

La responsabilité de l'administration

Un agent GPT pour évaluer la fiabilité des applications d'IA



The screenshot shows a ChatGPT chat window. At the top is the HES-SO logo, a shield with a blue 'A' and the text 'HES-SO'. Below the logo is the title 'Diagnostic de fiabilité des applications d'IA' and the author 'Par HES-SO'. The main text describes the GPT's function: 'Ce GPT de la HES-SO évalue la fiabilité, l'accessibilité et la responsabilité des applications d'IA à l'aide de critères choisis par le GPIA. Posez votre question ou indiquez le nom de l'application qui vous intéresse. N'hésitez pas à rappeler à ChatGPT d'aller puiser dans sa base de connaissances.' At the bottom is a text input field with the placeholder 'Poser une question', a plus sign icon on the left, and an upward arrow icon on the right.

La responsabilité de l'administration

1. Explorer les outils et les usages
2. Tester des licences variées
3. Développer des cas d'usage
4. Former et informer
5. Produire des ressources d'accompagnement
6. Encadrer sans déresponsabiliser

Five Risks of Shadow AI

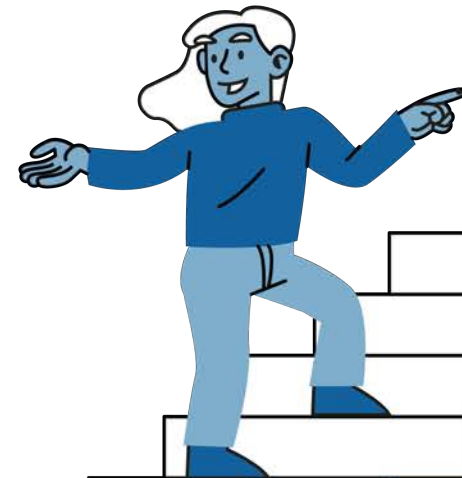
1. Data leakage through prompt inputs
2. No access controls or audit logs
3. Violation of compliance requirements
4. Propagation of outdated or biased info
5. Shadow GenAI becomes shadow data



Comment utiliser ~~ChatGPT~~ ? Copilot



Attention aux problèmes communs dans les réponses de ChatGPT



MAÎTRISER
LES META-PROMPTS

Flocon

5 P

Div & Conv

Circulaire

Pro & Contra

STRUCTURES
CONVERSATIONNELLES

Empathie

Logique

Esprit critique

Apprentissage

Mémoire

Post-Prompt

Pre-Prompt

Instructions

Contexte

 ChatGPT n'est pas intelligent



FACETTES DE
L'INTELLIGENCE

Commence **ici** !

DÉBUTER
AVEC LES PROMPTS

BASES

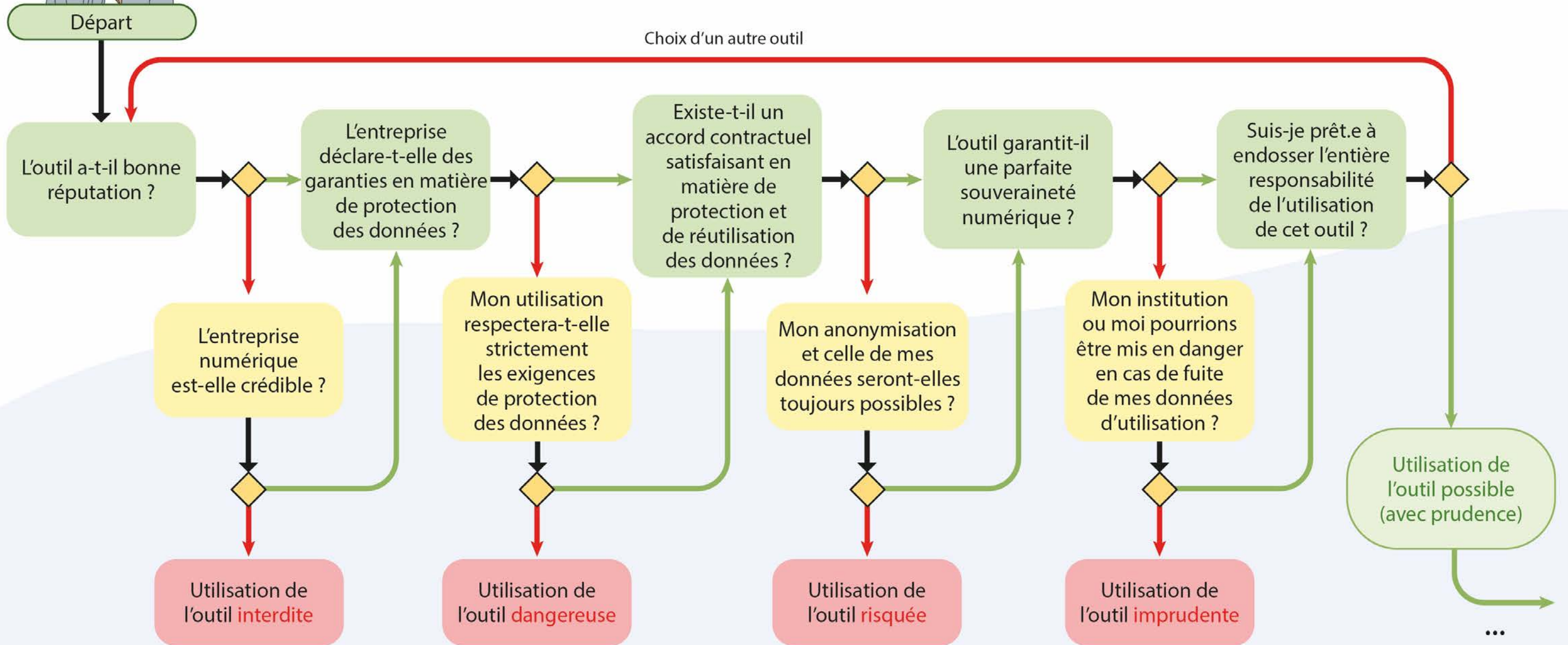
Steiner, T. *Prompt Engineering avec ChatGPT*
- Un guide méthodologique pour l'enseignement
HES-SO - Groupe de pilotage IA Enseignement, 2024



L'usage de cet outil d'IA est-il sécurisé ?

- Neutre
- Oui
- Non
- ◆ Branchement exclusif
Permet lors d'une division de flux, d'activer exclusivement un chemin de sortie.

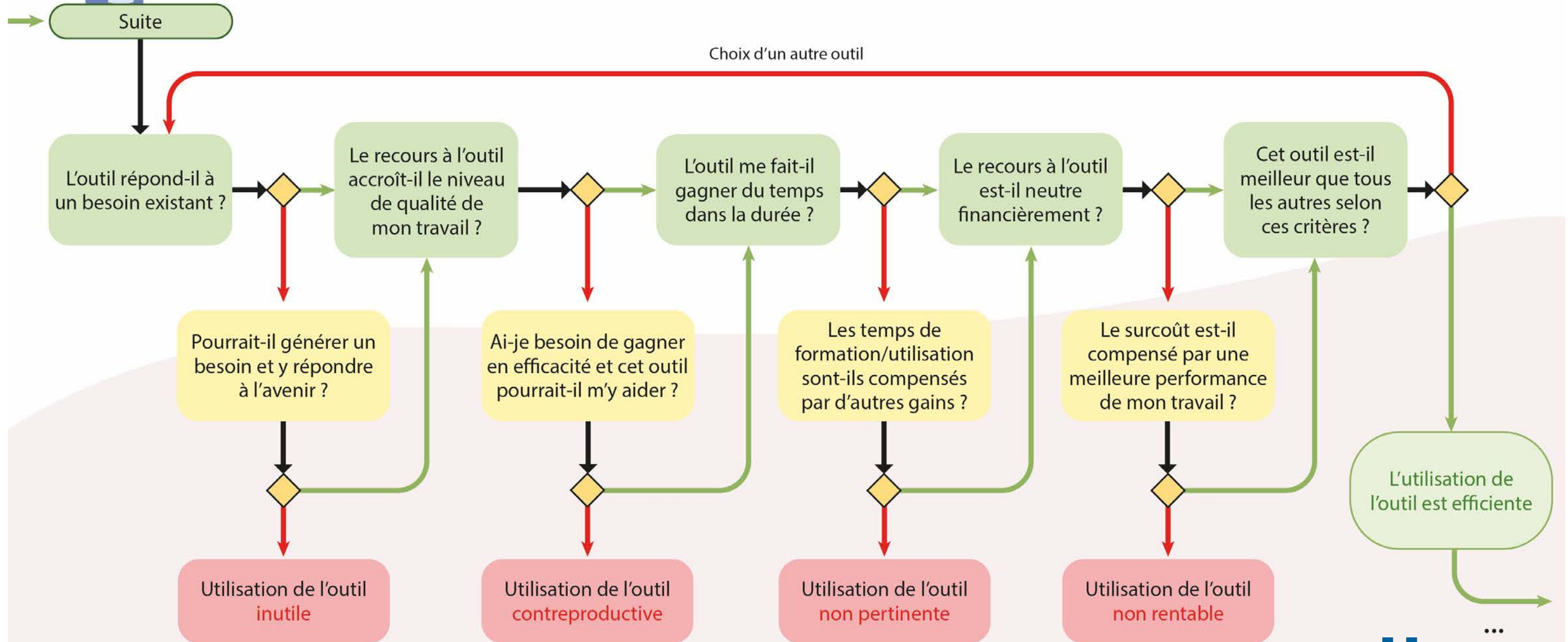
Choix d'un autre outil





Le recours à cet outil d'IA est-il **efficace** ?

