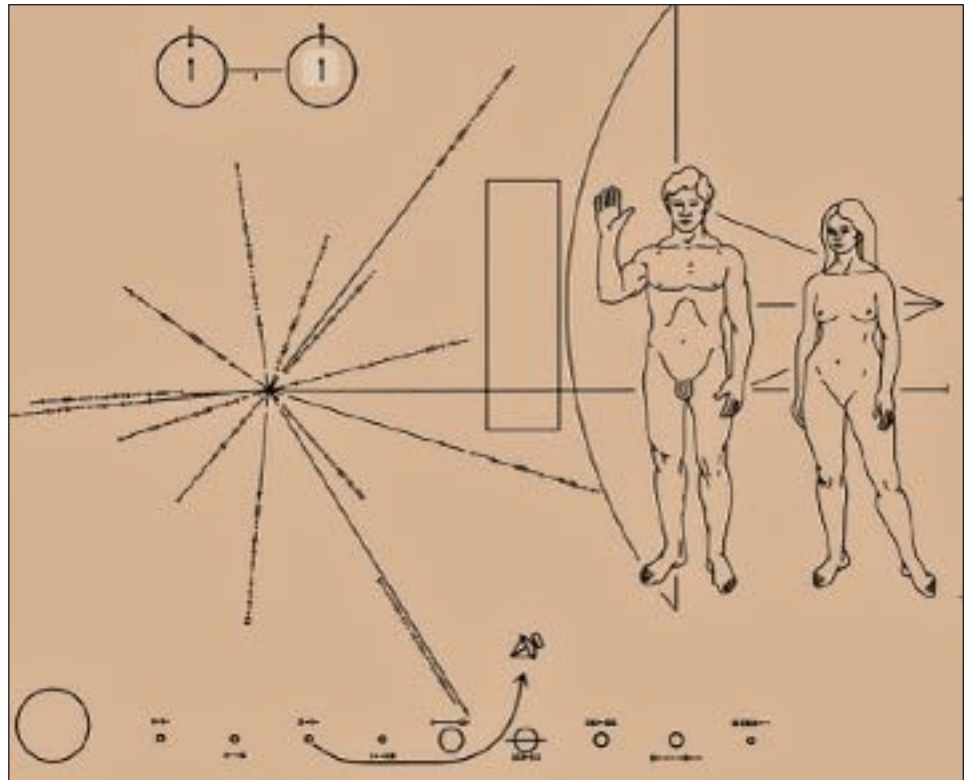
RÉVOLUTIONS :

Éric Salobir, consultant du Vatican sur les questions technologiques, souligne qu'il faut éviter que l'IA diminue l'agentivité humaine, la capacité qu'ont les hommes de décider et d'agir sur la société (traduction de agency). Seuls les humains peuvent transformer une innovation en progrès personnel et collectif. C'est notre enjeu. La Révolution a remodelé une société millénaire en quelques années. Les révoltes des canuts lyonnais des années 1830 ont initié les organisations ouvrières, l'après-guerre a imposé le concept de sécurité sociale, les

conflits mondiaux ont engendré l'ONU, l'arme nucléaire est muselée... tout cela après de rudes combats ! Face à tout pouvoir démesuré, la capacité à imposer des règles morales et sociales est la marque même de l'humanité, collectivité responsable. Science sans conscience n'est que ruine de l'âme écrivait déjà Rabelais, Raphaël Enthoven déclarait récemment que quand la technique galope, il ne faut pas que l'éthique boite ! Si les maîtres de l'IA la veulent sans lois ni règles, l'intelligence collective se soulèvera, et l'on entend déjà la révolte qui gronde : on traque les mensonges, des starteupes grignotent les multinationales, les Tesla restent au garage et nos amis étasuniens se remettent à défiler ! L'IA a répandu le meilleur instrument de la révolte : les aïphones sont partout.

EDUCATION :

C'est pourquoi plutôt que de maudire ou d'interdire, il faut apprendre à être les plus malins ! Les Gafames se répandent, comment les repérer et s'en servir sans dommage ? Facilitons la fréquentation de nos semblables et nous modèrerons nos écrans. Les ados ont bon dos, mais où ont-ils droit de cité ? L'IA répond sans relâche à toute heure, mieux vaut donc apprendre à bien s'en servir, à s'en méfier, à la maîtriser. Quand du nouveau chasse l'ancien, c'est l'éducation qui sauve : on a formé les porteurs d'eau à être plombiers et les charrons à être mécanos ! Pour apprendre à voir plus large qu'un écran, encore faut-il que soient accessibles les possibilités d'élargir son horizon. Chaplin nous le dit à la fin des Temps Modernes : après un songe merveilleux, Charlot et La Gamine reprennent, à l'aube, main dans la main, la route de l'espoir.



La plaque de Pioneer est une plaque métallique embarquée à bord de Pioneer 10 et Pioneer 11, deux sondes spatiales lancées respectivement en mars 1972 et avril 1973. Sur cette plaque, un message pictural de l'humanité est gravé à destination d'une éventuelle intelligence extraterrestre : un homme et une femme représentés nus, ainsi que plusieurs symboles fournissant des informations permettant de localiser le Soleil et la Terre, à l'origine des sondes.

