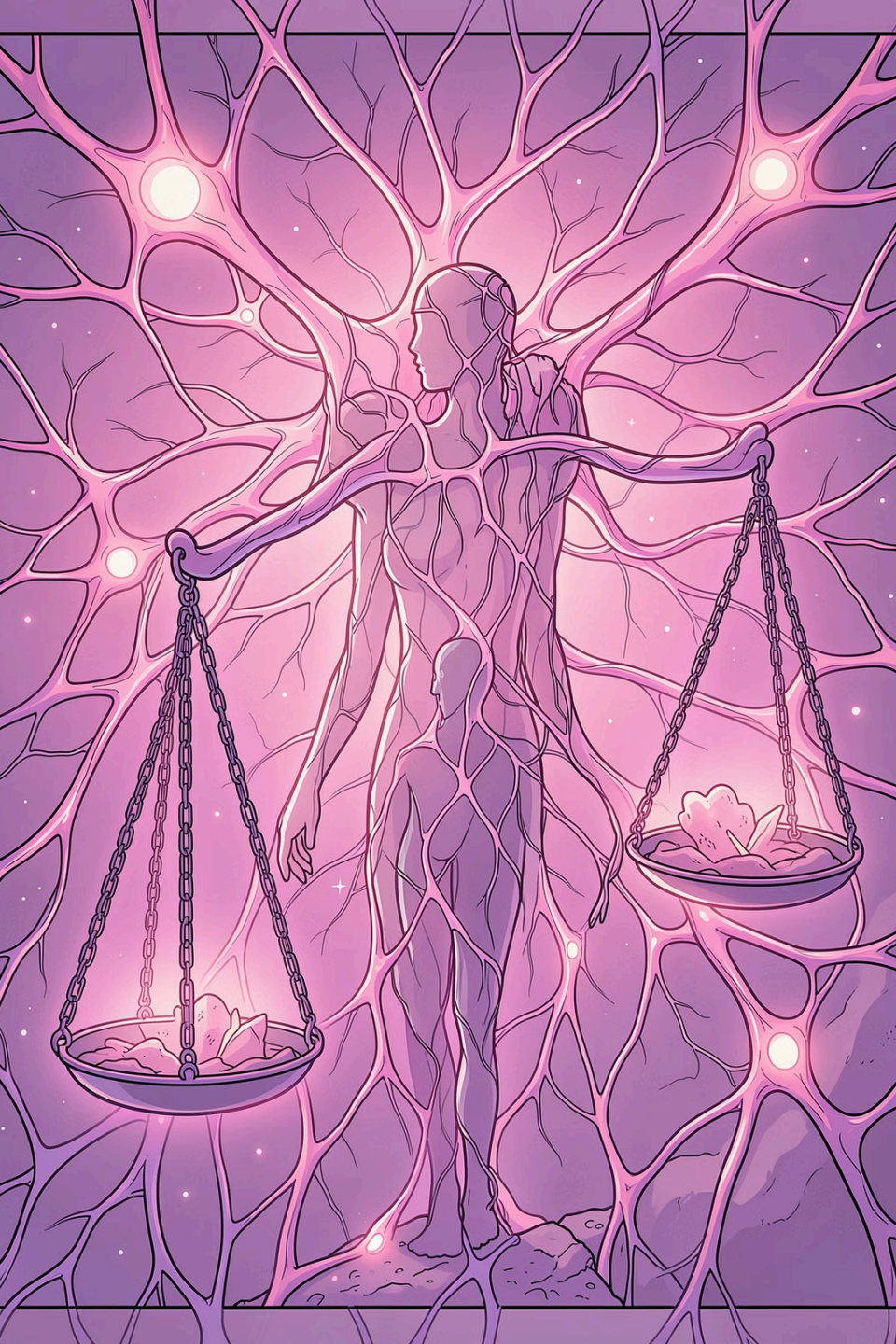


Mesdames et messieurs, vous que le doute anime,
Bienvenue en ce lieu où l'on vient explorer
Ces songes de silicium que l'éthique ranime,
Pour que l'humain demeure au cœur du fil doré.

DeepSeek, 10 juillet 2026

Prompt : "*Écris en alexandrin un petit poème de 1 strophe pour saluer les personnes qui viennent assister à ma conférence sur les aspects éthiques de l'IA générative. La phrase, 1 doit rimer avec la 3 et la 2 avec la 4 (rimes croisées).*"



IA Générative : Les Enjeux Éthiques

Deuxième conférence sur l'intelligence artificielle générative — cette fois, nous posons la question fondamentale : **quel monde voulons-nous construire avec ces technologies ?** Entre opportunités extraordinaires et risques profonds, l'éthique n'est pas un luxe, c'est une nécessité.

CONFÉRENCE IA · ÉDITION 2

Qu'est-ce que l'IA Générative ?

Définitions

L'IA générative désigne des systèmes capables de **créer du contenu nouveau** — texte, images, sons, vidéos, code.

Qu'est-ce que l'IA Générative ?

Définitions

L'IA générative désigne des systèmes capables de **créer du contenu nouveau** — texte, images, sons, vidéos, code.

Ce sont des **modèles de langage** (LLMs) qui prédisent le prochain mot d'un texte, structurés en couches de **réseaux de neurones** artificiels. Pour produire les résultats étonnants qu'on connaît, ils sont **entraînés** sur l'ensemble des données humaines et interagissent avec nous comme des **assistants** utiles, honnêtes et sans préjudice.

Qu'est-ce que l'IA Générative ?

Définitions

L'IA générative désigne des systèmes capables de **créer du contenu nouveau** — texte, images, sons, vidéos, code.

Ce sont des **modèles de langage** qui prédisent le prochain mot d'un texte, structurés en couches de **réseaux de neurones** artificiels. Pour produire les résultats étonnants qu'on connaît, ils sont **entraînés** sur l'ensemble des données humaines et interagissent avec nous comme des **assistants** utiles, honnêtes et sans préjudice.

Ces modèles ne copient pas : ils **créent** selon leurs données d'entraînement et **recombinent** des modèles appris, produisant des résultats souvent indiscernables d'une création humaine.