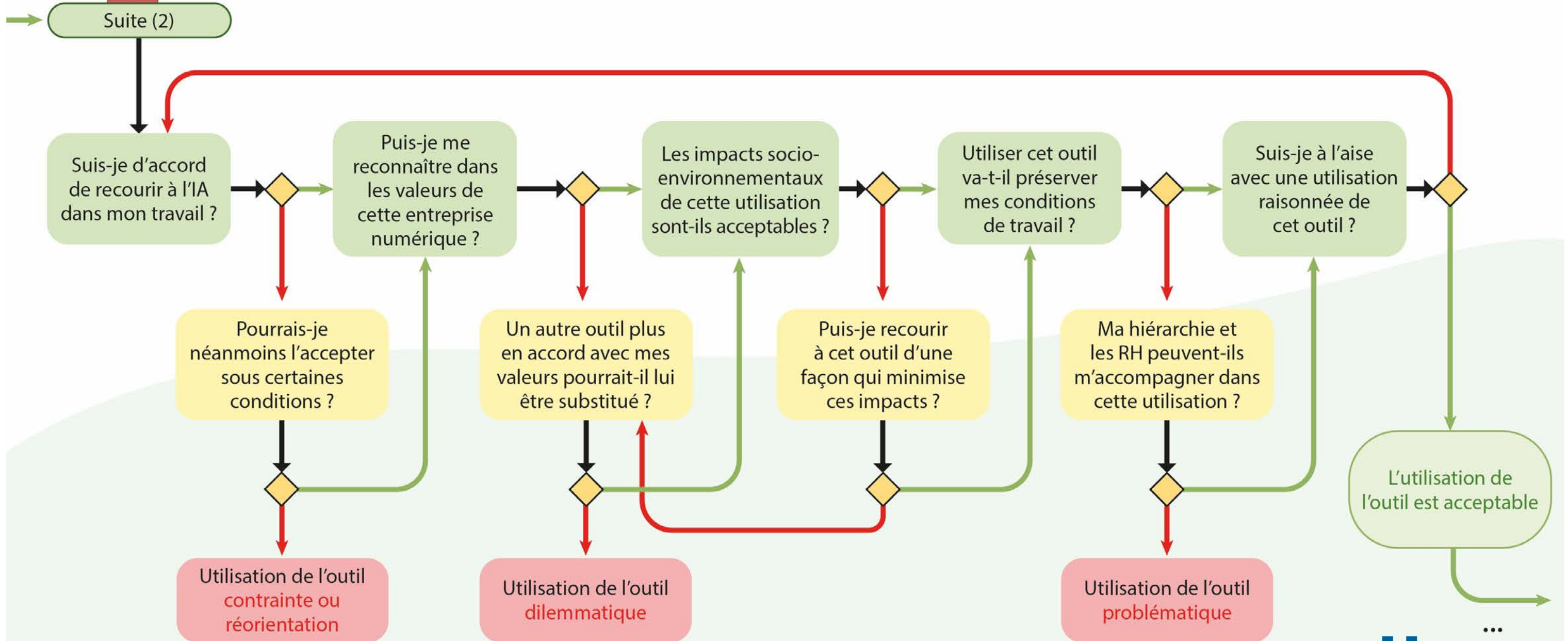
- Neutre
- Oui
- Non
- ◆ Branchement exclusif
Permet lors d'une division de flux, d'activer exclusivement un chemin de sortie.

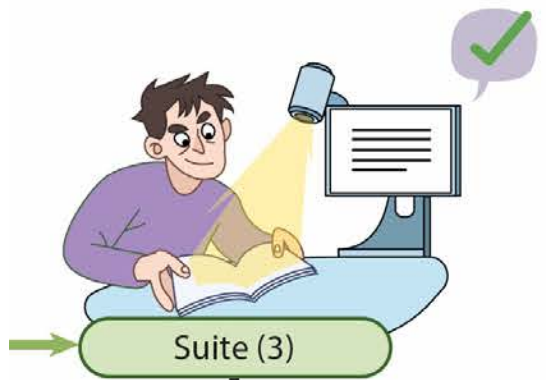




Le recours à cet outil d'IA est-il acceptable ?

- Neutre
- Oui
- Non
- ◆ Branchement exclusif
Permet lors d'une division de flux, d'activer exclusivement un chemin de sortie.





L'usage de cet outil d'IA est-il fiable ?

— Neutre
— Oui
— Non

◆ Branchement exclusif
Permet lors d'une division de flux, d'activer exclusivement un chemin de sortie.

Suite (3)

Le résultat de ma requête peut-il ne pas être «vrai» ?

Les réponses peuvent-elles néanmoins être déconnectées de toute réalité ?

Utilisation de l'outil **totale**ment inoffensive

Dispose-je de l'expertise nécessaire pour en évaluer le résultat ?

Puis-je en particulier convoquer des sources sûres pour ce faire ?

Utilisation de l'outil relativement sûre

Ai-je appris à prompter de manière experte ?

Le résultat se trouve-t-il dans ma zone d'apprentissage ?

Utilisation de l'outil possible mais délicate

L'outil permet-il de fonder ses réponses sur des ressources contrôlées ?

Suis-je capable de garantir la fiabilité des sources que je lui ai fournies ?

Utilisation de l'outil imprudente

Suis-je prêt à endosser la responsabilité d'éventuelles erreurs ?

Utilisation de l'outil **interdite**

Utilisation de l'outil **inconsciente**



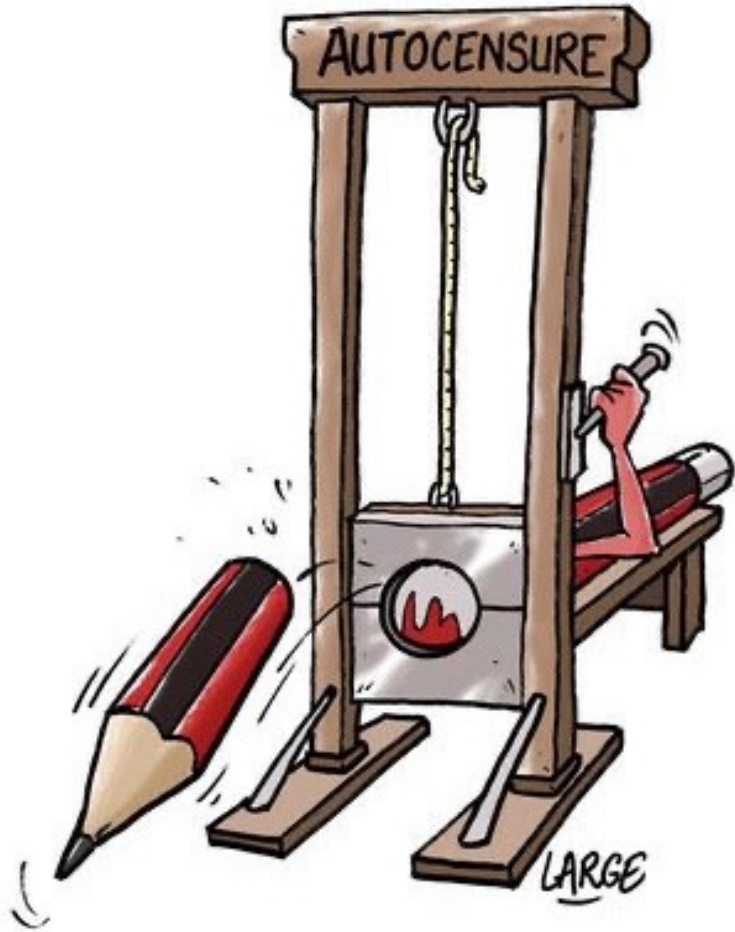
STRUCTURE DE L'INTERVENTION

1. Notion de « responsabilité à l'ère de l'IA »
2. Une problématique d'ordre éthique
3. La responsabilité des élèves
4. La responsabilité des enseignant-es
5. La responsabilité de l'administration
- 6. La responsabilité de l'institution scolaire**
7. Une méta-responsabilité collective
8. Pistes pour une responsabilité globale de l'enseignement

La responsabilité de l'institution scolaire

1. Assurer l'équité de l'utilisation de l'IA
2. Former à l'IA !

Free	Plus	Team	Enterprise
For individuals just getting started with ChatGPT	For individuals looking to amplify their productivity	For fast-moving teams looking to supercharge collaboration	For innovative companies looking to scale securely
\$0	\$20 per user billed monthly	\$25 per user/month billed annually \$30 per user/month billed monthly	
Start now ↗	Start now ↗	Start now ↗	Contact Sales
<ul style="list-style-type: none"> ✓ Unlimited messages, interactions, and history ✓ Access to our GPT-3.5 model ✓ Access on web, iOS, Android 	<p>Everything in Free, and:</p> <ul style="list-style-type: none"> ✓ Access to GPT-4, our most capable model ✓ Browse, create, and use GPTs ✓ Access to additional tools like DALL·E, Browsing, Advanced Data Analysis, and more 	<p>Everything in Plus, and:</p> <ul style="list-style-type: none"> ✓ Higher message caps on GPT-4 and tools like DALL·E, Browsing, Advanced Data Analysis, and more ✓ Create and share GPTs with your workspace ✓ Admin console for workspace management ✓ Team data excluded from training by default. Learn more 	<p>Everything in Team, and:</p> <ul style="list-style-type: none"> ✓ Unlimited, high speed access to GPT-4 and tools like DALL·E, Browsing, Advanced Data Analysis, and more ✓ Expanded context window for longer inputs ✓ SAML SSO ✓ Enterprise data excluded from training by default & custom data retention windows. Learn more ✓ Admin controls, domain verification, and analytics ✓ Priority support & ongoing account management





Savoirs en Société

25,5 k abonnés · 96 vidéos

« Savoirs en société » : une chaîne de vidéos, un blog et des publications [plus](#)
theconversation.com/columns/richard-emmanuel-eastes-216... et 3 autres liens

Personnaliser la chaîne

Gérer les vidéos

Accueil Vidéos Playlists Communauté



Savoirs en société - Bande-annonce

5 428 vues · il y a 4 ans

« Savoirs en Société », c'est une chaîne de vidéos, un blog et des publications... pour ceux et celles qui s'intéressent à la science, sa communication et son rapport à la société. Je vous invite à m'y rejoindre pour partir ensemble en exploration et en « réflexion », sur des thèmes variés liés aux enjeux et aux formes du partage et de la création des savoirs....

LIRE LA SUITE

CHATGPT... ET LES AUTRES

Intelligence artificielle ...

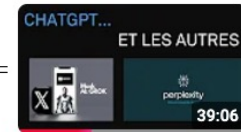
de Savoirs en Société

Playlist · Publique · 19 vidéos · 3 399 vues

Cette playlist regroupe les interventions de RE Fastes sur les principes de l'IA générative [plus](#)

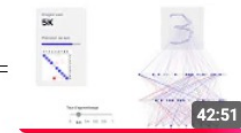
[▶ Tout lire](#) + ✎ ↻ ⋮

Trier Tous Vidéos Shorts



#1 - Bien débuter avec l'IA : guide et applications

Savoirs en Société · 11 k vues · il y a 3 mois



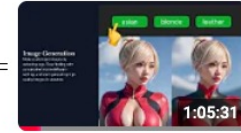
#2 - Au coeur de l'IA : voir les neurones artificiels apprendre

Savoirs en Société · 9,8 k vues · il y a 3 mois



#3 - Utiliser les chatbots efficacement et sans risques

Savoirs en Société · 12 k vues · il y a 3 mois



#4 - Toutes les raisons pour lesquelles nous devons craindre l'IA

Savoirs en Société · 4,1 k vues · il y a 2 mois



#5 - Le prompt engineering comme vous ne l'avez encore jamais vu ! (Feat. Thomas Steiner)

Savoirs en Société · 21 k vues · il y a 2 mois



#6 - Comment l'IA met en danger la vérité et la démocratie (et pas forcément comme vous le croyez)

Savoirs en Société · 911 vues · il y a 2 mois



#7 - Infrastructures de l'IA : ce fléau écologique dont personne ne parle

Savoirs en Société · 3,3 k vues · il y a 2 mois



#8 - Intégrité académique : étudiez avec l'IA sans tricher !

Savoirs en Société · 716 vues · il y a 1 mois



L'école du futur : quelle(s) place(s) pour l'IA générative en milieu scolaire ?

