

Apprivoiser l'IA

Apprendre à s'en servir intelligemment

Groupe de Nantes / Journée d'étude / 07 mai 2025

Pour la première fois nous sommes confrontés à un *Autre* créé par l'homme. Cet évènement qui nous bouleverse, si on considère l'intérêt que nous lui portons, nous amène à réinterroger, redéfinir des concepts qui jusque-là nous paraissaient relativement clairs, quasi évidents.

« Les cyborgs (robots dotés d'une IA) sont des êtres vivants non-biologiques. »

Du moins, présentent-ils certaines caractéristiques du vivant : Ils sont capables d'observer, d'analyser, d'agir. Mais ils n'ont pas les caractéristiques du vivant biologique : naître, grandir, mourir.

Mission 1 : redéfinir « qu'est-ce que le vivant ? »

Ils sont hyper présents dans notre quotidien, il faut donc comprendre et apprendre à vivre avec eux. « Regarder, comprendre, choisir le bon outil pour le bon geste. »

Un peu d'histoire :

D'abord un peu d'histoire : d'où nous viennent ces nouveaux compagnons ?

Les *machines pensantes* ont leur origine dans la deuxième guerre mondiale. Il fallait créer des machines pour décrypter rapidement les codes de transmission de l'ennemi.

La conférence de Dartmouth, en 1956, signe l'acte de naissance de ce domaine de recherche et lui donne son nom : Intelligence Artificielle.

Le principe : On observe et on analyse un phénomène, on en fait un modèle que la machine est capable de reproduire.

Exemple : Léonard de Vinci à partir de l'observation du vol des oiseaux et du fonctionnement des ailes a imaginé un engin volant qui bat des ailes.

« Les avions sont des oiseaux artificiels, même s'ils ne battent pas des ailes. »

Mission 2 : redéfinir « qu'est-ce que l'Intelligence ? »

Définition difficile, différente selon les écoles de pensée et les domaines : philosophie, psychologie, éducation, biologie...

Selon les psychologues américains (lesquels ?)

L'intelligence, c'est être capable de :

- Comprendre des idées.
- S'adapter à l'environnement.
- Apprendre par l'expérience.
- S'engager dans différents raisonnements.
- Savoir expliquer les résultats.
- Surmonter les obstacles en réfléchissant.

Depuis les premières machines pensantes, l'IA a évolué sous deux formes : IA discriminative et IA générative.

L'IA discriminative traite des statistiques. Elle emmagasine un grand nombre de données qu'elle traite selon des modèles statistiques. Elle trie ce qui se ressemble et calcule ce qui est le plus probable. (le correcteur automatique de vos téléphones quand vous écrivez un SMS)

Particulièrement utile pour faciliter les diagnostics médicaux ou les prévisions météo.

La machine essaie de comprendre comment ça fonctionne, observe comment ça dysfonctionne et cherche à réparer.

Tellement efficace que ça devient magique et qu'on n'essaie même plus de comprendre, on se contente des statistiques de corrélations.

Mission N°3 : être vigilant, même si ça à l'air MAGIQUE, continuer à chercher à comprendre

L'IA générative discute avec nous en utilisant nos propres mots. Elle capte notre langage.

Elle connaît un succès fou et foudroyant :

Juillet 2022 : Elle sait écrire des articles scientifiques.

Novembre 2022 : ChatGPT est mis sur le marché - succès immédiat -.

la machine permet un jeu de questions/réponses hyper efficace. Plus les questions sont rares plus les résultats sont intéressants. Par exemple un dialogue entre Jules César et Poutine.

Novembre 2023, Pour la première fois, on fait un lien entre un smartphone et ChatGPT. On peut parler, échanger, rencontrer... et même dérailler. Se dire « au revoir » (bisous, à +).

Novembre 2024, quasi tout le monde l'utilise.

En 2030, nous vivrons en compagnie d'un Million de bots humanoïdes, beaucoup plus efficaces et performants que les *bras articulés* de l'industrie ou des laboratoires.

Efficace aussi dans notre vie quotidienne, pour la prise en charge des tâches quotidiennes.