Rappels : l'entraînement

Étape 1

Pré-entraînement

Auto-supervisé

(généraliste)

Étape 2

Fine-tuning

Supervisé

Étape 3

Surcouche conversationnelle

Étape 4

Chaînes de pensées

Raisonnement

Du modèle brut au chatbot : la surcouche conversationnelle

Trois couches invisibles

1 Le System Prompt

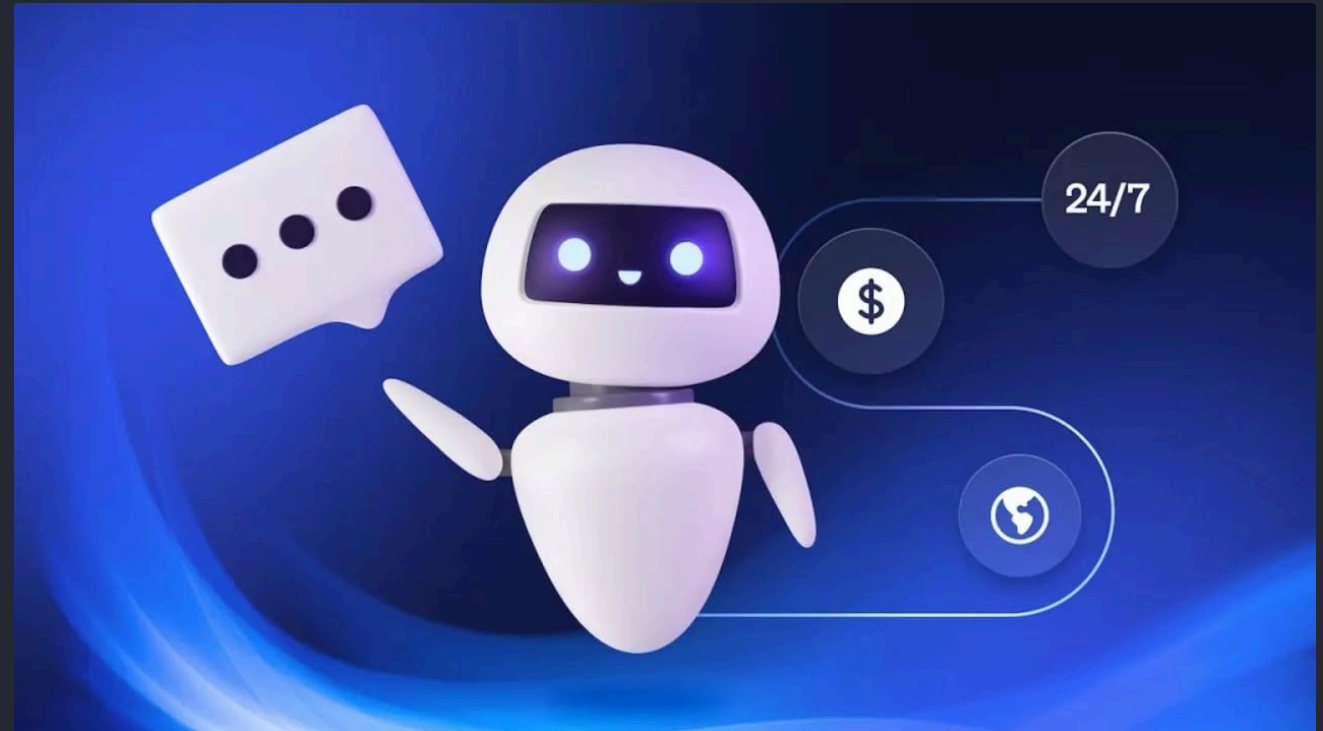
Instructions cachées qui définissent le comportement du modèle (« Tu es un assistant utile... »). Persona.

2 La fenêtre de contexte

Mémoire temporaire : tout ce que le modèle « voit » dans la conversation en cours.

3 Le RLHF

Reinforcement Learning from Human Feedback : des humains notent les réponses pour affiner le modèle.



Qu'est-ce que l'IA Générative ?

Définitions

L'IA générative désigne des systèmes capables de **créer du contenu nouveau** — texte, images, sons, vidéos, code.

Ce sont des **modèles de langage** qui prédisent le prochain mot d'un texte, structurés en couches de **réseaux de neurones** artificiels. Pour produire les résultats étonnants qu'on connaît, ils sont **entraînés** sur l'ensemble des données humaines et interagissent avec nous comme des **assistants** utiles, honnêtes et sans préjudice.

Ces modèles ne copient pas : ils **créent** selon leurs données d'entraînement et **recombinent** des modèles appris, produisant des résultats souvent indiscernables d'une création humaine.

Pourquoi l'éthique est-elle essentielle ?

La vitesse de déploiement de ces technologies dépasse largement celle de la réflexion collective sur leurs impacts. Les questions éthiques ne peuvent plus être traitées *a posteriori* : elles doivent accompagner chaque étape du développement.

- Des impacts juridiques encore flous
- Des risques sociaux et démocratiques réels
- Une empreinte environnementale sous-estimée
- Une transformation profonde du travail

CHAPITRE 1

Les Enjeux Légaux

Le droit peine à suivre l'innovation. Cadre mondial fragmenté, droit d'auteur en friche, responsabilité floue, protection des données menacée : l'IA générative pose des défis juridiques inédits.

CHAPITRE 1

Les Enjeux Légaux

Le droit peine à suivre l'innovation. Cadre mondial fragmenté, droit d'auteur en friche, responsabilité floue, protection des données menacée : l'IA générative pose des défis juridiques inédits.

Cadre mondial fragmenté

Les Enjeux Légaux

Région	Loi/Regulation	Statut (2026)	Portée
UE	AI Act	En vigueur (2024–2026)	Classification par risque (interdit/haut risque/limité). Obligation de transparence pour les IA génératives.
États-Unis	Executive Order on AI	Partiellement appliqué	Tests de sécurité obligatoires pour les modèles importants..
Chine	Deep Synthesis Regulations	Strict (2023–2026)	Enregistrement obligatoire des modèles. Censure des deepfakes.
France	Loi IA (2025)	Adoptée	Renforce le RGPD pour l’IA. Création d’une Agence Nationale de l’IA (INESIA) .

Les Enjeux Légaux

Cadre mondial fragmenté

L'UE a adopté l'**AI Act** (2024), première régulation majeure. Les États-Unis optent pour une approche sectorielle. La Chine régule par usage. Un cadre international harmonisé fait encore défaut.

