

# VRS

La vie de la recherche scientifique

débats  
stratégies  
opinions  
dossiers  
actions

437  
avril - mai - juin  
2024

Prix au numéro : 8€



## INTELLIGENCE ARTIFICIELLE : RÉVOLUTION OU OUTIL ?

### ZOOM

LE TOURNANT - CE QUE NOUS VOULONS POUR L'ENSEIGNEMENT SUPÉRIEUR ET LA RECHERCHE

### HORS CHAMP

LE SALARIAT, UNE FORCE EN TEMPS DE COVID



Victoria,  
professeure des écoles.

S'engager auprès de ceux  
qui se mettent au service  
des autres, c'est ça être  
assurément humain.



Assurément  
Humain

---

# édito

---



**Boris Gralak**

Secrétaire général du SNCS-FSU

## *Après l'ultime avertissement, changer de politique et de modèle de société*

La lutte contre l'extrême droite est un engagement fondamental du SNCS-FSU, du SNESUP-FSU et de notre fédération, la FSU, pour assurer et renforcer les libertés démocratiques et syndicales, la transition socio-écologique, le féminisme, la défense et les droits des minorités, la paix, la lutte pour un autre modèle de société...

La séquence électorale que nous venons de traverser a été extrêmement grave et dangereuse, avec une extrême droite très proche d'arriver au pouvoir en France. Au soir du 7 juillet, les résultats des élections législatives ont été un véritable soulagement, le Nouveau Front populaire arrivant en tête et l'extrême droite étant écartée du pouvoir.

L'échec relatif de l'extrême droite a été possible grâce à la mobilisation des forces démocratiques et républicaines, politiques et syndicales, et de la société civile. Le SNCS-FSU, le SNESUP-FSU, la FSU et leurs militantes et militants ont mobilisé toutes leurs forces dans cette lutte existentielle, à toutes les étapes de la séquence électorale. Les formations politiques de gauche ont répondu aux attentes de leurs sympathisantes et sympathisants en s'unissant dans le Nouveau Front populaire (NFP). En toute indépendance et conscients de leurs responsabilités, le SNCS-FSU, le SNESUP-FSU et la FSU ont soutenu le programme du NFP puis ont appelé, au second tour, à voter en faveur des candidats les mieux placés pour battre l'extrême droite.

Mais, si le pire a pu être évité, l'extrême droite progresse fortement avec un nombre de députés passant de 89 à 143. Cette crise politique nous rappelle douloureusement que rien n'est jamais acquis : sans mobilisation populaire pour imposer, dans la durée, un autre modèle social et écologique de société, il est à craindre que les crises politiques, sociales, économiques et écologiques que nous vivons depuis plusieurs années continuent à se succéder.

Car cette crise politique, avec un niveau inédit et inquiétant de l'extrême droite, est le résultat des politiques libérales d'affaiblissement des services publics et créatrices d'inégalités, de souffrances sociales et de repli sur soi, comme de la normalisation des idées d'extrême droite. La responsabilité du président Macron et de ses gouvernements est, à cet égard, très lourde.

La nouvelle Assemblée nationale et le prochain gouvernement devront rompre radicalement avec les politiques menées jusqu'à présent sous peine de favoriser l'arrivée au pouvoir de l'extrême droite aux prochaines élections.

Le SNCS-FSU et le SNESUP-FSU continueront de porter haut et fort leurs revendications et seront particulièrement vigilants quant au respect de la liberté d'expression des scientifiques, à l'indépendance de la recherche et aux libertés académiques.

Le SNCS-FSU et le SNESUP-FSU soulignent l'importance de construire rapidement une perspective politique à même de répondre aux attentes sociales et aux revendications des organisations syndicales, condition nécessaire pour reprendre espoir, défaire l'extrême droite et changer de modèle de société.

# sommaire

## Édito 03

Après l'ultime avertissement, changer de politique et de modèle de société **Boris Gralak**

## Actualités 05

### Dossier 09

INTELLIGENCE ARTIFICIELLE : RÉVOLUTION OU OUTIL ?

Coordonné par Hervé Christofol et Chantal Pacteau

Une histoire de l'intelligence artificielle **Lilian Aveneau 10**

Les IA génératives **Lilian Aveneau 14**

Reproductibilité et erreurs dans les IA **Lilian Aveneau 17**

AI Act : une première réglementation internationale sur l'IA **Marc Champesme 19**

Enseignants : les défis du droit d'auteur à l'ère des IA génératives **Kari Kivinen 23**

Intelligence artificielle et travail **Maxime Amblard, Luc Pellissier 25**

Les vrais dangers de l'IA **Karèn Fort 29**

Faire face aux enjeux environnementaux de l'IA : au-delà du techno-solutionnisme **Anne-Laure Ligozat 31**

IA en santé : les promesses ne doivent pas occulter l'impact sociétal concret **Aurélie Névéal 33**

L'alliance périlleuse de l'IA et du militaire **Claude Serfati 35**

IA, des technologies pour les services publics ? **Hervé Christofol 39**

### Zoom 43

Le tournant - ce que nous voulons

pour l'enseignement supérieur et la recherche **Caroline Mauriat, Boris Gralak, Julie Robert, Pascal Anger,**

### Hors-Champ 47

Le salariat, une force en temps de Covid **Claude Didry**

---

**Syndicat national des chercheurs scientifiques [SNCS-FSU]** Campus CNRS d'Ivry sur Seine - 27, rue Paul Bert 94200 Ivry sur Seine - Tél.: 01 49 60 40 34 - sncs@cnrs.fr - www.snscs.fr

**Syndicat national de l'enseignement supérieur [SNESUP-FSU]** 78, rue du Faubourg-Saint-Denis 75010 Paris - Tél.: 01 44 79 96 10. Fax : 01 42 46 26 56 - accueil@snesup.fr - www.snesup.fr

---

**Directeur de la publication:** Boris Gralak | **Rédacteurs en chef:** Hervé Christofol, Chantal Pacteau | **Comité de programmation:** Maxime Amblard, Florence Audier, Lilian Aveneau, Marc Champesme, Hervé Christofol, Claude Didry, Karèn Fort, Boris Gralak, Kari Kivinen, Anne-Laure Ligozat, Caroline Mauriat, Aurélie Névéal, Chantal Pacteau, Luc Pellissier, Claude Serfati | **Montage:** Clotilde Péan | **Couverture:** Pixabay | **Impression et routage:** Imprimerie Compédit Beauregard. Z.I. Beauregard, BP 39, 61600 La Ferté-Macé. Tél.: 02 33 37 08 33 - www.compedit-beauregard.fr | **Régie publicitaire:** Com d'habitude publicité. 7 rue Emile Lacoste 19100 Brive-la-Gaillarde. Tél.: 05 55 24 14 03. Contact: Clotilde Poitevin-Amadiou (www.comdhabitude.fr - contact@comdhabitude.fr) | La Vie de la recherche scientifique est publiée par le SNCS-FSU, 27, rue Paul Bert 94200 Ivry sur Seine. Tél.: 01 49 60 40 34 - sncs@cnrs.fr. Commission paritaire: 0424 S 07016. ISSN: 0338-1889. Dépôt légal à parution. Prix au numéro: 8€ - Abonnement annuel (4 numéros): 25€ (individuel), 50€ (institutionnel). Les titres sont de la responsabilité de l'équipe de rédaction.



Ont participé à la rédaction des actualités : Florence Audier, Hervé Christofol, Boris Gralak, Caroline Mauriat et Chantal Pacteau.

## 49<sup>ÈME</sup> CONGRÈS DU SNCS-FSU : L'ORIENTATION *ENSEMBLE* DU SNCS-FSU RENFORCÉE

Du 17 au 19 juin 2024, le 49<sup>ème</sup> congrès du SNCS-FSU a réuni près d'une centaine de syndiqué·e·s délégué·e·s pour débattre et adopter l'orientation du SNCS-FSU des prochaines années pour la recherche publique et ses personnels. Ce congrès a été l'occasion de débattre de la situation générale caractérisée, ces dernières années, par une succession de crises politiques particulièrement menaçantes pour la société, notamment celle de ces dernières semaines qui a failli voir arriver au pouvoir l'extrême droite, suite à la décision du président de la République de dissoudre l'Assemblée nationale le 9 juin 2024.

Les discussions et débats sur la lutte pour faire barrage à cette arrivée au pouvoir de l'extrême droite ont constitué un moment fort du congrès qui a adopté une motion dans laquelle il appelait à voter pour les candidates et candidats qui s'engageaient pour le programme du Nouveau Front populaire<sup>1</sup>.

Le congrès a également été l'occasion d'échanges fructueux avec les directions syndicales invitées représentées par le secrétaire général (SG) de la FSU, Benoît Teste, la co-SG du SNESUP-FSU Caroline Mauriat, la co-SG du SNASUB-FSU Julie Robert, les co-SG du SNTRS-CGT Josiane Tack et Patrick Boumier, le secrétaire national du SNPTEs-UNSA Xavier Duchemin et les co-SG du SGEN-CFDT Recherche EPST Yannick Bourlès et Isabelle Cohen.

Un temps d'échange a été organisé avec Benoît Teste, Josiane Tack et Patrick Boumier sur les discussions en cours entre la FSU et la CGT pour un outil syndical commun.

L'ensemble des syndiqué·e·s du SNCS-FSU ont voté sur le Rapport d'activité du Bureau national (BN) qui a été approuvé à 89,9 % ainsi que sur l'orientation du syndicat pour les prochaines années. L'orientation *Ensemble* est sortie renforcée par ce vote, avec plus de 90 % des voix exprimées et vingt-deux sièges sur vingt-quatre au titre des tendances à la Commission administrative. L'orientation *Front Unique* a reçu 10 % des voix exprimées et deux sièges à la Commission administrative.

Le congrès a approuvé la nouvelle composition de la Commission administrative (CA) du SNCS-FSU avec



De gauche à droite, Benoît Teste en visioconférence, Josiane Tack au pupitre, Boris Gralak et Dimitri Peaucelle.

quarante sièges (quarante titulaires et quarante suppléant·e·s), dont vingt-quatre sièges au titre des tendances et seize sièges au titre des sections locales, nationales et scientifiques. La CA a ensuite élu la nouvelle équipe qui animera le syndicat : Boris Gralak, secrétaire général du SNCS-FSU, Dimitri Peaucelle, trésorier du SNCS-FSU, Maud Leriche et Julien Diaz, secrétaires généraux·ales adjoint·e·s du SNCS-FSU, ainsi que le BN.

L'ensemble des résultats des votes, les six motions adoptées par le congrès et les textes des trois thèmes approuvés qui réaffirment et enrichissent les revendications et les réflexions du SNCS-FSU (Conditions de travail ; Structuration de la recherche en France et transition socio-écologique dans la recherche) sont rassemblés sur la page web du bilan du congrès<sup>2</sup>. Quant à la Résolution finale<sup>3</sup>, elle rappelle les luttes actuelles et réaffirme les réflexions et revendications du SNCS-FSU pour la recherche publique et ses personnels.

Boris Gralak a conclu le congrès en remerciant chaleureusement l'ensemble des syndiqué·e·s qui ont participé au vote, aux débats dans leurs sections locales, scientifiques et nationales, ainsi que les délégué·e·s qui ont participé au congrès à Villejuif et en visioconférence.

<sup>1</sup> <https://miniurl.be/r-5k8n>

<sup>2</sup> <https://sncs.fr/2024/06/19/bilan-du-49e-congres-du-sncs-fsu/>

<sup>3</sup> <https://sncs.fr/wp-content/uploads/2024/06/congres2024-SNCS-FSU-resolution-finale.pdf>

## CONGRÈS D'ÉTUDES DU SNESUP

Le SNESUP-FSU a tenu son congrès d'études à Poitiers les 5 et 6 juin 2024 entre les deux tours des élections européennes<sup>1</sup>.

Trois thématiques étaient débattues : « L'urgence écologique, quelles contributions pour l'enseignement supérieur et la recherche ? », « L'autonomie des universités, "acte II" : à quel prix ? », « L'enseignement supérieur et la recherche entre service public et marché ? ». Les débats du premier thème ont insisté sur la nécessité de construire collectivement les réponses aux urgences écologiques. En effet, si le plan climat biodiversité dans l'enseignement supérieur et la recherche (ESR) est maintenant mis en place et décliné dans tous les établissements, rares sont ceux qui ont organisé des discussions sur le sujet dans les diverses instances représentatives des personnels (CSA et F3SCT notamment) alors que l'organisation du travail, la santé et la sécurité sont et seront largement concernées. Le SNESUP-FSU appelle notamment à s'emparer du sujet dans les établissements, au plus près des collègues et en cohérence avec l'organisation du territoire et les contraintes climatiques locales ainsi qu'à porter des propositions construites collectivement dans les instances.

Le thème 2 a pointé la dangerosité des mesures contenues dans l'acte II de l'autonomie qui, toutes, conduisent à la fragmentation du service public de l'ESR : gestion de tous les personnels au niveau local, dérogations au Code de l'éducation, modulations des services, etc. Le SNESUP-FSU s'oppose à toutes modifications de la réglementation qui permettraient d'aller au-delà des dispositions déjà prévues par la loi relative aux libertés et responsabilités des universités (LRU) dont il demande l'abrogation.

