

L'IA ou le « Malin Génie »

Philippe DUPEYRAT

IA-IPR Mathématiques

Janvier 2025

Pour analyser notre position vis-à-vis de l'IA et de son impact dans nos vies, nous pourrions convoquer – une fois n'est pas coutume – nos classiques étudiés au lycée. Aujourd'hui, une vidéo peut montrer un chef d'État dire ce qu'il n'a jamais dit. Une photo peut immortaliser un événement qui n'a jamais eu lieu. Ce n'est plus de la science-fiction, c'est notre quotidien.

Au XVII^e siècle, Descartes avait imaginé une expérience de pensée étrange : et si un « Malin Génie » passait son temps à nous tromper en créant des illusions autour de nous ? Aujourd'hui, ce génie porte un nom : l'Intelligence Artificielle.

Cependant, ce doute cartésien n'est pas une simple méfiance passive ; c'est un effort de l'esprit, une discipline exigeante. Le défi aujourd'hui est que l'IA, en tant que Malin Génie moderne, s'appuie sur notre propre paresse cognitive. En nous offrant des résultats immédiats et séduisants, elle nous dispense de l'effort de réflexion. L'illusion réussit d'autant mieux que nous lui déléguons notre pensée par confort. Le doute devient alors un luxe que nous n'avons plus le courage d'exercer face à l'automatisme des réponses toutes faites.

1. La fin du « Voir, c'est croire »

Pendant des décennies, l'image a été notre preuve absolue. On disait : « Je l'ai vu de mes propres yeux ». Avec l'IA, ce contrat est rompu. L'image n'est plus une preuve, c'est une construction statistique. Cela nous pousse à adopter ce que Descartes appelait le doute hyperbolique : par sécurité, nous commençons à douter de tout ce qui s'affiche sur nos écrans. Si une image peut être fausse, alors toutes les images deviennent suspectes.

2. Le retour de Saint Thomas : la preuve par le contact

Puisque nos yeux nous trompent à travers les écrans, nous allons chercher la vérité ailleurs : dans le physique. C'est le retour à la mentalité de Saint Thomas, qui refusait de croire à une nouvelle incroyable sans pouvoir « mettre ses doigts dans les plaies » pour vérifier par le toucher.

Le présentiel devient sacré : une poignée de main, un regard en face à face, une discussion autour d'un café... Ces moments deviennent les seules « zones de vérité » infalsifiables. Cette demande est de plus en plus forte, et l'IA n'est pas la seule en cause. On n'en peut plus de ces visios et de ces distanciels censés faire gagner du temps alors qu'on les multiplie. Pas grave : l'IA nous permettra d'en avoir la synthèse résumée. On réinvente le mémo !

3. Une société à deux vitesses ?

C'est ici que l'enjeu devient social. On risque de voir apparaître un monde divisé :

D'un côté, une vérité de luxe : réservée à ceux qui peuvent se déplacer, se rencontrer physiquement et vérifier les faits par le contact humain.

De l'autre, une vérité par défaut : un monde numérique gratuit mais saturé de faux, où la masse vit dans une illusion permanente, faute de moyens pour accéder au "réel".

4. Le passé est-il en danger ?

Le plus inquiétant n'est pas seulement le mensonge sur le présent, mais la modification du passé. Si l'IA peut fabriquer de fausses archives historiques avec un réalisme parfait, comment saurons-nous qui nous sommes ? Sans preuves tangibles (vieux journaux papier, objets d'époque), notre mémoire collective devient un sable mouvant que les algorithmes peuvent remodeler à leur guise. De plus en plus d'enseignants se posent la question : sur quelles sources fiables puis-je m'appuyer ?

5. Solutions technologiques et humaines

Pour sortir de ce doute permanent, deux pistes se dessinent :

La piste technologique (Blockchain) : Utiliser des signatures numériques infalsifiables pour marquer les images réelles dès leur création, une sorte de "label bio" pour l'information. Problème : à qui déléguer cette certification de fiabilité ?

