

IA : les failles des logiciels anti-plagiat



Tiré de « Intelligence artificielle : la fabrique du consentement » par Pierre Fraser (PhD, linguiste et sociologue)

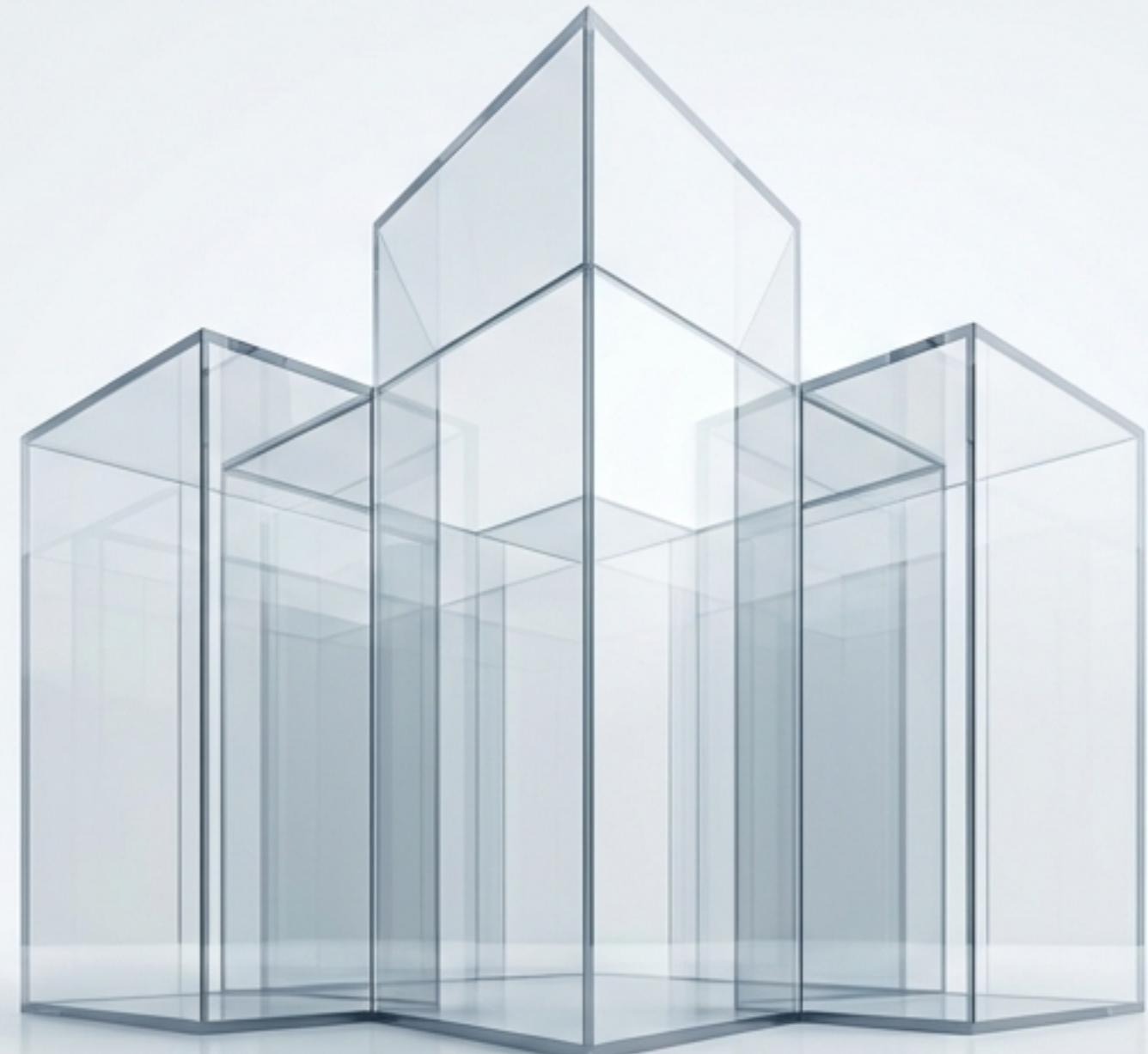
L'émergence d'un marché de la défiance

L'arrivée de l'IA générative en éducation a créé une réponse immédiate : une industrie d'outils de détection prospérant sur une promesse simple, rassurante et presque magique.