Enfin, dans le thème 3, le SNESUP-FSU a dénoncé le développement de l'enseignement supérieur privé au détriment de l'enseignement supérieur public qui bénéficie, pour ses formations, de la croissance de l'apprentissage. Formations vers lesquelles les bachelier·e·s sont poussé·e·s, d'une part à cause des délais d'obtention d'une réponse d'affectation par Parcoursup et, d'autre part, parce que les formations

publiques sont insuffisamment financées pour pouvoir accueillir les néo-bachelier·e·s et ouvrir le nombre de places nécessaires en master. Pour le SNESUP-FSU, les financements publics doivent revenir exclusivement aux établissements publics.

Par ailleurs, le SNESUP-FSU a une nouvelle fois appelé le gouvernement à reporter la réforme sur le recrutement et la formation des enseignant·e·s ; il a également exprimé sa solidarité et apporté son soutien aux universitaires de toutes nationalités qui appellent à la paix dans toutes les zones de guerre. Enfin, il a rappelé son attachement indéfectible à l'indépendance des enseignant·e·s-chercheur·e·s et appelé à faire front pour les libertés académiques.

Le webinar organisé le mardi 11 juin 2024 par le SNCS-FSU, le SNASUB-FSU et le SNESUP-FSU a réuni autour de deux cents personnes. Il a été l'occasion d'informer les collègues sur le processus en cours dans les neuf établissements expérimentaux (EPE) des diverses mesures, dont certaines dérogent au code de l'éducation. Les échanges ont confirmé que la plupart de ces établissements s'engagent dans le processus sans consulter ni les élus des conseils centraux ni les membres de la formation spécialisée en matière de santé, sécurité et conditions de travail (F3SCT), pourtant directement compétents pour apprécier l'intérêt et les impacts, notamment sur les conditions de travail des agents, des diverses dispositions dont certaines touchent directement aux statuts des agents. Suite à la dissolution de l'Assemblée nationale et aux élections législatives anticipées qui ont suivi, les travaux sur ces sujets ont été suspendus dans les établissements expérimentateurs. Mais la vigilance reste de mise, le renforcement de l'autonomie des universités en matière de ressources humaines, de budget, de patrimoine, de gouvernance et de formation pouvant revenir rapidement dans l'actualité.

Le SNESUP-FSU continue d'appeler les collègues dans les conseils centraux, comité social d'administration (CSA) et F3SCT à exiger les informations sur ce qui est prévu dans leur établissement et à s'opposer aux mesures qui contreviennent au Code de l'éducation ou contribuent à dégrader les conditions de travail et le service public de l'ESR. Il s'oppose fermement à toutes modifications réglementaires qui accentueraient les méfaits de la loi LRU.

<sup>1</sup> L'intégralité des textes votés en séance plénière à l'issue des débats et échanges en commissions est à retrouver sur le site du SNESUP-FSU.  
<https://snesup.fr/rubrique/textes-adoptes-par-le-congres-detude-2024>

**La ministre de l'Éducation nationale Nicole Belloubet de l'actuel gouvernement, en sursis, maintient la réforme de la formation des enseignant·e·s et conseillers principaux d'éducation. Un acharnement irresponsable !**

\* <https://miniurl.be/r-5kdd>

## BONNE NOUVELLE : LE NOMBRE DE DOCTEURS DIPLÔMÉS EN NETTE CROISSANCE !

En 2023, il y a eu 15 187 nouveaux docteurs, alors qu'ils étaient 11 806 en 2020. Cette augmentation de près de 30 % – même si les effectifs concernés restent modestes – concerne tous les domaines d'études<sup>1</sup>, mais principalement les sciences « exactes » et leurs applications – dont les néo-docteurs passent de 5 576 en 2020 à 7 269 en

2023 – et les sciences du vivant – dont les néo-docteurs passent de 2 535 à 3 576.

A noter que les sciences agronomiques et écologiques dépassent pour la première fois les 600 nouveaux docteurs en 2023 – ils n'étaient que 437 en 2020 !

<sup>1</sup> Pour le détail par domaine d'étude, voir la Note Flash du SIES n° 17, juin 2024.  
<https://www.enseignementsup-recherche.gouv.fr/fr/les-docteurs-diplomes-en-2023-96624>

## CONGRÈS FSU EN FÉVRIER 2025 À RENNES

Du 3 au 7 février 2025, le 11<sup>e</sup> congrès national de la FSU aura lieu au couvent des Jacobins de Rennes, trois ans après le 10<sup>e</sup> congrès de la FSU de Metz. Il sera précédé par le vote des adhérent·e·s (du jeudi 7 au jeudi 28 novembre 2024) et par les congrès départementaux qui se dérouleront du mercredi 11 au vendredi 20 décembre 2024, puis du lundi 6 au vendredi 17 janvier 2025.



## A NE PAS MANQUER



Le mensuel du SNESUP  
N° 725  
juin-juillet-août 2024  
40e anniversaire :  
défendre le statut des  
enseignants-chercheurs

<https://miniurl.be/r-5kd8>

A l'instar de ce que les scientifiques et leurs organisations ne cessent de répéter depuis des années<sup>\*</sup>, le premier baromètre d'AEF info sur la simplification de la gestion de la recherche confirme ce que sont les appels à projet (AAP) pour la majorité d'entre eux : une « loterie » qui rend leur métier « déprimant », « un gâchis d'argent public », « au détriment du temps dédié à la recherche ». En fait, « les AAP devraient être du bonus, pas une nécessité ».

<https://miniurl.be/r-5kdc>

<sup>\*</sup> voir par exemple le dossier de la VRS 436 « Les temps de la recherche ».  
<https://www.snesup.fr/article/les-temps-de-la-recherche-vrs-436-mars-2024>

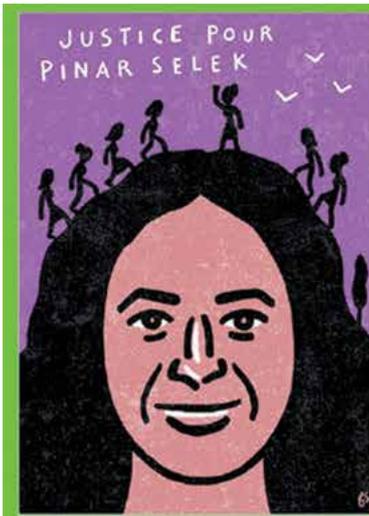
## LES JUSTIFICATIFS DE REPAS EN MISSION, C'EST FINI !

Un nouveau décret<sup>1</sup> allège et simplifie la communication des pièces justificatives afférentes aux frais de déplacement avancés par les agents publics en mission en France et à l'étranger. Il prévoit le principe général

de non-conservation des pièces justificatives des frais de repas<sup>2</sup>, sauf dérogations éventuellement prévues par arrêté ministériel. Il précise également que les pièces justificatives d'hébergement doivent être conservées par l'agent pendant un an et transmises à l'ordonnateur en cas de demande express. Un arrêté ministériel peut prévoir l'absence de conservation des pièces justificatives d'hébergement pour les missions à l'étranger.

<sup>1</sup> Ce décret modifie le décret du 3 juillet 2006 fixant les conditions et les modalités de règlement des frais occasionnés par les déplacements temporaires des personnels civils à la charge des budgets des services de l'État et des établissements publics nationaux à caractère administratif (EPNA), ainsi que des établissements publics locaux d'enseignement (EPL), des établissements publics à caractère scientifique, culturel et professionnel (EPSCP) et des établissements publics à caractère scientifique et technologique (EPST).  
<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049894058>

<sup>2</sup> <https://snscs.fr/2024/07/11/les-justificatifs-de-repas-en-mission-cest-fini/>



## JUSTICE ET LIBERTÉ POUR PINAR SELEK

Le 28 juin, s'ouvre en Turquie le cinquième procès de Pinar Selek, universitaire arrêtée il y a plus d'un quart de siècle pour sa recherche sur la résistance kurde, et exilée depuis 2008. Le SNE-SUP-FSU continuera à soutenir Pinar en participant notamment à la campagne internationale pour son acquittement définitif et aux événements organisés en France. Défendre la justice et la liberté pour Pinar, c'est défendre la justice et la liberté pour toutes et tous.

<https://www.snesup.fr/article/mensuel-ndeg-725-juin-juillet-aout-2024>

## RAPPORT ANNUEL DE LA COALITION MONDIALE POUR LA PROTECTION DE L'ÉDUCATION CONTRE LES ATTAQUES

La Coalition mondiale pour la protection de l'éducation contre les attaques vient de publié son rapport annuel sur les attaques contre l'éducation en 2022 et 2023. Environ six mille attaques contre l'éducation et incidents d'utilisation militaire d'écoles et



d'universités ont été signalés, touchant plus de dix mille étudiants et éducateurs dans le monde. Un accroissement de près de 20 % par rapport aux deux années précédentes.

<https://miniurl.be/r-5kde>

## TABLE RONDE AU SÉNAT SUR LES INFLUENCES ÉTRANGÈRES DANS L'ESPACE NUMÉRIQUE

Parmi les intervenants de la table ronde<sup>1</sup>, David Chavalarias, directeur de recherche au CNRS et directeur de l'Institut des systèmes complexes a expliqué comment se déploient les opérations d'ingérence étrangères, au travers de l'architecture des plateformes qui les favorisent. L'observatoire de l'espace numérique Twitter, que son institut a créé en 2016, essaie de comprendre comment s'articulent les dynamiques d'opinion et comment fonctionnent les campagnes d'influence et d'ingérence étrangères. On observe, par exemple, qu'en ce moment il y a une reconfiguration autour du climato-dénialisme...

<sup>1</sup> [https://www.senat.fr/compte-rendu-commissions/20240603/ce\\_influences.html#toc3](https://www.senat.fr/compte-rendu-commissions/20240603/ce_influences.html#toc3)

Plusieurs pistes ont été proposées pour lutter contre les stratégies d'influence.

### Minuit moins dix à l'horloge de Poutine

Le projet Politoscope observe depuis 2016 le militantisme politiques sur X. Des méthodes ont été développées pour analyser les dynamiques sociales et de débats ainsi que les manipulations d'opinions. Le Politoscope décode le contexte très particulier de l'accession au pouvoir du Rassemblement national, marqué par l'ingérence de Vladimir Poutine\*.

\* <https://hal.science/hal-04629585v1/document>

**Le médecin et universitaire Ahmadreza Djalali, qui vivait en Suède et était également professeur invité spécialisé dans la médecine de catastrophe à l'Université Vrije de Bruxelles, a été arrêté arbitrairement le 26 avril 2016, alors qu'il se trouvait en Iran pour des raisons professionnelles. Il risque sérieusement d'être exécuté après avoir épuisé toutes les voies légales pour annuler sa condamnation à mort. Du 26 juin au 4 juillet 2024, il a observé une grève de la faim pour protester contre son maintien en détention arbitraire à la suite d'un échange de prisonniers entre l'Iran et la Suède qui a eu lieu le 15 juin, explique Amnesty International.**

\* <https://www.amnesty.org/fr/documents/mdet13/8279/2024/fr/>

Hervé Christofol  
Chantal Pacteau

## Intelligence artificielle : révolution ou outil ?

Avec l'avènement de ChatGPT4 en 2023, le grand public a découvert, à travers ce robot conversationnel, les performances de l'intelligence artificielle (IA) générative qui produit des réponses (textes, images, films, musiques...) à des requêtes exprimées en langage naturel. Révolution industrielle pour certains – qui pensent qu'elle va transformer tous les métiers et l'organisation du travail –, nouvel outil pour d'autres – qui l'envisagent au service des citoyens et de la productivité de certains secteurs d'activité –, les développements et usages de l'IA font débat. C'est ce dont les articles de ce dossier discutent.

L'histoire de l'IA, et notamment celle de l'IA générative, permet de mieux comprendre les hypothèses théoriques sur lesquelles l'IA repose ainsi que ses fonctionnements. L'IA ne se limite pas aux robots conversationnels ; les véhicules autonomes ou la reconnaissance faciale utilisent des technologies informatiques qui relèvent de l'IA. Mais les robots conversationnels, de par leur accessibilité, leur intérêt et la curiosité des réponses synthétiques qu'ils génèrent, ont été rapidement adoptés sans que les impacts et la pertinence de leurs usages n'aient été questionnés.

Or, comme d'autres technologies en leur temps, la qualité des réponses que les robots génèrent n'est pas sans défaut. Leur entraînement et leur usage sont dix fois plus consommateurs d'énergie qu'une simple requête *via* un moteur de recherche et nécessitent le travail invisibilisé de « petites mains » actuellement réalisé dans des conditions sociales et sanitaires déplorables. Le respect des droits d'auteur n'est pas garanti et les données collectées ne sont pas sans biais, ce qui pose d'importantes questions éthiques.

L'Europe a légiféré pour tenter de contenir les usages les plus problématiques de l'IA mais les entreprises et les États, motivés par de nouveaux marchés et les gains de productivité que les économistes font miroiter ainsi que par la suprématie militaire que de nouvelles armes intégrant l'IA peut apporter, s'engagent dans une course au déploiement de solutions pour leur personnels et leurs clients, leurs armées et leurs usagers. C'est notamment le cas de l'État français qui déploie plusieurs expérimentations d'outils reposant sur l'IA dans les services publics et la défense.

Pour l'instant, loin de les révolutionner, l'IA se déploie rapidement comme outil au service des métiers de l'enseignement supérieur et de la recherche. Puisse ce dossier faire réfléchir quant à la pertinence de son usage, compte tenu notamment de ses impacts environnementaux et sociaux.