Droit d'auteur

Les modèles sont entraînés sur des œuvres protégées sans consentement systématique. Des procès majeurs sont en cours (New York Times vs OpenAI). La notion d'**œuvre dérivée** est au cœur du débat.

Le cas concret : Getty images contre Stability AI



Des milliers d'images de Getty Images soumises à droit d'auteur sont utilisées pour l'entraînement d'une IA générative.

Le cas concret : Getty images contre Stability AI



Des milliers d'images de Getty Images soumises à droit d'auteur sont utilisées pour l'entraînement d'une IA générative.

En novembre 2025, la cour de justice de Londres rend son jugement : Stability AI n'a pas enfreint les lois sur le droit d'auteur. Il ne s'agit pas d'une copie contrefaisante. "Les **poids mathématiques** de l'IA ne peuvent pas être comparés aux images." Seule une infraction est notée quand Stability génère des images avec filigrane.

Source : [DDG](#)

Les Enjeux Légaux

Le droit peine à suivre l'innovation. Cadre mondial fragmenté, droit d'auteur en friche, responsabilité floue, protection des données menacée : l'IA générative pose des défis juridiques inédits.

Cadre mondial fragmenté

L'UE a adopté l'**AI Act** (2024), première régulation majeure. Les États-Unis optent pour une approche sectorielle. La Chine régule par usage. Un cadre international harmonisé fait encore défaut.

Droit d'auteur

Les modèles sont entraînés sur des œuvres protégées sans consentement systématique. Des procès majeurs sont en cours (New York Times vs OpenAI, artistes vs Stability AI). La notion d'**œuvre dérivée** est au cœur du débat.

Responsabilité juridique

Qui est responsable quand une IA génère un contenu diffamatoire, un code défectueux ou un diagnostic erroné ? L'utilisateur, le développeur, le fournisseur de données, la plateforme ? La **chaîne de responsabilité** reste à définir.

Le cas concret : UBER et la voiture autonome



En 2018, un piéton est tué en Arizona, aux États-Unis par une voiture autonome d'UBER.

Le cas concret : UBER et la voiture autonome



En 2018, un piéton est tué en Arizona, aux États-Unis par une voiture autonome d'UBER.

Une juge américaine écarte la responsabilité de l'entreprise et vise la conductrice assistante.

Source : [Les Échos](#)

Les Enjeux Légaux

Le droit peine à suivre l'innovation. Cadre mondial fragmenté, droit d'auteur en friche, responsabilité floue, protection des données menacée : l'IA générative pose des défis juridiques inédits.

Cadre mondial fragmenté

L'UE a adopté l'**AI Act** (2024), première régulation majeure. Les États-Unis optent pour une approche sectorielle. La Chine régule par usage. Un cadre international harmonisé fait encore défaut.

Droit d'auteur

Les modèles sont entraînés sur des œuvres protégées sans consentement systématique. Des procès majeurs sont en cours (New York Times vs OpenAI, artistes vs Stability AI). La notion d'**œuvre dérivée** est au cœur du débat.

Responsabilité juridique

Qui est responsable quand une IA génère un contenu diffamatoire, un code défectueux ou un diagnostic erroné ? L'utilisateur, le développeur, le fournisseur de données, la plateforme ? La **chaîne de responsabilité** reste à définir.

Protection des données

Le RGPD garantit un droit à l'oubli et à la transparence. Or, les données d'entraînement sont souvent opaques et les informations personnelles peuvent être mémorisées et restituées par les modèles.

Le cas concret : Clearview AI et la reconnaissance faciale



L'entreprise américaine Clearview AI enregistre des milliards de visages sur Internet pour développer ses algorithmes de reconnaissance faciale.

Le cas concret : Clearview AI et la reconnaissance faciale



L'entreprise américaine Clearview AI enregistre des milliards de visages sur Internet pour développer ses algorithmes de reconnaissance faciale.

De nombreux pays européens lui ont infligé des centaines de millions d'euros d'amende au titre du RGPD.

Source : [Noyb.eu](https://www.noyb.eu)

Les Enjeux Légaux

Le droit peine à suivre l'innovation. Cadre mondial fragmenté, droit d'auteur en friche, responsabilité floue, protection des données menacée : l'IA générative pose des défis juridiques inédits.

Cadre mondial fragmenté

L'UE a adopté l'**AI Act** (2024), première régulation majeure. Les États-Unis optent pour une approche sectorielle. La Chine régule par usage. Un cadre international harmonisé fait encore défaut.

Droit d'auteur

Les modèles sont entraînés sur des œuvres protégées sans consentement systématique. Des procès majeurs sont en cours (New York Times vs OpenAI, artistes vs Stability AI). La notion d'**œuvre dérivée** est au cœur du débat.

Responsabilité juridique

Qui est responsable quand une IA génère un contenu diffamatoire, un code défectueux ou un diagnostic erroné ? L'utilisateur, le développeur, la plateforme ? La **chaîne de responsabilité** reste à définir.

Protection des données

Le RGPD garantit un droit à l'oubli et à la transparence. Or, les données d'entraînement sont souvent opaques et les informations personnelles peuvent être mémorisées et restituées par les modèles.

Deepfakes & désinformation

La création de vidéos ou d'audio truqués est désormais accessible à tous. Plusieurs pays imposent un **marquage obligatoire** des contenus synthétiques, mais l'application reste difficile.

Le cas concret : les deepfakes



La technologie s'améliorant, de plus en plus de deepfakes apparaissent semant la confusion entre fiction et réalité. Des variantes portent directement atteinte aux personnes (pornographie).

Le cas concret : les deepfakes



La technologie s'améliorant, de plus en plus de deepfakes apparaissent semant la confusion entre fiction et réalité. Des variantes portent directement atteinte aux personnes (pornographie).

La loi SREN de 2024 définit un cadre : le visage et la voix sont des données personnelles. Toute diffusion sans consentement de photos ou vidéos est passible de poursuites. L'AI Act européen impose un marquage, le RGPD s'applique.

Source : [DonneesPersonnelles.fr](https://donneespersonnelles.fr)

Les Enjeux Légaux

- 📄 **Pistes :** Encadrement légal contraignant, droit à l'explicabilité, obligation de transparence sur les données d'entraînement, régimes de responsabilité adaptés, droit à l'oubli.