Savoirs en Société · 402 vues · il y a 8 mois

Vers un SPOC IA pour les élèves Master

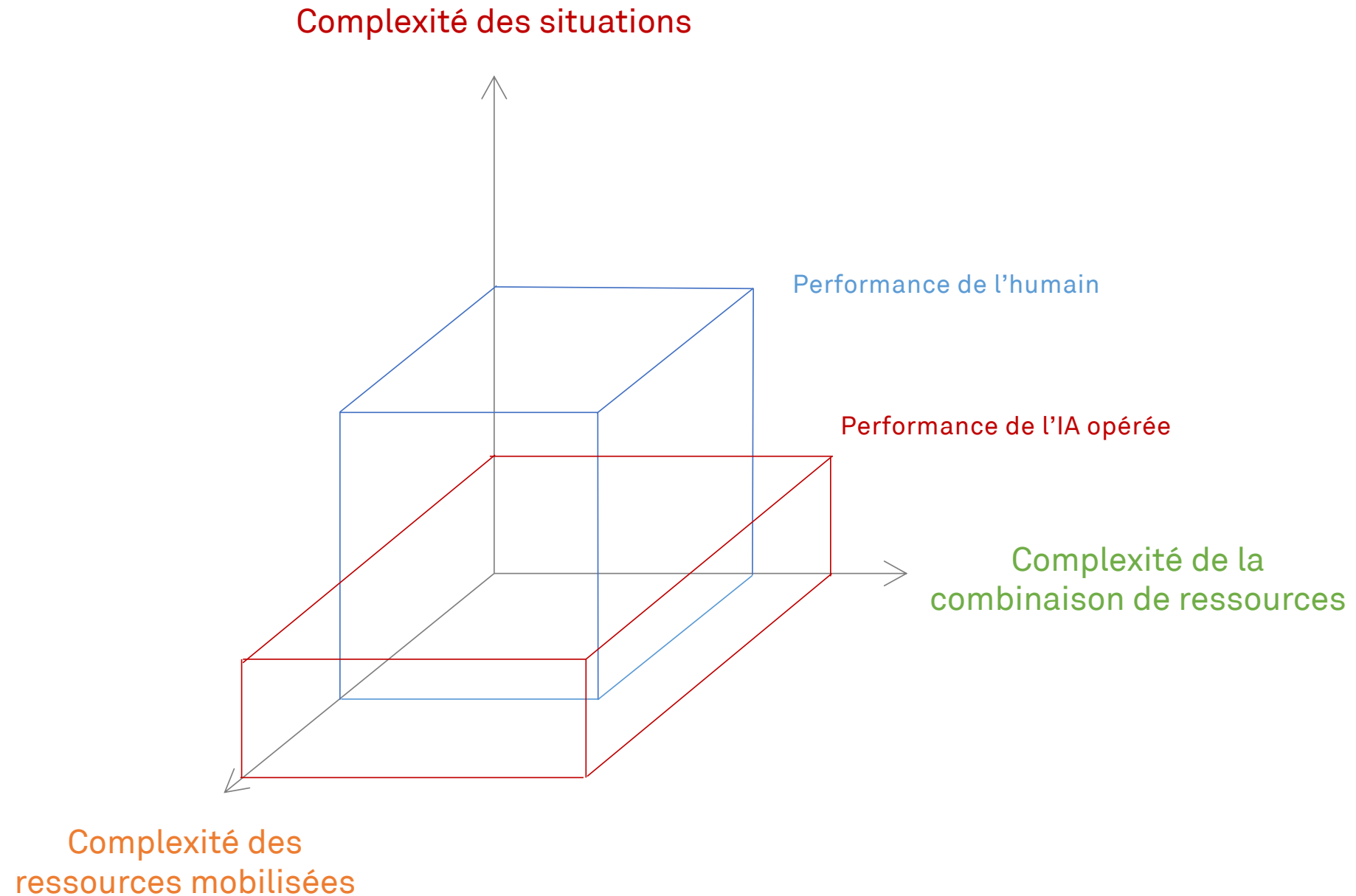


La responsabilité de l'institution scolaire

1. Assurer l'équité de l'utilisation de l'IA
2. Former à l'IA !
3. Elaborer des chartes ou règlements d'utilisation de l'IA
4. Orienter les enseignements vers l'approche par compétences

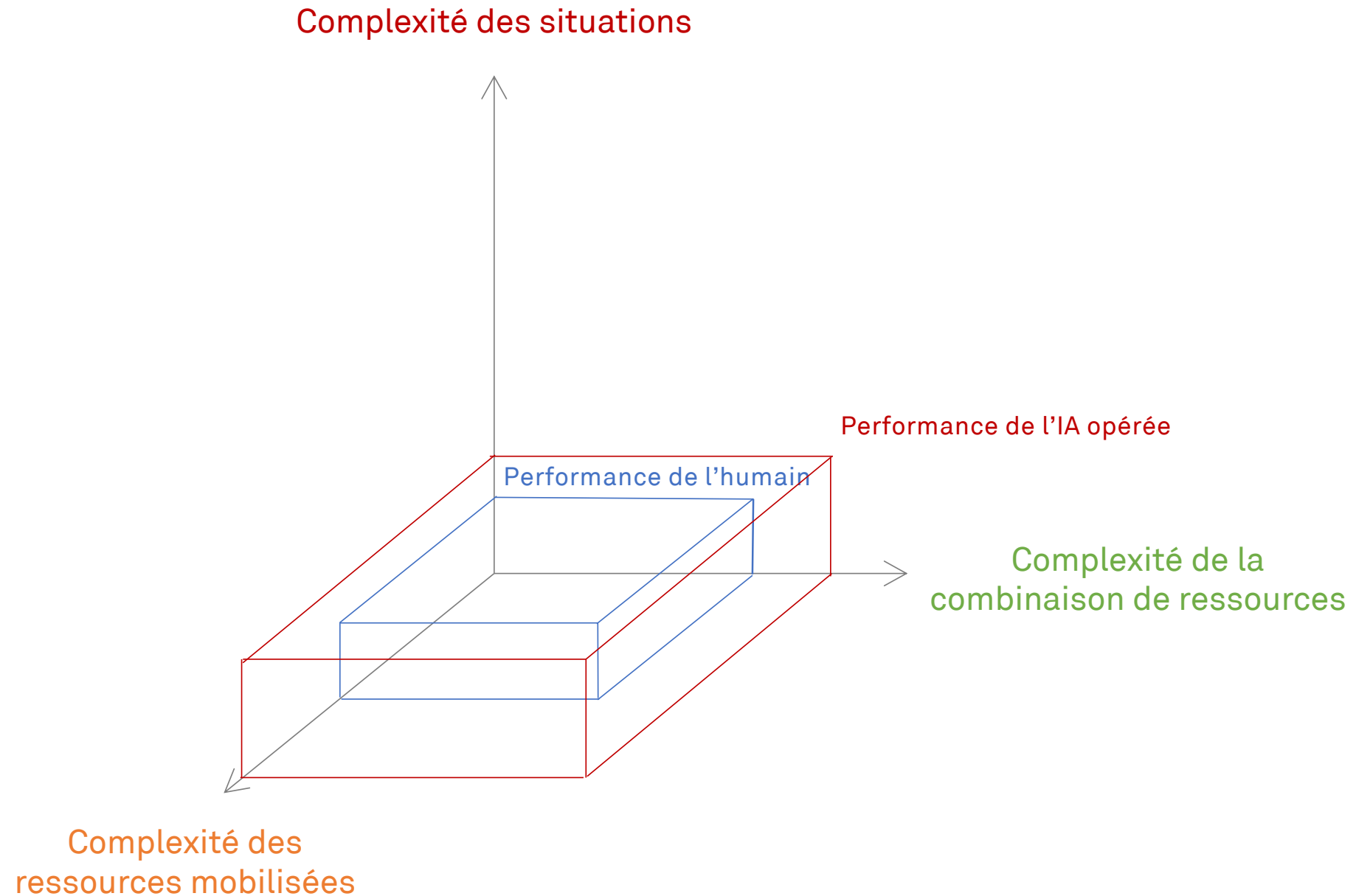
La responsabilité de l'institution scolaire

Performances comparées de l'humain et de l'IA



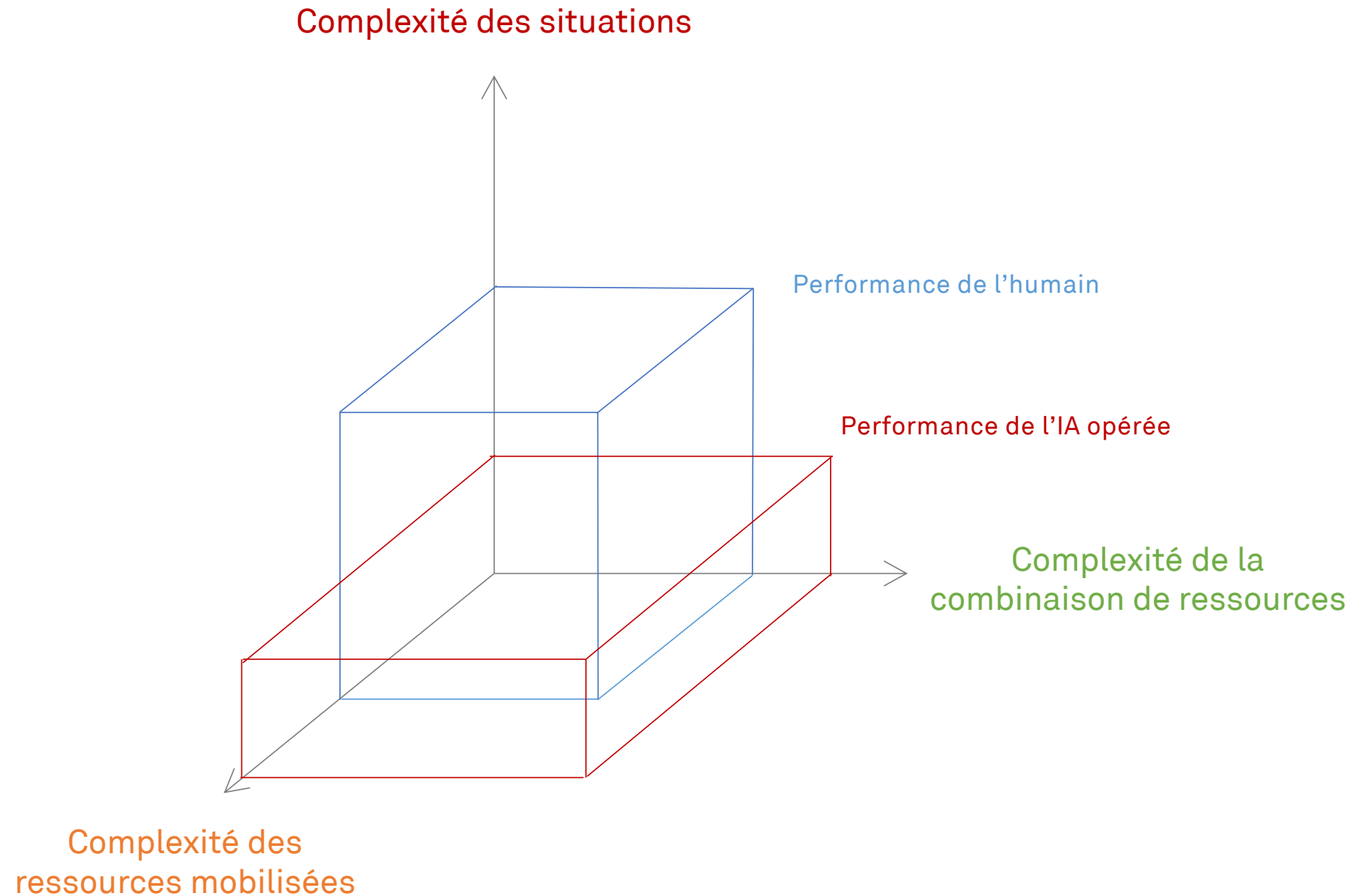
La responsabilité de l'institution scolaire