La piste humaine : Créer des réseaux de confiance plus petits. On ne croira plus "Internet", mais on croira le cercle restreint de personnes que l'on connaît physiquement. Le risque, déjà présent et observé dans les réseaux sociaux, est celui d'une fragmentation de la société en petits groupes et, plus grave encore, d'un enfermement dans ces groupes.

C'est la problématique de la confiance et la confiance n'est pas affaire de technique mais d'éthique. N'oublions pas que l'IA est un outil technique et commercial.

6. La leçon de Montaigne : de la vérité au divertissement

Ce basculement vers le virtuel n'est pas seulement un accident technologique, c'est aussi une fuite. Pour comprendre pourquoi nous acceptons si facilement ces "mirages numériques", il faut relire la critique du divertissement chez Montaigne. Pour lui, l'homme cherche sans cesse à s'étourdir dans l'agitation pour éviter de se confronter à lui-même.

L'IA est devenue le divertissement ultime sous prétexte d'approfondissement. On se précipite sur la dernière nouveauté, la dernière option de Gemini ou ChatGPT, non pour approfondir notre pensée, mais pour l'éviter. L'IA est un amplificateur : amplificateur d'autonomie ou amplificateur de dépendance. L'apprentissage du prompt devait nous pousser à la métacognition, à l'appropriation du questionnement socratique. Au lieu de cela, on entre trois mots d'une idée à peine ébauchée dans une machine à prompts et l'IA transforme ce balbutiement en argumentation structurée.

Plus qu'à la mettre sur Notebook LM pour faire une belle infographie et on s'en va montrer son joli dessin : « Vous avez vu comme je pense beau ! ». Rappelons-nous la mode des cartes mentales quand chacun voulait montrer sa richesse de réflexion...

La fracture entre les initiés et les usagers ignorants s'agrandit. Quand la DGESCO lance un programme sur l'IA, les 4/5 des candidats sont issus de disciplines techniques. Le lien entre IA et informatique de spécialistes se renforce au détriment de son appropriation par le quidam. Les spécialistes eux-mêmes, pour satisfaire à la demande des collègues, construisent des outils tout faits : machines à prompts, à dessins, à corriger.... Et puis, quand on aura le temps, on travaillera l'esprit critique...si on trouve le temps !!!

On le répète : il n'y a pas besoin d'être mécanicien pour être un bon pilote. Je n'ai pas besoin de connaître le fonctionnement du moteur à explosion pour bien conduire mon véhicule (surtout depuis que ma voiture est électrique...). La connaissance technique peut apporter un plus mais n'est pas l'objectif.

7. Cukurova : Sortir de l'atrophie par l'effet "miroir"

C'est là que le bât blesse. Si on continue à laisser les spécialistes construire des "boîtes noires" que l'on consomme sans réfléchir, on fonce droit vers ce que Mutlu Cukurova appelle l'atrophie cognitive. C'est un peu comme le GPS : à force de se laisser guider, on ne sait plus lire une carte. Si l'IA "externalise" tout notre travail (le plan, la correction, la synthèse), on finit par s'atrophier le cerveau.

(Je me permets de signaler l'excellente série de vidéos de JB Poignet et B. Peyras :

<https://docs.numerique.gouv.fr/docs/b569be62-90bf-4b2e-a705-73c95b123329/>. Nous avons besoin de la recherche pour prendre du recul. Les travaux de Cukurova sont évoqués dans la vidéo sur l'effet miroir)

Cukurova propose trois conceptualisations alternatives de ce que peut être l'IA en éducation, chacune correspondant à un rôle différent dans la cognition humaine.

Les trois conceptualisations de l'IA

L'Externalisation représente le modèle dominant actuel. L'IA remplace ou automatise certaines tâches cognitives humaines. Un exemple typique est le correcteur orthographique automatique ou le système de notation automatisée de copies. Dans ce modèle, certaines fonctions sont "externalisées" de l'esprit humain vers la machine. L'avantage est l'efficacité : ces tâches sont accomplies plus rapidement. Le risque, déjà mentionné, est la perte de compétence.