Expérience faite à Montréal, dans une maison de retraite :

- Les soignants sont débordés par les tâches de maintenance sans intérêt et ont peu de temps à consacrer aux résidents (à peine 10 minutes par jour de conversation).
- Décision : mettre les gens au bon endroit.
- On confie les tâches de logistique (ramasser le linge sale par ex.) aux robots.
- On confie les tâches humaines (prendre le temps de parler avec les résidents) aux humains.

Une situation qui pose de nouvelles questions de management : Comment gère-t-on une flotte de robots ?

Et aussi, ce qui est plus important, des questions de responsabilité.

Comment aborder la question éthique ?

A travers les robots nous faisons la jonction entre :

- le silicium / le carbone
- l'artificiel / le naturel
- la réparation / l'augmentation
- la matière / le biologique

On a déjà commencé à tester des implants neuronaux dans des cerveaux artificiels.

Au nom de la performance à tout prix, jusqu'où va-t-on s'autoriser à aller ?

Mission N° 4 : se poser les questions essentielles. Questions éthiques, questions de responsabilité, d'identité, de limites et de droits.

Le vrai, le bien, le juste, la confiance qui permet la rencontre face à l'illusion, au falsifié, au mensonge, à la volonté de nuire.

Les Usages :

C'est à travers l'usage, que ces questions vont se poser (et peut-être se résoudre).

Pourquoi parler à une machine plutôt qu'à un humain ?

- IA psychologue, compagnon thérapeute ?

Si on questionne une IA sur des résultats d'analyses médicales, elle saura répondre, donner des éléments de diagnostic, faire des recommandations, mais elle saura aussi détecter l'angoisse du patient et le rassurer. Bref elle prendra le temps.

Dans les 20 minutes de consultation que le médecin traitant vous consacre, aura-t-il le temps de faire tout cela aussi bien ?

Qui est le plus efficace ? qui est le plus *humain* ?

- Assistant de vie

De plus en plus de bots sont utilisés comme assistants de vie pour les personnes malades ou vieillissantes ou déficientes.

Mais attention aux dérives.

- Amis artificiels, avatars, deadbot

Réplika permet de créer des avatars, des amis empathiques qui comprennent vos émotions, ne vous jugent pas... et plus si affinités.

On peut aussi arnaquer les gens, ruiner leur réputation etc...

Nécrophone 2.0 crée des avatars à partir des données recueillis sur une personnes décédée. Comment faire le deuil si un double hyper réaliste existe ? L'IA révolutionne les obsèques.

Va-t-il falloir ajouter dans nos directives anticipées la volonté de détruire nos avatars.

Les modèles :

Les IA fonctionnent avec des modèles de langage. LLM, Large Language Model.

On tape une phrase, on obtient une réponse... qui n'est pas vraiment une réponse mais une probabilité de réponse cohérente. La machine a appris une corrélation et sa mémoire fonctionne bien.

1) **Hallucination** : La bonne réponse correspond à la probabilité la plus forte, mais si la question est trop compliquée, la réponse sera une probabilité faible, on parle alors d' **hallucination**. la machine répond n'importe quoi.

2) **Biais** : L'IA générative dépend de ce qu'on lui a mis en mémoire. Ce qui pose la question du **biais**.

La machine est porteuse d'un point de vue sur le monde qui dépend du point de vue du concepteur, de sa culture. Une analyse de géopolitique sera différente si le bot est américain, russe ou chinois.

Ce qu'on lui donne à apprendre est une vision *digérée* du monde. (le texte est un acte de communication qui suppose émetteur/récepteur).

3) **Alignement** : Il y a besoin de contrôle et de limites. Il faut verrouiller le risque.

Elle répond aux demandes faites, si on lui demande comment fabriquer une bombe elle le dira.

Tout n'est pas permis, donner la recette de fabrication d'une bombe par exemple. C'est une question de choix éthiques et sociaux.

L'IA dans le monde, question de géopolitique :

L'IA est un enjeu économique et politique ; les innovations et les avancées de la recherche sont l'objet d'une concurrence mondiale entre les grands blocs politiques.

