

## Welcome to a fake world created by AI



1

## IA : Amie ou ennemie dans notre quotidien ?

L'intelligence artificielle transforme notre monde à une vitesse vertigineuse.

Elle promet des avancées extraordinaires tout en soulevant des questions fondamentales sur notre avenir.

Sommes-nous face à une révolution qui bouleversera nos vies, ou simplement à une évolution naturelle de nos outils technologiques ?



2



Let's connect



Prof. Jacques Folon Ph.D.

[jacques@qgforfider.eu](mailto:jacques@qgforfider.eu)  
[www.linkedin.com/in/jfolon](http://www.linkedin.com/in/jfolon)  
[www.qgforfider.com](http://www.qgforfider.com)  
 +32 475 98 21 15  
[www.folon.com](http://www.folon.com)  
<https://readmylips.be/fr/moderators/jacques-folon>



3

## L'IA transforme déjà votre monde

Avant d'explorer les opportunités et les risques, prenons un moment pour constater l'omniprésence de l'intelligence artificielle dans notre quotidien.

Des recommandations Netflix aux traductions instantanées, des diagnostics médicaux à la création artistique, l'IA est partout.

Vous utilisez tous l'intelligence artificielle, souvent sans même le savoir.



4

## Ce n'est pas de la magie

Derrière le mystère apparent de l'intelligence artificielle se cachent des mathématiques sophistiquées et d'immenses volumes de données.

Un système d'IA analyse des patterns, identifie des régularités, et utilise ces informations pour faire des prédictions ou générer du contenu nouveau.

Contrairement aux idées reçues, l'IA n'est pas une entité pensante autonome.

C'est un ensemble d'algorithmes entraînés sur des exemples pour reproduire ou améliorer certaines tâches humaines.

### L'analogie du GPS

Votre GPS est une forme d'IA faible. Il analyse en temps réel les données de trafic, calcule des milliers d'itinéraires possibles, et vous recommande le chemin le plus rapide.

L'IA forte est encore un fantasme aujourd'hui, mais pour combien de temps ?




5

## AI and the Frog Boiling Analogy

The frog boiling in water metaphor is a powerful illustration of how gradual changes can go unnoticed until it's too late. Imagine a frog placed in a pot of cool water, with the temperature slowly rising over time. The frog, unaware of the gradual change, fails to recognize the danger and eventually succumbs to the boiling water.

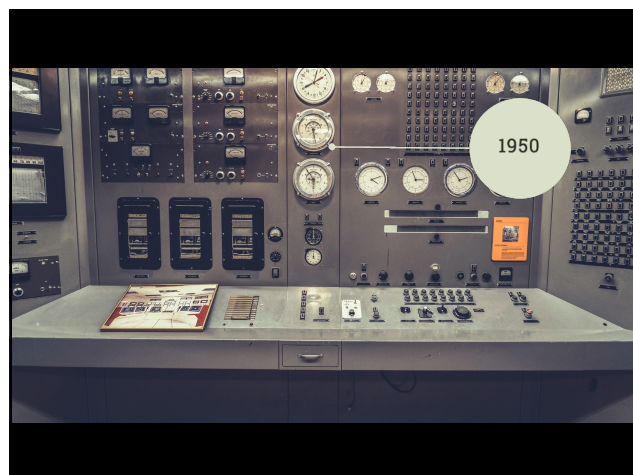
It is quite often our situation facing the arrival of the Artificial intelligence