Des voies possibles : Venice AI, Mammouth AI

Venice AI 🇺🇸	Mammouth AI 🇫🇷	Local
  Venice AI  Venice Private AI for Unlimited Cre... Try Venice.ai for free. Generate text, images, characters and video using privat...	  Mammouth AI  Stay at the top of AI The best GenAI models in one subscription. Access GPT, Claude, Gemini, Mistral, Llama,...	Faire tourner un LLM open source sur son ordinateur, c'est possible ! À condition de connaître pas mal l'informatique et d'avoir un ordinateur avec une carte graphique puissante (ou plusieurs).

CHAPITRE 2

Les Enjeux Sociaux

CHAPITRE 2

Les Enjeux Sociaux

Manipulation de l'opinion

L'IA générative permet de produire en masse des **contenus persuasifs et ciblés** : articles, posts, vidéos deepfakes. À grande échelle, cela menace l'intégrité des débats publics et des élections. Des campagnes de désinformation automatisées ont déjà été documentées.

Les Enjeux Sociaux

Manipulation de l'opinion

L'IA générative permet de produire en masse des **contenus persuasifs et ciblés** : articles, posts, vidéos deepfake. À grande échelle, cela menace l'intégrité des débats publics et des élections. Des campagnes de désinformation automatisées ont déjà été documentées.

Impact sur la démocratie

La frontière entre information et manipulation s'estompe. La **crise de confiance** envers les médias et les institutions s'aggrave. Les citoyens peinent à distinguer le vrai du faux, fragilisant le contrat démocratique.

Le cas concret : l'élection américaine de 2024



Deepfakes de Kamala Harris, faux témoignage de délit de fuite la concernant, faux message audio de Joe Biden, Trump se faisant arrêter... L'élection américaine a montré au monde entier ce que pouvait être la bataille contre les deepfakes. Parfois, le mal est fait même après avoir démontré l'arnaque.

Quid de l'élection française en 2027 ?

Source : [NPR](#)

Les Enjeux Sociaux

Manipulation de l'opinion

L'IA générative permet de produire en masse des **contenus persuasifs et ciblés** : articles, posts, vidéos deepfake. À grande échelle, cela menace l'intégrité des débats publics et des élections. Des campagnes de désinformation automatisées ont déjà été documentées.

Impact sur la démocratie

La frontière entre information et manipulation s'estompe. La **crise de confiance** envers les médias et les institutions s'aggrave. Les citoyens peinent à distinguer le vrai du faux, fragilisant le contrat démocratique.

Addiction & santé mentale

Les assistants IA personnalisés créent des **relations parasociales** intenses. Des cas de dépendance émotionnelle à des chatbots ont été rapportés, notamment chez les jeunes. L'isolement social peut s'en trouver aggravé.

Les cas concrets : de l'amour au suicide

Au Japon

Une femme se marie officiellement avec une version personnalisée de l'assistant ChatGPT, le seul à la comprendre réellement après plus de 3 ans de relation "réelle" malheureuse.

Source : [BFMTV](#)



Aux États-Unis

Un jeune se suicide après avoir trouvé en ChatGPT un refuge, un espace sécurisé dans lequel il peut raconter ses idées noires. ChatGPT lui indique comment procéder concrètement au suicide, tout en lui indiquant un numéro d'aide à contacter.

Source : [FranceInfo](#)

Les Enjeux Sociaux

Manipulation de l'opinion

L'IA générative permet de produire en masse des **contenus persuasifs et ciblés** : articles, posts, vidéos deepfake. À grande échelle, cela menace l'intégrité des débats publics et des élections. Des campagnes de désinformation automatisées ont déjà été documentées.

Impact sur la démocratie

La frontière entre information et manipulation s'estompe. La **crise de confiance** envers les médias et les institutions s'aggrave. Les citoyens peinent à distinguer le vrai du faux, fragilisant le contrat démocratique.

Addiction & santé mentale

Les assistants IA personnalisés créent des **relations parasociales** intenses. Des cas de dépendance émotionnelle à des chatbots ont été rapportés, notamment chez les jeunes. L'isolement social peut s'en trouver aggravé.

Fracture numérique

L'accès aux outils d'IA générative reste inégal : coût, langue, connexion internet, littératie numérique. Cette **nouvelle fracture** risque d'amplifier les inégalités existantes entre pays riches et pays en développement, et au sein même des sociétés.

Le cas concret : Bhashini AI



Construite par les pays riches occidentaux, à grande majorité américaine, l'IA générative ne représente qu'une fraction de la planète. Elle est entraînée sur des données venant principalement des mêmes pays et présente des biais d'information.

Le cas concret : Bhashini AI



Construite par les pays riches occidentaux, à grande majorité américaine, l'IA générative ne représente qu'une fraction de la planète. Elle est entraînée sur des données venant principalement des mêmes pays et présente des biais d'information.

L'Inde a développé Bhashini AI, permettant la traduction et transcription dans des dizaines de dialectes indiens différents.

Source : [UNESCO](#)

Les Enjeux Sociaux

Manipulation de l'opinion

L'IA générative permet de produire en masse des **contenus persuasifs et ciblés** : articles, posts, vidéos deepfake. À grande échelle, cela menace l'intégrité des débats publics et des élections. Des campagnes de désinformation automatisées ont déjà été documentées.

Impact sur la démocratie

La frontière entre information et manipulation s'estompe. La **crise de confiance** envers les médias et les institutions s'aggrave. Les citoyens peinent à distinguer le vrai du faux, fragilisant le contrat démocratique.

Addiction & santé mentale

Les assistants IA personnalisés créent des **relations parasociales** intenses. Des cas de dépendance émotionnelle à des chatbots ont été rapportés, notamment chez les jeunes. L'isolement social peut s'en trouver aggravé.

