



Intelligence artificielle : les impacts sociétaux et environnementaux

Dominique Desbois, Paris-Saclay Applied Economics

INRAE/AgroParisTech

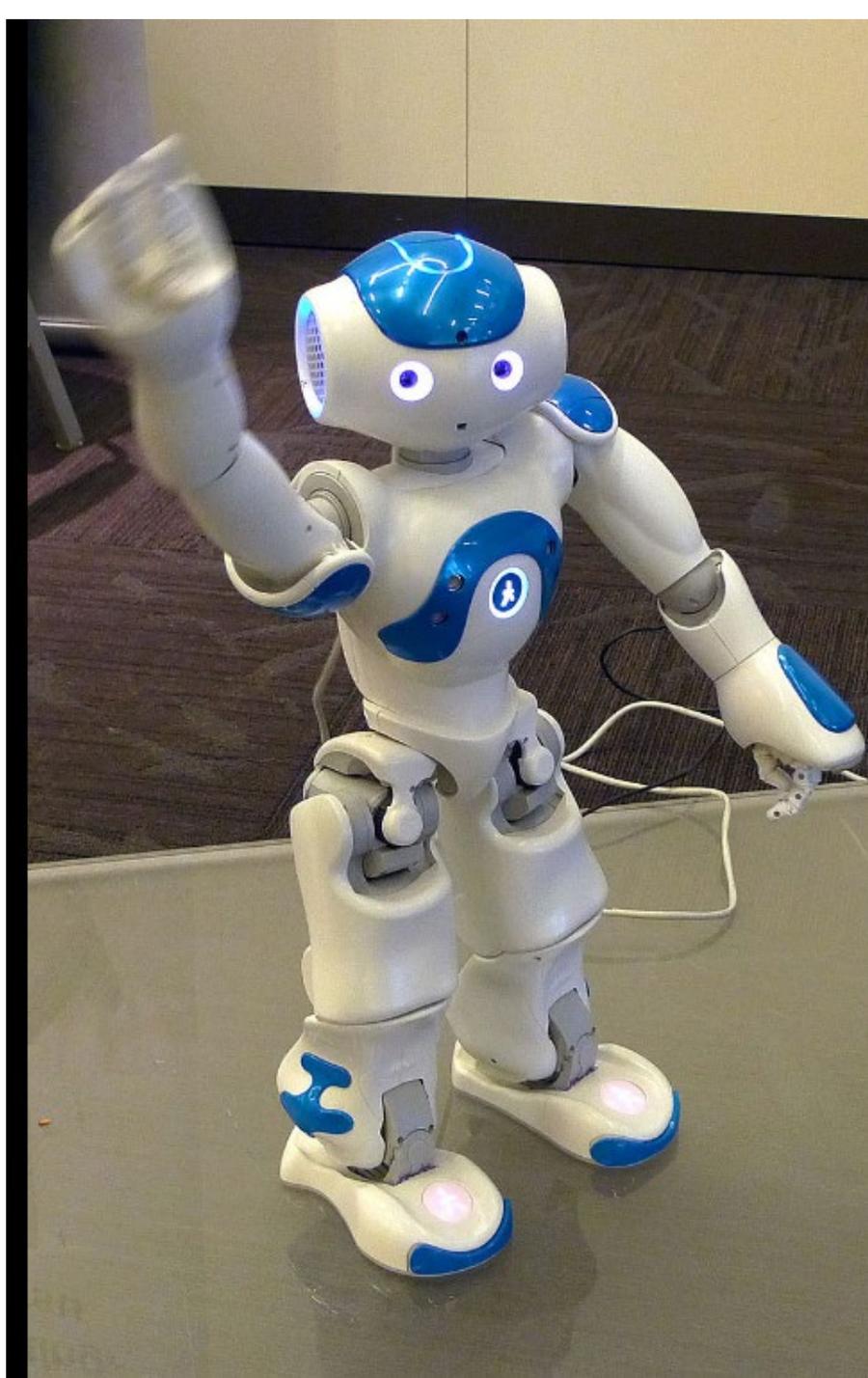
(dominique.desbois@agroparistech.fr)

FECODD : former au monde de demain

INSPÉ de Clermont Ferrand (France), 9-11/07/2024

IA faible : les champs d'application

- Analyse des images satellitaires
(localisation des zones de désertification);
- Agents conversationnels (« *chatbots* ») ;
- Robotique (interaction homme-machine) ;
- Avionique (pilotage automatique « intelligent ») ;
- Informatique (*AutoML* topologie de réseaux neuronaux) ;
- Formation
(système de diagnostic de pannes *SHERLOCK* de la *DARPA*) ;
- Finance (trading algorithmique, *Kensho* de Goldman Sachs) ;
- Santé (diagnostics, imagerie médicale, création de médicaments) ;
- Police (prévention des crimes et délits au Royaume-Uni) ;
- Armée (guidage de missiles, ciblage zones, *Gospel*) ;
- Ressources humaines (recrutement Unilever, *TextRecruit*)



Les problématiques sociétales d'application de l'intelligence artificielle

- Ouvrir la boîte noire des réseaux de neurone multi-couches (légitimation des algorithmes);
- Eviter la surcharge cognitive dans les interfaces homme-machine (alertes inopportunes) ;
- Détecter les biais cognitifs a priori dans la sélection des bases d'apprentissage ;
- Eviter le renforcement des biais par le sur-apprentissage (compromis biais-variance) ;
- Renforcer la vigilance sur les effets sociétaux induits (politique d'évaluation et de test).