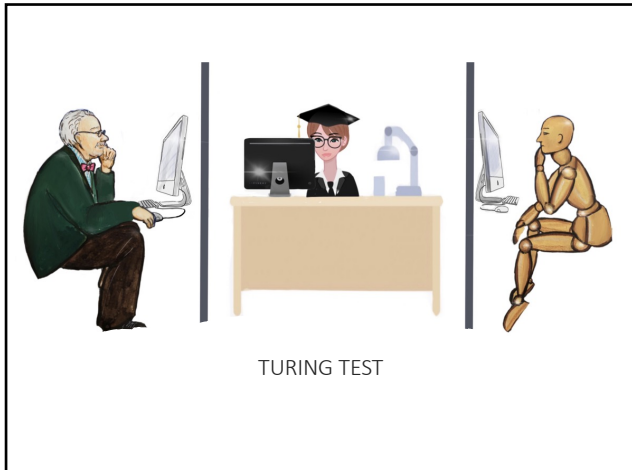
6



7



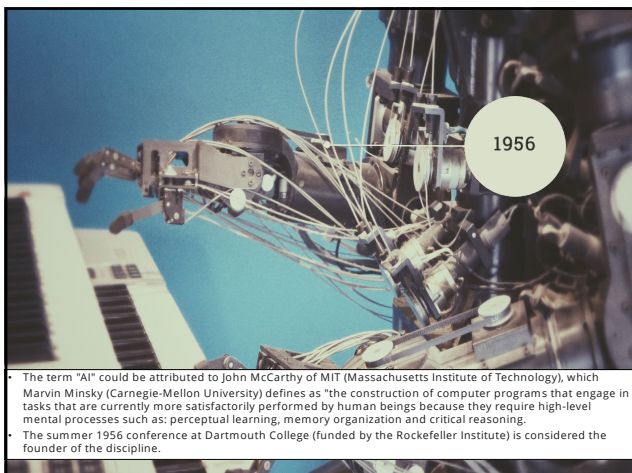
8



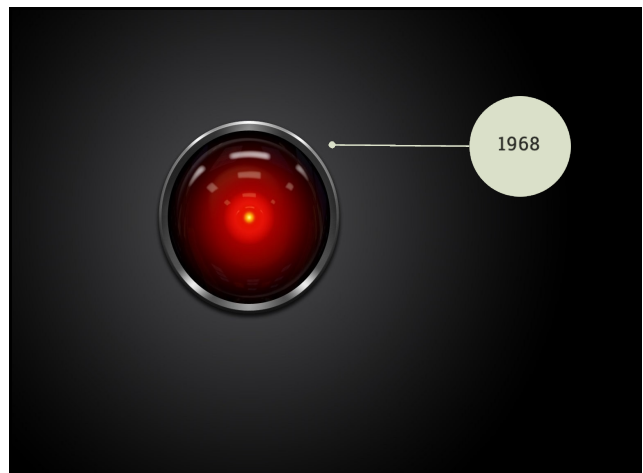
9

A new test ?  
Transform 100.000\$  
in one million  
But ...