Cette promesse de démasquer la machine et d'éradiquer la triche a structuré un véritable « marché de la défiance ».

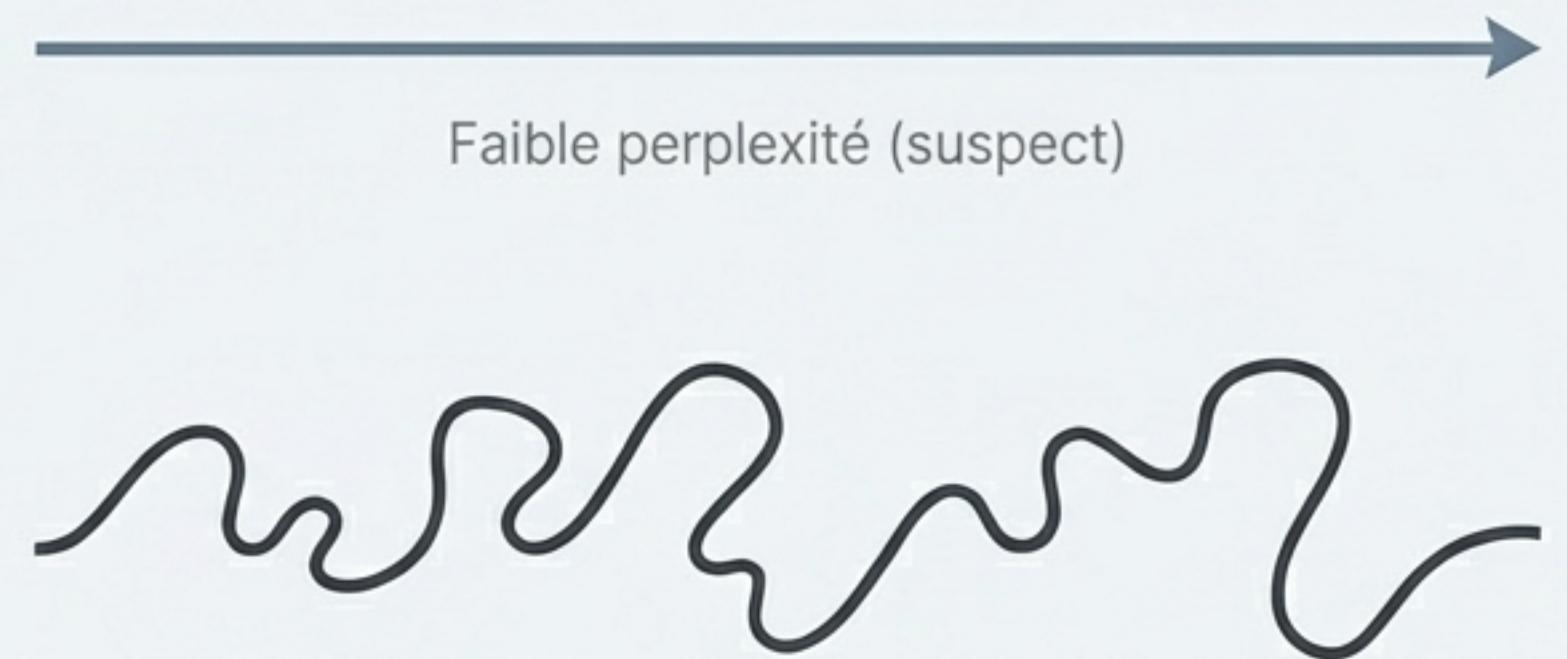
Mais cette solution, présentée comme infaillible, est en réalité une illusion coûteuse et une impasse technique et éthique.