IA: notre ambition pour la France



IA, notre ambition pour la France (1)

- Six propositions du rapport :

► **lancer immédiatement un plan de sensibilisation et de formation de la nation** : animation de débats publics en continu sur les impacts économiques et sociétaux de l'IA au plus près des lieux du quotidien, structuration de l'offre de formation d'enseignement supérieur, massification de la formation continue aux outils d'IA, intégration de l'IA comme objet et outil du dialogue social ; **(ODD 4)**



► **réorienter structurellement l'épargne vers l'innovation et créer, à court terme, un fonds « France & IA » de 10 Md€**, pour financer l'émergence de l'écosystème d'IA et la transformation du tissu économique français ; **(ODD9)**

► **faire de la France un pôle majeur de la puissance de calcul** : approvisionnement collectif sécurisé d'ampleur nationale et européenne, appel à projets d'implantation de centres de calcul avec garantie publique d'utilisation et simplification des procédures, crédit d'impôt IA pour l'entraînement de modèles ; **(ODD9)**



IA, notre ambition pour la France (2)

► **faciliter l'accès aux données en matière de données à caractère personnel**, modernisation du mandat de la Commission nationale de l'informatique et des libertés (Cnil) et de son collègue, suppression de certaines procédures d'autorisation préalable d'accès aux données de santé et réduction des délais de réponse ; en matière culturelle, mise en place de l'infrastructure technique favorisant l'entraînement des modèles d'IA dans le respect des droits de propriété intellectuelle **(ODD3)** ;



► **assumer le principe d'une « exception IA » dans la recherche publique** : libération des chercheurs des contraintes administratives, revalorisation de leur rémunération, doublement des moyens de la recherche publique spécialisée en IA ; **(ODD9)**



► **promouvoir une gouvernance mondiale de l'IA** : création d'une Organisation mondiale de l'IA pour évaluer et encadrer les systèmes d'IA, d'un Fonds international pour l'IA au service de l'intérêt général et d'un mécanisme de solidarité « 1 % IA » pour les pays en voie de développement **(ODD3)**



Règlement européen sur l'IA (*IA Act*)

Risque inacceptable

Les systèmes d'IA à risque inacceptable sont des systèmes considérés comme une menace pour les personnes et seront interdits. Ils comprennent :

- la **manipulation cognitivo-comportementale** de personnes ou de groupes vulnérables spécifiques : par exemple, des jouets activés par la voix qui encouragent les comportements dangereux chez les enfants
- le **score social** : classer les personnes en fonction de leur comportement, de leur statut socio-économique, de leurs caractéristiques personnelles
- la **catégorisation** et l'**identification biométriques** des personnes
- les **systèmes d'identification biométrique** en temps réel et à distance, tels que la **reconnaissance faciale**.

Règlement européen sur l'IA (*IA Act*)

- **Risque élevé**

Les systèmes d'IA qui ont un impact négatif sur la sécurité ou les droits fondamentaux seront considérés comme à haut risque et seront divisés en deux catégories.

1. Les systèmes d'IA qui sont utilisés dans les produits relevant de la [législation de l'UE sur la sécurité des produits](#). Cela comprend les **jouets**, l'**aviation**, les **voitures**, les **dispositifs médicaux** et les **ascenseurs**.

2. Les systèmes d'IA relevant de **domaines spécifiques** qui devront être enregistrés dans une base de données de l'UE :

- la gestion et l'exploitation des infrastructures critiques ;
- l'éducation et la formation professionnelle ;
- l'emploi, la gestion des travailleurs et l'accès au travail indépendant ;
- l'accès et la jouissance des services privés essentiels et des services et avantages publics ;
- les forces de l'ordre ;
- la gestion de la migration, de l'asile et du contrôle des frontières ;
- l'aide à l'interprétation juridique et à l'application de la loi.

Tous les systèmes d'IA à haut risque seront **évalués avant leur mise sur le marché et tout au long de leur cycle de vie**.

Projet de Règlement européen sur l'IA (3)

IA généraliste et générative

L'IA générative, comme ChatGPT, devrait se conformer aux exigences de transparence :

- indiquer que le **contenu a été généré par l'IA**;
- concevoir le modèle pour **l'empêcher de générer du contenu illégal** ;
- publier des **résumés des données protégées par le droit d'auteur** utilisées pour la formation.