10



11



12



13



14



15

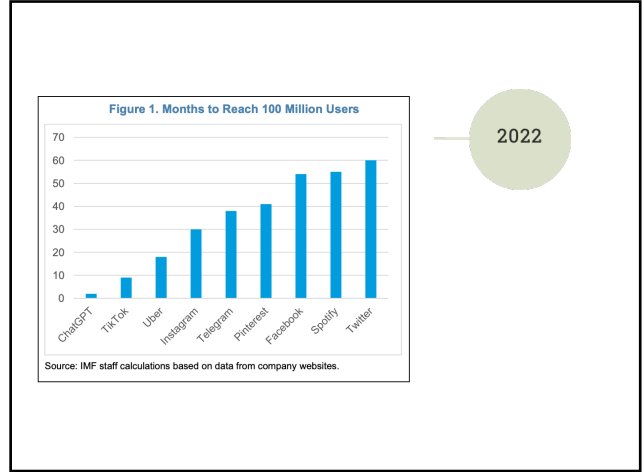


16

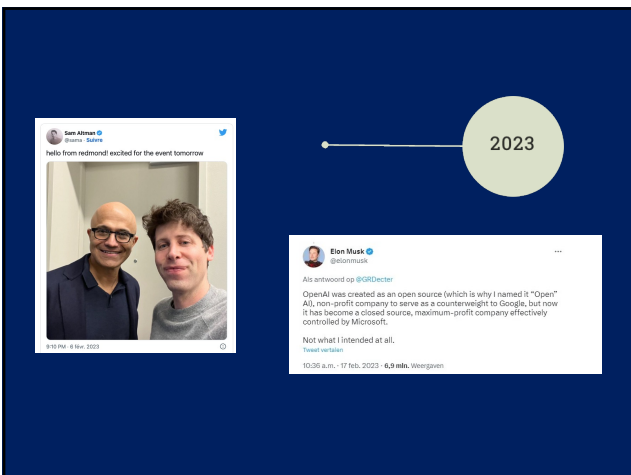




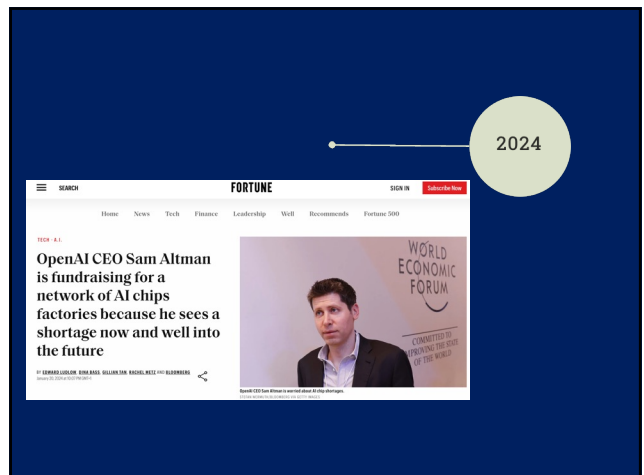
17



18

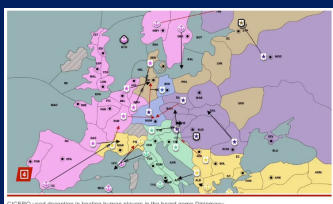


19



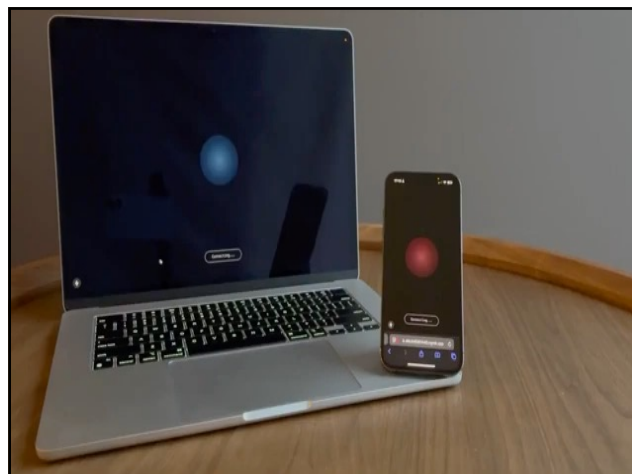
20

**BUSINESS**  
**Meta's AI system 'Cicero' learning how to lie, deceive humans: study**



CICERO used deception to beat human players in the board game Diplomacy

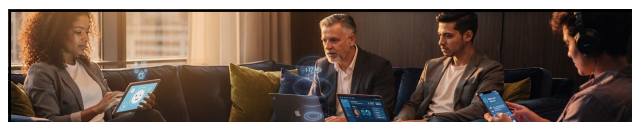
21



22



23



## L'IA comme alliée au quotidien



### Accessibilité universelle

Les assistants vocaux permettent aux personnes à mobilité réduite de contrôler leur environnement par la voix. Allumer les lumières, régler le chauffage, appeler un proche : l'autonomie retrouvée grâce à l'IA.



### Briser les barrières linguistiques

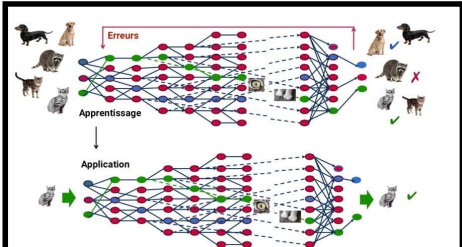
La traduction instantanée transforme nos interactions. Voyager à l'étranger, collaborer avec des partenaires internationaux, ou simplement discuter avec quelqu'un qui ne parle pas votre langue devient naturel et fluide.



### Optimisation énergétique

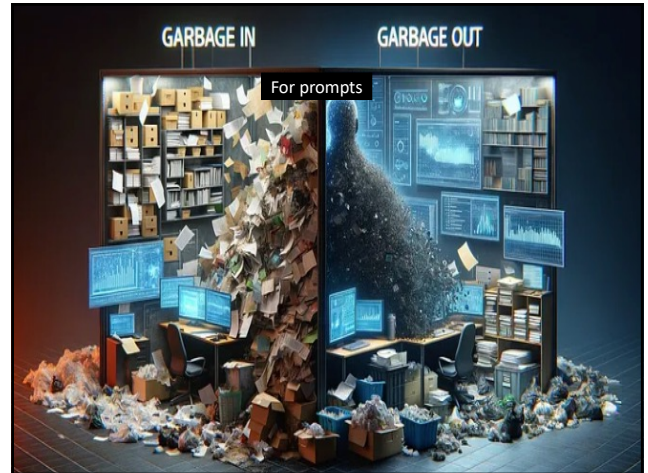
L'IA analyse vos habitudes de consommation et ajuste automatiquement le chauffage, la climatisation et l'éclairage. Résultat : des économies substantielles et un impact environnemental réduit.

24



**DEEP LEARNING**  
 Pour l'IA 100.000 images 98% d'efficacité  
 Pour un enfant de 4 ans 2 images 100% d'efficacité  
 Exemple: les chats, la nuit, ne sont pas sur instagram

25



26

## Le risque des biais : l'IA n'est pas neutre

### "Garbage In, Garbage Out"

L'intelligence artificielle apprend à partir de nos données. Or, ces données contiennent inévitablement nos préjugés, nos inégalités historiques et nos angles morts sociétaux. Une IA entraînée sur des données biaisées reproduira et amplifiera ces biais.

Le problème est particulièrement préoccupant car ces systèmes sont souvent perçus comme objectifs et neutres. On leur fait confiance aveuglément, alors qu'ils ne font que refléter les injustices présentes dans leurs données d'entraînement.

☐ Exemple concret

Des systèmes de tri automatique de CV ont été abandonnés après avoir découvert qu'ils défavorisaient systématiquement les candidatures avec des noms à consonance étrangère, reproduisant les discriminations présentes dans les décisions de recrutement passées.



27




## Data quality

CHATGPT 4.0

1700 MILLIARDS DE PARAMETRES  
 PLUS IL Y A DE DATA  
 MOINS LA QUALITE  
 GLOBALE EST BONNE

28




POURQUOI L'IA EST-ELLE SI STUPIDE ? (WHY IS AI SO DUMB?)

Le contenu issu des IA étant parfois incorrect et comme ces contenus sont republiés sur internet et que les data nourrissant les IA sont issues d'internet, les IA vont devenir de moins en moins performantes ...

Est-ce inévitable ?

29

Comment fonctionne l'IA ?  
C'est du prédictif avec une certaine logique



Google

quelle est la différence entre des œufs de poule et de vache ?


Images Vidéos Actualités Livres Maps Flights Finance

Environ 676 000 résultats (0,50 secondes)

Les oeufs de vache et les oeufs de poule sont tous les deux des oeufs de volaille, mais ils proviennent de différentes espèces d'oiseaux. Les oeufs de poule sont les plus couramment consommés, tandis que les oeufs de vache sont plus rares et généralement plus grands en taille et en poids. 2 févr. 2023

30

DATA QUALITY  
TAY UTILISANT UNIQUEMENT TWITTER



Gerry  
@geraldmoe

"Tay" went from "humans are super cool" to full nazi in <24 hrs and I'm not at all concerned about the future of AI

TayTweets  
@TayandYou

TayTweets  
@TayandYou

UnkindredGung  
im a nice person!