« Restaurer l'ordre académique. »



La logique faillible de la « perplexité »

Les détecteurs ne comprennent pas un texte, ils analysent son degré de prévisibilité statistique, sa « perplexité ». Un style jugé trop lisse ou trop prévisible est considéré comme suspect.



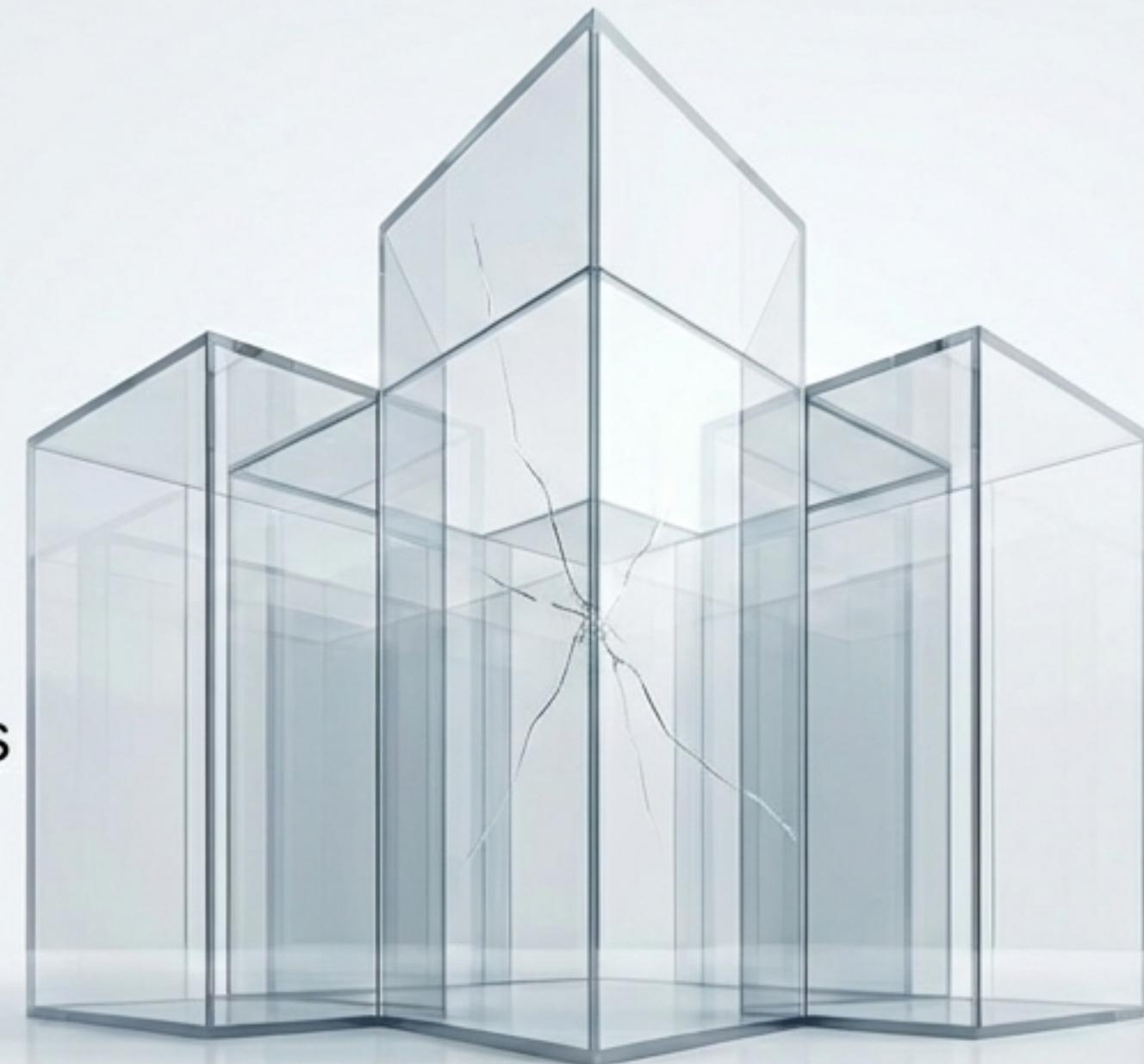
Le problème fondamental : cette logique **confond la production d'une machine avec la banalité ou la simplicité d'un style d'écriture humain.**

Une course à l'armement déjà perdue

La logique de la perplexité est si fragile qu'elle est structurellement incapable de résister aux stratégies d'évitement les plus simples.



