



Intelligences artificielles *en sociétés*

2. Et si ça marche !

Politiques de l'émergence et de l'accélération

Au delà de la prédiction...

Multimodalité textes, sons et images

→ Interaction plus large et « modèles de mondes plus complexes »

Génération à enrichissement contextuel Retrieval-Augmented Generation - RAG

→ Adaptation à des documents précis

Recherche

→ Actualisation des données utilisées

Thinking Chain-of-Thought et apprentissage de raisonnements

→ Capacité de « raisonnement » exploratoire et structuré

Tools calcul, code...

→ Utilisation d'outils spécifiques lorsque c'est plus efficace (calculatrice, code Python,...)

Agents

→ Décomposition d'une tâche en sous-tâches spécialisées

Model Context Protocol MCP

→ Connexion à d'autres services (bases de données, fichiers, services en ligne...)

Distillation

→ Apprentissage d'un petit modèle à partir d'un grand modèle (données « synthétiques »)

Se rassurer avec des mains à six doigts ?



Se rassurer avec des mains à six doigts ?



Création personnelle - Gemini Nano Banana 2 - "Image of a goldfish inside a "Cola-Coca" bottle on a woden table" -

L'asymétrie de la critique

Confabulations

Connaissances

Rapidité

Biais d'apprentissage

Discrimination

Perte de contexte

Fiabilité

Imitation

Interprétabilité

...

→ **Test de turing inversé**

Des perroquets très intelligents ?

Complexité
Prédiction
Expérience
Alignement
Émergence
Accélération

→ Une analogie problématique

Politique

Ambivalence de la critique

Être discriminé par erreur ou intentionnellement ?

→ Il est temps d'avoir une politique de l'IA

Politiques des IA

Epuisement des sources et des ressources
Fragilisation de la sphère publique
Fragilisation des modalités de l'apprentissage - *Deskilling*
Captation privée de la recherche et de l'innovation
Fragilisation de l'environnement de travail
Concentration du pouvoir
Concentration des gains de productivité
Automatisation des discriminations
Délégation cognitive incontrôlée
Fragilisation de la sécurité informatique
Renforcement des capacités militaires et sécuritaires

Transition écologique accélérée et biens communs
Amélioration de la veille et du fact-checking
Individualisation de l'apprentissage - *Upskilling*
Accélération de la découverte scientifique
Redistribution du temps de travail
Renforcement du contrôle démocratique
Identification et correction des discriminations
Redistribution des gains de productivité
Réflexivité et explicitation augmentées
Renforcement de la sécurité informatique
Construction d'une souveraineté numérique et défensive

→ **Sortir les lumières de l'obscurité**

Politiques des IA - Des lumières sombres

Épuisement des sources et des ressources

→ Extraction de la valeur des sources (presse, Wikipédia et autres plateformes collaboratives) menaçant leur renouvellement, conjuguée à une croissance rapide de la consommation d'énergie, d'eau et de matières premières.

Fragilisation de la sphère publique

→ Production de contenus synthétiques (deepfakes, astroturfing) à coût marginal quasi nul, érosion des signaux de crédibilité et ciblage individuel de plus en plus contextuel.

Fragilisation des modalités de l'apprentissage – Deskilling

→ Transformation des modalités, de la motivation et de l'organisation de la formation initiale et continue ; risque d'atrophie des compétences cognitives par délégation systématique aux systèmes d'IA.

Captation privée de la recherche et de l'innovation

→ Concentration des moyens de calcul, des données et des talents dans un petit nombre d'acteurs privés, marginalisant la recherche publique et orientant les agendas scientifiques vers des intérêts commerciaux.

Fragilisation de l'environnement de travail

→ Reconsidération des compétences, hausse des attentes de productivité, surveillance accrue et précarisation des collectifs de travail.

Concentration du pouvoir

→ Dépendance croissante vis-à-vis d'un nombre très restreint d'acteurs (infrastructures, modèles, plateformes) et perte de souveraineté économique, politique et culturelle.

Concentration des gains de productivité

→ Captation des gains potentiellement massifs par les détenteurs du capital et des infrastructures, au détriment du travail et des finances publiques.

Délégation cognitive incontrôlée

→ Transfert croissant de décisions à des systèmes dont le fonctionnement n'est ni compris ni auditable, avec normalisation progressive de cette perte de contrôle.

Automatisation des discriminations

→ Reproduction et amplification à grande échelle de biais présents dans les données d'entraînement, opacifiant des décisions discriminatoires dans l'embauche, le crédit, la justice ou l'accès aux services publics.

Fragilisation de la sécurité informatique

→ Abaissement du seuil de compétence pour conduire des cyberattaques, automatisation de la découverte de vulnérabilités et accès facilité à des savoirs opérationnels dangereux.