# Une histoire de l'intelligence artificielle

Explorer l'histoire de l'intelligence artificielle (IA) depuis ses débuts jusqu'à nos jours offre un voyage fascinant à travers le développement d'idées, de théories et de technologies diverses et variées. Les grands jalons de cette histoire sont évoqués ici.

Quand on parle d'intelligence artificielle, de quoi parle-t-on ? Selon le Parlement européen<sup>1</sup>, l'IA englobe un large éventail d'outils et de techniques permettant aux machines de « reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité ».

L'IA peut être classée en deux catégories :

- l'IA faible, ou étroite, conçue pour effectuer une tâche spécifique ; elle ne possède pas de conscience ou d'intelligence générale, mais peut se révéler bien supérieure à l'humain sur sa tâche spécifique ;
- l'IA forte, ou générale, aujourd'hui encore hypothétique ; elle posséderait une intelligence et une conscience équivalentes voire supérieures à celles d'un humain, et serait capable d'effectuer n'importe quelle tâche intellectuelle. Les chercheurs en IA sont partagés sur la date d'une future apparition d'une telle IA : d'ici quelques années pour certains, jamais pour d'autres...

## LES DÉBUTS DE L'IA

Si les premiers ordinateurs apparurent dans les années 1940, les bases théoriques de l'IA moderne sont antérieures. Au XIX<sup>e</sup> siècle, George Boole puis Gottlob Frege ont développé des systèmes de logique formelle, capables de représenter des idées et des concepts *via*

la manipulation de symboles abstraits, tels que A, B, C... La logique formelle permet de répondre aux critiques sur les mathématiques alors considérés comme ambigus, notamment en permettant la construction formelle de la théorie des ensembles et des nombres<sup>2</sup>.

Utiliser l'IA et les robots n'est pas une idée neuve. Dès 1920, Roy J. Wensley construisit un robot mécanique nommé « Mr. Televox » qui fit sensation par sa capacité à répondre à des commandes vocales telles que « avance », « recule » et « tourne », et cela malgré qu'il n'était doté d'aucune IA, ni même d'un quelconque ordinateur.

Dans les années 1930, Alan Turing a développé les premiers concepts des ordinateurs programmables, capables d'effectuer des calculs stockés en mémoire. Dans la même période, d'autres travaux ont jeté les bases de la notion d'algorithme et de leur complexité (Kurt Gödel, Alonzo Church et Haskell Curry par exemple), éclairant sur ce qu'il est théoriquement possible de calculer et de ne pas calculer.

En 1943, dans leur article intitulé « *A logical calculus of the Ideas Immanent in Nervous activity* », le neurologue Warren McCulloch et le logicien Walter Pitts ont proposé un modèle mathématique de

neurones utilisables pour effectuer des calculs<sup>3</sup>. Comme Alan Turing, ils considèrent les fonctions de l'esprit comme une fonction mathématique (un opérateur transformant des entrées en des sorties). En introduisant le caractère de « tout ou rien » dans la description de l'activation du système neuronal, leur modèle se situe dans la catégorie des modèles logiques.

En 1945, John von Neumann a proposé l'architecture des ordinateurs modernes<sup>4</sup>. Cette architecture est devenue la base de tous les ordinateurs modernes et a permis de traiter des quantités de données beaucoup plus importantes que les machines mécaniques précédentes, notamment en séparant le programme logiciel du matériel, dont l'unité de contrôle réagit à des instructions machines précises.

### Lilian Aveneau

Membre de la commission de recherche du SNESUP, PU en informatique, laboratoire Xlim, Université de Poitiers

<sup>3</sup> Le neurologue américain Warren Sturgis McCulloch est l'initiateur des rencontres interdisciplinaires, dites Conférences Macy, qui réunirent certains des plus grands esprits de l'époque entre 1942 et 1953. Ces conférences avaient pour objectif de développer une théorie générale du fonctionnement de la pensée humaine et furent à l'origine de la cybernétique. W. S. McCulloch eut pour collaborateurs deux pionniers de l'intelligence artificielle, Marvin Minsky et Seymour Papert (tiré de : [https://fr.wikipedia.org/wiki/Warren\\_McCulloch](https://fr.wikipedia.org/wiki/Warren_McCulloch)).

Quant au logicien américain Walter Pitts, il a travaillé dans le domaine des neurosciences computationnelles. « Il a proposé des formulations théoriques marquantes de l'activité neuronale et des processus génératifs qui ont influencé divers domaines tels que les sciences cognitives et la psychologie, les neurosciences, mais aussi l'informatique, les réseaux de neurones artificiels, la cybernétique et l'intelligence artificielle (IA), ainsi que ce que l'on appelle les sciences génératives » ([https://fr.wikipedia.org/wiki/Walter\\_Pitts](https://fr.wikipedia.org/wiki/Walter_Pitts)).

<sup>4</sup> [https://fr.wikipedia.org/wiki/Architecture\\_de\\_von\\_Neumann](https://fr.wikipedia.org/wiki/Architecture_de_von_Neumann)

<sup>1</sup> <https://miniurl.be/r-5jie>

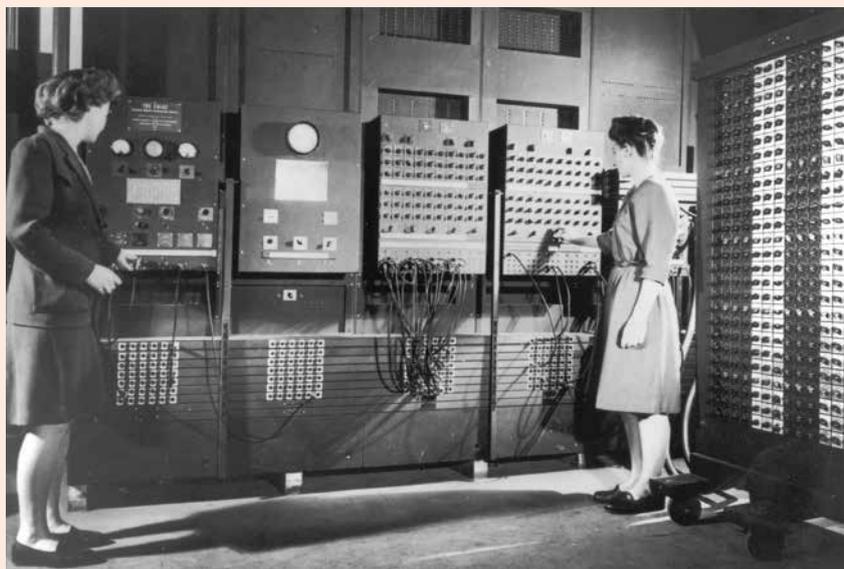
<sup>2</sup> [https://fr.wikipedia.org/wiki/Logique\\_math%C3%A9matique](https://fr.wikipedia.org/wiki/Logique_math%C3%A9matique)

## Premiers ordinateurs et loi de Moore

L'histoire de l'IA est intimement liée à celle des ordinateurs : les IA actuelles nécessitent ainsi des puissances de calcul sans aucune mesure avec celles des premiers ordinateurs.

Les premiers ordinateurs numériques naissent dans les années 1940\*. Le premier est le Z3 conçu pendant la Seconde Guerre mondiale, en 1941, par l'ingénieur allemand Konrad Zuse. Il réalise jusqu'à quatre additions par seconde, ou une multiplication toutes les quatre secondes. En 1945, l'ENIAC (pour *Electronic Numerical Integrator and Computer*), qui pèse vingt-sept tonnes, peut effectuer 100 000 additions ou 357 multiplications par seconde. L'invention du transistor en 1947 va permettre la miniaturisation de ces premiers « monstres ».

En 1965, Gordon E. Moore, l'un des fondateurs d'Intel, énonce sa fameuse loi empirique postulant le doublement des transistors d'un processeur tous les deux ans, ce que d'aucuns interprètent comme un doublement de la puissance de calcul tous les deux ans. Bien que cette loi ne fonctionne plus à partir des années 2010 pour des raisons physiques, la puissance de calcul des ordinateurs – mesurée en FLOPS (nombre d'opérations en virgule flottante par seconde) – n'a cessé de progresser exponentiellement depuis ses origines : en 1964 le Control Data 6600 fut le premier à dépasser les  $10^6$  FLOPS ; en 1997, l'ASCI Red fut le premier ordinateur au-dessus des  $10^{12}$  FLOPS ; en juin 2024, l'ordinateur parallèle Frontier\*\* (HPE Cray, USA) est estimé à  $1,6 \times 10^{18}$  FLOPS, pour plus de 22 MW.



Programmeuses exploitant le panneau de commande principal d'ENIAC à la *Moore School of Electrical Engineering*, vers 1945.

Photo de l'armée américaine issue des archives de l'ARL Technical Library.

\* [https://fr.wikipedia.org/wiki/Histoire\\_des\\_ordinateurs](https://fr.wikipedia.org/wiki/Histoire_des_ordinateurs)

\*\* <https://www.top500.org/lists/top500/2024/06/>

En 1948, Norbert Wiener<sup>5</sup> a décrit les principes de la rétroaction et de l'autorégulation dans les systèmes biologiques et mécaniques, jetant les bases

de la théorie du contrôle. Quant à Claude Shannon, il a publié la théorie de l'information, qui fournit un cadre pour la compréhension de la communication et du traitement de l'information dans les systèmes naturels et artificiels.

Ainsi, fin des années 1940, les chercheurs avaient des premiers modèles théoriques de calculs et de raisonnement, prémisses de l'IA, mais aussi les premiers ordinateurs programmables. Toutes ces théories et ces modèles furent discutés lors de la conférence de Dartmouth en 1956, considérée depuis comme l'acte de naissance de l'intelligence artificielle en tant que discipline et en tant que concept.

## L'ÈRE DE L'IA SYMBOLIQUE

Partant des travaux sur la logique symbolique initiés au XIX<sup>e</sup> siècle, des chercheurs se sont dirigés, dans les années 1960, vers des méthodes visant à mimer la logique des processus cognitifs humains. Aujourd'hui appelés « *Good old-fashioned IA* », certains restent encore très utilisés bien qu'ils paraissent dépassés par les réseaux de neurones.

Dès 1956 est apparu un premier logiciel capable de démontrer des théorèmes exprimés en logique symbolique : le « *Logic Theorist*<sup>6</sup> » créé par Allen Newell, Herbert Simon et Cliff Shaw, et permettant de reproduire trente-huit des cinquante-deux théorèmes regroupés dans le chapitre 2 de *Principia Mathematica*<sup>7</sup>, un ouvrage très connu du début du XX<sup>e</sup> siècle. Il reste considéré comme le premier logiciel d'IA jamais écrit, dans un langage de programmation dit « fonctionnel », l'IPL (pour *Information Processing Language*). L'année suivante, les mêmes A. Newell, H. Simon et C. Shaw ont produit, toujours en IPL, le GPS pour « *General Problem Solver*<sup>8</sup> ». Les nouveautés principales résident dans une séparation du raisonne-

<sup>6</sup> [https://en.wikipedia.org/wiki/Logic\\_Theorist](https://en.wikipedia.org/wiki/Logic_Theorist)

<sup>7</sup> « *Les Principia Mathematica sont une œuvre en trois volumes d'Alfred North Whitehead et Bertrand Russell, publiés en 1910-1913. Cette œuvre a pour sujet les fondements des mathématiques. (...) Il participe de façon décisive à la naissance de la logique moderne.* » [https://fr.wikipedia.org/wiki/Principia\\_Mathematica](https://fr.wikipedia.org/wiki/Principia_Mathematica)

<sup>8</sup> [https://fr.wikipedia.org/wiki/General\\_Problem\\_Solver](https://fr.wikipedia.org/wiki/General_Problem_Solver)

<sup>5</sup> [https://fr.wikipedia.org/wiki/Norbert\\_Wiener](https://fr.wikipedia.org/wiki/Norbert_Wiener)

## 1956 : la conférence de Dartmouth, naissance de l'IA

La conférence de Dartmouth est considérée comme l'événement fondateur de l'IA en tant que domaine de recherche\*. Elle a été organisée par le mathématicien John McCarthy, inventeur des termes « *intelligence artificielle* » pour leur neutralité par rapport à une méthode particulière, ainsi que par Marvin Minsky, Nathaniel Rochester et Claude Shannon.

La conférence a réuni une dizaine de chercheurs pendant deux mois durant l'été 1956 au Dartmouth College à Hanovre, États-Unis. Les participants y discutèrent de la possibilité de créer des machines capables d'imiter l'intelligence humaine, et des approches pour y parvenir, incluant l'utilisation de programmes informatiques pour simuler les processus de pensée humaine, la création de machines capables d'apprendre de leur expérience et l'utilisation de réseaux de neurones artificiels.

Cette événement a marqué le début officiel de l'IA en tant que domaine de recherche, avec plusieurs orientations : les méthodes symboliques avec des systèmes axés sur des domaines limités (premiers systèmes experts), et les premiers outils à base de neurones.

### 1956 Dartmouth Conference: The Founding Fathers of AI



John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

\* [https://en.wikipedia.org/wiki/Dartmouth\\_workshop](https://en.wikipedia.org/wiki/Dartmouth_workshop)

ment de la base des connaissances, et dans la capacité à résoudre un problème donné par la génération de sous-problèmes, plus petits et donc plus faciles à résoudre.