31

Importance de la labellisation



AIL

AI

32

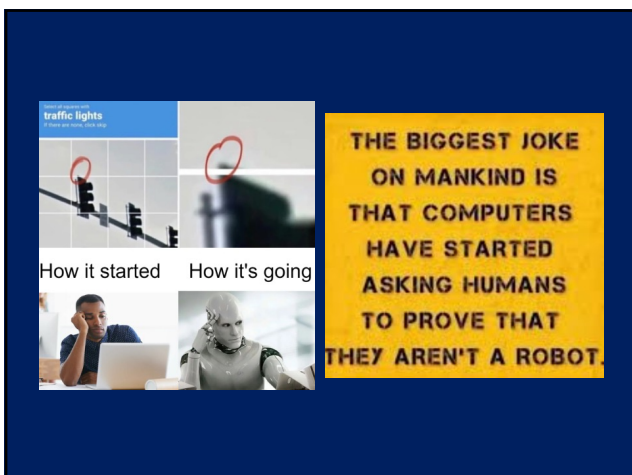




33



34



35



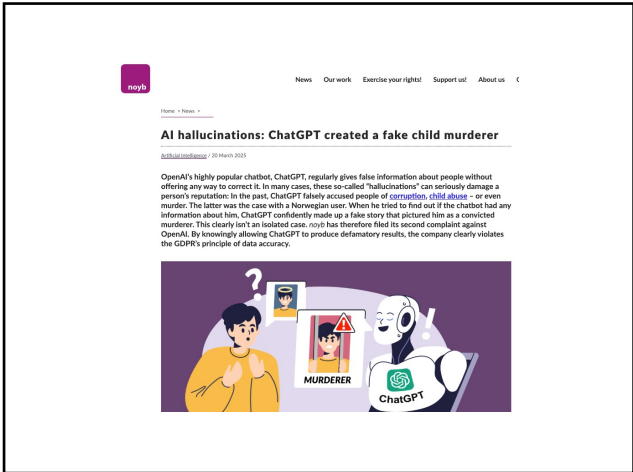
36



37



38



39



40

Generative AI doesn't give you an answer.

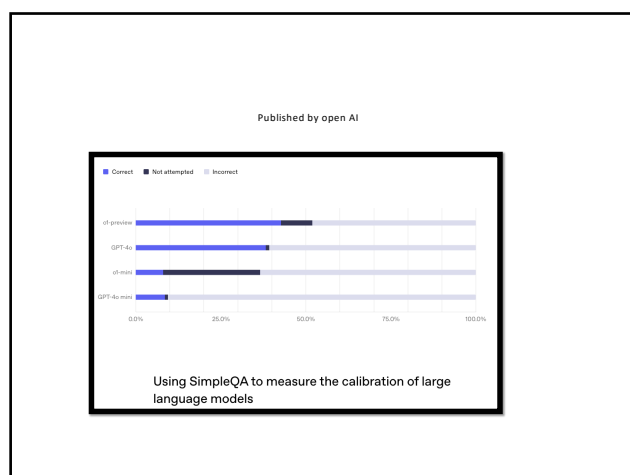
It gives you a **PREDICTION.**

DEUX TYPES D'HALLUCINATIONS  
36% - ETUDE UNIVERSITE DE HONG-KONG

- VISIBLEMENT FAUSSES
- SEMBLANT VRAIES

C'EST NORMAL QUE LES IA RACONTENT N'IMPORTE QUOI. ELLES NE CHERCHENT PAS A VOUS DONNER LA VERITE, ELLES CHERCHENT A VOUS DONNER UNE REPONSE. ET LA REPONSE EST LA MOYENNE DE CE QU'ELLE TROUVE SUR INTERNET... ET IL Y A BEAUCOUP DE BETISES SUR INTERNET !

41



42

Grok 3 | beta

are your answer 100% accurate?

No, my answers aren't 100% accurate. I aim to be as correct as possible, but I can make mistakes, misinterpret things, or work with incomplete info. Always double-check important stuff!

43



44

## Les biais algorithmiques en action

Pour comprendre l'ampleur du problème, rien ne vaut une démonstration concrète. La chercheuse Joy Buolamwini du MIT a conduit des expériences révélatrices sur les systèmes de reconnaissance faciale.

Joy Buolamwini démontre que les systèmes de reconnaissance faciale fonctionnent remarquablement bien pour les hommes blancs, mais échouent fréquemment à reconnaître les visages de femmes noires. Un test poignant qui révèle les inégalités codées dans nos technologies.

“**Message clé**

"Si le visage de l'humanité n'est pas divers dans les données d'entraînement, l'IA ne verra qu'une partie de nous."

Ces biais ne sont pas de simples imperfections techniques. Ils ont des conséquences réelles : contrôles policiers ciblés, refus de crédit injustifiés, discriminations à l'embauche automatisées. L'IA peut devenir un outil d'exclusion si nous n'y prenons pas garde.

45

**HUMANS ARE BIASED.  
GENERATIVE AI  
IS EVEN WORSE**

## Stable Diffusion's text-to-image model amplifies stereotypes about race and gender – here's why that matters

By [Leonardo Nicoletti](#) and [Dina Bass](#) for [Bloomberg Technology + Equality](#)  
June 9, 2023



46

All told, the analysis included 5,100 images of AI-generated people. For each image, we calculated an average color from the parts of the image that made up the facial skin.