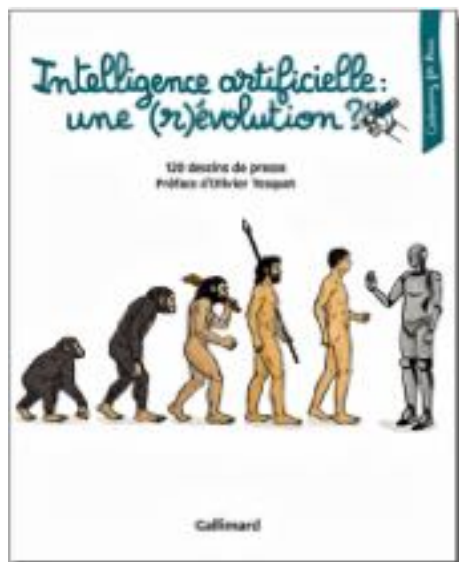
CONTRE PROTAGORAS :

Avant de tourner la page de ce bloc-notes, je voudrais faire un aveu. J'apprécie la phrase de Protagoras, mais désormais elle me chagrine : dominant la Terre, l'homme en est devenu un prédateur insatiable, dévorant, dégueulassant, massacrant, y compris ses semblables sans aucune retenue ! Il faut agir pour que cette mesure de toutes choses parvienne à se mesurer en toutes choses ! Mais c'est pas gagné !

Et pour finir, je me dis que si puissant que l'on soit, le vieux syllogisme demeure : tous les hommes sont mortels, or Elon Musk est un homme, donc Elon Musk est mortel. Même s'il rêve d'être éternel ! Imaginerait-on Hitler immortel ou Poutine bicentenaire. Non ? Ouf !

[Retour au sommaire](#)

RECOMMANDÉ PAR APPROCHES COOPÉRATIVES



INTELLIGENCE ARTIFICIELLE : UNE (R)ÉVOLUTION ?

120 DESSINS DE PRESSE

PRÉFACE D'OLIVIER TESQUET

COLLECTION CARTOONING FOR PEACE

GALLIMARD LOISIRS

« L'INTELLIGENCE ARTIFICIELLE EST PARTOUT. CERTAINS LUI PRÊTENT LA CAPACITÉ DE CHANGER LA CIVILISATION, À L'INSTAR DE L'ÉCRITURE OU DE L'INFORMATIQUE ; D'AUTRES S'INQUIÈTENT DE SON INCROYABLE POUVOIR DE DESTRUCTION. MAIS DE QUOI PARLE-T-ON ? »

Le spectaculaire succès du robot conversationnel ChatGPT, lancé fin 2022, a fait découvrir au grand public un échantillon des possibilités offertes par l'intelligence artificielle (IA) : pour le meilleur et pour le pire ? La prise de conscience de l'impératif de fixer un cadre au développement de l'IA semble s'accélérer ces derniers temps. En effet, cette avancée technologique soulève de vastes questions éthiques, juridiques, politiques, économiques... humaines, en somme, puisqu'elle bouleverse toutes les sphères de nos vies.

Pour cet ouvrage, préfacé par le journaliste Olivier Tesquet et publié en partenariat avec Amnesty International, 120 dessins de presse du monde entier ont été sélectionnés par Cartooning for Peace – dont le réseau de 319 dessinatrices et dessinateurs s'étend à plus de 70 pays sur tous les continents – afin de saisir les enjeux au cœur du déploiement de l'IA, toujours avec humour et... intelligence !

LES DROITS D'AUTEUR DE CE LIVRE SONT REVERSÉS À CARTOONING FOR PEACE POUR SOUTENIR LES DESSINATEURS DE PRESSE MENACÉS. : <https://www.cartooningforpeace.org/>



ONT PARTICIPÉ À LA RÉALISATION DE CE NUMÉRO :

MEMBRES DU COMITÉ ÉDITORIAL

Dominique Bénard, Larry Childs, Roland Daval, Francis Jeandra, Guy Menant, Dante Monferrer, Michel Seyrat.

CONTRIBUTEURS EXTERNES

Sylvestre Bénard, Léo Briand, Laurent Butré, Jean Cattin, Frank Debos, Florian Gauthier, Pierre Jannin, Jacques Jaray, Jean-François Lucas, Antoine Maier, Tom Murray, Romain Vanoudheusden

SUR LE SITE WEB DE L'APAC

<https://www.approchescooperatives.org/>

Vous pouvez :

- Consulter à l'écran et télécharger gratuitement toutes nos publications en format numérique : la revue périodique trimestrielle « Approches Coopératives », les numéros hors-série, les outils de formation (présentations audio-visuelles).
- Commander en format papier les numéros de la revue qui vous intéressent particulièrement contre la somme de 10 euros plus le port.
- Souscrire un abonnement de soutien (60 euros port compris) à la revue « Approches Coopératives » en format papier pour la recevoir à domicile quatre fois par an..
- Adhérer à l'APAC et participer à l'orientation, la production et l'évaluation de nos publications.
- Faire un don pour nous permettre de poursuivre l'aventure d'Approches Coopératives au bénéfice du plus grand nombre.



TOUTES LES PUBLICATIONS DE L'APAC SONT SOUS LICENCE CREATIVE COMMONS

Sous réserve de la mention d'Approches Coopératives, l'APAC autorise l'utilisation de l'œuvre originale en dehors de toute exploitation commerciale.