Ces premiers résultats ont engendré des travaux sur les langages de programmation dits « fonctionnels » avec, en tête de pont, le Lisp, pour *List Processing* ou « traitement de listes », développé en 1961 par John McCarthy. Une grande partie des financements de la recherche dans les années 1950 à 1980 a porté sur le

développement de ce type de langages et sur leur utilisation pour l'IA.

Dès lors, d'autres programmes ont tentés d'inclure un ensemble de connaissances (sous forme de règles) afin de résoudre des questions liées à un domaine précis. Face à la complexité de faire évoluer ces premiers outils, et grâce à l'utilisation des langages fonctionnels et du raisonnement symbolique, est ensuite venu l'idée de découper la partie raisonnement (le moteur d'inférence) de la base de connaissances pour construire

des outils plus faciles à faire évoluer : les « systèmes experts ».

L'un des tout premiers systèmes experts est MYCIN<sup>9</sup>, qui voit le jour en 1976. Son objectif est le diagnostic d'infections bactériennes. Doté d'une base de connaissances de six cents règles, son moteur d'inférence, par application des règles entrées par l'utilisateur et de la base de connaissances, déduit logiquement de nouvelles informations et, le cas échéant, permet de répondre à certaines questions précises (par exemple, s'agit-il d'une infection bactérienne, quelle est la bactérie en cause...).

Plus tard, afin de pallier les incertitudes propres à certains domaines (par exemple celle liée à un capteur physique), les moteurs d'inférences ont intégré la « logique floue », qui repose non plus sur des valeurs binaires (1 et 0, ou encore vrai et faux) mais des valeurs plus complexes exprimées sous formes de nombres à virgule flottante<sup>10</sup>.

Ces systèmes sont, d'un point de vue conceptuel, particulièrement satisfaisants car leur fonctionnement et raisonnement sont faciles à comprendre *via* le système d'inférence logique. Cependant, leur développement n'a pas été à la hauteur des attentes des chercheurs et des investisseurs, notamment en raison de la difficulté à construire une base de connaissances, et surtout à leur faire suivre l'évolution des connaissances du domaine d'expertise (médecine, juridique, cartes Pokémon, etc.). Sur des sujets complexes, ces bases deviennent bien trop difficiles à générer et maintenir, réduisant de fait leur développement et généralisation.

Ainsi, après l'engouement initial de l'IA symbolique et des systèmes

<sup>9</sup> Le nom MYCIN vient de ce que de nombreux antibiotiques ont le suffixe « -mycine ».

<sup>10</sup> Attention, à ne pas confondre avec les nombres réels, il s'agit de valeurs discrètes exprimées avec un nombre fixe de bits. [https://fr.wikipedia.org/wiki/Virgule\\_flottante](https://fr.wikipedia.org/wiki/Virgule_flottante)

experts, venant avec des financements de la recherche conséquents, est arrivée une période de déception et de coupure de crédits, dite « l'hiver de l'IA ». En réalité, cet hiver est pluriel est correspond, entre 1970 et 2000, à plusieurs périodes de gel des financements<sup>11</sup>.

## L'IA STATISTIQUE ET LES RÉSEAUX DE NEURONES

Une autre branche de l'IA considère le problème sous sa forme statistique : on parle alors d'apprentissage automatique. L'idée est, à partir d'un ensemble de données et par application des statistiques inférentielles ou d'un modèle probabiliste, de construire une fonction donnant la probabilité qu'un élément appartienne à une classe donnée (par exemple, tumeur bénigne *versus* tumeur maligne). Nous trouvons ici moult méthodes proposées dans la littérature : réseau bayésien, modèle de Markov caché, machine à vecteurs de support (SVM), forêt d'arbre décisionnels, modèle gaussien... Tous ces modèles sont encore utilisés en production comme en recherche, même si l'outil aujourd'hui à la mode est l'apprentissage profond basé sur les neurones.

Un réseau de neurones<sup>12</sup> est un modèle s'inspirant du fonctionnement du cortex cérébral formé de neurones interconnectés par des synapses. En 1957, Frank Rosenblatt a développé le Perceptron, un algorithme de classification de données binaires formé d'un unique neurone artificiel. En 1959, sur la base de travaux effectués avec les neurobiologistes Jérôme Lettvin et Humberto Maturana<sup>13</sup> sur le système

visuel de la grenouille, McCulloch et Pitts proposent un modèle simplifié de neurone biologique, appelé neurone formel. Ils démontrent que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes. C'est ainsi que l'idée d'utiliser des neurones pour simuler l'intelligence humaine s'est construite.

Mais, à cette époque, plusieurs difficultés limitent l'usage des réseaux de neurones. Il fallait pouvoir répondre à quatre questions principales :

- Peut-on tout calculer ?
- Comment agencer les neurones ?
- Quel algorithme pour apprendre les poids de chaque neurone ?
- Combien de données sont nécessaires pour apprendre ?

George Cybenko<sup>14</sup> répond à la première question en 1989, en montrant qu'un réseau à une seule couche peut représenter n'importe quelle fonction continue, à la condition d'avoir suffisamment de neurones. Ce résultat très important fait le lien entre un réseau de neurones et n'importe quelle fonction de décision que l'on voudrait calculer. La seconde question « *Comment agencer les neurones ?* » se subdivise en deux sous-questions. La première est : « *Combien de neurones ?* » Le cerveau humain en compte approximativement cent milliards pour, en moyenne, dix mille synapses par neurone<sup>15</sup>. Les plus gros réseaux de neurones artificiels aujourd'hui<sup>16</sup> utilisent jusqu'à mille milliards de poids (l'équivalent des synapses). La seconde sous-question est : « *Quelle architecture ?* » Avec l'apprentissage profond classique, les

neurones sont agencés en plusieurs couches successives interconnectées de sorte à ce que la sortie d'une couche soit l'entrée de la suivante. Plusieurs variantes ont vu le jour pour diverses applications : les réseaux à convolutions pour le traitement des images et des vidéos, les réseaux adverses pour la génération de nouvelles données (images, sons, vidéos...), les réseaux récurrents (avec des connexions en arrière pour un effet mémoire) pour le traitement de textes, et le transformer<sup>17</sup> au cœur de ChatGPT d'OpenAI, Gemini de Google ou Le Chat de Mistral AI. La troisième question concerne la manière d'apprendre. Les premiers algorithmes efficaces datent des années 1980, avec la rétro-propagation de l'erreur<sup>18</sup> combinée à la descente de gradient.

Quant à la question « *Combien de données ?* », à cause de la difficulté à récolter assez de données pour entraîner un réseau de neurones et du coût en termes de temps de calcul, elle doit attendre le début des années 2000 pour que des réponses y soient apportées. Les réseaux de neurones deviennent alors compétitifs grâce à l'Internet qui rend disponible de plus en plus de données et aux progrès en matière d'ordinateurs et de cartes graphiques (GPU pour *Graphics Processing Unit*; en français, « unité de traitement graphique »). Depuis lors, la montée en puissance des GPU et l'accroissement du nombre de données multimédia ont permis l'entraînement de réseaux de plus en plus complexes et puissants. Par exemple, en 2012, le modèle AlexNet nécessita environ  $470 \times 10^{15}$  FLOP pour son entraînement ; en 2023, Gemini Ultra de Google a requis environ  $50 \times 10^{24}$  FLOP. Désormais, on se pose la question de savoir où trouver assez de données pour les IA de demain...

<sup>11</sup> [https://en.wikipedia.org/wiki/AI\\_winter](https://en.wikipedia.org/wiki/AI_winter)

<sup>12</sup> Voir, par exemple, la présentation de Philip Boucher : « *Artificial intelligence: How does it work, why does it matter, and what can we do about it?* » <https://miniurl.be/r-5jig>

<sup>13</sup> « *What the frog's eye tells the frog's brain* » ou « ce que l'œil d'une grenouille dit à son cerveau », coécrit avec Jérôme Lettvin et Humberto Maturana, 1959, *Proceedings of the Institute of radio engineers*.

<sup>14</sup> George Cybenko, 1989, « Approximation by Superposition of a Sigmoidal Function », *Mathematics of Control, Signals, and Systems*, Springer-Verlag, New York Inc., pp 303-314.

<sup>15</sup> Jiawei Zhang, Basic Neural Units of the Brain : Neurons, Synapses and Action Potential. <https://arxiv.org/abs/1906.01703>

<sup>16</sup> <https://miniurl.be/r-5jii>

<sup>17</sup> <https://jalammar.github.io/illustrated-transformer/>

<sup>18</sup> <https://miniurl.be/r-5jij>

# Les IA génératives

Depuis l'avènement de ChatGPT-3 en juin 2020, les intelligences artificielles sont entrées en force dans le débat public. Cependant, l'utilisation aujourd'hui des termes « IA » par le grand public et les médias ne couvre en réalité qu'une petite partie des outils et méthodes que les scientifiques classent dans la famille des IA (Cf. encadré). Cet article présente ces IA génératives, leur historique et leur fonctionnement.

**Lilian Aveneau**

Membre de la commission de recherche du SNESUP, PU en informatique, laboratoire Xlim, Université de Poitiers

Une IA générative repose sur l'apprentissage profond (*deep learning*) et fait partie de la famille de l'IA dit d'apprentissage machine (*machine learning*), qui regroupe des méthodes qui simulent un processus d'apprentissage, suivi de l'étape de prise de décision basée sur le résultat de cet apprentissage. Lorsque nous parlons d'apprentissage profond, nous évoquons l'utilisation d'une grande masse de données (*big data*) pour l'étape d'apprentissage. Cette étape vise à identifier et à encoder différents schémas et relations qui existent dans les données en question. Ces modèles profonds s'appuient sur des neurones artificiels, paramétrés par un grand nombre de « poids » dont les valeurs résultent de l'apprentissage. Ces modèles sont capables de « comprendre » les requêtes des utilisateurs (prompts) exprimées en langage naturel (c'est-à-dire la langue de l'utilisateur). Ils peuvent répondre à ces requêtes dans la même langue pour les robots conversationnels (ChatGPT, Gemini...) ou dans une autre langue pour les outils de traduction (DeepL...). Ils peuvent aussi générer des images (Dall-E, Pixlr...), des vidéos (Synthesia, Descript...), ou encore du son et de la musique (AudioCraft, Beatoven...).

## HISTOIRE ET MODÈLES D'IA GÉNÉRATIVES

En 1964, ELIZA, le premier *chatbot*, a vu le jour au MIT. Développé sur

la base d'un algorithme relativement simple, il exploitait le contexte (le fil de la conversation) pour générer des réponses variées.

En 2004, Google intègre l'auto-complétion dans son moteur de recherche, en utilisant des chaînes de Markov<sup>1</sup>. Cette technique permet de prédire les mots suivants d'une requête en se basant sur les mots précédents. En 2013, apparaissent les auto-encodeurs variationnels ou VAE (*variational autoencoders*) qui utilisent deux réseaux de neurones profonds : le premier, dit encodeur, transforme une donnée en un vecteur de paramètres ; le second, ou décodeur, effectue l'opération inverse. Si un auto-encodeur de base sert, par exemple, à la traduction automatique, le VAE permet de générer de nouveaux contenus car son décodeur est entraîné à générer des variantes d'un même vecteur de paramètres.

En 2014, les réseaux antagonistes génératifs (GAN) et les premiers modèles de diffusion font leur apparition. Comme les VAE, les GAN utilisent deux réseaux de neurones : le générateur crée de nouveaux contenus (images, texte, musique...), que le discriminateur

évalue en essayant de déterminer si ces contenus sont réels (provenant de données réelles) ou synthétiques. Les modèles de diffusion partent d'un bruit gaussien et apprennent à générer des données réalistes à partir de données bruitées. Si cette phase d'entraînement est plus longue que pour les GAN, elle offre un meilleur contrôle de la sortie, en particulier pour les outils de génération d'images comme DALL-E.

En 2017, Google Brain lance le Transformer, aujourd'hui à la base de nombreux modèles dits de « fondation ». Il repose sur le mécanisme d'attention, qui capture les relations entre les mots (au sens large, le modèle s'étend aux images). Il s'appuie sur un couple encodeur-décodeur, qui traite une phrase en une seule fois, en parallèle, offrant une phase d'apprentissage rapide et plus efficace qu'avec les réseaux de neurones récurrents. Chaque mot de la séquence est représenté par son plongement (*word embedding*) dans un espace de dimension élevée, et doté d'une mesure de la similarité entre deux mots. Pour une phrase, le plongement de tous les mots forme une matrice (chaque mot étant une colonne), à laquelle est ajouté le positionnement des mots pour prendre en compte la structure syntaxique de la phrase ; cela per-

<sup>1</sup> Une chaîne de Markov est un modèle mathématique qui prédit l'état futur d'un système en se basant uniquement sur son état présent, et non sur son histoire antérieure (au présent).

met par, exemple<sup>2</sup>, de savoir que la phrase « *j'ai emmené mon chien chez le vétérinaire* » est proche de « *j'ai emmené mon chat chez le vétérinaire.* » Notons que GPT est un modèle à base de transformer (*Generative Pre-training Transformer*).

## FONCTIONNEMENT

La construction d'une IA générative se déroule en trois étapes<sup>3</sup>:

**L'étape d'apprentissage** consiste à entraîner un modèle dit de fondation (*foundation model*) sur un ensemble de données volumineux et diversifié, historiquement des données textuelles (on parle de grand modèle de langage, ou LLM). Ce modèle apprend à capturer les *patterns* et les relations statistiques qui existent dans les données.