Les modèles d'IA à usage général et à fort impact susceptibles de présenter un risque systémique, tels que le modèle d'IA plus avancé GPT-4, devront faire l'objet d'évaluations approfondies et signaler tout incident grave à la Commission.

Règlement européen sur l'IA (*IA Act*)

Risque limité

Les systèmes d'IA à risque limité doivent respecter des exigences de transparence minimales qui permettront aux **utilisateurs** de prendre des **décisions éclairées**.

Après avoir interagi avec les applications, l'utilisateur peut alors décider s'il souhaite continuer à l'utiliser.

Les **utilisateurs doivent être informés lorsqu'ils interagissent avec une IA**.

Cela inclut les systèmes d'IA qui génèrent ou manipulent du contenu image, audio ou vidéo (par exemple, les *deepfakes*, contenus faux rendus crédibles par l'IA).

Chronologie

- Le 9 décembre 2023, [le Parlement est parvenu à un accord provisoire avec le Conseil sur la loi sur l'IA](#). Le texte convenu devra maintenant être formellement adopté par le Parlement et le Conseil pour entrer en vigueur.
- Les [commissions du marché intérieur](#) et celle des [libertés civiles](#) du Parlement voteront sur l'accord lors d'une prochaine réunion.
- **Adoption par les députés européens le 12/03/2024 (523 voix contre 46)**, approbation par le Conseil européen en avril, publication au JO de l'UE en juin.

Lutter contre les effets pervers de l'IA

- Informer et former les professionnels à l'éthique

Repères

Le code d'éthique et de conduite professionnelle de l'IFIP

Quelques éléments de contexte, de contenu et de réflexion.

Dominique Desbois

<https://doi.org/10.4000/terminal.6402>

132-133 | 2022

Quand la gouvernance d'Internet fait controverse

Repères

Code d'éthique et de conduite professionnelle de l'international federation for information processing

Traduction de Dominique Desbois et Cedric Gossart

<https://doi.org/10.4000/terminal.8463>



terminal

Lutter contre les effets pervers de l'IA

- *Guide juridique pour une IA de confiance*

<https://www.cadrescfdt.fr/actualites/un-guide-juridique-pour-une-ia-de-confiance>

Éditions Lavoisier

UN GUIDE JURIDIQUE POUR UNE IA DE



ACTUALITÉS

PUBLICATIONS

EN RÉGION

MÉTIERS

QUI SOMMES-NOUS ?

COUP DE POUCE JEUNES



[Bon de commande - Le guide de l'IA au travail.pdf \(cadrescfdt.fr\)](#)



INTELLIGENCE ARTIFICIELLE :

*L'Humain aux
commandes*



IA, les propositions de la CFDT-Cadres

« *IA : Notre ambition pour la France* », rapport de la **Commission Intelligence Artificielle**, a été rendu public ce mercredi 13 mars 2024. Parmi les quinze experts ayant contribué à sa réalisation, **Franca Salls-Madinier**, secrétaire nationale CFDT Cadres et membre du Comité économique et social européen, a avancé plusieurs propositions :

- 1 **Prévoir un effort de formation sans précédent.**
- 2 **Organiser un dialogue social et une démarche de co-construction.**
- 3 **Anticiper l'impact de l'IA sur les emplois.**
- 4 **Garantir la transparence et l'explicabilité des algorithmes.**
- 5 **Évaluer les algorithmes employés sur le lieu de travail.**
- 6 **Installer une gouvernance ouverte avec un Conseil de l'IA.**

4 ÉDUCATION
DE QUALITÉ



9 INDUSTRIE,
INNOVATION ET
INFRASTRUCTURE



8 TRAVAIL DÉCENT
ET CROISSANCE
ÉCONOMIQUE



15 VIE
TERRESTRE



« Il m'est apparu que nous étions ici en présence d'une autre potentialité sociale d'une importance inouïe pour le meilleur et pour le pire »,
Norbert Wiener,
Cybernetics or Controls and Communication in the Animal and the Machine (1948).

Dossier "Pour qui sonne l'IA ?"

Revue *Cadres* n°501.

Extrait du sommaire :

Edito : L'intelligence artificielle comme fait social,

Dominique Desbois

La technologie évacue le travail réel,
entretien avec

Marc-Eric Bobillier Chaumon

Ce que l'IA fait au travail humain,

Dominique Desbois

Inclusion numérique et coopération,

Elisabeth Le Faucheur

Contre l'organisation cybernétique du travail,

Romain Bertrand

Prévenir les risques liés au numérique,

Carole Desmarais

Plus d'informations sur :

<https://www.larevuecadres.fr/numeros/pour-qui-sonne-l-ia/6959>



Cadres

Lire le travail

Juillet 2024 / n°501

Débordés ? Ne culpabilisez pas

Nouvelles normes de durabilité

Compétences psychosociales

L'implantation de l'extrême droite

Pour qui sonne l'IA ?