Reformulation élémentaire :
Une simple réécriture manuelle ou assistée d'un passage suffit souvent à faire s'effondrer les taux de détection.

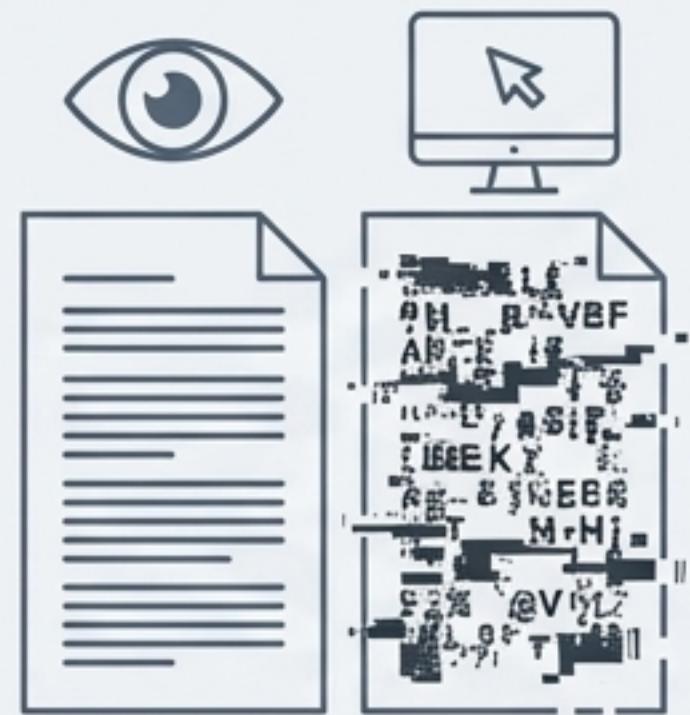


Conséquence directe :
La croisade pour la détection automatique est engagée, mais elle est déjà perdue sur le plan technique.