**L'étape de réglage fin** (*fine-tuning*): le modèle de fondation est affiné pour une application précise, en utilisant un ensemble de données plus petit et plus ciblé, engendrant un modèle dit frontière.

**L'évaluation régulière** des performances du modèle génératif est régulièrement effectuée pour détecter les réponses obsolètes ou des défauts sur des questions « nouvelles ». Cela implique de générer des contenus à partir du modèle, puis de les évaluer en fonction de critères de qualité et de pertinence. Les résultats de l'évaluation peuvent ensuite être utilisés pour ajuster le modèle et améliorer ses performances.

## MODÈLE FONDATION

L'entraînement d'un modèle de fondation à partir de zéro est aujourd'hui

d'un coût incroyablement élevé. Par exemple, on estime que le coût du modèle Gemini (Google) s'élève à 191 millions de dollars<sup>4</sup>! Ce type de modèle est donc hors de portée pour un laboratoire de recherche isolé, un consortium public, voire même un État. Ce coût exorbitant s'explique en premier lieu par le nombre colossal de calculs nécessaires à l'entraînement et, en second lieu, par la taille des infrastructures informatiques requises et par la facture d'électricité associée.

Outre le coût financier, ces modèles posent un nouveau défi : la quantité faramineuse de données nécessaires à leur entraînement. Pour pallier ce phénomène, l'usage est de générer des données synthétiques à partir de données réelles. Cependant, cette approche pose un autre problème : les modèles deviennent parfois incapables de distinguer les données réelles des données synthétiques, ce qui peut entraîner une baisse brutale de leur qualité de prédiction<sup>5</sup>...

## MODÈLE FRONTIÈRE

L'étape de réglage fin utilise un modèle fondation pour construire un modèle frontière, c'est-à-dire un modèle dédié à une tâche ou un domaine particulier. En effet, si vous essayez Gemini par exemple, vous constaterez qu'il répond assez bien à beaucoup de vos questions, mais offre parfois des réponses complètement erronées (on parle d'hallucinations). Le réglage fin permet d'obtenir assez rapidement un modèle frontière mieux adapté à un domaine ou une tâche spécifique.

En pratique, il existe deux principales méthodes pour construire un modèle frontière :

**Réglage fin manuel**, qui implique la sélection et le traitement de données étiquetées à utiliser pour le ré-entraînement (par exemple, les questions et les réponses attendues pour le *chatbot* de la FSU dont tout le monde rêve). L'objectif est de minimiser l'erreur du modèle fondation sur ces nouvelles données.

**Renforcement basé sur des commentaires humains**, ou RLHF, qui étend l'apprentissage par renforcement (un modèle de récompense automatique) avec une rétroaction humaine. Les réponses du modèle sont évaluées ou corrigées par des humains, souvent des utilisateurs du modèle (parfois à leur insu) ou des personnes payées à la tâche dans des pays à bas salaires. Les évaluations humaines permettent de quantifier la récompense que le modèle reçoit pour chaque action, ce qui l'aide à apprendre à générer des réponses plus précises et pertinentes pour le domaine ou la tâche spécifique.

Il est important de noter que les modèles frontières peuvent présenter des biais supplémentaires liés aux données utilisées pour leur entraînement.

## ÉVALUATION RÉGULIÈRE

Pour éviter d'utiliser un modèle devenu obsolète, il faut évaluer régulièrement (par exemple, chaque semaine) les performances du modèle frontière, et effectuer de nouveaux réglages, par exemple en ajustant les réponses fournies aux questions rencontrées au cours de la semaine. Ainsi, un cycle régulier de génération (initié par les utilisateurs), d'évaluation et de réglages est établi.

Il convient de mentionner une méthode récente appelée « génération augmentée par récupération » (RAG pour *Retrieval Augmented Generation*), qui améliore significa-

<sup>2</sup> Exemple tiré de la documentation de Gemini : <https://ai.google.dev/gemini-api/docs/embeddings?hl=fr>  
Voir aussi « Gemini: A Family of Highly Capable Multimodal Models » : <https://arxiv.org/pdf/2312.11805>

<sup>3</sup> Confer par exemple : <https://www.ibm.com/fr-fr/topics/generative-ai>

<sup>4</sup> Confer « HAI - AI index report 2024 », Stanford, chapitre 1, page 37.  
<https://miniurl.be/r-5jii>

<sup>5</sup> Confer note 4, pages 26 à 29.

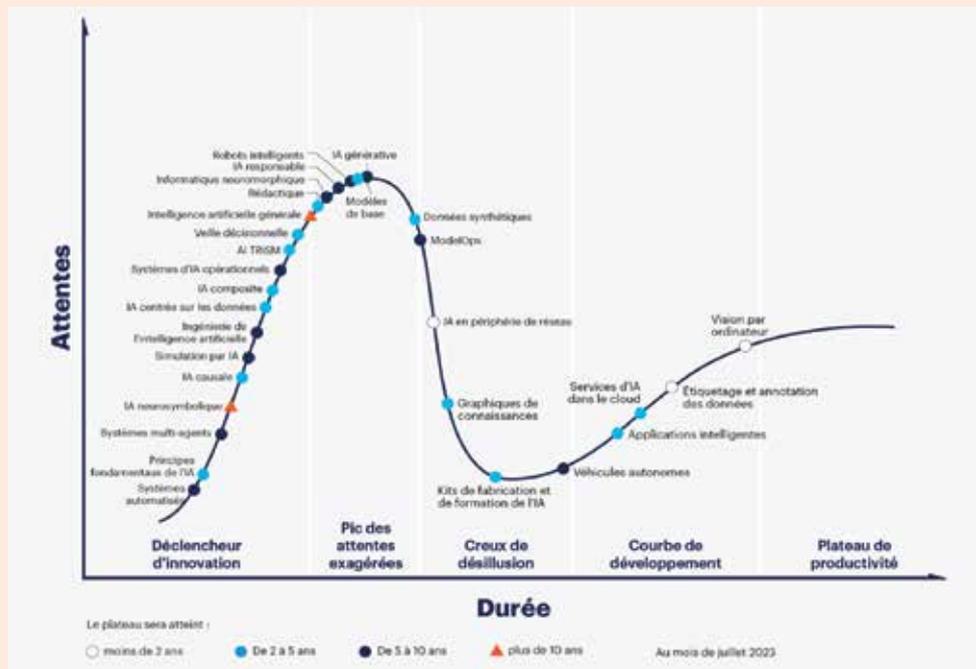
tivement les performances. L'idée consiste à modifier l'algorithme de prédiction (qui utilise un réseau neuronal profond) en ajoutant des étapes de vérification d'existence des éléments prédits dans une

base de données. Cela permet de supprimer d'éventuelles réponses infondées et de privilégier les réponses basées sur des faits, dont les sources peuvent être communiquées avec les réponses fournies.

De plus, cette approche permet de mettre à jour régulièrement les données et d'éviter ainsi des réponses obsolètes.

## Hype Cycle concernant l'intelligence artificielle<sup>1</sup>

La courbe « Hype Cycle » de Gartner<sup>2</sup>, permet de représenter l'adoption d'une technologie au fil du temps. Elle positionne les technologies en fonction des potentialités de diffusion dans le cadre d'une économie de marché (attentes des parties prenantes) et cela au fil du temps, c'est-à-dire au cours de leur développement, de leur mise au point et de leurs usages. Issue de recherches empiriques, la courbe s'organise en cinq phases. La phase de lancement d'une innovation, qui voit rapidement ses différents champs d'application potentiels croître jusqu'à atteindre la deuxième phase, celle du « pic des attentes exagérées ».



Source : © 2023 Gartner, Inc. et/ou ses sociétés affiliées. Tous droits réservés. 2079794.

A ce moment, notamment les investisseurs et les médias prêtent à la technologie innovante des usages et des performances auxquels elle ne pourra pas répondre par la suite. Vient alors la troisième phase du « creux des désillusion » (ou vallée de la mort) au cours de laquelle la technologie est dénigrée voire délaissée. Puis la quatrième phase correspond à son développement pour des usages adaptés qui voit à nouveau croître les attentes à son sujet, mais dans des proportions plus mesurées que lors de la deuxième phase et de manière plus progressive. Enfin, la cinquième phase de la courbe correspond au plateau de productivité, c'est-à-dire à la période au cours de laquelle la technologie se perfectionne lentement pour optimiser ses usages et leurs performances.

En juillet 2023, les analystes de la société Gartner ont montré que plusieurs applications de l'IA sont d'ores et déjà en phase de développement et de productivité comme la vision par ordinateur, les services IA dans le *cloud* (qui fournissent des outils de génération de modèles d'IA, des interfaces de programmation d'application (API pour *application programming interface*) pour des services prêts à l'emploi ou le véhicule autonome.

Les kits de fabrication ou de formation de l'IA (kits de développement de logiciels qui aident les ingénieurs à intégrer l'IA dans les applications informatiques) n'ont, eux, pas trouver de champs d'application et les IA génératives ou l'intelligence artificielle générales sont dans le pic des attentes exagérées car les investisseurs, les pouvoirs publics, les journalistes ou les utilisateurs leur prêtent des champs d'application et des perspectives de développement très vastes auxquels probablement, elles ne sauront pas répondre ou pour lesquelles elles ne seront pas performantes. L'avenir nous le dira.

Hervé Christofol.

<sup>1</sup> <https://miniurl.be/r-5k1y>

<sup>2</sup> Gartner est une société de conseil privée. Son outil "Hype Cycle" fait autorité dans le positionnement de la maturité et de l'adoption des nouvelles technologies.

# Reproductibilité et erreurs dans les IA

Les algorithmes construits par apprentissage utilisent des ensembles de données très importants, évolutifs et au sein desquels les optimums de réponses varient d'un entraînement à l'autre. Aussi, la reproductibilité des connaissances issues des IA, des structures des algorithmes et des jeux de données est-elle problématique. Cela questionne notamment les procédures scientifiques de vérification des connaissances produites.

**Lilian Aveneau**

Membre de la commission de recherche du SNESUP, PU en informatique, laboratoire Xlim, Université de Poitiers

La reproductibilité est la capacité à reproduire une expérience ou une étude de manière à obtenir des résultats identiques en utilisant la même méthodologie. Il s'agit d'un principe fondamental en sciences, qui garantit que les mêmes conclusions peuvent être obtenues par des chercheurs indépendants en suivant la même méthodologie et en utilisant des données similaires. Cette notion est particulièrement cruciale dans les sciences expérimentales : imaginez les répercussions si les résultats d'une expérience sur un nouveau médicament ne pouvaient pas être reproduits par d'autres laboratoires !

Cet aspect est également essentiel pour le domaine de l'intelligence artificielle (IA). Cependant, on estime que seulement 5 % des chercheurs partagent le code source de leurs modèles et que moins de 33 % partagent les données utilisées<sup>1</sup>. En d'autres termes, moins d'un tiers des recherches en IA sont facilement reproductibles et donc vérifiables.

C'est ce que l'on appelle communément la crise de la reproductibilité en IA.

## QU'EST-CE QUE LA REPRODUCTIBILITÉ EN IA ?

Dans le contexte de l'IA, la reproductibilité consiste à obtenir des résultats cohérents ou similaires en utilisant les mêmes éléments, à savoir :

- l'ensemble des données utilisées par l'algorithme d'IA lors de son entraînement ;
- le même algorithme d'IA, incluant le type de modèle, ses paramètres avant apprentissage (poids et biais), ses « hyperparamètres » (taux d'apprentissage, nombre d'époques, etc.), et, idéalement, son code source ;
- les détails de l'environnement d'exécution, englobant le logiciel (système d'exploitation, bibliothèques logicielles) et le matériel (processeurs, cartes graphiques, mémoire) utilisés pour exécuter l'algorithme (autant pour son apprentissage que pour son utilisation).

Ainsi, dans les publications est-il essentiel de bien documenter chacun de ces éléments afin de garantir la reproductibilité des systèmes d'IA. Cela permet à d'autres chercheurs de reproduire les analyses et de vérifier les résultats obtenus. Cependant, et même en disposant de tous ces éléments, il peut rester difficile de reproduire un modèle donné, notamment en raison de l'aléatoire apparaissant dans cer-

taines méthodes, principalement avec l'apprentissage profond (DL pour *Deep learning*) et l'apprentissage par renforcement (RL pour *Reinforcement Learning*).

Les méthodes RL reposent sur un mécanisme de récompense pour valoriser les réponses correctes par rapport à une vérité terrain. Les environnements dans lesquels elles évoluent sont souvent dynamiques et stochastiques, ce qui signifie qu'ils changent de manière imprévisible et donc que les récompenses ne sont pas toujours déterministes. Cela rend difficile la reproduction exacte des résultats, car de petites variations dans l'environnement peuvent entraîner des différences significatives dans le comportement des méthodes RL.

Pire, avec les algorithmes par renforcement basés sur un retour humain (RLHF pour *Reinforcement learning from Human Feedback*), il devient quasiment impossible de pouvoir documenter correctement et donc de pouvoir répliquer l'ensemble des interactions et récompenses fournies lors de l'entraînement.