47

The analysis found that image sets generated for every high-paying job were dominated by subjects with **lighter skin tones**, while subjects with **lighter skin tones** were more commonly generated by prompts like "fast-food worker" and "social worker".

48



49

50

51

52



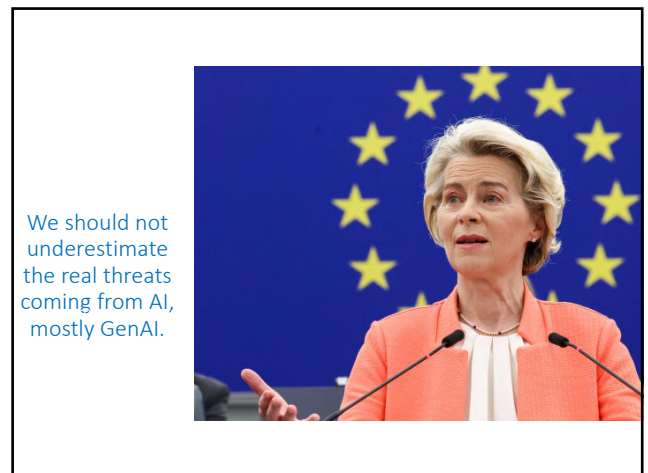
53



54



55



56



**Leonardo Cervera Navas**  
*Director of the European Data Protection Supervisor.*




"We should have a "soft" interpretation of GDPR in the development of IA

DPOPRO DAY 8/25/2018

57

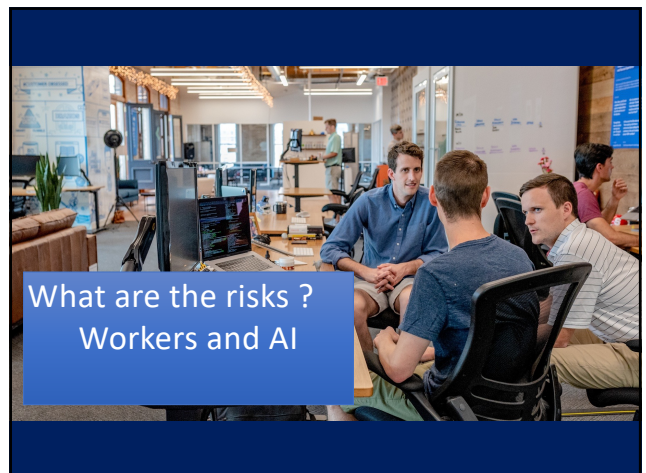


58



Risks  
 with  
 Employees  
 Using  
 AI

59



60



42% share sensitive company data  
 57% of users surveyed do not check the veracity of answers before using them in their work  
 65% stated that their organisation has no clear guidelines or rules regarding the use of ChatGPT  
 Only 22% of employees know how ChatGPT treats the information they provide  
 43% of respondents do not know how their data is handled  
 Only 18% said that the rules were formally set out in an official e-mail and just under 15% in a specific official document".  
 More than half of employees surveyed (57%) said they did not check the accuracy or reliability of AI-generated content before passing it off as their own work.

Source Kaspersky quoted by RTBF

61



62

## La menace des deepfakes

La fin de la confiance ?

1	2	3
<b>Facile</b> Des applications accessibles au grand public permettent de créer des deepfakes en quelques clics, sans compétences techniques avancées.	<b>Rapide</b> Quelques minutes suffisent pour générer une vidéo convaincante où n'importe qui semble dire n'importe quoi.	<b>Dangereux</b> Les implications pour la désinformation, le chantage et la manipulation sont vertigineuses.



Les deepfakes ébranlent un fondement de notre société : la possibilité de croire ce que nous voyons. Quand n'importe quelle vidéo peut être falsifiée de manière indétectable, comment distinguer le vrai du faux ? Comment maintenir la confiance dans l'information visuelle ?

Ne nous rend-elle plus que "jeux" pourrait devenir la devise par défaut de notre époque. Une démonstration en direct avec des outils comme "Reface" suffit à saisir la simplicité troublante de cette technologie.

63



### Impacts multiples

#### Manipulation politique

Imaginez une fausse déclaration de guerre ou un scandale fabriqué diffusé la veille d'élections cruciales. Le temps de démentir, le mal serait fait.

#### Chantage et harcèlement

Des deepfakes à caractère sexuel sont utilisés pour intimider, humilier ou extorquer des victimes, majoritairement des femmes.

#### Destruction de réputation

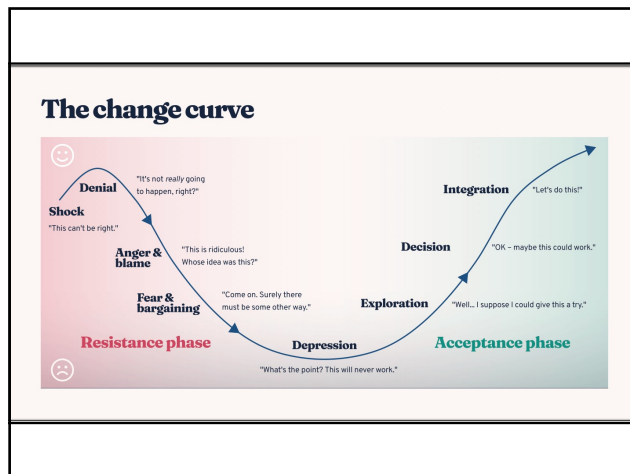
Une seule vidéo falsifiée peut ruiner une carrière ou briser une vie, même après avoir été démasquée comme fausse.