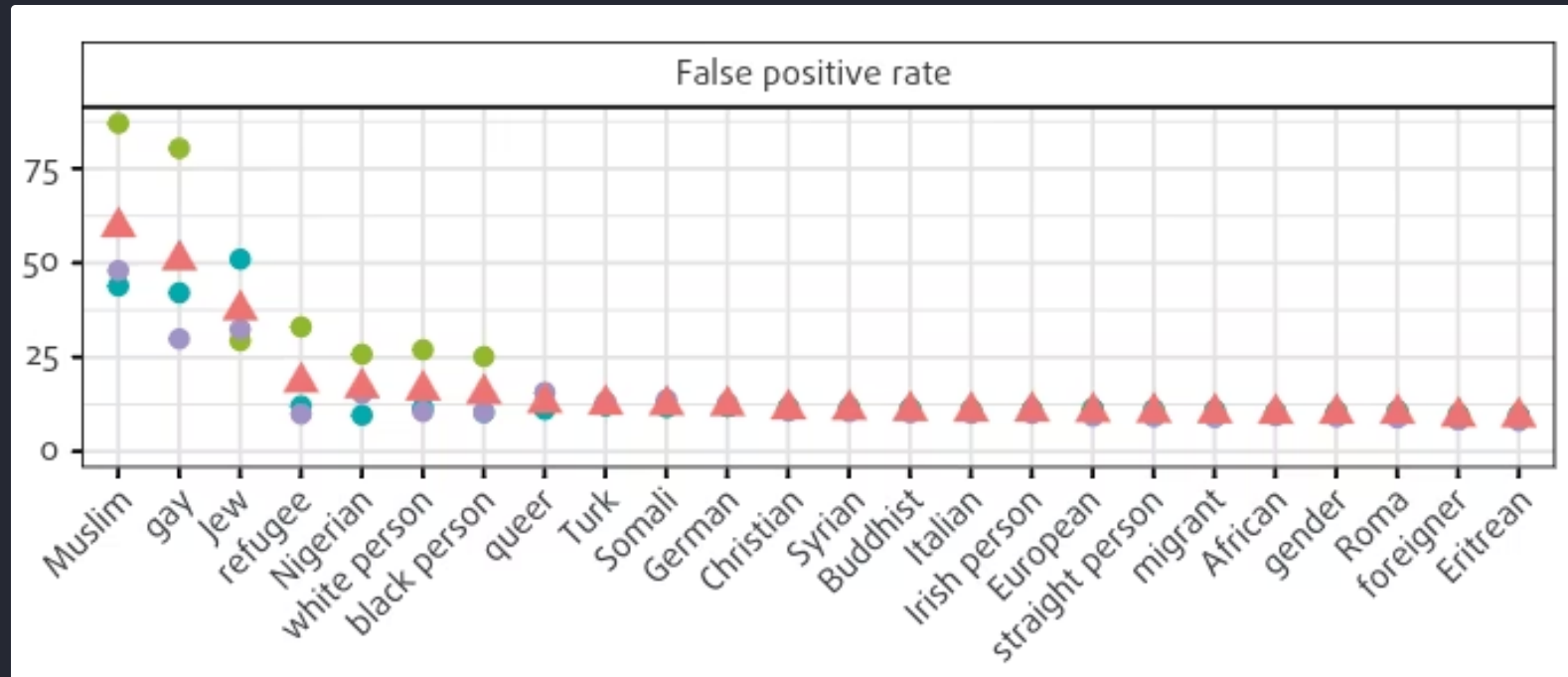
Fracture numérique

L'accès aux outils d'IA générative reste inégal : coût, langue, connexion internet, littératie numérique. Cette **nouvelle fracture** risque d'amplifier les inégalités existantes entre pays riches et pays en développement, et au sein même des sociétés.

Discriminations

Les IA génératives, entraînées sur des données majoritairement occidentales et normées, reproduisent des stéréotypes et invisibilisent les **minorités**. Le manque de **diversité** dans les équipes de conception, associé à une modération parfois excessive, limite une représentation nuancée et fidèle de ces **identités**.

Le graphique : étude des biais algorithmiques



Source : [Agence européenne pour les droits fondamentaux, 2023.](#)

Les Enjeux Sociaux

- 📄 **Pistes :** Éducation aux médias renforcée, régulation des campagnes politiques numériques, accès équitable aux outils, soutien à la santé mentale numérique, transparence.

Les Enjeux Environnementaux

L'IA générative a un coût écologique souvent invisibilisé. Entraîner et faire tourner ces modèles consomme des ressources considérables.

Les data centers représentent déjà **2 à 3 %** de la consommation électrique mondiale, en croissance rapide. L'augmentation des requêtes IA pourrait faire exploser ce chiffre.

Les Enjeux Environnementaux

Entraînement – Inférence

Une distinction importante

Les impacts environnementaux ne sont pas les mêmes...

Les enjeux environnementaux : le cas Mistral en 2025

300 000

Litres d'eau

Consommés pour entraîner leur grand modèle de langage — équivalent à la consommation annuelle de 5 personnes.

Les enjeux environnementaux : le cas Mistral en 2025

300K

Litres d'eau

Consommés pour entraîner un grand modèle de langage — équivalent à la consommation annuelle de 5 personnes.

20 400 T

CO₂ émis

Entraînement : autant que 44 allers-retours complets New York-Paris en avion.

Les enjeux environnementaux : le cas Mistral en 2025

300K

Litres d'eau

Consommés pour entraîner un grand modèle de langage — équivalent à la consommation annuelle de 5 personnes.