Renforcement des capacités militaires et sécuritaires

→ Automatisation et réduction du contrôle humain dans le renseignement, la planification, la guerre informationnelle et l'armement.

Politiques des IA - Des lumières renouvelées

Transition écologique accélérée et biens communs

→ Découverte de nouveaux procédés de fabrication, de recyclage et de captation, optimisation des flux logistiques et énergétiques, modélisation plus fine des écosystèmes ; mécanismes de contribution obligatoire des opérateurs aux communs numériques (presse, Wikipédia, bases ouvertes) pour restaurer la boucle contributive.

Amélioration de la veille et du fact-checking

→ Détection à grande échelle de contenus douteux, recoupement automatisé de sources, traçabilité des affirmations et outillage des journalistes et chercheurs pour identifier rapidement manipulations et désinformation.

Individualisation de l'apprentissage – Upskilling

→ Personnalisation des parcours de formation, tutorat adaptatif à grande échelle, accès élargi à des ressources pédagogiques de qualité, y compris dans des langues et des contextes jusqu'ici mal couverts.

Accélération de la découverte scientifique

→ Exploration systématique d'espaces combinatoires jusqu'ici inaccessibles dans la conception de médicaments, de matériaux et de procédés énergétiques, réduisant le temps entre hypothèse et validation expérimentale.

Redistribution du temps de travail

→ Si les gains de productivité sont politiquement orientés vers la réduction du temps de travail plutôt que vers la seule compression des effectifs, possibilité d'un rééquilibrage historique entre travail, formation et vie personnelle.

Renforcement du contrôle démocratique

→ Outillage des citoyens, des journalistes et des institutions de contrôle pour analyser les flux financiers, les décisions publiques et les discours des gouvernants, réduisant l'asymétrie d'information entre gouvernants et gouvernés.

Identification et correction des discriminations

→ Détection systématique de biais dans les décisions administratives, judiciaires, financières ou d'embauche, rendant visibles des inégalités de traitement que l'analyse humaine peine à repérer à grande échelle.

Redistribution des gains de productivité

→ Si la fiscalité, le droit du travail et les régimes de propriété des modèles et des données sont adaptés, possibilité d'un partage plus large des gains entre travail, capital, finances publiques et usagers.

Réflexivité et explicitation augmentées

→ Aide à la formulation, à l'argumentation et à la mise à l'épreuve des raisonnements, permettant aux individus comme aux collectifs d'explicitier leurs présupposés, de tester des contre-arguments et d'améliorer la qualité du débat.

Renforcement de la sécurité informatique

→ Détection automatisée de vulnérabilités, durcissement des défenses, accélération de la réponse aux incidents et accessibilité des bonnes pratiques de cybersécurité auprès des organisations les moins dotées.

Construction d'une souveraineté numérique et défensive

→ Développement de modèles, d'infrastructures et de jeux de données sous contrôle public ou coopératif, réduisant la dépendance vis-à-vis d'un oligopole privé étranger et permettant une gouvernance alignée sur des choix démocratiques.

Politiques des IA - Cinq axes de tension

01

Nécessité - délibération

Qui détient le rythme du déploiement ?

L'inévitable

Le rythme est dicté par la course technologique, le marché, la géopolitique. La politique suit.

L'instituable

Tout déploiement est un choix collectif, discutable, suspendable, réorientable.

02

Extraction - contribution

Que devons-nous à ce dont nous vivons ?

Prélèvement sans dette

L'IA prélève sur les communs informationnels, l'attention, les écosystèmes, sans avoir à les reconstituer.

Obligation contributive

Tout système qui prélève sur un commun lui doit réparation : presse, savoirs, environnement, la recherche, la formation.

03

Dépossession - libération

Que faisons-nous du travail humain ?

Le travail capté

Les gains de productivité remontent vers le capital ; l'IA intensifie, surveille et déqualifie le travail restant. Le travailleur perd la maîtrise de son activité.

Le travail libéré

Les gains financent la réduction du temps de travail et la délégation des tâches pénibles ou répétitives. L'IA rend au travail vivant ce qui ne peut être délégué : le jugement, le sens, la relation.

04

Substitution - capacitation

Que reste-t-il de notre puissance d'agir ?

Le sujet déchargé

L'IA décide, crée, juge à notre place. La dépendance s'installe, l'atrophie s'étend.

Le sujet outillé

L'IA augmente la capacité à comprendre, argumenter, délibérer, contester.

05

Indistinction - véridiction

Reste-t-il une base commune pour s'accorder sur ce qui est ?

L'équivalence généralisée

Le synthétique sature le champ informationnel ; les signaux du vrai s'effondrent.

L'institution du vrai

L'IA outille les institutions de véridiction : presse, science, justice, archives.

→ *Ce que l'IA fera de nous dépend de ce que nous ferons d'elle.*