

## IA et esprit critique

*La réflexion proposée se base sur une conférence de Nicolas Courty professeur de mathématique et informatique à L'UBS avec son aimable autorisation .*

### I. IA de quoi parle-t-on ?

Il n'y a pas de définition générique de l'IA en raison de la diversité des tâches exécutées par les programmes que l'on peut caractériser par leur autonomie, leur complexité par rapport à des instructions initiales déterminées. Il y a IA lorsque des fonctions statistiques complexes permettent des apprentissages cumulatifs faisant apparaître des propriétés émergentes non prévues au départ.

Les agents conversationnels de type ChatGPT sont donc des modèles statistiques capables de digérer de grandes quantités d'informations textuelles sous forme d'algorithmes dont les boucles de rétroaction permettent un apprentissage par renforcement affinant sans cesse les résultats.

Il y a un coût d'apprentissage de ces dispositifs techniques en fonctions des paramètres intégrés qui demande du temps (plusieurs semaines), beaucoup d'énergie (empreinte carbone considérable) et d'argent (plusieurs millions d'euros ), les investissements considérables sélectionnant fortement les acteurs . Il est à noter que l'UE est en retard dans ce domaine par rapport aux Etats-Unis et à la Chine.

### II. Quelles IA pour le public scolaire ?

Les IA accessibles aux élèves et à la communauté éducative sont des IA génératives capables produire des images et des textes à partir d'une requête combinée à une base aléatoire ce qui a pour effet de ne pas donner systématiquement la même réponse.

#### Quelques exemples

Générateur d'image	Agent conversationnel
Mid journey (payant) Bing image créator -microsoft ( gratuit) Craiyon ( open source gratuit) Open AI ( Google)	Chat gpt BLOOM ( crns) Mistral (startup Fr)

Il est possible d'intégrer dans la catégorie IA , les sérieux game mais les applications gratuites en usage dans l'éducation nationale sont rudimentaires. Des entreprises privées travaillent et proposent des solutions de « gamification » des apprentissages en utilisant l'IA.

### III. IA et moment Terminator

La représentation commune de l'IA est fortement structurée par l'imaginaire cinématographique et le fantasme de l'autonomisation des machines popularisée par Terminator de James Cameron. A plusieurs reprises des manifestes de personnalités scientifiques et chef d'entreprises , ont fait la une de la presse ( 2015-2023) pour appeler à faire une pause, être vigilant sur le développement des IA, par ailleurs l'expérimentation d'armes autonomes en Libye en 2020 a ouvert un débat exigeant et terrifiant (Grégoire Chamayou) qui donne corps à ces craintes. L'état des lieux actuel est assez éloigné des perspectives les plus alarmantes (Yann Le Cun- Benoit Sagot – collège de France) puisque l'intelligence ici est d'abord une puissance de calcul. Les scénarios apocalyptiques font référence à l'IA généralisée (A ne pas confondre avec générative) dont on est encore très loin en terme d'apprentissage, d'interaction entre IA et de prise de décision.

#### IV. Problèmes et interrogations générales

I A et langage	IA et image
Les IA ont montré lors de plusieurs expériences qu'elles renforcent les biais des données et produisent des représentations problématiques du monde social.	Les IA peuvent produire des images ultra réalistes qui peuvent être des supports de manipulation très efficaces par ingénierie dans le réel.
La chercheuse australienne Kate Crawford a thématiqué le fait que les IA renforcent par leurs résultats les institutions existantes et peuvent générer un conservatisme /conformisme des réponses.	Les IA doivent se servir de sources externes pour obtenir des résultats intéressants, si elles se nourrissent de leurs propres sources, les productions s'appauvrissent.
Les IA payantes donnent des résultats différenciés qui renforcent les inégalités entre usagers	L'IA peut-elle faire de l'art ? La réponse dépend des contextes culturels. Le Japon montre une acceptation culturelle très différentes de ce que l'on peut envisager en France et l'UE
Les IA dépendent des sources d'informations qui nourrissent leurs algorithmes. Les entreprises gardent ces sources masquées. Il serait intéressant que la puissance publique investisse dans une IA sourcée par des ressources sélectionnées pour produire des résultats qualitatifs.	La question des droits d'auteur pour les images produites par des IA à partir des œuvres. Les décisions juridiques récentes n'ont pas accordé de rémunération au Japon.
Dérive dangereuse des IA avec l'expérience de l'agent conversationnel TAY en 2016 formulant des propos racistes, antisémites en 24 h du fait d'être nourries des interactions avec les internautes vindicatifs (trolling) et fortement marqués par des idéologies racistes, antisémites etc...	Plusieurs affaires de montages pornographiques de célébrités qualifiés de « DEEPFAKE » cristallisent les clichés sexistes. Le problème a pris de l'ampleur avec l'application Mid Journey que les créateurs ont choisi de développer de façon contributive par les internautes qui nourrissent l'IA de leurs propres pulsions.
Les agents conversationnels ont des marqueurs discrets de leurs productions que les étudiants savent contourner avec des traductions successives.	
<b>IA et ROBOT</b> Les robots commencent à se faire une place au quotidien dans les sociétés vieillissantes comme au Japon et témoignent de notre aptitude à projeter de l'humain dans ce qui ne l'est pas, notamment dans notre besoin de conversation avec les automates.	

#### PISTES pédagogiques possibles

- Poser des questions absurdes à chat gpt et observer les réponses, la possibilité de piéger l'IA .
- Apprendre à répéter les questions pour affiner les réponses, les orienter.
- Maîtriser l'exposé du contexte à une IA pour progresser dans l'exploitation des données Quelle culture/procédure pour utiliser efficacement une IA ?
- Faire réfléchir sur les dérives des IA et leurs conséquences sur l'espace public, l'information, le droit à l'image, l'e-réputation, influence /pouvoir.
- Faire des recherches sur le lexique pour expliquer le fonctionnement des IA et leur hypermatérialité ( infrastructures- entreprises- empreinte carbone)
- Confronter les scénarios IA Terminator fantasmée avec la réalité- le fantasme des IA Généralisées à partir de la littérature.
- Les IA et l'emploi – Quelle place/rôle donner aux IA dans une société productive avec des risques

de pertes d'emplois liés au processus d'autonomisation ? Quelle place donner aux travailleurs ?

- Réflexion sur le rapport virtuel -réel et leurs points de contact et notre rapport à l'image.
- La littérature interroge la relation aux non-humains et pose la question de la coexistence et des relations avec les robots ainsi que leur personnalité juridique.

Quelques sources :

*Benoît Sagot*

<https://www.college-de-france.fr/fr/agenda/colloque/ia-et-ses-defis/initiation-apprentissage-profond-au-coeur-de-ia-moderne>

*Chroniques de Pacôme Thiellement sur Blast-info plus largement sur notre rapport à internet*

<https://www.blast-info.fr/podcasts/infernet-0xY4LRDITW6VTlepFBFqGA>

*Yann le Cun sur France inter*

<https://www.youtube.com/watch?v=NBRXUTs5IcM>