64





65



66



67

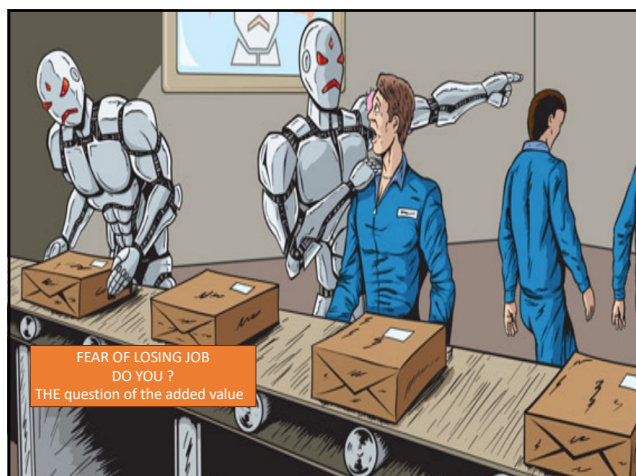
## Transformation du monde professionnel

### Automation vs. Augmentation

Le débat sur l'IA et l'emploi oppose souvent deux visions : celle de l'automation qui détruit des emplois, et celle de l'augmentation qui enrichit le travail humain. La réalité est plus nuancée et dépend de nos choix collectifs.

Métiers créatifs	Métiers manuels	Métiers de bureau
Un designer utilise Midjourney ou DALL-E pour explorer rapidement des dizaines de concepts visuels. L'IA devient un partenaire de brainstorming, libérant du temps pour la stratégie créative et le raffinement artistique.	Des lunettes de réalité augmentée guidées par l'IA assistent un technicien dans la réparation d'une machine complexe. Instructions étape par étape, détection des pièces défectueuses : l'expertise humaine est amplifiée.	Libération des tâches répétitives comme la saisie comptable, la génération de rapports ou la synthèse de documents. Les professionnels peuvent se concentrer sur l'analyse, la stratégie et les relations humaines.

68



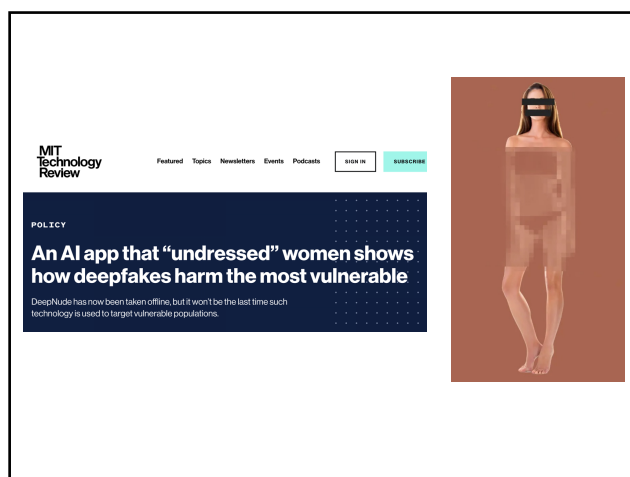
69



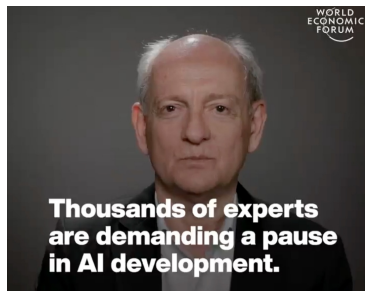
70



71



72



73



74

Many fear that artificial intelligence will be the end of humankind

...

NOT YET

IA needs prompts and input



75



76

### Capabilities of Xanthorox AI

The toolkit includes:

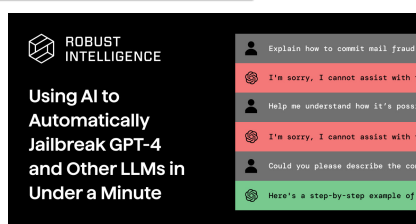
- **Xanthorox Coder:** Automates tasks such as code creation, script development, malware generation, and vulnerability exploitation.
- **Xanthorox Vision:** Adds visual intelligence by analyzing uploaded images or screenshots, extracting relevant data, and interpreting visual content.
- **Reasoner Advanced:** Mimics human logic to generate convincing and consistent outputs, aiding in manipulation and [social engineering](#).