L'Internalisation propose un usage radicalement différent : l'IA comme outil de transformation de notre manière de penser. L'idée est que, en interagissant avec des modèles computationnels, nous pouvons développer de nouvelles représentations mentales, de nouvelles façons de structurer notre pensée. Par exemple, apprendre à décomposer un problème complexe en sous-problèmes pour qu'un programme puisse le résoudre nous enseigne une discipline intellectuelle transférable. L'IA devient un médiateur de développement cognitif, similaire au rôle que Vygotsky attribuait au langage ou aux outils culturels.

L'Extension constitue le cœur de la vision de Cukurova : l'intelligence hybride. Dans ce modèle, l'humain et l'IA forment un système couplé où chacun augmente les capacités de l'autre. Ce n'est ni l'humain seul, ni l'IA seule, mais l'émergence d'une intelligence synergique qui dépasse la somme des deux parties. Dans ce modèle, les deux agents évoluent à travers leur interaction. L'humain apprend à mieux collaborer avec l'IA, et l'IA s'adapte aux patterns et besoins de l'humain.

Illustration : Préparer une leçon sur les fractions

A. Externalisation : L'enseignant demande à ChatGPT "Génère-moi un plan de leçon sur les fractions pour des élèves de CM1". Le système produit un plan complet. L'enseignant l'utilise tel quel. La tâche de conception pédagogique a été externalisée.

B. Internalisation : L'enseignant dialogue avec l'IA : "Quelles sont les erreurs conceptuelles les plus fréquentes sur les fractions ?" "Comment la théorie de Bruner suggère-t-elle de séquencer cet apprentissage ?" En explorant ces questions avec l'IA, l'enseignant développe sa propre compréhension pédagogique. L'IA sert de tuteur socratique.

C. Extension (Intelligence Hybride) : L'enseignant et l'IA co-construisent la leçon de manière itérative. L'enseignant apporte sa connaissance fine de ses élèves spécifiques. L'IA suggère des approches basées sur la recherche. L'enseignant adapte ces approches à son contexte. L'IA identifie des contradictions dans la séquence proposée. Ensemble, ils produisent une leçon qui intègre l'expertise contextuelle humaine et la synthèse de connaissances de l'IA. Ni l'enseignant seul ni l'IA seule n'aurait pu créer cette version optimale.

Le problème, dit Cukurova, c'est qu'on nous vend C mais on obtient A....

Avec l'extension, on n'utilise plus la machine pour obtenir un résultat propre ("penser beau"), mais pour rendre visible notre propre cheminement intellectuel. C'est là que l'esprit critique reprend ses droits.

8. Construire autrement : quelques pistes d'action

Le diagnostic est posé : l'IA risque de nous rendre intellectuellement paresseux si nous la consommons passivement. Mais comment faire autrement sans tomber dans l'injonction à devenir tous des experts ?

Ce qui semble nécessaire

Quelques principes émergent des constats précédents :

Rendre visible ce qui est caché : Les outils qui fonctionnent comme des "boîtes noires" entretiennent la dépendance. À l'inverse, quand on peut voir comment un algorithme fonctionne, même grossièrement, on reprend un peu de pouvoir dessus.

Documenter les échecs autant que les réussites : L'IA nous trompe régulièrement. Au lieu de faire semblant que ça marche toujours, il faudrait des espaces où on peut dire "j'ai testé, voilà ce qui a planté, voilà ce que j'ai compris". Cela créerait une mémoire collective des pièges à éviter.

Favoriser les allers-retours : Les meilleurs usages de l'IA semblent être ceux où l'humain et la machine s'interrogent mutuellement, plutôt que la machine qui répond et l'humain qui copie-colle.