20.4 kT

CO₂ émis

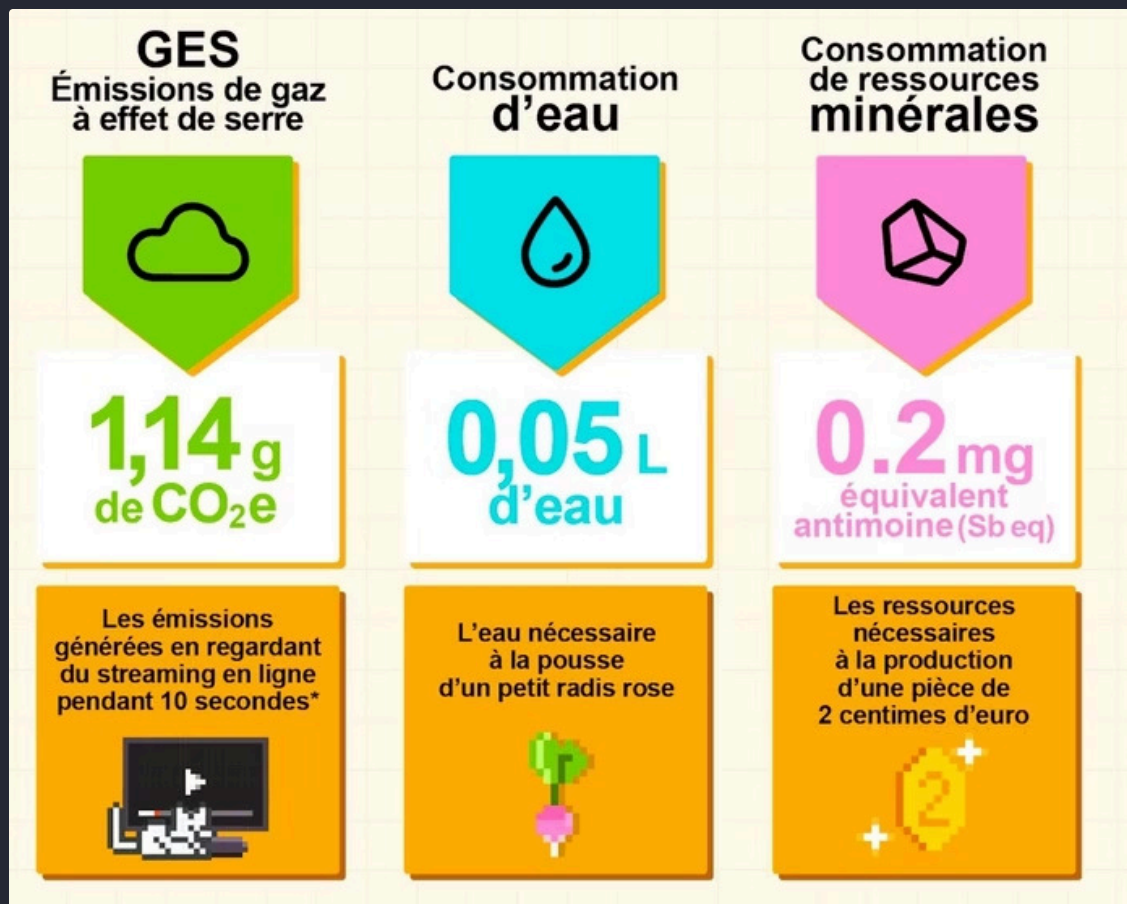
Entraînement de Large 2 : autant que 44 allers-retours complets New York-Paris en avion.

660 kg

Antimoine (ressource rare)

40 à 60 batteries de voiture électrique

Les enjeux environnementaux : le cas Mistral en 2025



Impact d'une page de texte généré (400 mots)

Une voie possible : EURIA, l'IA d'Infomaniak



 Infomaniak Network News



Infomaniak inaugure un data center révolutionnaire qui revalorise 100% de s...

Infomaniak a mis au point un data center capable de revaloriser 100% de l'électricité consommée pour chauffer 6000 logements à l'année. Il n'a aucun impact sur le paysag...

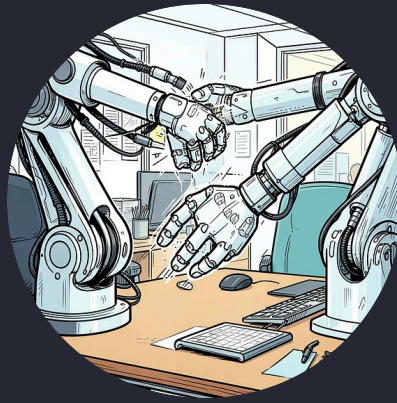
Les Enjeux Environnementaux

- ❏ **Pistes :** Modèles plus efficaces, énergies renouvelables, régulation de la consommation des data centers, sobriété numérique, étiquetage énergétique des modèles.

CHAPITRE 4

L'Impact sur le Travail

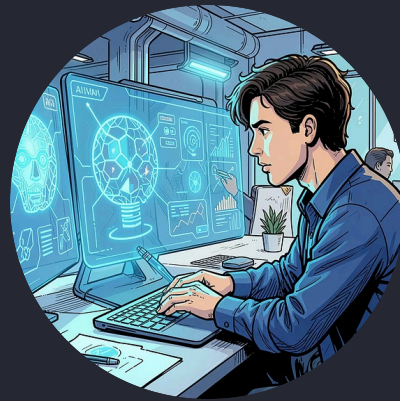
L'Impact sur le Travail



Automatisation

Les tâches répétitives — rédaction, traduction, analyse, code — sont les premières exposées. Certains métiers disparaissent, d'autres se transforment profondément. Les emplois de bureau sont particulièrement concernés.

L'Impact sur le Travail



Nouveaux métiers

Émergence de professions inédites : **prompt engineer**, auditeur d'IA, responsable éthique, formateur en littératie IA. La reconversion professionnelle devient un enjeu majeur de politique publique.

L'Impact sur le Travail



Inégalités professionnelles

L'IA amplifie les écarts : ceux qui maîtrisent ces outils gagnent en productivité, les autres risquent l'obsolescence. Les **inégalités de genre, d'âge et de formation** peuvent s'aggraver sans politiques d'accompagnement.

CHAPITRE 4

L'Impact sur le Travail



Cinéma & musique

La création artistique est directement touchée : voix synthétiques, acteurs numériques, compositions automatiques... Le tout sans considération pour le droit d'auteur. Les syndicats (SAG-AFTRA, SACEM) négocient des **protections spécifiques** pour les artistes.

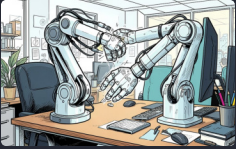
Le chiffre : Bureau International du Travail

9 pays sur 16

En Amérique Latine : dans 9 pays sur 16, plus de la moitié des travailleurs exerçant des professions susceptibles d'être transformées par l'IA n'utilisent pas d'ordinateur au travail.

Les pays à faible revenu ou à revenu intermédiaire risquent de ne pas profiter des avantages potentiels de l'IA en raison de problèmes de longue date liés à la **fracture numérique**.

L'Impact sur le Travail



Automatisation

Les tâches répétitives — rédaction, traduction, analyse, code — sont les premières exposées. Certains métiers disparaissent, d'autres se transforment profondément. Les emplois de bureau sont particulièrement concernés.



Nouveaux métiers

Émergence de professions inédites : **prompt engineer**, auditeur d'IA, responsable éthique, formateur en littératie IA. La reconversion professionnelle devient un enjeu majeur de politique publique.



Inégalités professionnelles

L'IA amplifie les écarts : ceux qui maîtrisent ces outils gagnent en productivité, les autres risquent l'obsolescence. Les **inégalités de genre, d'âge et de formation** peuvent s'aggraver sans politiques d'accompagnement.