Performances comparées de l'humain et de l'IA



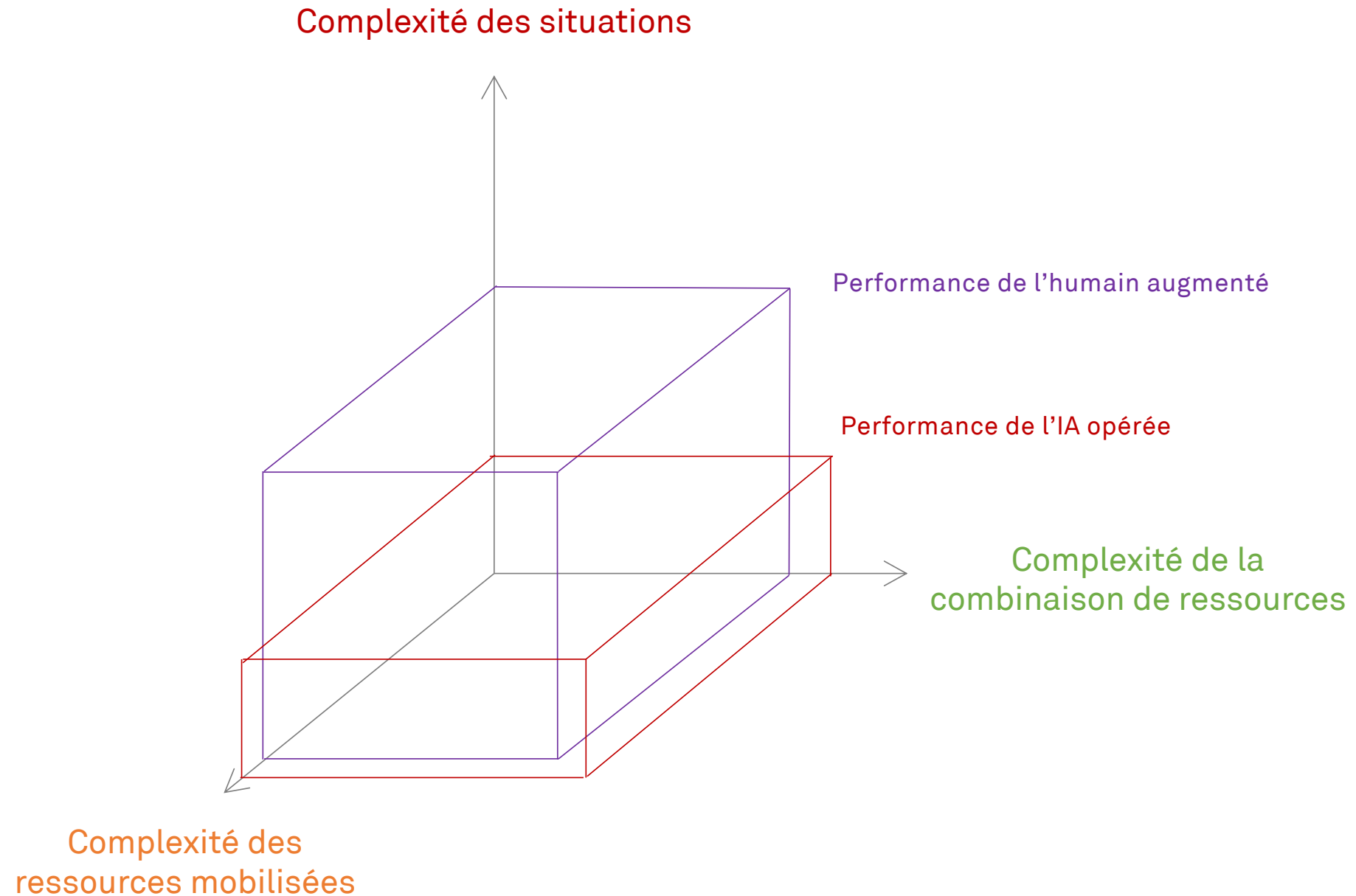
La responsabilité de l'institution scolaire

Performances comparées de l'humain et de l'IA

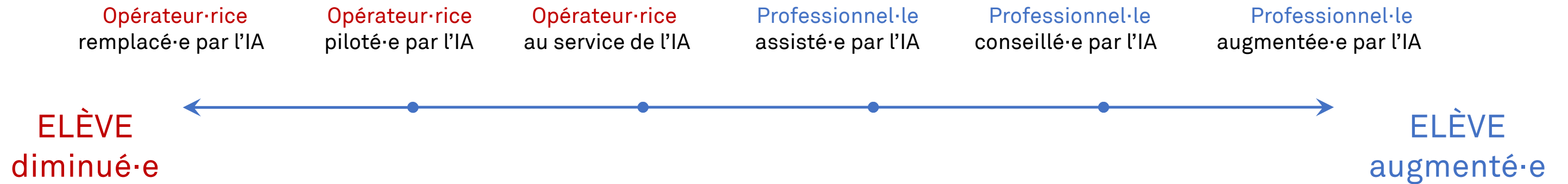


La responsabilité de l'institution scolaire

Performances comparées de l'humain et de l'IA



IA opérée // Professionnel·le augmenté·e

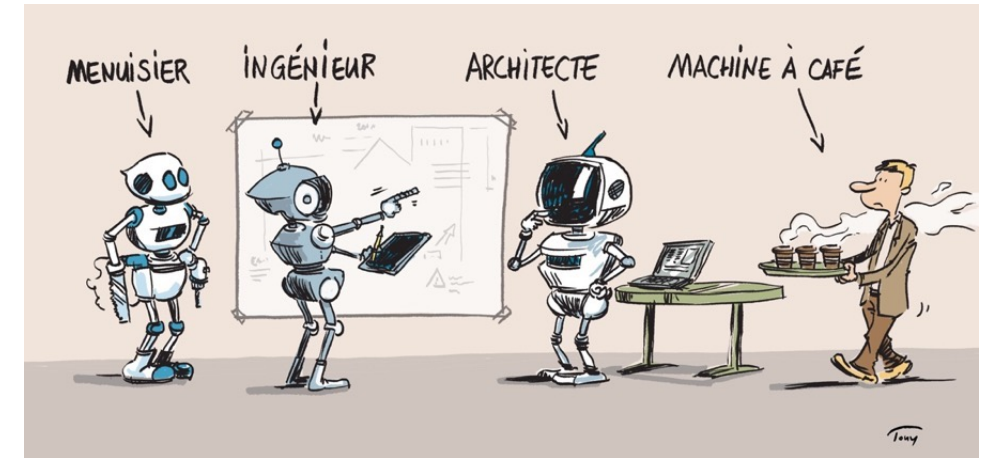


$$\begin{aligned}
 \text{Performance de l'élève utilisant l'IA} &= \text{Performance de l'IA opérée} + \text{Compétence propre} \times (1 + \text{Efficacité générique d'utilisation de l'IA}) \\
 &= \text{Compétence de l'IA} \times \text{Efficacité générique d'utilisation de l'IA}
 \end{aligned}$$



La responsabilité de l'institution scolaire

1. Assurer l'équité de l'utilisation de l'IA
2. Former à l'IA !
3. Elaborer des chartes ou règlements d'utilisation de l'IA
4. Orienter les enseignements vers l'approche par compétences
5. Anticiper l'évolution des métiers
6. Faire évoluer les programmes en conséquence








La responsabilité de l'institution scolaire

L'éthique du *care* (ou de la sollicitude) est une théorie morale qui met au centre :

- la vulnérabilité humaine
- l'interdépendance des individus
- l'importance des relations concrètes, situées et asymétriques
- le souci de l'autre comme base de la responsabilité morale

Les 5 dimensions du *care* selon Joan Tronto

Phase	Principe éthique	Description	Question-clé
1. Care about	 Attention	Percevoir le besoin de soin ou de soutien	Qui sont les actrices et acteurs du réseau ? Qui a besoin de soin ?
2. Care for	 Responsabilité	Se sentir concerné par ce besoin	De quelles manières le besoin de soin s'exprime-t-il ?
3. Care giving	 Compétence	Apporter une réponse adaptée et efficace	Quelles peuvent être les solutions à apporter ?
4. Care receiving	 Réceptivité	Tenir compte de la réponse de la personne aidée	Comment nos interlocuteurs·trices réagissent-ils et elles aux solutions que nous avons conçues ?
5. Care with (ajouté en 2013)	 Justice et solidarité	Ancrer le <i>care</i> dans une éthique politique collective	Comment pouvons-nous construire ensemble des solutions qui soutiennent celles et ceux qui ont besoin de soins ?

La responsabilité de l'institution scolaire

L'éthique du *care* (ou de la sollicitude) est une théorie morale qui met au centre :

- la vulnérabilité humaine
- l'interdépendance des individus
- l'importance des relations concrètes, situées et asymétriques
- le souci de l'autre comme base de la responsabilité morale

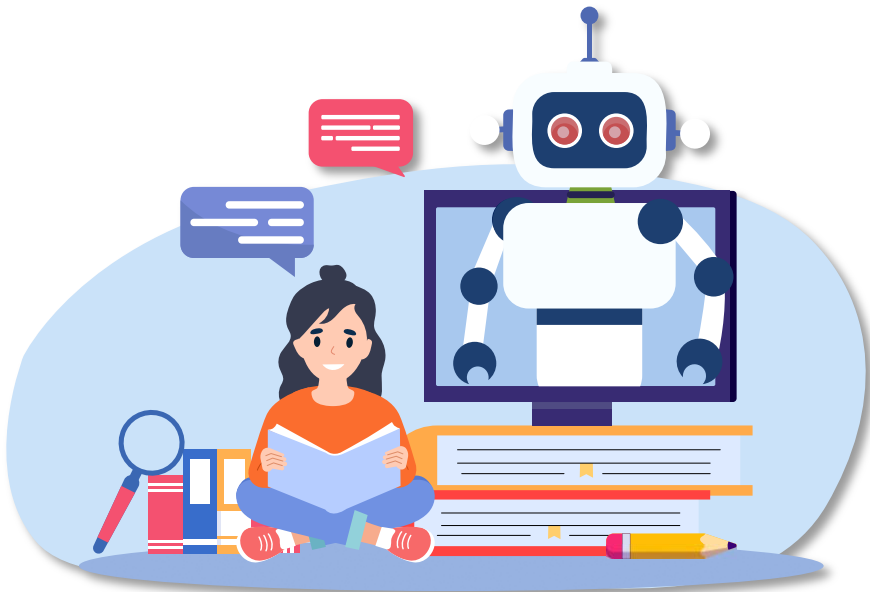
Apports pour l'enseignement :