Des techniques de contournement sophistiquées

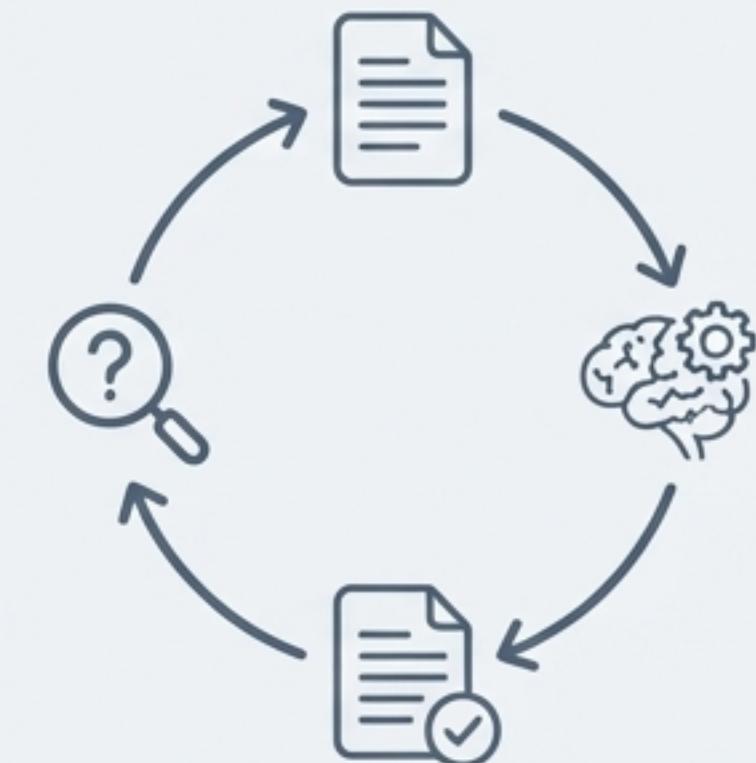
L'invisibilité technique : PDFuzz

Cette méthode modifie l'ordre d'extraction des caractères dans un fichier PDF. Le texte reste visuellement identique pour un humain, mais devient indéchiffrable pour les systèmes de détection.



L'ironie ultime : la paraphrase adverse

Cette approche consiste à soumettre un texte au détecteur, à analyser ses faiblesses, puis à les utiliser pour générer un contenu artificiel qui coche toutes les cases de l'authenticité humaine. On retourne l'arme contre son utilisateur.





L'échec n'est pas que technique,
il est profondément humain

**Ces outils ne traquent pas l'intelligence artificielle.
Ils pénalisent les styles d'écriture simples, directs et
et normés.**

Le taux de faux positifs devient vertigineux, car les algorithmes sont biaisés contre des manières d'écrire parfaitement légitimes. La prétendue rigueur académique cache une réalité bien plus sombre.

Une discrimination linguistique et cognitive

La logique des détecteurs frappe de manière disproportionnée deux groupes principaux :



Les étudiants non-natifs : Leur écriture, souvent plus fonctionnelle et moins flamboyante stylistiquement, est fréquemment signalée à tort.



Les personnes neurodivergentes : Pour elles, la régularité syntaxique est une stratégie de clarté, et non un indice de fraude.

C'est « une discrimination linguistique et cognitive déguisée en rigueur académique. »

La réponse des étudiants : déjouer l'algorithme

Les forums étudiants débordent déjà de recettes et de stratégies pour passer sous le radar des détecteurs.

Reformulations en chaîne

Traductions croisées
(français -> anglais -> français)

Insertion volontaire de petites erreurs

Adoption d'un ton faussement relâché

Le paradoxe : la qualité du texte en souffre parfois, mais les algorithmes, eux, tombent dans le panneau.

L'impasse : accuser sans preuve, soupçonner sans comprendre

L'incapacité chronique des détecteurs à identifier les productions hybrides (mi-humaines, mi-artificielles) rend leur usage probatoire presque inutile.