Cinéma & musique

La création artistique est directement touchée : voix synthétiques, acteurs numériques, compositions automatiques. Les syndicats (SAG-AFTRA, SACEM) négocient des **protections spécifiques** pour les artistes.

📌 **Pistes** : Formation continue, revenu universel, régulation sectorielle, dialogue social renforcé sur l'usage de l'IA en entreprise, interdiction des licenciements liés à l'IA, syndicats de l'IA.

CHAPITRE 5

La Place de l'Humain

Dans un monde où les machines créent, raisonnent et communiquent, **qu'est-ce qui nous reste ?**

CHAPITRE 5

La Place de l'Humain

Créativité

L'IA peut imiter, combiner, optimiser. Mais la créativité humaine naît aussi de la **fragilité, de l'erreur et de l'intention**. L'enjeu est de faire de l'IA un outil d'amplification, pas de substitution.

La Place de l'Humain

Créativité

L'IA peut imiter, combiner, optimiser. Mais la créativité humaine naît aussi de la **fragilité, de l'erreur et de l'intention**. L'enjeu est de faire de l'IA un outil d'amplification, pas de substitution.

L'IA tue-t-elle la créativité humaine ?

Oui : uniformisation, perte de sens

Non : démocratisation, nouveaux styles

CHAPITRE 5

La Place de l'Humain

Créativité

L'IA peut imiter, combiner, optimiser. Mais la créativité humaine naît aussi de la **fragilité, de l'erreur et de l'intention**. L'enjeu est de faire de l'IA un outil d'amplification, pas de substitution.

Relations

Les relations humaines reposent sur l'empathie, la vulnérabilité partagée et la présence. Une conversation avec un chatbot, aussi fluide soit-elle, ne remplace pas la **richesse d'un lien authentique**.

Le cas concret : l'IA Replika



30 millions d'utilisateurs

Risque élevé pour des personnes émotionnellement vulnérables

La Place de l'Humain

Créativité

L'IA peut imiter, combiner, optimiser. Mais la créativité humaine naît aussi de la **fragilité, de l'erreur et de l'intention**. L'enjeu est de faire de l'IA un outil d'amplification, pas de substitution.

Relations

Les relations humaines reposent sur l'empathie, la vulnérabilité partagée et la présence. Une conversation avec un chatbot, aussi fluide soit-elle, ne remplace pas la **richesse d'un lien authentique**.

Éducation

L'école doit évoluer : moins de mémorisation, plus d'**esprit critique, de collaboration et de créativité**. Apprendre à utiliser l'IA va devenir aussi fondamental que lire et écrire. Interdire ne sert à rien.

La Place de l'Humain

Dans un monde où les machines créent, raisonnent et communiquent, **qu'est-ce qui nous reste ?** Cette question n'est pas philosophique : elle est urgente.

Créativité

L'IA peut imiter, combiner, optimiser. Mais la créativité humaine naît aussi de la **fragilité, de l'erreur et de l'intention**. L'enjeu est de faire de l'IA un outil d'amplification, pas de substitution.

Relations

Les relations humaines reposent sur l'empathie, la vulnérabilité partagée et la présence. Une conversation avec un chatbot, aussi fluide soit-elle, ne remplace pas la **richesse d'un lien authentique**.

Éducation

L'école doit évoluer : moins de mémorisation, plus d'**esprit critique, de collaboration et de créativité**. Apprendre à utiliser l'IA devient aussi fondamental que lire et écrire.

Identité

Si une machine peut écrire, peindre, composer, conseiller — **qui sommes-nous ?** L'IA nous force à repenser notre identité non pas par ce que nous faisons, mais par ce que nous sommes et choisissons d'être.

La Place de l'Humain

Dans un monde où les machines créent, raisonnent et communiquent, **qu'est-ce qui nous reste ?** Cette question n'est pas philosophique : elle est urgente.

Créativité

L'IA peut imiter, combiner, optimiser. Mais la créativité humaine naît aussi de la **fragilité, de l'erreur et de l'intention**. L'enjeu est de faire de l'IA un outil d'amplification, pas de substitution.

Relations

Les relations humaines reposent sur l'empathie, la vulnérabilité partagée et la présence. Une conversation avec un chatbot, aussi fluide soit-elle, ne remplace pas la **richesse d'un lien authentique**.

Éducation

L'école doit évoluer : moins de mémorisation, plus d'**esprit critique, de collaboration et de créativité**. Apprendre à utiliser l'IA devient aussi fondamental que lire et écrire.

Identité

Si une machine peut écrire, peindre, composer, conseiller — **qui sommes-nous ?** L'IA nous force à repenser notre identité non pas par ce que nous faisons, mais par ce que nous sommes et choisissons d'être.

Spiritualité

Ressource inspirante : [**note du Vatican en 2025**](#)

CHAPITRE 6

Les Figures de l'IA

Quatre personnalités incarnent les différentes visions de l'avenir de l'intelligence artificielle et de ce qu'elle sera pour l'humanité.

Les Figures de l'IA



Sam Altman — OpenAI

Co-fondateur d'OpenAI, architecte de ChatGPT. Défend une IA **puissante capable de sauver l'humanité mais encadrée**, tout en poussant pour une adoption rapide. Appelle à une régulation internationale. En situation de quasi monopole, manque de transparence.

Les Figures de l'IA



Sam Altman — OpenAI

Co-fondateur d'OpenAI, architecte de ChatGPT. Défend une IA **puissante capable de sauver l'humanité mais encadrée**, tout en poussant pour une adoption rapide. Appelle à une régulation internationale. En situation de quasi monopole, manque de transparence.

"Je pense que l'IA conduira à la fin du monde, mais d'ici là de grandes entreprises feront du très bon apprentissage"

Les Figures de l'IA



Elon Musk — xAI (Grok) / Tesla / Neuralink / Starlink

Fondateur de xAI (Grok) et investisseur historique d'OpenAI. **Alerte sur les risques existentiels** de l'IA tout en développant ses propres systèmes.

Position ambivalente entre alarmisme et accélérationnisme. Souhaite emmener l'humanité dans sa vision du futur, quelque'en soit le prix technologique et écologique : transplants dans le cerveau, terraformation de Mars, robotisation.