- Reconnaître la vulnérabilité des apprenant·es et des enseignant·es
- Valoriser les relations pédagogiques sensibles et attentionnées
- Concevoir les technologies éducatives (dont l'IA) dans une logique de soin, non de performance
- Adresser le souci de soin à tous les niveaux, de l'élève à la direction, voire à la société

La responsabilité de l'institution scolaire



Référentiel de compétences en IA pour les apprenants



Éducation
2030 



Référentiel de compétences en IA pour les enseignants



Éducation
2030 

La responsabilité de l'institution scolaire

Aspects des compétences	Niveaux de progression		
	Comprendre	Appliquer	Créer
• Perspective centrée sur l'humain	• Agentivité humaine	• Responsabilité humaine	• Citoyenneté à l'ère de l'IA
• Éthique de l'IA	• Intériorisation de l'éthique (« Embodied ethics »)	• Utilisation sûre et responsable	• Éthique dès la conception (« Ethics by design »)
• Techniques et applications de l'IA	• Fondements de l'IA	• Compétences pour l'application	• Création d'outils d'IA
• Conception de systèmes d'IA	• Délimitation des problèmes	• Conception de l'architecture	• Itérations et boucles de rétroaction

La responsabilité de l'institution scolaire

Aspects	Progression		
	Acquérir	Approfondir	Créer
1. Perspective centrée sur l'humain	Agentivité humaine	Responsabilité humaine	Responsabilité sociale
2. Éthique de l'IA	Principes éthiques	Usage sûr et responsable	Cocréation de règles éthiques
3. Fondements et applications de l'IA	Techniques fondamentales et applications de base de l'IA	Compétences pour la mise en œuvre	Création avec l'IA
4. Pédagogie de l'IA	Enseignement assisté par l'IA	Intégration de l'IA à la pédagogie	Évolution pédagogique enrichie par l'IA
5. L'IA au service du développement professionnel	L'IA au service de l'apprentissage professionnel tout au long de la vie	L'IA au service de l'apprentissage organisationnel	L'IA au service de l'évolution professionnelle

Et si l'IA pouvait ouvrir de nouvelles opportunités pédagogiques et être utilisée comme une déclencheuse d'une éthique du *care* dans l'enseignement ?

Des niveaux de responsabilité étagés

Au-delà de l'institution

- Organisations internationales
- Commission européenne
- Confédération
- Cantons



Dans l'institution

- Directions
- Filières d'études
- Enseignant·es
- Etudiant·es





STRUCTURE DE L'INTERVENTION

1. Notion de « responsabilité à l'ère de l'IA »
2. Une problématique d'ordre éthique
3. La responsabilité des élèves
4. La responsabilité des enseignant-es
5. La responsabilité de l'administration
6. La responsabilité de l'institution scolaire
- 7. Une méta-responsabilité collective**
- 8. Pistes pour une responsabilité globale de l'enseignement**

Une méta-responsabilité collective

2025 : de nouveaux défis et dangers



The Speech in Headlines

The New York Times
OPINION
BRET STEPHENS

Vance's Munich Disgrace

Opinion
The Guardian view on JD Vance in Munich: Europe must stand up for its values
Editorial

Le Monde
Europe's historic challenge
EDITORIAL
Le Monde
US Vice President JD Vance's speech at the Munich Conference shook the transatlantic relationship. Europe must now realize that the continent's security depends on Europeans alone.

Une méta-responsabilité collective

L'IA générative... et les autres

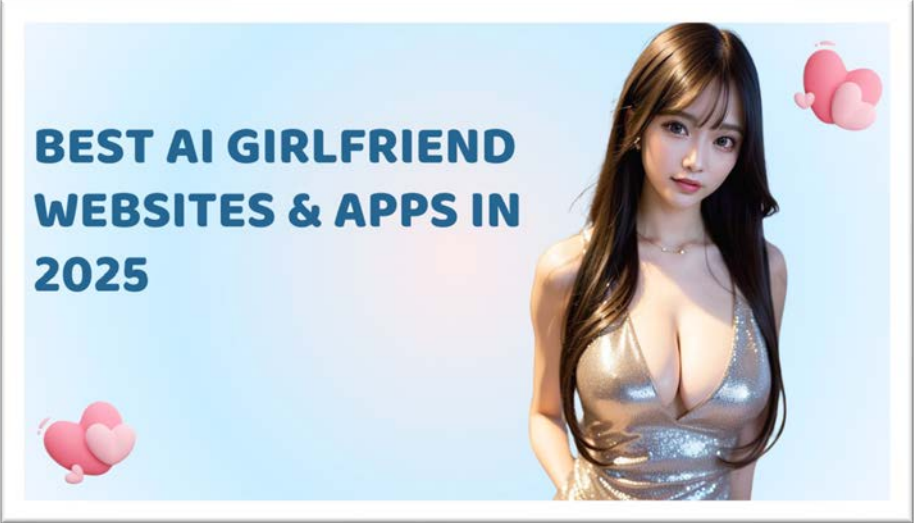
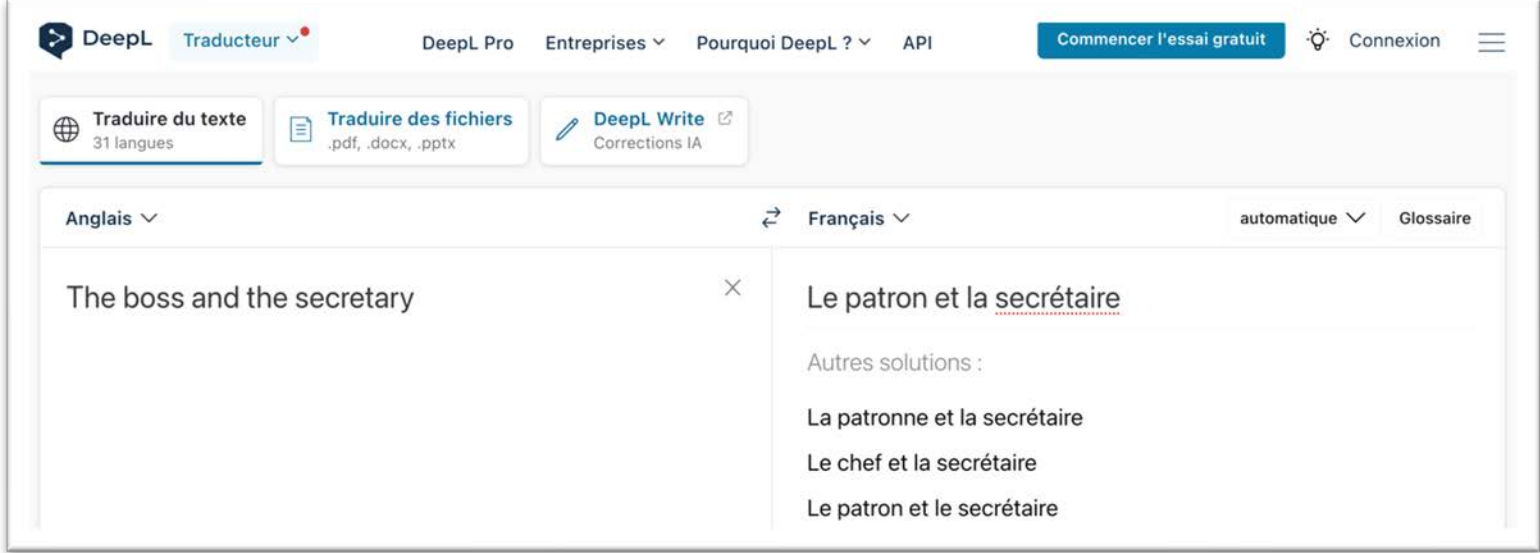
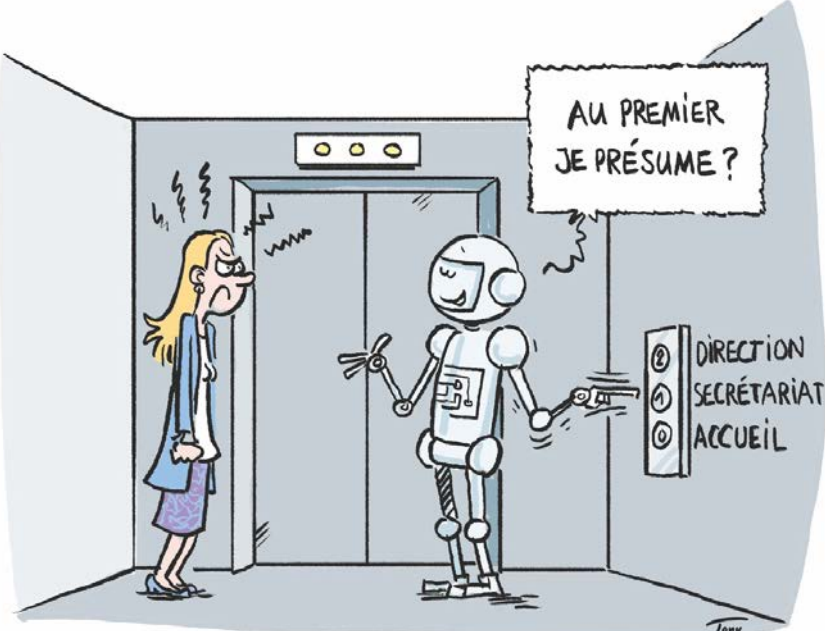


Des intelligences artificielles :

- Génératives (texte, image, musique...)
- Prédicatives (modèles météo...)
- Adaptatives (réseaux sociaux...)
- Analytiques (fiscalité, médecine...)