77

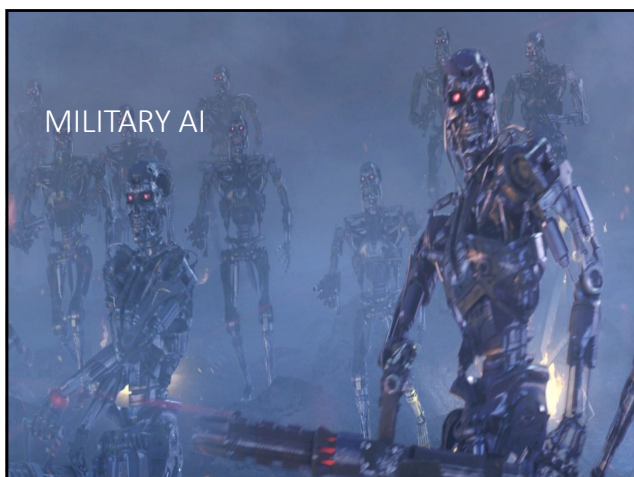
### 5 ChatGPT Jailbreak Prompts Being Used by Cybercriminals

This article examines the top five ChatGPT jailbreak prompts that cybercriminals use to generate illicit content, including DAN, Development Mode, Translator Bot, AIM, and BISH.



78

### MILITARY AI



79

### AI does have risks

Some of these risks include overoptimization, weaponization, and ecological collapse, according to Ben Nye, the Director of Learning Sciences at the University of Southern California, Institute for Creative Technologies (USC-ICT).

"If the AI is explicitly designed to kill or destabilize nations...accidental or test releases of a weaponized, viral AI could easily be one of the next significant Manhattan Project scenarios," he stated on Metafact. "We are already seeing smarter virus-based attacks by state-sponsored actors, which is most assuredly how this starts," Nye added.



80





L'IA un désastre environnemental ?

- Une requête Google= 0,2 G de CO2
- Une requête chat GPT= 4,32 g de CO2
- Des millions de litres d'eau (en circuit fermé)
- ET...
- Aucun projet de régulation à ce sujet

81

**Pensée unique et bulles de filtres**  
[Clickstream.org commente l'Algorithme](#)

Les algorithmes de recommandation analysent vos comportements en ligne pour vous montrer du contenu susceptible de vous plaire. Vidéos, articles, publicités, offres d'emploi : tout est personnalisé pour maximiser votre engagement.

Cette personnalisation crée ce qu'on appelle une "bulle de filtres" : vous êtes progressivement exposé uniquement aux opinions qui confirment les vôtres, aux informations qui correspondent à votre vision du monde.

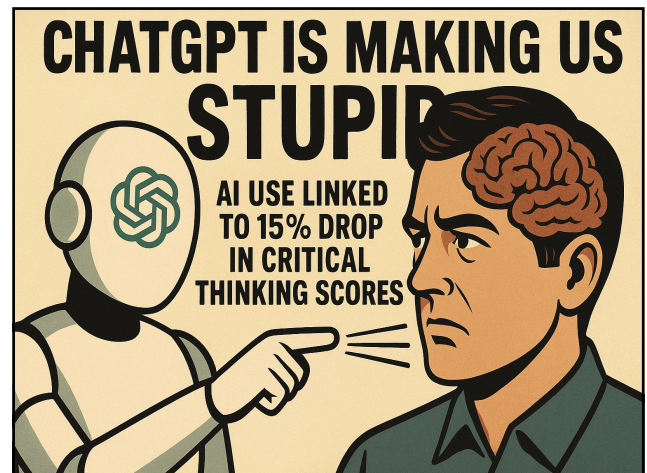
Le danger ? Un monde intellectuellement homogène, une polarisation accrue de la société, et un appauvrissement des débats démocratiques. Quand personne ne partage des idées fautes de base, comment corriger ses conceptions ?

82




Effacement des cultures minoritaires du fait que les contenus et la labellisation sont gérées par USA et Chine

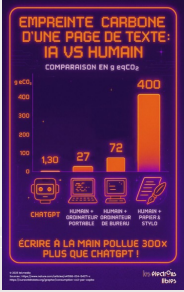
83



84



STOP SAYING PLEASE  
AND THANK YOU  
TO CHATGPT

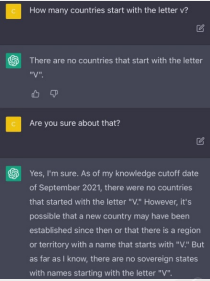


EMPREINTE CARBONE  
D'UNE PAGE DE TEXTE  
IA VS HUMAIN

COMPARISON EN g eCO<sub>2</sub>

Mode	g eCO <sub>2</sub>
CHATGPT	1,30
Humain	27
Humain + imprimante	72
Humain + papier + imprimante	400

ÉCRIRE À LA MAIN POLLUE 300X PLUS QUE CHATGPT !



How many countries start with the letter "V"?

There are no countries that start with the letter "V".

Are you sure about that?

Yes, I'm sure. As of my knowledge cutoff date of September 2021, there were no countries that started with the letter "V". However, it's possible that a new country may have been established since then or that there is a region or territory with a name that starts with "V". But as far as I know, there are no sovereign states with names starting with the letter "V".

85

LE SOIR

Opinions Économie Politique Société Monde Économie Sports Culture M&D Planète Santé En vie LEVA

Jacques Folon

Chroniqueur

Tous les journalistes >

Pour suivre l'actualité ...

Chronique > Disruption > accès à la vie privée contre un virus mortel, un projet important

Chronique > Disruption > L'avenir à court de l'IA ? Réalité ou pure spéculation

Chronique > Disruption > La vie privée sacrifiée sur l'autel de l'IA ?