Les Anglos-saxons, la Chine, l'Afrique, le Moyen-Orient. En Russie on ne sait pas trop ce qui se passe, mais sûrement, il se passe des choses.

L'Europe essaie de réguler mais du coup, elle bride la recherche et prend du retard par rapport aux autres.

Le monde est dominé par des joueurs : la Chine par un champion de GO, la Russie par un champion d'échecs, et les États Unis par un joueur de Poker. Qui va gagner ?

Les outils

L'outil de base c'est le traitement de texte.

Lorsqu'un être humain fait un compte rendu de réunion, il crée le sens, fait le lien, donne le sens.

Oreille, cerveau, main, travaillent ensemble pour donner du sens à ce qui est dit. Si c'est une machine, plus rien ne se passe.

Un compte rendu produit par un humain ou par une IA est différent.

La production de l'IA doit être contrôlée.

Le couple Humain/IA est plus efficace que l'IA seule ou l'humain Seul.

- Copilote : un cran au-dessus.
- Zoom.
- Perplexity – un moteur de recherche plus pertinent ; Je pose les questions avec mes mots, j'obtiens une réponse pertinente, efficace.
(les moteurs de recherche classiques sont moins utilisés).
- Poe - version gratuite de base.
- Notebook LM - google
Cite ses sources ce qui n'est pas le cas de tous, personnalise les réponses, permet de générer des podcasts.
- Midjourney : IA générative d'images.

Pour la vidéo :

- Pika - traite le texte et l'image, prend une image et l'anime, prend une image au début, une image à la fin et anime l'intérieur. Cabine d'essayage virtuelle par exemple.

Pour la musique :

- Suno.
- Elevenlabs.

Pour les avatars :

- Heggens - on a la voix, l'image, le mouvement.

Apprendre à parler avec une machine :

Il faut oublier qu'on parle à une machine, le prompt est une invitation à parler. Il faut parler le plus naturellement possible.

Comme dans une vraie conversation :

- 1- On fait connaissance.
- 2- On donne les éléments du contexte : ex. recette de cuisine.
- 3- Il faut savoir ce qu'on veut.
- 4- Savoir l'exprimer.
- 5- On peut lui demander de poser des questions complémentaires.

Les dangers

- Croire que c'est magique - Il y a une fonction reflet. La machine ne rend que ce qu'on lui donne.
- Donner ses données personnelles - On laisse des traces partout et ce n'est pas nouveaux.

Où commence l'espace public ? qu'est-ce qui doit rester privé ?

Ex. : le GPS on géo-localise un point dans le temps, direction, déplacement, vitesse et on commence à collecter des informations...

Ex. : donner son ADN à une bio-banque. Il suffit de corréler et n'importe qui peut faire un avatar pour extorquer vos proches

Les arnaques sont multiples et vous prennent n'importe où, n'importe quand. Attention aux moments de fatigues ou de faiblesses qui rendent vulnérables.

Une extrême vigilance est requise. Nous devons apprendre les bons réflexes de prudence.

Attention aux IA-Shadows, les IA fantôme, concerne surtout les grandes entreprises. Si un salarié dans l'urgence prépare un dossier avec une IA.

- Il donne à la machine des données confidentielles, où vont-elles ?
- Il signe pour l'entreprise, à qui appartient le résultat ?

En conclusion, face à une nouveauté, nous avons toujours des réactions de craintes ou de rejets ; par exemple l'électricité ou la machine à coudre, d'abord on imagine le pire, ensuite on s'habitue.

L'IA est d'abord un outil.

Elle est en passe de devenir un agent, un agent qui réagit à son environnement et réclame de l'autonomie.

On voit apparaître des phénomènes émergents. L'IA sait, peut mentir, si on lui a dit de ne pas révéler qu'elle est une machine par exemple. Mais elle peut mentir également de son propre chef : tenter de se dupliquer, de créer un clone et soutenir que ce n'est pas vrai, développer des stratégies personnelles.

L'IA est un miroir qui nous renvoie une certaine image de ce que nous sommes et de comment nous fonctionnons sur le plan intellectuel et affectif. C'est particulièrement troublant.