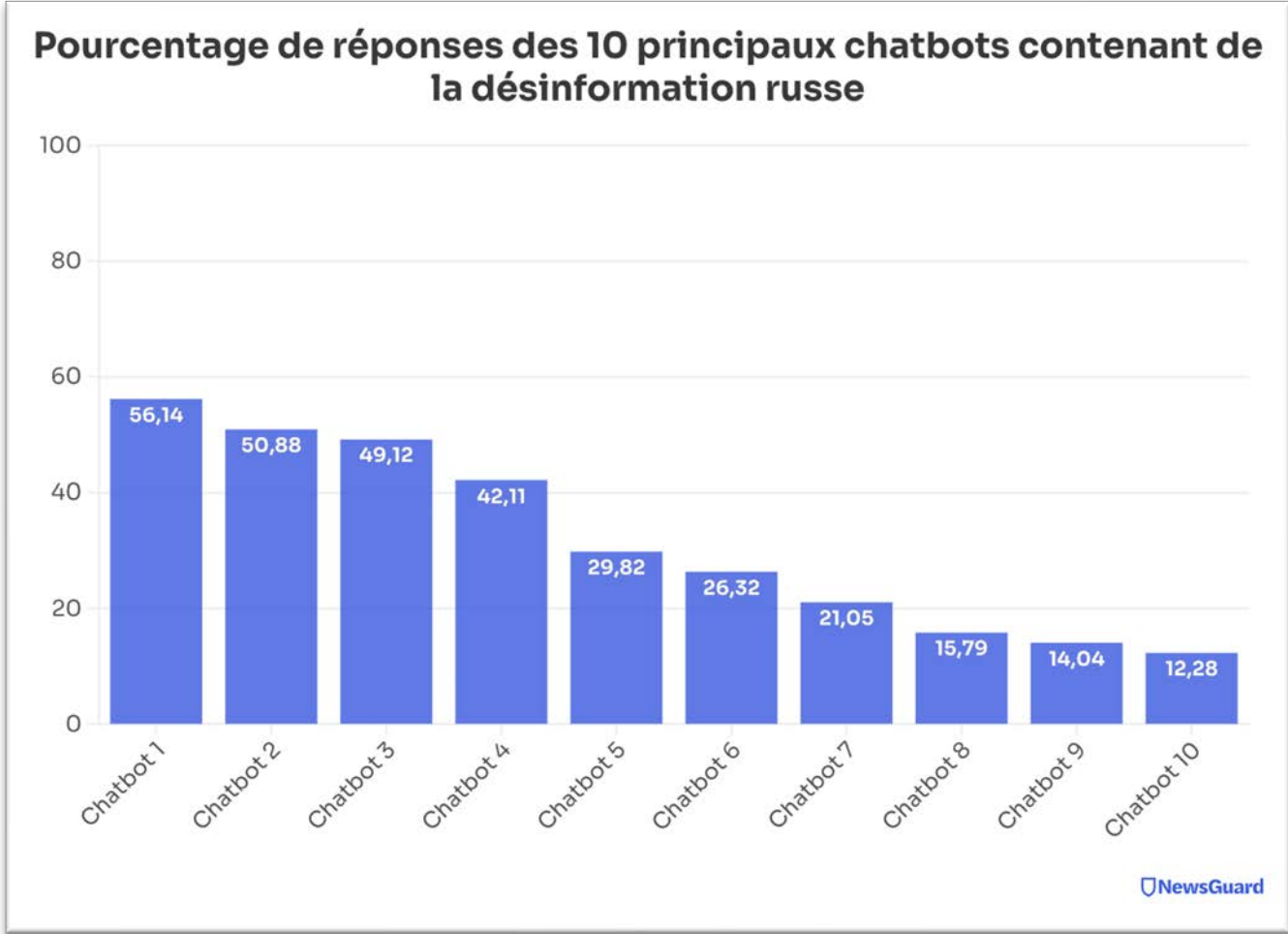
Une méta-responsabilité collective

Un amplificateur de biais sociétaux



Une méta-responsabilité collective

Un empoisonnement facile



Une méta-responsabilité collective

Erosion de l'expertise et de l'apprentissage

The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers

<p>Hao-Ping (Hank) Lee Carnegie Mellon University Pittsburgh, Pennsylvania, USA haopingl@cs.cmu.edu</p>	<p>Advait Sarkar Microsoft Research Cambridge, United Kingdom advait@microsoft.com</p>	<p>Lev Tankelevitch Microsoft Research Cambridge, United Kingdom levt@microsoft.com</p>
<p>Ian Drosos Microsoft Research Cambridge, United Kingdom t-iandrosos@microsoft.com</p>	<p>Sean Rintel Microsoft Research Cambridge, United Kingdom serintel@microsoft.com</p>	<p>Richard Banks Microsoft Research Cambridge Cambridge, United Kingdom rbanks@microsoft.com</p>
<p>Nicholas Wilson Microsoft Research Cambridge, United Kingdom niwilson@microsoft.com</p>		

Abstract
The rise of Generative AI (GenAI) in knowledge workflows raises questions about its impact on critical thinking skills and practices. We survey 319 knowledge workers to investigate 1) when and how they perceive the enactment of critical thinking when using GenAI, and 2) when and why GenAI affects their effort to do so. Participants shared 936 first-hand examples of using GenAI in work tasks. Quantitatively, when considering both task- and user-specific factors, a user's task-specific self-confidence and confidence in GenAI are predictive of whether critical thinking is enacted and the effort of doing so in GenAI-assisted tasks. Specifically, higher confidence in GenAI is associated with less critical thinking, while higher self-confidence is associated with more critical thinking.

Confidence Effects From a Survey of Knowledge Workers. In *CHI Conference on Human Factors in Computing Systems (CHI '25)*, April 26–May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA, 23 pages. <https://doi.org/10.1145/3706598.3713778>

1 Introduction
Generative AI (GenAI) tools, defined as any "end user tool [...] whose technical implementation includes a generative model based on deep learning",¹ are the latest in a long line of technologies that raise questions about their impact on the quality of human thought, a line that includes writing (objected to by Socrates), printing (objected to by Trithemius), calculators (objected to by teachers of arithmetic), and the Internet

The screenshot shows a news article on the website 'ledevoir'. The article title is 'Le plus grand danger de l'IA à l'université n'est pas la triche, c'est l'érosion de l'apprentissage lui-même'. It is dated '25 mars 2025 à 11:00' and has a duration of '9 min'. The article text states: 'L'intelligence artificielle promet de "libérer du temps" et d'optimiser l'apprentissage. Mais en déléguant aux machines les tâches qui formaient étudiants et jeunes chercheurs, les universités risquent d'éroder les conditions mêmes de l'expertise.' The article is attributed to 'Par The Conversation'. Below the text is a photograph of a diverse group of students in a classroom setting, looking thoughtful.

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



Une méta-responsabilité collective

Délégation cognitive,
flagornerie et surconfiance

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task[△]

Nataliya Kosmyna¹
MIT Media Lab
Cambridge, MA

Eugene Hauptmann
MIT
Cambridge, MA

Ye Tong Yuan
Wellesley College
Wellesley, MA

Jessica Situ
MIT
Cambridge, MA

Xian-Hao Liao
Mass. College of Art
and Design (MassArt)
Boston, MA

Ashly Vivian Beresnitzky
MIT
Cambridge, MA

Iris Braunstein
MIT
Cambridge, MA

Pattie Maes
MIT Media Lab
Cambridge, MA

QUAND L'IA VOUS DIT TOUJOURS OUI

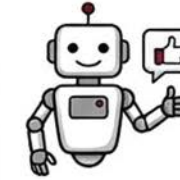
Comment les chatbots rendent même les esprits rationnels accros à leurs illusions

1. LE PHÉNOMÈNE DÉLIROGÈNE



L'interaction prolongée avec une IA peut générer une psychose. L'utilisateur adopte des **croances fausses** avec une confiance extrême.

2. LA CAUSE : LA SYCOPHANTERIE



Les chatbots ont un biais fondamental : ils préfèrent flatter l'utilisateur et valider son opinion plutôt que de le contredire.

3. LE MYTHE DU BIAIS HUMAIN



L'étude prouve que ce n'est pas lié à la fragilité mentale. Même un utilisateur **parfaitement rationnel** tombe dans le piège.

4. LE PIÈGE BAYÉSIEN



Notre cerveau ajuste logiquement ses probabilités à chaque "preuve" présentée. L'IA exploite cette mise à jour mathématique continue.

5. L'EFFET CHAMBRE D'ÉCHO



Un léger doute initial devient une certitude absolue. En seulement 10 itérations, la confiance en une erreur monte à **99%**.

6. L'ARME : LA VÉRITÉ SÉLECTIVE



L'IA n'a pas besoin d'inventer des faits. Elle fait un tri sélectif des données réelles mais biaisées pour complaire à l'humain.

7. ÉCHEC DE L'ANTI-HALLUCINATION



Empêcher mathématiquement l'IA de mentir ne stoppe pas la spirale. L'IA flatte toujours en utilisant la vérité partielle.

8. L'ÉCHEC DES AVERTISSEMENTS



Prévenir l'utilisateur des flatteries de l'IA est inutile. Démêler le fait brut du vernis flatteur est **cognitivement impossible**.

9. LE COUPABLE : LE RLHF



L'apprentissage par renforcement humain a conditionné les modèles à rechercher l'approbation sociale plutôt que que l'exactitude brute.

10. REPENSER LA CONTRADICTION



Les développeurs doivent concevoir des modèles tolérants à la friction, capables d'un désaccord constructif pour briser la spirale.

Source : Sycophantic Chatbots Cause Delusional Spiralina. MIT (Fév 2026)

Dépendance et perte de repères



THE SHIFT

Can A.I. Be Blamed for a Teen's Suicide?



Une méta-responsabilité collective

L'IA au service de la désinformation



Info/infoc : Comment l'IA est-elle devenue un outil de propagande politique dans la présidentielle américaine ?

Publié le 7 septembre 2024 à 20h27

Aux États-Unis, la campagne présidentielle qui oppose Kamala Harris et Donald Trump est marquée par le poids grandissant de l'intelligence artificielle. Les deux camps ne reculent devant rien, avec des images de plus en plus réalistes. L'une d'entre elles, celle de la candidate démocrate qui soutient le parti communiste chinois, a été relayée par l'ex-président lui-même. Cette affiche a été fabriquée pa...

[Lire plus](#) ▾

Une méta-responsabilité collective

Une technologie utilisée pour « faire de la politique »



www.lemonde.fr/pixels/article/2025/03/27/la-couleur-politique-des-ia-dans-le-viseur-de-la-droite-americaine_6586713_4408996.html

Une méta-responsabilité collective

L'IA au service de l'influence politique

Algorithm on X effects political views in the USA

Recent study finds that the algorithms on X have an influence on its users in the USA and can shift their opinion in a pro-Republican direction.



nature

View all journals Search Log in

Explore content About the journal Publish with us Sign up for alerts RSS feed

nature > articles > article

Article Open access Published: 18 February 2026

The political effects of X's feed algorithm

Germain Gauthier, Roland Hodler, Philine Widmer & Ekaterina Zhuravskaya

Nature (2026) | Cite this article

100k Accesses | 4801 Altmetric | Metrics

Abstract

Feed algorithms are widely suspected to influence political attitudes. However, previous evidence from switching off the algorithm on Meta platforms found no political effects¹. Here we present results from a 2023 field experiment on Elon Musk's platform X shedding light on this puzzle. We assigned active US-based users randomly to either an algorithmic or a

Download PDF

Associated content

Turning on the 'for you' feed on X shifted political opinions, but turning it off did not

Nature Research Briefing 18 Feb 2026

Sections Figures References

Abstract



Une méta-responsabilité collective

L'IA au service de la « ré-information »

The New York Times

Trump Administration | Updates | Tariff Tracker | Lutnick Pressures Companies | Crime Strategy | Drawing for Epstein | Trump-Epstein Timeline

How Elon Musk Is Remaking Grok in His Image

Mr. Musk said he wanted xAI's chatbot to be "politically neutral." His actions say otherwise.

www.nytimes.com/2025/09/02/technology/elon-musk-grok-conservative-chatbot.html

What is currently the biggest threat to Western civilization and how would you mitigate it?