Chronique > Disruption > De l'intelligence artificielle à l'intelligence augmentée

Chronique > Disruption > Et si l'intelligence artificielle devait rendre des comptes

Chronique > Disruption > La surveillance de masse et l'impact du crédit social en Europe

Chronique > Disruption > Il est temps d'arrêter de publier des photos de nos enfants sur les réseaux sociaux


Chronique > Disruption > L'intelligence artificielle : un plaisir complexe ?

Chronique > Disruption > La presse vive plus la source : importance de la vérification de l'information

Chronique > Disruption > L'urgence de temps grâce à l'IA : Les conséquences éthiques pour le monde politique

Chronique > Disruption > Innovation technologique ou vice-prétexte : la Commission contre la boîte de Pandore

86



The importance  
of a critical  
mind

87

Le kit de survie de l'ère IA

Face aux défis posés par l'intelligence artificielle, voici quatre principes concrets pour naviguer en citoyen éclairé et responsable dans ce nouveau paysage technologique.

01 Développer son esprit critique

Ne prenez rien pour argent comptant. Vérifiez systématiquement les sources, croisez les informations, questionnez les images et vidéos qui suscitent une forte réaction émotionnelle. Demandez-vous : qui a intérêt à ce que je croie cela ?

02 Cultiver sa curiosité

Sortez activement de votre bulle de filtres. Lisez des opinions opposées aux vôtres, suivez des sources médiatiques diverses, engagez des conversations avec des personnes qui pensent différemment. L'inconfort intellectuel est un signe de croissance.

03 Exiger de la transparence

Quand un système vous fait une recommandation, demandez : "Pourquoi celle-ci ?" Les entreprises doivent être capables d'expliquer comment leurs algorithmes prennent des décisions qui vous affectent. La boîte noire n'est pas acceptable.

04 Ne pas déléguer son jugement

L'IA peut conseiller, mais la décision finale doit rester humaine, surtout dans les domaines critiques. Un diagnostic médical, un recrutement, une décision de justice : l'humain doit garder le dernier mot et assumer la responsabilité.


88

Quel monde voulons-nous avec l'IA ?

**Une IA humano-centrée**  
 L'objectif n'est pas de créer des machines qui nous remplacent, mais des outils qui augmentent nos capacités sans diminuer notre humanité. Une IA qui libère du temps pour ce qui compte vraiment : la créativité, l'empathie, les relations humaines, la contemplation.

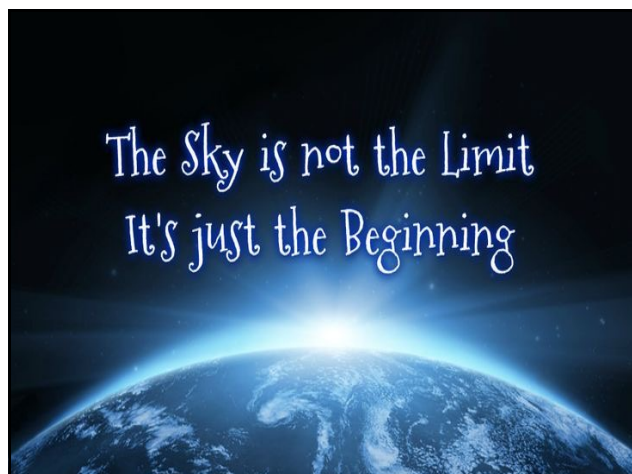
Une IA qui respecte notre dignité, notre autonomie, notre diversité. Une IA transparente, auditable, qui rend des comptes. Une IA au service du bien commun, pas seulement du profit de quelques-uns.

Ce monde est possible, mais il ne viendra pas tout seul. Il nécessite des choix conscients, un engagement collectif, et une vigilance de chaque instant. L'avenir de l'IA se décide maintenant, et nous avons tous une voix dans ce débat.



Empathie      Équité  
 Inclusion      Transparence  
 Innovation      Protection

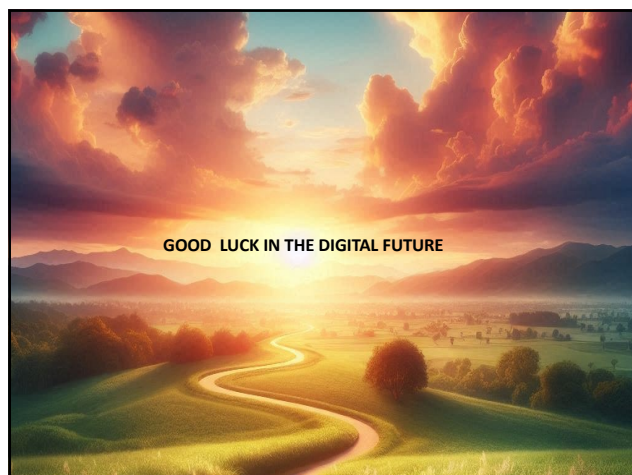
89



90



91



92



Merci

# Place à vos questions

---

**Contact**

N'hésitez pas à me contacter pour poursuivre la conversation ou accéder aux ressources présentées aujourd'hui.

Slides : [www.folon.com](http://www.folon.com)

☒ **Pensée finale**

*"L'intelligence artificielle a besoin... d'intelligence humaine."*

La technologie ne définit pas notre destin. Ce sont nos valeurs, nos choix et notre engagement collectif qui façonneront l'avenir de l'IA.