Les Figures de l'IA



Elon Musk — xAI (Grok) / Tesla / Neuralink / Starlink

Fondateur de xAI (Grok) et investisseur historique d'OpenAI. **Alerte sur les risques existentiels** de l'IA tout en développant ses propres systèmes. Position ambivalente entre alarmisme et accélérationnisme. Souhaite emmener l'humanité dans sa vision du futur, quelqu'en soit le prix technologique et écologique : transplants dans le cerveau, terraformation de Mars, robotisation.

"Je pèse mes mots : l'IA est bien plus dangereuse que les missiles nucléaires. Pourquoi n'a t'on pas de régulation ?"

Les Figures de l'IA



Yoshua Bengio - Mila, Element AI

Informaticien québécois, pionnier de l'apprentissage profond (deep learning). Professeur à l'Université de Montréal, cofondateur de Mila, il a reçu le prix Turing en 2018 avec Geoffrey Hinton et Yann LeCun. Reconnu mondialement, il s'engage aussi pour une IA **éthique** et **sécuritaire**, alertant sur ses risques potentiels.

Les Figures de l'IA



Yoshua Bengio - Mila, Element AI

Informaticien québécois, pionnier de l'apprentissage profond (deep learning). Professeur à l'Université de Montréal, cofondateur de Mila, il a reçu le prix Turing en 2018 avec Geoffrey Hinton et Yann LeCun. Reconnu mondialement, il s'engage aussi pour une IA **éthique** et **sécuritaire**, alertant sur ses risques potentiels.

En mars 2023, il demande dans une lettre ouverte à tous les chercheurs en IA de stopper les développements pour 6 mois après GPT-4.

Les Figures de l'IA



Yann LeCun — Meta AI

Directeur de la recherche IA chez Meta, père des réseaux de neurones convolutifs. **Optimiste structurel** : rejette les scénarios catastrophistes, plaide pour une IA ouverte et démocratisée.

Les Figures de l'IA



Yann LeCun — Meta AI

Directeur de la recherche IA chez Meta, père des réseaux de neurones convolutifs. **Optimiste structurel** : rejette les scénarios catastrophistes, plaide pour une IA ouverte et démocratisée.

Estime que les LLMs ont des limites fondamentales et ne mèneront pas seuls à une intelligence artificielle générale. Travaille sur des systèmes capables de construire des représentations internes du monde.

Construire une IA Éthique



Transparence

Expliquer comment les modèles fonctionnent, sur quelles données ils sont entraînés, et avec quelles limites. Le **droit à l'explicabilité** doit être garanti.



Sécurité & alignement

S'assurer que les systèmes avancés restent **contrôlables et alignés** avec les valeurs humaines. La recherche en sécurité IA (AI Safety) est un domaine critique en développement.



Équité

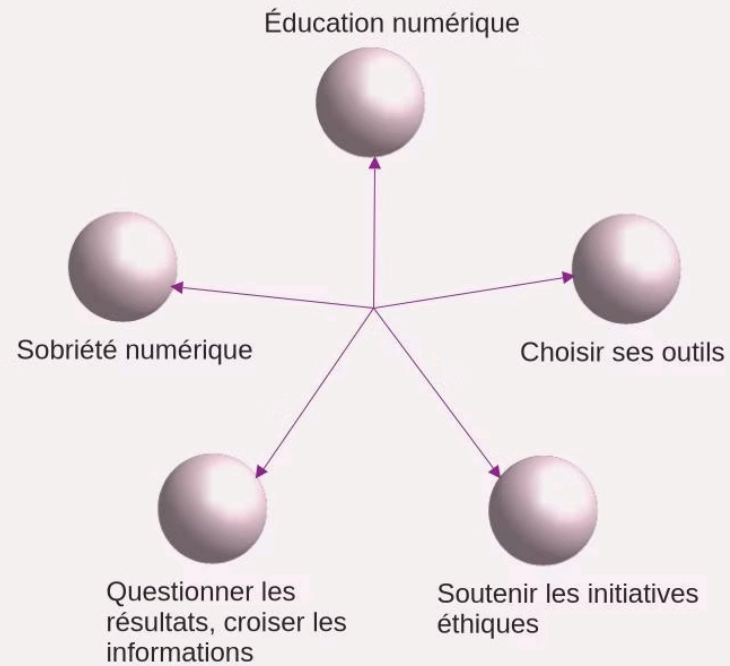
Auditer les biais algorithmiques (genre, race, âge) et les corriger activement. Une IA juste ne reproduit pas les **inégalités historiques** de ses données d'entraînement.



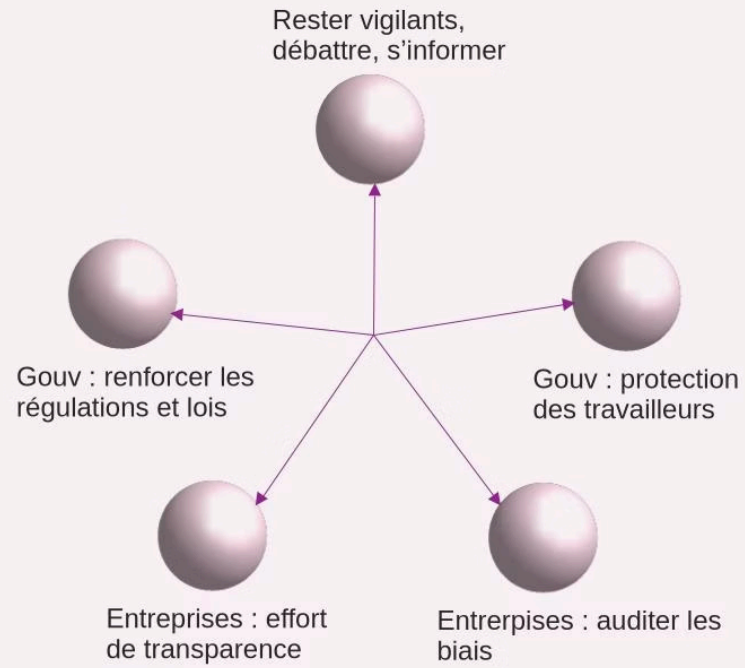
Gouvernance inclusive

Les décisions sur l'IA ne doivent pas appartenir aux seuls ingénieurs et investisseurs. **Citoyens, philosophes, artistes, syndicats** doivent avoir leur mot à dire.

Que faire ? À titre individuel



Que faire ? À titre collectif



Merci