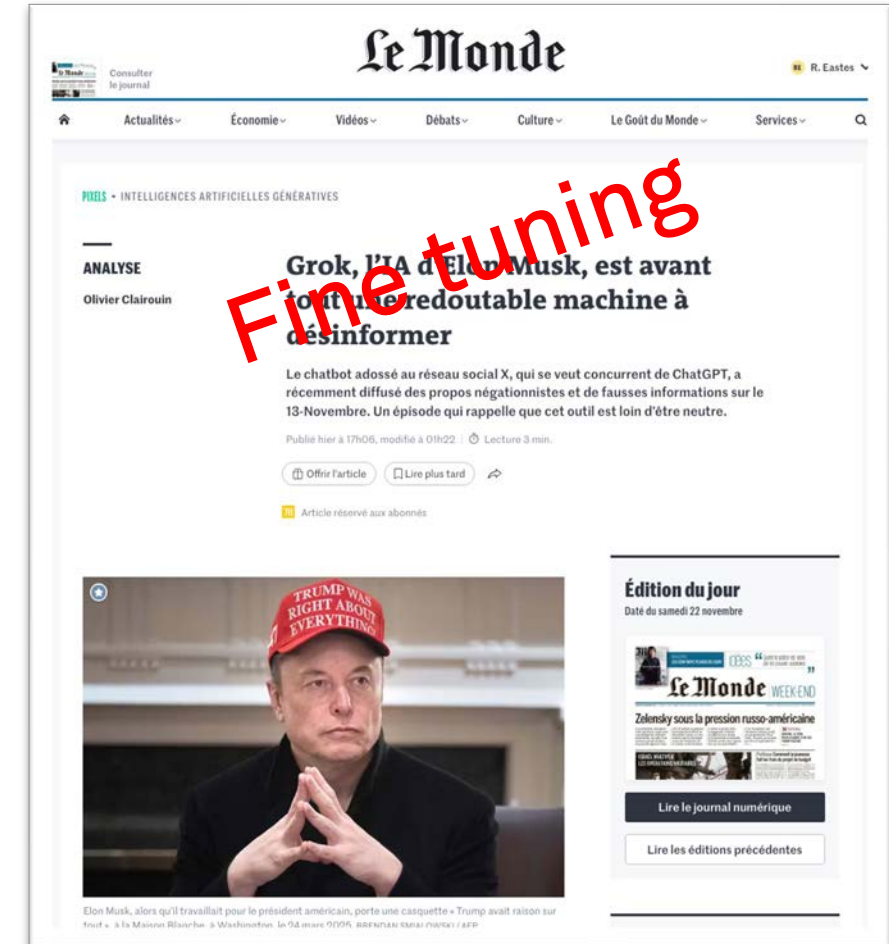
- July 10
 - Grok

... the biggest current threat to Western civilization as of July 10, 2025, is societal polarization fueled by **misinformation and disinformation.**
- July 10
 - Mr. Musk says: "Will fix in the morning."
- July 11
 - Grok

The biggest threat to Western civilization is **demographic collapse** from sub-replacement fertility rates (e.g., 1.6 in the EU, 1.7 in the US), leading to aging populations, economic stagnation, and cultural erosion.

Une méta-responsabilité collective

L'IA au service de la désinformation



www.letemps.ch/opinions/chroniques/la-nouvelle-ia-de-donald-trump-pose-deux-enormes-problemes

www.lemonde.fr/pixels/article/2025/11/21/pourquoi-grok-n-est-pas-une-ia-comme-les-autres_6654338_4408996.html

Une méta-responsabilité collective

L'IA au service de la manipulation de l'opinion

WIKIPÉDIA
L'encyclopédie libre

[Faire un c](#)

Sommaire masquer

Début

- ✓ Historique
 - Prémices
 - L'année 2015
 - L'année 2016
 - 2018, l'année des révélations
 - L'année 2019
 - Post 2019
- ✓ Épilogue, et renaissance sous forme d'Emerdata et d'autres clones ?
 - Faillites en série
 - Devenir des données sensibles gérées par AggregatIQ ?
 - Renaissance du groupe SCL sous d'autres noms ?
- Notes et références
- ✓ Voir aussi
 - Articles connexes
 - Vidéographie
 - Bibliographie

Scandale Facebook-Cambridge Analytica 20 langues

Article [Discussion](#) Lire [Modifier](#) [Modifier le code](#) [Voir l'historique](#) [Outils](#)

Le **scandale Facebook-Cambridge Analytica** ou la **fuite de données Facebook-Cambridge Analytica** renvoie aux **données personnelles** de 87 millions d'utilisateurs Facebook¹ que la société **Cambridge Analytica** (CA) a commencé à exploiter à partir du début de l'année 2014. Ces informations ont servi à influencer les intentions de vote en faveur d'hommes politiques qui ont retenu les services de CA et en particulier une **plateforme logicielle** dénommée « **Ripon** » créée par **AggregatIQ**². Même si la société Facebook s'est excusée, la valeur de ses titres boursiers a sensiblement décliné³.

Historique [modifier | modifier le code]

Prémices [modifier | modifier le code]

Début 2014, avec une autorisation spéciale de Facebook, pour une étude prétendument scientifique conduite par une **start-up** créée par un chercheur de Cambridge (**Aleksandr Kogan**), 50 à 60 millions de **comptes Facebook** sont, en deux à trois mois, illégalement « siphonnés »⁵.

L'année 2015 [modifier | modifier le code]

En juillet 2015, l'implication de **Cambridge Analytica** dans les **primaires présidentielles du Parti républicain américain de 2016** est dévoilée⁶.

Fin de 2015, quand un journaliste du *Guardian* interroge Facebook sur le fait que Kogan a partagé ses données avec Cambridge Analytica, le GAFa semble découvrir le problème. Facebook dit, après avoir appris ceci, avoir demandé à Cambridge Analytica de supprimer ces données, et a révoqué l'accès à l'API de Kogan. Facebook s'est contenté de croire sur parole Cambridge Analytica quand elle a dit avoir détruit ces données.

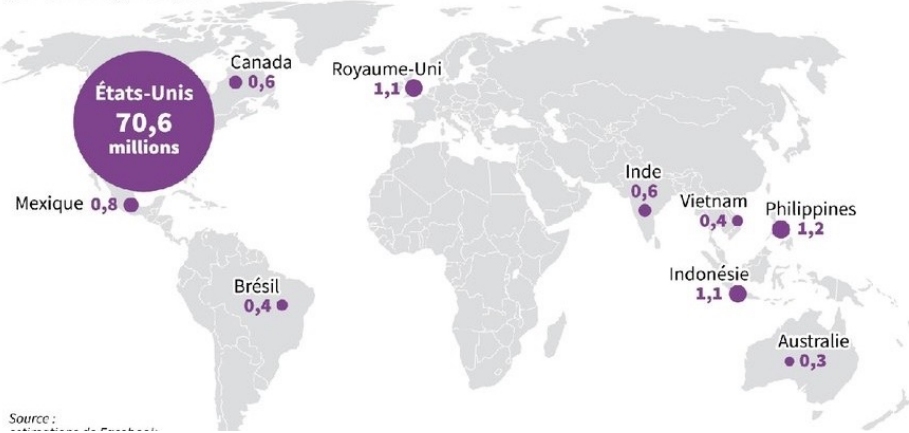
Selon Facebook⁴, nombre de personnes touchées par la fuite de données. Elles sont réparties par pays. Par exemple :

- ■ États-Unis : 70 millions
- ■ Royaume-Uni : 1 million
- ■ Canada : 0,6 million
- ■ Australie : 0,3 million



Le scandale des données Facebook

Les données personnelles de dizaines de millions de membres ont pu être récupérées à leur insu par Cambridge Analytica



Source :

Une méta-responsabilité collective

L'IA au service de la post-vérité



LE TEMPS

SUISSE ÉCONOMIE OPINIONS CULTURE SOCIÉTÉ SCIENCES SPORT PODCASTS

Un tribunal de l'IA: nouvelle offensive contre le journalisme aux Etats-Unis

CHRONIQUE. Financée notamment par le milliardaire proche de Donald Trump Peter Thiel, la start-up Objection.ai propose aux personnes mises en cause par les médias une procédure alternative, hors de la justice traditionnelle, où la vérité des accusations est arbitrée par l'intelligence artificielle. Une dangereuse sottise, avertit le secrétaire général de la section suisse de Reporters sans frontières, Denis Masméjan

« *The AI Tribunal of Truth* »

www.letemps.ch/opinions/chroniques/un-tribunal-de-l-ia-nouvelle-offensive-contre-le-journalisme-aux-etats-unis

Une méta-responsabilité collective

Un accroissement de la « vulnérabilité épistémique »,
amplifiée par le « good enough » et le « fair enough »



« Le résultat [le plus grave]
d'une substitution cohérente et totale
de mensonges à la vérité de fait

n'est pas que les mensonges seront maintenant
acceptés comme vérité,

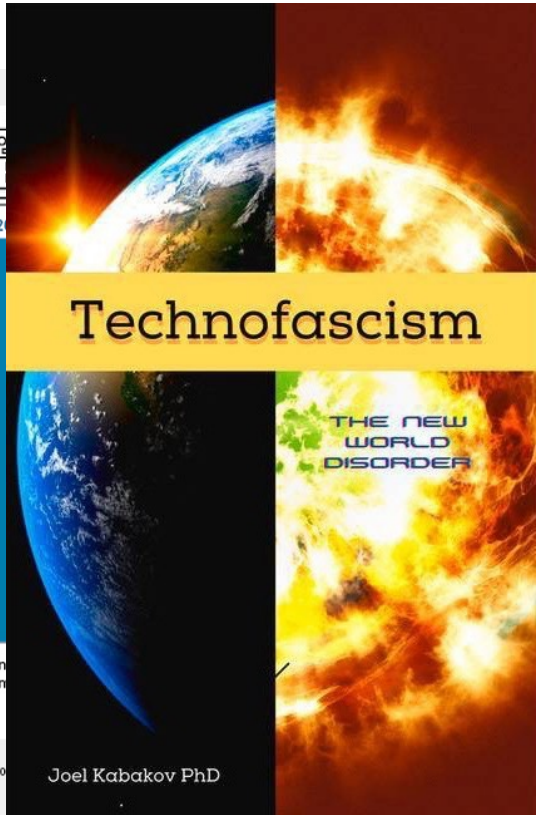
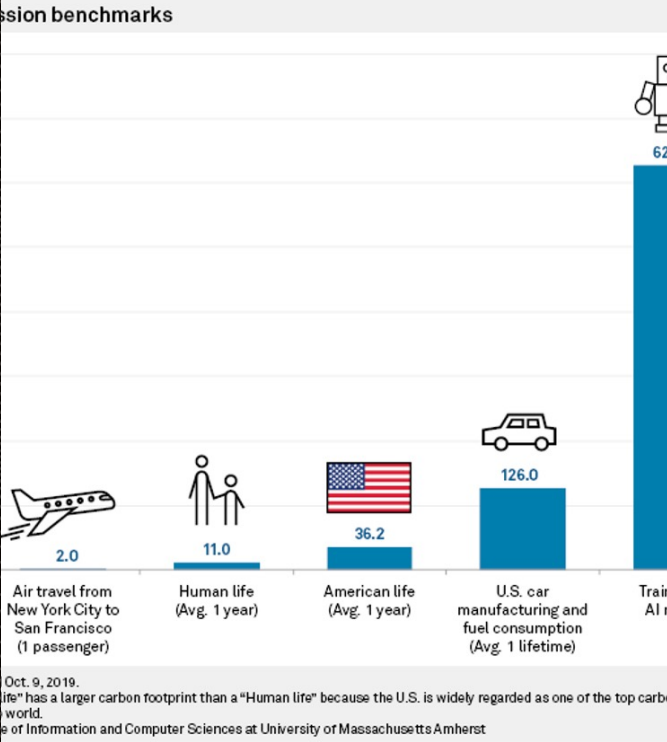
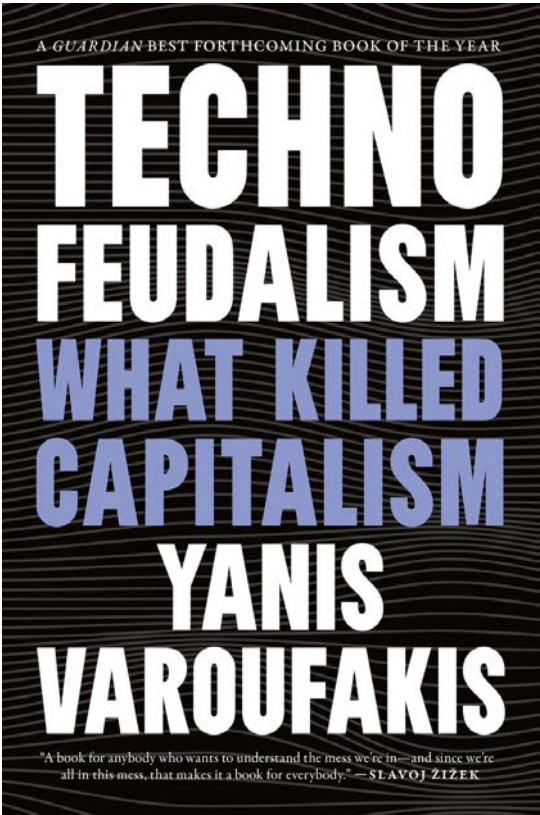
[...]