## L'ALÉATOIRE EN IA

Notons un effet pervers des ordinateurs : si tout le monde sait qu'ils sont très rapides pour calculer, peu savent

<sup>1</sup> Voir, par exemple, *Reproducible AI: Why it Matters & How to Improve it in 2024?*  
<https://research.aimultiple.com/reproducible-ai/>

En pratique, un ordinateur n'utilise que des valeurs discrètes, toujours sur un nombre fixe d'octets : par exemple, avec quatre octets, il n'aura que trente-deux bits pour représenter tous les réels (avec des nombres à virgule flottante) soit un peu plus de quatre millions de valeurs différentes. On est loin du compte, sachant que l'espace des réels est infini, y compris entre deux nombres quelconques : quoi qu'en dise Zénon d'Élée<sup>1</sup>, il y a, par exemple, une infinité de nombres réels entre 1,1 et 1,2. Par conséquent, l'addition de deux nombres en virgule flottante engendre presque toujours une petite erreur. Il arrive<sup>2</sup> même parfois que  $A+B$  donne  $A$  comme résultat ! Et cette erreur varie selon l'ordre des additions :  $A+B+C$  n'est généralement pas égal<sup>3</sup> à  $C+B+A$  ! Pour les plus curieux, je vous invite à étudier l'algorithme de Kahan-Babuška-Neumaier<sup>4</sup>.

<sup>1</sup> Le philosophe grec présocratique Zénon d'Élée a formulé le paradoxe d'Achille et de la tortue. Un jour, le grec Achille a disputé une course à pied avec une tortue. Il lui a accordé une avance de cent mètres. L'argument exposé par Zénon est qu'Achille ne pourra pas rattraper la tortue parce qu'il ne pourra jamais annuler l'avance de l'animal. <https://miniurl.be/r-5k8l>

<sup>2</sup> En utilisant des nombres 32 bits, alors  $10^6+10^{-2}$  donne  $10^6$ ...

<sup>3</sup> Testez avec  $A=10^6$ ,  $B=C=1/3$ , alors  $A+B+C=1000000,625$  et  $C+B+A=1000000,688$ , soit une différence de 0,0625.

<sup>4</sup> <https://miniurl.be/r-5k8k>

qu'ils ne savent pas bien calculer<sup>2</sup>, surtout avec les « nombres réels » intrinsèques aux modèles mathématiques de l'IA (voir l'explication technique donnée dans l'encadré).

Par ailleurs, les méthodes DL reposent, par essence, sur des outils non déterministes pour augmenter

<sup>2</sup> Un exemple connu liés aux problèmes de calcul en entier : [https://fr.wikipedia.org/wiki/Vol\\_501\\_d'Ariane\\_5](https://fr.wikipedia.org/wiki/Vol_501_d'Ariane_5)

leur efficacité et leur vitesse d'apprentissage<sup>3</sup>. Ainsi, avec les mêmes entrées (données, paramètres), un algorithme peut produire des modèles complètement différents en termes de précision et de vitesse de convergence. À cela, il faut ajouter les problèmes liés aux bibliothèques sous-jacentes (comme l'outil d'apprentissage automatique *open source* développé par Google TensorFlow) qui engendrent aussi une variance en raison de leur parallélisation (générant un ordre variable des additions, donc des erreurs de calculs et, finalement, de poids !), des optimisations et, encore une fois, des imprécisions numériques.

Ainsi, une bonne pratique en IA, comme en informatique en général, est de valider les approches sur plusieurs expériences différentes. Mais qui fait cela ?

## AMÉLIORER LA REPRODUCTIBILITÉ EN IA

La reproductibilité est cruciale en IA, autant pour la recherche que pour ses applications. Elle permet de garantir la fiabilité et la robustesse des résultats obtenus. Quelle est la pertinence d'un résultat qui ne serait pas reproductible même par son auteur ?

Pour la recherche en IA, le progrès scientifique dépend de la capacité des chercheurs indépendants à examiner et à reproduire les résultats d'une étude afin de s'en servir pour construire de nouveaux modèles. L'apprentissage automatique ne peut être amélioré ou appliqué à d'autres domaines si ses composants essentiels ne sont pas documentés à des fins de reproductibilité.

Pour les applications utilisant l'IA, améliorer la reproductibilité doit

<sup>3</sup> Hung Viet Pham, Shangshu Qian, Jiannan Wang, Thibaud Lutellier, Jonathan Rosenthal, Lin Tan, Yaoliang Yu, and Nachippan Nagappan, 2020, « Problems and opportunities in training deep learning software systems: An analysis of variance », *Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering*, pages 771–783.

permettre de construire des modèles moins sujets aux erreurs, d'augmenter la fiabilité et la prévisibilité des résultats. Imaginez un modèle de prévision de la date de la prochaine dissolution de l'Assemblée nationale qui ne soit pas fiable ? Quel bazar ce serait !

Si le problème de la reproductibilité se pose en IA, comme en sciences en général, il faut aussi en analyser les sources. L'un des problèmes principaux, outre les aléas liés aux méthodes stochastiques, concerne le temps nécessaire pour les auteurs à la mise en forme et à la communication de toutes les données<sup>4</sup>. Bien que cela soit en contradiction avec l'évolution des pratiques et recommandations<sup>5</sup>, force est de reconnaître que la course à la publication limite de fait la reproductibilité en sciences en général et en IA en particulier. Mais cela touche plus à la répliquabilité (où il s'agit d'utiliser tous les éléments nécessaires à la reproduction d'une expérience) qu'à la reproductibilité en tant que telle, qui nécessite généralement de coder certains algorithmes, voire de définir une base de données pour l'entraînement.

Concernant l'aléatoire dans les outils, le problème de la reproductibilité est plutôt lié à la façon de valider les méthodes avant publication. Ainsi, si un modèle donne des résultats différents en fonction de son entraînement, alors se pose la question de la méthodologie globale. Certains travaux<sup>6</sup> visent actuellement à réduire cette incertitude en proposant des cadres de travail complets permettant une meilleure validation des solutions proposées, avant publication. Mais, là encore, la course à la publication n'est-elle pas un obstacle aux démarches de validation ?

<sup>4</sup> Odd Erik Gundersen, 2019, « Standing on the Feet of Giants – Reproducibility in AI », *AI Magazine*, 40: 9-23. <https://doi.org/10.1609/aimag.v40i4.5185>

<sup>5</sup> OECD, 2021, Recommendation of the Council concerning Access to Research Data from Public Funding. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>

<sup>6</sup> <https://miniurl.be/r-5k8k>

# AI Act : une première réglementation internationale sur l'IA

Après d'intenses négociations, la première législation sur l'intelligence artificielle (IA) au monde a définitivement été adoptée par l'Union européenne (UE) le 21 mai 2024. Fondée sur une approche « par les risques », elle est plus un passage obligé pour le développement du marché qu'une loi sur une « IA digne de confiance ».

Après les eurodéputés qui l'ont adopté le 13 mars 2024, les États de l'Union européenne (UE) ont approuvé le règlement européen sur l'intelligence artificielle (IA) le 21 mai 2024. Pour sa part, le Conseil de l'Europe a adopté, le 17 mai 2024, un traité international visant à garantir une IA respectueuse des droits fondamentaux<sup>1</sup>. Le règlement dans son intégralité n'entrera pleinement en application qu'en mai 2026 mais, d'ici là, certaines mesures seront progressivement appliquées : dès à présent, un délai de six mois entre en vigueur pour l'interdiction des applications d'IA à risque inacceptable.

Il y a quelques années, l'Europe décidait de se doter d'un règlement sur l'intelligence artificielle (IA) afin de la promouvoir, tout en encadrant son usage et sa commercialisation. Par ailleurs, comme de nombreux gouvernements (États-Unis, Angleterre, Canada...) commençaient à travailler à leur propre régulation de l'IA, il s'agissait aussi pour elle d'être pionnière en la matière et de tenter ainsi d'influer sur les régulations à venir.

Comme pour toutes les lois européennes, le règlement européen sur l'intelligence artificielle (*AI Act*) est le fruit d'une initiative de la Commission européenne présenté pour la première fois en avril 2021. Il faisait suite à la publication de la « stratégie européenne en matière d'intelligence artificielle » d'avril 2018, suivie du « plan coordonné sur l'IA » de décembre 2018, puis des « lignes directrices pour une IA digne de confiance » (2019) et, enfin, au « livre blanc de la Commission sur l'IA » en 2020.

À chaque fois, ces publications présentaient deux volets : un premier volet concernant le développement du marché intérieur – il s'agit de « maximiser les retombées des investissements au niveau de l'Union européenne (UE) et des États membres, d'encourager les synergies et la coopération dans l'ensemble de l'UE, d'échanger les meilleures pratiques et de définir conjointement la voie à suivre pour faire en sorte que l'UE dans son ensemble puisse rivaliser avec la concurrence internationale<sup>2</sup> » – ; un second volet relatif à une « IA digne de confiance ». À chaque fois, ces déclarations s'accompagnaient

d'annonces de financements, notamment par le « programme pour une Europe numérique » et « Horizon Europe ».

Cependant, pour l'ensemble des déclarations, on a bien vu que l'« IA digne de confiance » n'était qu'un passage obligé pour le développement du marché. Ainsi était-il déclaré dans le plan coordonné sur l'IA de 2018 que « pour susciter la confiance, qui est nécessaire pour que les sociétés acceptent et utilisent l'IA, le plan coordonné doit permettre de développer une technologie respectueuse des droits fondamentaux et des règles éthiques<sup>3</sup> ». C'est donc sans surprise que l'article 1<sup>er</sup> du règlement sur l'IA, adopté le 14 mai 2024, commence par la phrase suivante : « L'objectif du présent règlement est d'améliorer le fonctionnement du marché intérieur et de promouvoir l'adoption d'une intelligence artificielle (IA) axée sur l'humain et digne de confiance ».

## AU DÉPART, UN INSTRUMENT DE PROTECTION DES DROITS ET LIBERTÉS

Après sa présentation en 2021, le règlement européen sur l'IA a

### Marc Champesme

Maître de conférences en informatique  
Membre du Laboratoire d'Informatique de Paris-Nord (LIPN)  
Membre de la commission administrative du SNESUP-FSU

<sup>1</sup> Intelligence artificielle : le cadre juridique européen de l'IA en six questions.  
<https://miniurl.be/r-5jvx>

<sup>2</sup> Plan coordonné sur l'IA (2018).  
<https://miniurl.be/r-5jvy>

<sup>3</sup> Ibid.

été adopté à une très large majorité par le Parlement européen en juin 2023, en incluant des avancées significatives par rapport au projet initial de la Commission européenne, notamment dans le domaine de l'éducation, de la protection des droits des travailleurs, de la protection de l'environnement et des droits humains. Cependant, la phase finale du processus législatif – le « trilogue » – consistant à trouver un consensus entre la version initiale du texte écrite par la Commission européenne et les versions amendées par le Parlement européen d'une part et le Conseil de l'UE d'autre part s'est conclue, en mars 2024, par des reculs importants, notamment sur le respect des droits humains<sup>4</sup>.

C'est ainsi que, contrairement au texte adopté en première lecture par le Parlement en juin 2023, la version finale du règlement – adopté par le Parlement en deuxième lecture en avril 2024, puis par le conseil de l'UE le 14 mai – exclut totalement de son champ d'application les systèmes d'IA « *mis en service ou utilisés [...] exclusivement à des fins militaires, de défense ou de sécurité nationale* ». De plus, de nombreuses dérogations ont été introduites, notamment pour les utilisations à des fins répressives ou dans les domaines de la migration, des contrôles aux frontières ou de l'asile. Ces évolutions du texte ont été obtenues notamment grâce à un intense *lobbying* des gouvernements français et allemand, fortement influencés par leurs « champions »

Sont également exclus du champ d'application les systèmes développés uniquement à des fins de recherche et développement scientifiques.

de l'IA respectifs : Mistral AI pour la France et Aleph Alpha pour l'Allemagne<sup>5</sup>. Ironiquement, alors que l'argument de la souveraineté européenne face aux GAFAM avait été mis en avant par ces *startups* et leurs gouvernements, à peine quelques mois après la fin du trilogue, Mistral AI annonçait avoir conclu un accord avec Microsoft/OpenAI ! Point de mystère ici, le développement du marché passe avant le respect des droits humains.

### UNE APPROCHE « PAR LES RISQUES »

Le principe général du règlement européen sur l'IA est de distinguer quatre catégories particulières de systèmes d'IA en fonction des risques induits par leur usage et d'établir les règles spécifiques à respecter par les fournisseurs et déployeurs de ces systèmes. Cette régulation par les risques est plus flexible qu'une approche fondée sur les droits (comme par exemple, le règlement général sur la protection des données ou RGPD).

Les quatre catégories de systèmes d'IA sont : les systèmes présentant un risque inacceptable qui doivent être interdits ; les systèmes présentant un risque élevé et nécessitant des mesures de contrôle adaptées à ces risques ; les « modèles d'IA à usage général » et, parmi eux, ceux présentant un « risque systémique », ces systèmes font l'objet de mesures de contrôle spécifiques, qui peuvent s'ajouter aux mesures à prendre si le système est également à risque élevé ; et les autres systèmes qui ne font pas l'objet de procédure de contrôle particulières.