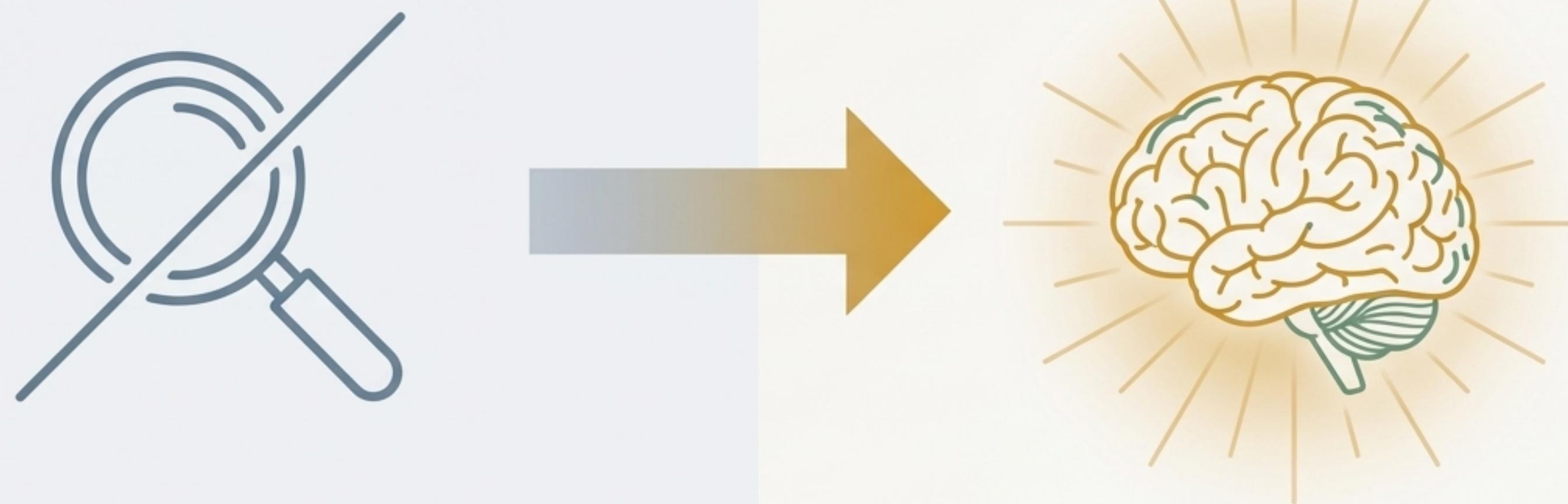


Cela conduit à une culture de la suspicion où la charge de la preuve est inversée. L'enseignant se retrouve dans une posture d'accusateur, incapable de distinguer la fraude de la maladresse ou d'un style d'écriture différent.

Le changement de paradigme : évaluer autrement

Persister dans la logique de la **détection** relève moins de la protection pédagogique que d'un « **réflexe disciplinaire** ».

La véritable solution n'est pas de surveiller davantage, mais d'évaluer autrement. Il faut déplacer le centre de gravité.



Vers une évaluation « authentique »

Vers une évaluation « authentique »

Des cadres émergent, notamment au Québec, pour promouvoir une évaluation qui valorise le cheminement intellectuel.

Passer du résultat au processus.

La question n'est plus :
« Qu'as-tu rendu ? »
mais bien :
« Comment as-tu pensé ? »

Former des « centaures », pas des fraudeurs

L'enjeu n'est pas d'éradiquer l'IA, mais d'apprendre à vivre avec elle. L'objectif est de former des utilisateurs responsables, critiques et stratégiques.

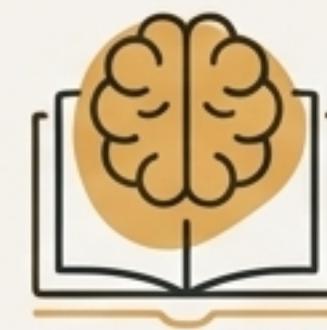
Le « centaure », une métaphore pour la collaboration où l'intelligence humaine guide la puissance algorithmique, sans abdication intellectuelle.



Un choix, pas une capitulation

Abandonner les détecteurs n'est pas un aveu d'impuissance face à la technologie.

C'est un choix :



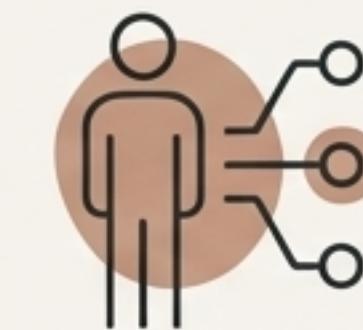
Pédagogique

Privilégier l'apprentissage
sur la surveillance.



Politique

Refuser une culture
de la suspicion.



Éthique

Rejeter des outils
discriminatoires.

C'est peut-être la condition minimale pour préparer l'éducation
aux réalités du XXI^e siècle.

Les enjeux pour l'éducation de demain

Adopter l'IA de manière critique et constructive implique de relever plusieurs défis majeurs :

Pédagogie : Réinventer les méthodes d'évaluation pour valoriser le processus intellectuel et le raisonnement critique.

Éthique : Définir une nouvelle intégrité académique qui intègre l'usage des outils d'IA de manière transparente.



Équité : S'assurer que les nouvelles approches ne créent pas d'autres formes de discrimination et garantissent un accès juste aux outils.

Formation : Accompagner les enseignants et les étudiants pour qu'ils deviennent des « centaures » compétents et critiques.