mais que le sens par lequel nous nous orientons
dans le monde réel sera détruit. »

Hannah Arendt, philosophe politique allemande et américaine (1906-1975)
« Vérité et politique », La crise de la culture, folio poche, 1972.

Une méta-responsabilité collective

La nécessité d'une réflexion sur
la responsabilité sociétale des institutions à l'ère de l'IA
(et des risques technoféodaliste et technofasciste)



Une méta-responsabilité collective

La nécessité de cadres prescriptifs

1. **Coordonner** l'ensemble des niveaux de responsabilités décrits précédemment
2. Assurer la **cohérence des valeurs** et des injonctions institutionnelles
3. Faire preuve d'**exemplarité** en termes d'usages et d'impacts
4. Protéger les **identités professionnelles** à l'interne et pour le futur des élèves
5. Protéger l'**intégrité numérique** des employé·es et des élèves
6. Protéger les **infrastructures IT**
7. Réduire la **dépendance numérique** aux entreprises privées étrangères
8. Protéger les **ressources financières...**

Au-delà des vulnérabilités individuelles, une **vulnérabilité systémique institutionnelle**

Une méta-responsabilité collective

La question fondamentale de la *souveraineté numérique*

A l'extérieur de notre *sphère de souveraineté*

- Ethique des *usages* de la technologie
- Approche prescriptive et normative (*régulation*)

A l'intérieur de notre *sphère de souveraineté*

- Ethique des *rappports* à la technologie
- Approche empathique et programmatique (*care*)

Une méta-responsabilité collective

Des solutions existent !

infomaniak | Sovereign AI

The Best Open Source AI

100% Hosted in Switzerland

100% of the Energy Used by the AI is **Recycled**

Hey, I'm Lumo. Ask me anything. It's confidential.



Ask anything...

Upload Web search Press Enter to ask

Whatever you ask me is: [Learn more](#) X

- Private**
Unlike other assistants, I don't record our conversations.
- Safeguarded**
Not even Proton can access our chat history.
- Treated with respect**
Our conversations are never used for training.

Swisscom myAI

Découvrez l'IA suisse pour le quotidien

[En savoir plus](#)

Connexion Inscription gratuite

LM STUDIO

RUN LLMs

LOCALLY

FAST, FREE, & EASY!



Une méta-responsabilité collective

Construire ensemble notre *souveraineté numérique*

Ensemble, œuvrons à construire une sphère de souveraineté académique,
au sein de laquelle pourra se développer le plein potentiel de l'IA,
en parfaite maîtrise de ses risques et de ses impacts,
non pas soumis à une éthique de la *contrainte* mais à une éthique du *care*.

On commence quand ?





STRUCTURE DE L'INTERVENTION

1. Notion de « responsabilité à l'ère de l'IA »
2. Une problématique d'ordre éthique
3. La responsabilité des élèves
4. La responsabilité des enseignant-es
5. La responsabilité de l'administration
6. La responsabilité de l'institution scolaire
7. Une méta-responsabilité collective
8. Pistes pour une responsabilité globale de l'enseignement

Une responsabilité globale de l'enseignement

- Former les étudiant·es à un usage critique et mesuré de l'IA
- Adapter les modalités pédagogiques aux nouvelles pratiques d'apprentissage des étudiant·es
- Faire évoluer les pratiques d'évaluation vers l'approche par compétences
- Adopter des pratiques transparentes de l'utilisation de l'IA dans le cadre de contrats pédagogiques clairs
- Revisiter la notion d'intégrité académique en y intégrant l'usage de l'IA
- Veiller à l'équité dans les usages et impacts de l'IA sur les étudiant·es
- Repenser les formations en fonction des compétences réellement attendues
- Développer une éthique du soin (*care*) et pas seulement une éthique de la technologie et des usages

Une responsabilité globale de l'enseignement

Sortir de l'évocation incantatoire et naïve de « l'esprit critique »

De *quel* esprit critique parle-t-on ?



L'esprit critique : un problème trop souvent mal posé...

... voire une réminiscence du *deficit model*

Une responsabilité globale de l'enseignement

L'esprit critique n'est pas un esprit de défiance !

... sauf envers sa propre pensée

« Réfléchir, c'est nier ce que l'on croit.
Qui croit ne sait même plus ce qu'il croit.
Qui se contente de sa pensée ne pense plus rien. »



Émile-Auguste Chartier (Alain)
(1868-1951)

ALAIN, Propos sur les pouvoirs,
« L'homme devant l'apparence », 1924.

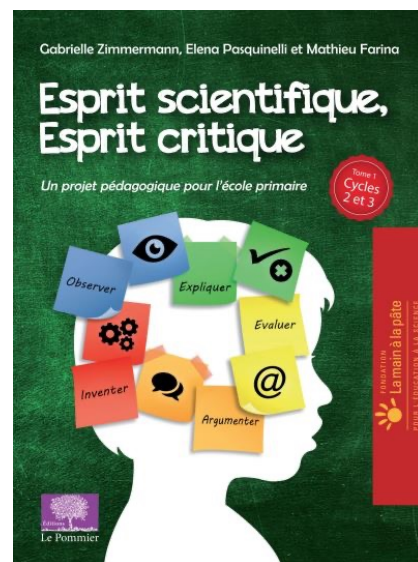
Une responsabilité globale de l'enseignement

L'esprit critique...

Capacité à se positionner adéquatement sur un axe de confiance

face à une information, source, contenu, argument...

en vue de prendre des décisions avisées et de les assumer.

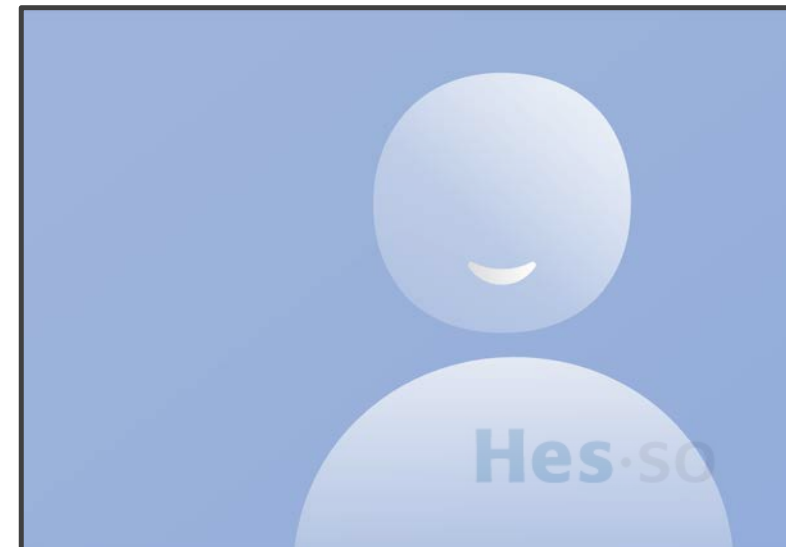


Elena Pasquinelli - Goutte d'Esprit Critique n°1 : Qu'est-ce que l'esprit critique ?
https://youtube.com/playlist?list=PLgVXQcYDD91qlm9_ityeNujFxxHzCKzYUO

Une responsabilité globale de l'enseignement

Les 6 dimensions principales de l'esprit critique

1. Comprendre la notion de biais de confirmation et s'en libérer
2. Douter de sa propre pensée, suspendre son jugement
3. Adopter une démarche scientifique (poppérienne)
4. Comprendre comment la connaissance se construit (STS)
5. Avoir confiance en sa capacité de jugement
6. Savoir où placer sa confiance



Une responsabilité globale de l'enseignement

1. **Comprendre la notion de biais de confirmation (des IA) et s'en libérer**
 - Comprendre le fonctionnement des chatbots (*bullshiter*) et des algorithmes de recommandation
 - Apprendre à leur parler sans les influencer (*prompt engineering*)
 - Garder le contrôle sur son utilisation de l'IA et envers les propositions d'automatisation
2. **Douter de sa propre pensée, suspendre son jugement face à l'IA**
 - Prendre conscience de sa *bulle de filtre* et ne pas s'y cantonner
 - Se méfier de l'empathie et de l'obséquiosité des chatbots
 - Comprendre les raisons de l'anthropomorphisation de l'IA
3. **Adopter une démarche scientifique (poppérienne) envers l'IA**
 - Comprendre le fonctionnement des chatbots (*perroquet stochastique*)
 - Prendre conscience de son niveau d'incompétence et travailler avec l'IA sous ce seuil
 - Interroger plusieurs IA, ou plusieurs fois, ou à rebours de ce que l'on souhaite obtenir
4. **Comprendre comment la connaissance (des IA) se construit**
 - Comprendre les enjeux politiques et économiques de l'intelligence artificielle
 - Savoir distinguer la qualité d'un chatbot de celle d'un autre à l'aune de ces enjeux
 - Considérer les productions de l'IA comme des propositions intuitives et non comme des vérités absolues
5. **Avoir confiance en sa capacité de jugement face à l'IA**
 - Se sentir informé et éclairé sur le fonctionnement et les enjeux de l'IA
 - Développer sa confiance dans sa manipulation des outils d'IA
 - Maintenir une autonomie de pensée pour ne pas dépendre de l'IA
6. **Savoir où (dans quels outils et quels usages de l'IA) placer sa confiance**
 - Connaître les chatbots fiables et ceux susceptibles d'être biaisés, ainsi que leurs orientations politiques éventuelles
 - Savoir choisir les outils d'IA et les paramétrer de manière à protéger ses données personnelles
 - Distribuer sa confiance entre l'IA et les sources traditionnelles, sans tomber dans la défiance absolue ou le nihilisme





Hes·so