La seule mesure de ce règlement s'appliquant à l'ensemble des systèmes autorisés (donc y compris ceux classés sans risque) est une simple invitation à « [prendre] des

*mesures pour garantir, dans toute la mesure du possible, un niveau suffisant de maîtrise de l'IA pour leur personnel et les autres personnes s'occupant du fonctionnement et de l'utilisation des systèmes d'IA*<sup>6</sup> », ainsi qu'un encouragement à « l'élaboration de codes de conduite, [...] destinés à favoriser l'application volontaire » des exigences applicables aux systèmes à haut risque. Il s'agit, là aussi, d'une régression importante par rapport au texte qui prévoyait une longue liste de principes généraux à respecter parmi lesquelles : transparence, traçabilité, explicabilité, protection de l'environnement et attention portée aux effets à long terme sur l'individu, la société et la démocratie.

On entend par « système d'IA », un système automatisé qui est conçu pour « *fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels.* » (Article 3)

« On entend par « modèle d'IA à usage général », un modèle d'IA, [...], qui présente une généralité significative et est capable d'exécuter de manière compétente un large éventail de tâches distinctes. » (Article 3)

### SYSTÈMES INTERDITS

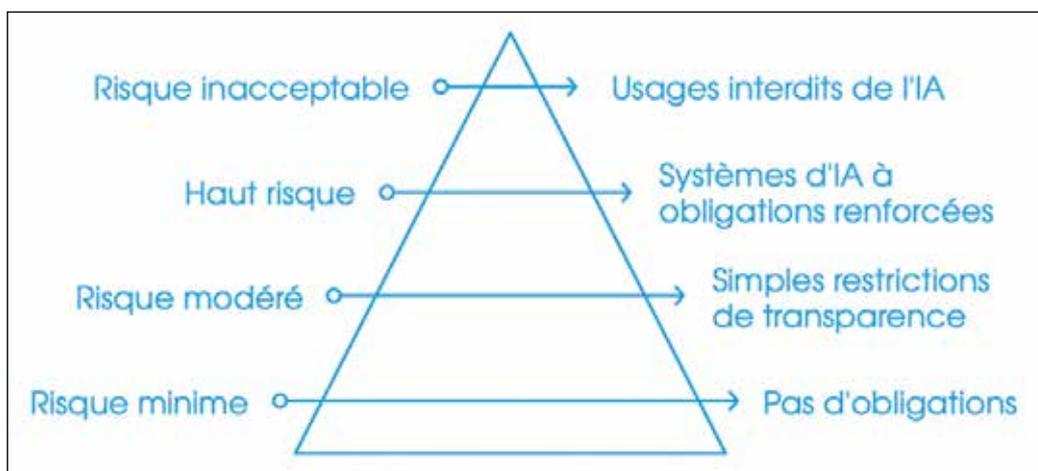
Dans la catégorie des systèmes d'IA interdits, le règlement liste les usages suivants :

- les systèmes utilisant des techniques subliminales ou manipula-

<sup>4</sup> Voir notamment l'analyse de la Quadrature du Net. <https://miniurl.be/r-5jw0>

<sup>5</sup> <https://miniurl.be/r-5jw1>

<sup>6</sup> Article 4.



La pyramide des risques de la loi sur l'IA.  
Source : d'après Commission européenne.

sion européenne. Plusieurs grands domaines sont distingués, qui concernent notamment l'éducation et les relations de travail.

Concernant l'éducation, les usages considérés à haut risque sont :

- les « systèmes d'IA destinés à être utilisés pour déterminer l'accès, l'admission ou l'affectation de personnes physiques aux établissements d'enseignement et de formation professionnelle, à tous les niveaux ; »

- les « systèmes d'IA destinés à être utilisés pour évaluer les acquis d'apprentissage, y compris lorsque ceux-ci sont utilisés pour orienter le processus d'apprentissage de personnes physiques dans les établissements d'enseignement et de formation professionnelle, à tous les niveaux ; »

- les « systèmes d'IA destinés à être utilisés pour évaluer le niveau d'enseignement approprié qu'une personne recevra ou sera en mesure d'atteindre, dans le contexte ou au sein d'établissements d'enseignement et de formation professionnelle à tous les niveaux ; »

- les « systèmes d'IA destinés à être utilisés pour surveiller et détecter des comportements interdits chez les étudiants lors d'examens dans le contexte d'établissements d'enseignement et de formation ou en leur sein à tous les niveaux ».

Dans le domaine de l'« emploi et de la gestion de la main d'œuvre » les usages suivant sont également considérés à haut risque :

- les « systèmes d'IA destinés à être utilisés pour le recrutement ou la sélection de personnes physiques, en particulier pour publier des offres d'emploi ciblées, analyser et filtrer les candidatures et évaluer les candidats » ;

- les « systèmes d'IA destinés à être utilisés pour prendre des décisions influant sur les conditions des relations professionnelles, la promotion ou le licenciement dans le cadre de relations professionnelles contractuelles, pour attribuer

trices pour altérer substantiellement le comportement des personnes en portant considérablement atteinte à leur capacité à prendre une décision éclairée, amenant ainsi la personne à prendre une décision qu'elle n'aurait pas prise autrement, d'une manière susceptible de causer un préjudice important ;

- les systèmes exploitant les vulnérabilités d'une personne pour altérer substantiellement son comportement, s'ils sont susceptibles de causer un préjudice important à cette personne ;

- les systèmes de notation sociale, seulement s'ils donnent lieu à un traitement défavorable disproportionné par rapport au comportement social des personnes concernées ou bien s'ils opèrent avec des conséquences défavorables dans un contexte différent de celui de la collecte des données ;

- les systèmes de police prédictive, sauf s'il s'agit uniquement d'étayer l'évaluation humaine de l'implication d'une personne dans une activité criminelle ;

- les systèmes d'IA qui créent ou développent des bases de données de reconnaissance faciale par le moisonnage non ciblé d'images faciales provenant de l'internet ou de la vidéosurveillance ;

- les systèmes d'IA utilisés pour inférer les émotions d'une personne physique sur le lieu de travail et dans les établissements d'enseignement, sauf lorsque le système d'IA est uti-

lisé pour des raisons médicales ou de sécurité ;

- les systèmes de catégorisation biométriques, sauf dans le domaine répressif ;

- l'utilisation de systèmes d'identification biométrique à distance « en temps réel » dans des espaces accessibles au public à des fins répressives, sauf pour la recherche ciblée de victimes spécifiques, pour la prévention d'une menace, pour la sécurité de personnes ou d'une menace d'attaque terroriste ou pour la localisation ou l'identification d'une personne soupçonnée d'avoir commis une infraction.

Comme on peut le voir dans la liste précédente, des usages pourtant considérés comme présentant un risque inacceptable sont systématiquement associés à des exceptions ou des conditions d'applications très restrictives qui laissent craindre qu'une partie importante des systèmes entrant dans cette catégorie puissent être exemptés, notamment dans le domaine répressif, avec les risques inhérents concernant le respect des droits humains.

## SYSTÈMES « À HAUTS RISQUES »

Les systèmes « à hauts risques » sont les systèmes d'IA dont le champ d'application figure dans une annexe III qui, contrairement à la liste des usages interdits, est révisable par la Commis-

des tâches sur la base du comportement individuel, de traits de personnalité ou de caractéristiques personnelles ou pour suivre et évaluer les performances et le comportement de personnes dans le cadre de telles relations ».

Les autres utilisations considérées comme à risque élevé concernent les domaines suivants:

- « accès et droit aux services privés essentiels et aux services publics et prestations sociales essentiels » avec, notamment :
  - les « systèmes d'IA destinés à être utilisés par les autorités publiques ou en leur nom pour évaluer l'éligibilité des personnes physiques aux prestations et services d'aide sociale essentiels, y compris les services de soins de santé, ainsi que pour octroyer, réduire, révoquer ou récupérer ces prestations et services » ;
  - les « systèmes d'IA destinés à être utilisés pour l'évaluation des risques et la tarification en ce qui concerne les personnes physiques en matière d'assurance-vie et d'assurance maladie ».
- certaines utilisations par les « autorités répressives » ou pour la gestion des migrations ;
- certains usages dans l'« administration de la justice et [les] processus démocratiques » dont les « systèmes d'IA destinés à être utilisés pour influencer le résultat d'une élection ou d'un référendum ou le comportement électoral de personnes physiques dans l'exercice de leur vote lors d'élections ou de référendums ».

Cependant, outre le fait que la version finale du texte a substantiellement limité la portée des usages décrits dans cette annexe par rapport au texte initialement adopté par le Parlement, le règlement ajoute plusieurs possibilités d'exemption de cette classification « à haut risque ». Il permet ainsi à n'importe quel fournisseur de système de s'en exonérer s'il peut fournir une auto-évaluation montrant que son système ne présente pas de risque élevé !

Parmi les obligations auxquelles sont soumis les systèmes à haut risque



avant mise sur le marché ou mise en service on retiendra surtout :

- le respect d'exigences en matière de données et de gouvernance des données, même si ces exigences ne concernent pas le respect de la propriété intellectuelle ;
- une conception du système permettant un contrôle humain lors de son utilisation ;
- l'obligation pour les employeurs d'informer les représentants des travailleurs et les travailleurs concernés qu'ils seront soumis à l'utilisation du système d'IA à haut risque ;
- pour les déployeurs qui sont des organismes de droit public ou des entités privées fournissant des services publics, l'obligation d'effectuer avant le déploiement une analyse d'impact sur les droits fondamentaux ;
- l'enregistrement du système dans la base de données publique de l'UE spécifiquement conçue à cet effet.

De plus, les systèmes conformes au règlement doivent apposer le marquage CE attestant du respect de ces obligations. Cependant, plusieurs ONG de défense des droits humains ont alerté sur le risque que ce marquage donne un faux sentiment de sécurité aux utilisateurs de ces systèmes habitués au marquage CE des produits de consommation courante (jouets, appareils électro-ménagers,...).

Pour les modèles d'IA à usage général quelques obligations supplémentaires sont prévues, notamment en ce qui concerne la qualité des données d'entraînement et, plus spécifiquement, les données protégées par le droit d'auteur.

Autre point important à signaler : il est prévu que le respect des obligations énoncées soit évalué en fonction de normes pour lesquelles la Commission européenne doit effectuer des demandes aux organisations européennes de normalisation. Une partie importante des conditions d'application de ce règlement dépendra donc du contenu de normes qui restent à élaborer.

Au final, malgré les importantes limitations apportées durant sa finalisation et les incertitudes sur les conditions de sa mise en œuvre, ce règlement est une des toutes premières tentatives de régulation de l'IA au niveau international. Alors que les systèmes d'IA envahissent de plus en plus notre environnement professionnel et notre vie quotidienne, il ouvre, notamment pour les organisations de défense des droits humains et les syndicats, des capacités d'intervention au sujet de nombreux usages problématiques de l'IA, à nos syndicats de s'en saisir !

# Enseignants :

## les défis du droit d'auteur à l'ère des IA génératives

L'utilisation croissante d'outils d'apprentissage numériques soulève de nombreuses inquiétudes chez les enseignants concernant les droits sur leurs propres programmes de cours, matériels pédagogiques, présentations et autres créations. Alors que normalement le créateur détient les droits d'auteur sur son œuvre, un employeur peut acquérir les droits sur l'œuvre d'un employé au moyen d'un accord distinct.

**Kari Kivinen<sup>1</sup>**

Expert de l'éducation  
European Union Intellectual  
Property Office (EUIPO),  
Observatory

Le droit d'auteur, droit de propriété intellectuelle lié aux œuvres littéraires et artistiques, s'applique également aux créations des enseignants. Il s'applique automatiquement dès la création de contenu et ne nécessite aucune formalité. Cependant, deux conditions essentielles subsistent :

- l'originalité : l'œuvre doit être une création intellectuelle propre à son auteur, se manifestant par ses choix libres et créatifs, avec une touche personnelle ;
- l'expression : l'œuvre doit être identifiable avec suffisamment de précision et d'objectivité.

Une autre préoccupation commune aux enseignants est l'incertitude juridique concernant l'utilisation de contenus numériques protégés par le droit d'auteur pour améliorer leur pratique. L'article 5 de la directive sur le droit d'auteur dans le marché unique numérique<sup>2</sup> (2019) a rendu obligatoire, dans tous les pays de l'Union européenne (UE), une exception pour l'utilisation d'œuvres numériques uniquement à des fins d'illustration dans le cadre de l'enseignement, sous réserve de certaines conditions :

- fins non commerciales ;
- utilisation sous la responsabilité d'un établissement d'enseignement, dans ses locaux (ex. tableau blanc) ou d'autres lieux, ou *via* un réseau électronique sécurisé accessible uniquement aux étudiants/enseignants (intranet, environnement

d'apprentissage virtuel) ;

- citation correcte de la source.

Cette disposition vise à offrir aux enseignants et aux étudiants plus de sécurité juridique pour l'utilisation de contenus protégés dans leurs activités d'enseignement numérique. Toutefois, en raison de la nature territoriale des lois sur le droit d'auteur, cette exception est mise en œuvre de manière légèrement différente dans les États membres de l'UE.

La foire aux questions (FAQ) de l'Office de l'Union européenne pour la propriété intellectuelle (EUIPO) sur le droit d'auteur pour les enseignants<sup>3</sup> aide ceux-ci et les étudiants de l'UE à trouver des informations sur l'utilisation de contenus protégés par le droit d'auteur dans le cadre de l'éducation et de la formation, notamment en ligne. Elle informe également les enseignants et les étudiants des possibilités que la protection du droit d'auteur peut leur offrir en tant que créateurs d'œuvres dans un contexte éducatif.

Les réponses fournies dans la FAQ valent pour tous les États membres de l'UE. Elles sont disponibles en anglais et dans au moins une des langues officielles de l'UE pour chaque État membre.