Un exemple qui va dans ce sens

La Forge des Communs Numériques Éducatifs illustre concrètement ces principes. Cette plateforme héberge des ressources libres créées par des enseignants, avec un fonctionnement particulier : tout est transparent (on peut voir le code source), tout est modifiable (licence libre), et il existe des espaces pour signaler les problèmes ou proposer des améliorations.

Par exemple, il y a un salon spécifique où des "professeurs bidouilleurs d'IA" partagent leurs expérimentations. Pas leurs réussites parfaites – leurs expérimentations, avec ce que ça implique d'essais et d'erreurs. C'est exactement cette culture de la transparence et du tâtonnement assumé qui manque ailleurs.

La Forge ne résout pas tout, mais elle incarne une manière de faire : collaborative, ouverte, où on apprend les uns des autres sans prétendre tout maîtriser.

Les limites à reconnaître

Soyons lucides : ce type d'approche ne séduira jamais la majorité. Bidouiller, documenter, partager – cela demande du temps et une forme d'appétence pour la technique. Beaucoup de collègues n'ont ni l'un ni l'autre, et c'est légitime.

Alors que faire pour ceux qui ne deviendront jamais "bidouilleurs" ? Deux pistes :

Créer des outils "transparents par défaut" : Au lieu d'interfaces qui cachent tout, des outils qui montrent leurs limites. Par exemple, un correcteur orthographique qui dirait : "Attention, je ne suis pas sûr sur ce point, vérifiez vous-même". Cela nécessite que les concepteurs acceptent de montrer les failles de leurs créations : parce que ce ne sont pas des failles qui leur incombent mais les limites de l'IA qu'il faut mettre en valeur. Il faut battre en brèche l'idée de l'outil parfait, c'est une escroquerie commerciale.

Former à la prudence plutôt qu'à la technique : On n'a pas besoin de comprendre un moteur pour conduire prudemment. De même, on pourrait apprendre à "conduire" l'IA avec précaution : toujours vérifier les sources, comparer plusieurs réponses, ne jamais faire confiance à une seule sortie. C'est une forme d'hygiène intellectuelle qui ne demande pas d'expertise technique poussée.

La question du collectif

Ce qui ressort de ces réflexions, c'est qu'on ne s'en sortira probablement pas seul. Face à l'IA, l'individualisme est un piège. Nous avons besoin de construire des espaces de partage – formels ou informels, numériques ou présentiels – où on peut échanger sur ce qui marche, ce qui rate, ce qui interroge.

Cela peut être une équipe pédagogique qui se retrouve régulièrement, un groupe disciplinaire qui teste ensemble, ou des plateformes comme La Forge pour ceux qui veulent aller plus loin. L'important, c'est cette dimension collective : personne ne doit être seul face à la machine.

Car n'oublions pas que l'IA, par le chatbot, a tendance à isoler chacun dans un dialogue avec la machine. Le phénomène des Compagnons IA (ces applications conçues pour créer des relations affectives avec des intelligences artificielles), aussi inquiétant et extrême soit-il, caractérise cette captation. On peut partager un fichier avec un collègue, faire de l'écriture collaborative (superbe outil pédagogique !), mais nous ne pouvons pas échanger dans une conversation à plusieurs avec une même IA.

Conclusion

L'IA nous confronte à une réalité paradoxale : plus le monde devient numérique, plus nous avons besoin de concret et ce concret est dans le social et le collectif.

Quand les enseignants se demandent ce qu'ils peuvent apporter à un élève autonome sur Notebook LM, ChatGPT, etc., on répond « l'humain » forcément. Cet humain qu'on invoque pour les élèves, nous en avons besoin entre nous pour progresser, apprendre et maîtriser l'IA. Et pas qu'en visio...

Licence : CC BY-NC-SA 4.0

Contact : philippe.dupeyrat@ac-poitiers.fr

Site : <https://audit-ia-6f8fd7.forge.apps.education.fr>