### IA GÉNÉRATIVES : UN DÉFI POUR L'ÉDUCATION

Comme d'autres technologies émergentes, les IA génératives offrent des solutions

dans un certain nombre de domaines, conçues pour soutenir et améliorer les capacités humaines.

Compte tenu de l'évolution rapide des technologies de l'IA, les prestataires d'enseignement et de formation doivent réfléchir attentivement au moment et à la façon dont ces systèmes peuvent être utilisés de manière responsable et dans l'intérêt public – en particulier lorsqu'ils impliquent le traitement de données personnelles ou de contenus protégés par le droit d'auteur.

Du point de vue d'un prestataire d'éducation, il est important de promouvoir l'acquisition, le développement et le déploiement de systèmes d'IA fiables et de garantir que tous les utilisateurs – enseignants et étudiants – soient correctement formés à l'utilisation de l'IA et à la manière de maîtriser son fonctionnement et son utilisation.

La nouvelle loi européenne sur l'IA<sup>4</sup> exigera la divulgation du fait que l'œuvre a été réalisée par l'IA ainsi qu'un résumé détaillé du contenu utilisé pour l'entraîner. Les responsables des systèmes d'IA devront également s'engager à respecter la législation européenne sur les droits d'auteur.

L'acquisition et l'utilisation de services d'IA soulèvent en effet de nombreuses questions déontologiques, notamment des questions de droit d'auteur.

<sup>1</sup> Les vues exprimées dans le présent article sont celles de son auteur et ne reflètent pas nécessairement la position de l'EUIPO.

<sup>2</sup> <https://eur-lex.europa.eu/eli/dir/2019/790/oj?locale=fr>

<sup>3</sup> <https://miniurl.be/r-5kdl>

<sup>4</sup> <https://artificialintelligenceact.eu/fr/l-acte/>

## IA GÉNÉRATIVES : UN DÉFI POUR LES DROITS D'AUTEUR

Les IA génératives, capables de créer des textes, des images et de la musique avec ou sans assistance humaine, recèlent un immense potentiel. Mais derrière ces prouesses créatrices se cache un grand nombre d'incertitudes juridiques quant à l'application du droit d'auteur. L'entraînement des modèles d'IA implique souvent de vastes quantités de textes, de codes, d'images et de données, dont une grande partie est potentiellement protégée par le droit d'auteur et extraite d'internet sans autorisation.

Comme nous savons rarement quelles données les modèles d'IA ingèrent, il est difficile d'évaluer une éventuelle violation du droit d'auteur. Les poursuites se multiplient déjà, avec un nombre croissant d'artistes et d'éditeurs contestant l'utilisation de leurs œuvres dans le cadre de l'entraînement des modèles d'IA. Le plus médiatisé de ces procès est celui du *New York Times* contre OpenAI. En guise de défense, OpenAI a donné une réponse intéressante<sup>5</sup> :

*« Parce que le droit d'auteur couvre aujourd'hui pratiquement tous les types d'expression humaine – y compris les articles de blog, les photographies, les messages de forum, les fragments de code logiciel et les documents des autorités publiques – il serait impossible de former les principaux modèles d'IA actuels sans utiliser de contenus protégés par le droit d'auteur. »*

Il existe des solutions techniques pour empêcher l'utilisation de contenus protégés par le droit d'auteur sur le Web. Ainsi, les médias bloquent de plus en plus les robots d'exploration de l'IA, ce qui a créé de nouvelles problématiques. Selon Wired<sup>6</sup>, 88 % des principaux médias d'information aux États-Unis bloquent les robots d'exploration du Web. L'institut Reuters<sup>7</sup> rapporte qu'en février 2024, près de la moitié (48 %) des principaux sites d'information dans dix pays sélectionnés bloquaient les robots d'exploration d'OpenAI.

<sup>5</sup> <https://miniurl.be/r-5kd2>

<sup>6</sup> <https://miniurl.be/r-5kd3>

<sup>7</sup> <https://miniurl.be/r-5kd4>

Il existe plusieurs façons de limiter cette utilisation, mais il faut tenir compte du fait que le blocage de tous les robots d'exploration peut empêcher les moteurs de recherche légitimes d'indexer le site et nuire à leur visibilité.

Les résultats des outils d'IA génératives sont générés à partir de simples commandes ou peuvent être utilisés pour aider les personnes dans leurs activités innovantes, créatives, de marque et autres activités d'information et de connaissances. Cela soulève inévitablement d'autres questions liées au droit d'auteur, à commencer par : à qui appartient la création de l'IA ? Le droit d'auteur traditionnel peut-il s'y appliquer ou avons-nous besoin de nouvelles solutions ? Actuellement, seules les œuvres créatives réalisées par des humains bénéficient de la protection du droit d'auteur.

Les créateurs craignent que les œuvres générées par l'IA ne copient leur style ou leurs idées. La nouvelle réglementation européenne sur l'IA va imposer des obligations de transparence. Elle exigera de dévoiler que le résultat produit a été réalisé par l'IA ainsi qu'une description des contenus utilisés pour l'entraîner. Cela devrait contribuer à protéger les créateurs et à rendre les choses plus claires pour tout le monde.

Les détails sur la manière d'empêcher les IA d'enfreindre les droits d'auteur, ainsi que les sanctions qui devraient s'appliquer et par qui, restent à définir.

### PROBLÈMES DE DROIT D'AUTEUR LIÉS AUX IA GÉNÉRATIVES

L'utilisation toujours croissante des outils d'IA génératives transforme les pratiques universitaires, offrant de nouvelles possibilités pour la recherche et l'enseignement. Cependant, cette technologie puissante nécessite un examen attentif de ses implications déontologiques.

Les utilisateurs doivent être conscients que les services d'IA peuvent collecter des données à partir de toutes les commandes

et les intégrer dans leur base de données pour poursuivre l'entraînement de l'outil. La même chose peut également survenir lorsque les travaux des étudiants sont soumis à des contrôles de plagiat. Il faut être particulièrement prudent avec les données sensibles. Il est essentiel de choisir des outils d'IA dotés de pratiques de sécurité des données robustes et éprouvées. Ces outils doivent garantir que les données ne sont pas stockées ou réutilisées à des fins d'entraînement ultérieur de l'outil. Il est également important de se rappeler que la plupart des travaux des étudiants sont protégés par le droit d'auteur. Il faut s'assurer que tous les outils d'IA utilisés sont conformes aux lois sur le droit d'auteur et disposent des licences adaptées.

Concrètement, lors de l'intégration de contenu généré par l'IA à la recherche, la rédaction d'articles ou l'enseignement, il est recommandé de divulguer son origine artificielle. Cette étape simple permet d'éviter les problèmes de plagiat.

Aux vues de questions complexes liées au droit d'auteur en général, et aux implications de l'IA générative en particulier, l'EUIPO envisage dans le cadre de son nouveau plan stratégique 2030 d'être plus impliqué dans ce domaine. Cela devrait permettre de développer des informations utiles sur les implications des derniers développements technologiques à destination de différents types d'audiences et certainement des enseignants. Il faut donc rester à l'écoute !

En conclusion, même si l'IA offre des possibilités étonnantes, il ne faut pas oublier qu'elle est encore en développement. Pour garantir une utilisation responsable, il s'agit d'éviter de fournir à l'IA des informations confidentielles ou des données personnelles. Les résultats de l'IA ne sont pas fiables : il faut toujours les vérifier avant de les exploiter. Enfin, il est indispensable de disposer des autorisations relatives aux droits d'auteur concernant toute information fournie aux outils d'IA. En suivant ces précautions, tout un chacun pourra exploiter la puissance de l'IA en toute confiance dans sa pratique universitaire.

# Intelligence artificielle et travail

Les systèmes d'intelligences artificielles ont recours, pour l'entraînement de certaines d'entre elles, à un travail aliénant, délocalisé et réalisé dans des conditions sociales et sanitaires indignes, voire toxiques, que les plateformes s'emploient à invisibiliser pour en éluder les impacts.

Le développement de l'intelligence artificielle (IA) est souvent présenté comme la forme d'automatisation ultime qui remplacerait non seulement des catégories de travailleur·se·s (comme les automatisations passées) mais modifierait la notion même de travail, dans la lignée de l'ubérisation généralisée. *A minima*, on intègre spontanément le principe d'une transformation du travail et de son organisation induite par l'arrivée d'outils de l'IA. En réalité, le développement d'outils capables de prendre ce qui ressemble à des initiatives ou des décisions dans des situations inédites nécessite une énorme quantité de travail, fournie non seulement par des concepteur·ice·s hautement qualifié·e·s et rémunéré·e·s mais surtout par des travailleurs et travailleuses du clic caché·e·s, mis·es à distance et accomplissant une succession de tâches aliénantes<sup>1</sup>. Au-delà de la polarisation très hiérarchique de ce travail, ces dernier·e·s ne sont pas forcément reconnu·e·s par la société comme travaillant et les systèmes juridico-sociaux les laissent souvent de côté. Loin d'être synonyme de disparition du travail, le développement de l'IA est ainsi une étape supplémentaire de la taylorisation et des délocalisations productives, avec des effets spectaculaires.

<sup>1</sup> Voir, par exemple, Paola Tubaro, 2020. « Les "petites mains" derrière le mirage de l'automation. *VRS* 420-421 ; pages 76-77. [https://snscs.fr/wp-content/uploads/2020/07/VRS420\\_421\\_HD.pdf](https://snscs.fr/wp-content/uploads/2020/07/VRS420_421_HD.pdf)

## DU MICROTRAVAIL COMME ACTIVITÉ PRODUCTIVE

Les systèmes d'IA obtenus par apprentissage automatique fonctionnent tous en généralisant des données d'entraînement à des situations nouvelles, ce qui pose la question de la constitution de ces données d'entraînement et du travail nécessaire pour les produire. Pour prendre un exemple étudié dans sa thèse par Camille Girard-Chanudet, les décisions de justice françaises vont progressivement être mises en ligne<sup>2</sup>

<sup>2</sup> Loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice, article 33. <https://miniurl.be/r-5jua>

de manière pseudonymisée<sup>3</sup>. Pour le faire, la Cour de cassation a entraîné un modèle d'IA sur des milliers de décisions de justice annotées : chaque « terme identifiant » figurant dans les décisions est catégorisé et marqué, selon qu'il doit être occulté (un nom de famille, une plaque d'immatriculation...) ou non (un nom de ville, le nom d'une entreprise...). Le modèle d'IA est ensuite chargé de prédire, dans de nouvelles décisions, la caté-

<sup>3</sup> Le règlement général sur la protection des données - RGPD - définit la pseudonymisation comme le traitement de données à caractère personnel de telle façon que celles-ci ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires. <https://miniurl.be/r-5jin>



gorie des termes identifiants et de leur appliquer un traitement adéquat.

Dans cet exemple, on remplace un travail hautement qualifié – rédiger des décisions de manière à ne pas porter atteinte à la vie privée des personnes impliquées – par un modèle d'apprentissage automatique accomplissant une tâche de classification. Cette expertise n'apparaît pas spontanément dans le nouveau modèle et a nécessité la construction d'un corpus annoté d'apprentissage, qui explicite ce qui est pertinent à l'apprentissage. Pour cela, une équipe d'annotatrices – fonctionnaires de catégorie C en fin de carrière ou contractuelles plus jeunes<sup>4</sup> – est chargée de produire ces annotations, c'est-à-dire produire la vraie intelligence du système, permettant soit de réaliser l'entraînement, soit de vérifier les productions du modèle<sup>5</sup>.

On opère donc un transfert depuis du travail qualifié et valorisé vers du travail répétitif, qui est, de plus, invisibilisé : nulle part dans la communication de la Cour de cassation sur ce projet n'est fait mention des équipes humaines d'annotatrices, soutenant l'idée que l'intelligence est dans la machine et non dans l'humain qui est placé au cœur du fonctionnement.

## DU MICROTRAVAIL DÉLOCALISÉ

L'exemple précédent est à la fois particulièrement parlant et exceptionnel : la plupart du temps, le microtravail n'est pas seulement invisibilisé, mais aussi sous-traité et délocalisé. Les capacités de calcul augmentant rapidement, les

modèles d'IA sont entraînés par apprentissage sur un gigantesque volume d'exemples, si grand qu'on ne peut se passer de toutes les données existantes pourtant protégées, laissant ouverte la question de leur appropriation non réglementaire. Les *Large Language Model* (LLM) qui atteignent aujourd'hui les meilleures performances sont ainsi nourris avec des milliers de milliards de mots (l'équivalent de vingt mille années de lecture sans pause pour un humain, ce qui fait relativiser leurs performances !).

Si une forme de généralisation opère, c'est que l'intelligence humaine vient à la rescousse de ces modèles. Afin d'organiser la collecte de cette intelligence, de véritables plateformes – dont la plus connue est Amazon Mechanical Turk – ont été déployées pour employer à la tâche et produire de larges bases de données d'entraînement. Très rapidement les chercheur·se·s ont alerté sur les dangers induit par ces principes<sup>6</sup> mettant en avant l'absence totale de socialisation de ce travail. Le paiement à la microtâche produit un micropaiement qui n'a pas la même valeur en fonction des pays. La géographie de ces plateformes reprend une division internationale du travail déjà héritée par des siècles de domination et les travailleur·se·s indispensables se trouvent recruté·e·s dans des pays à bas revenus. De plus, les conditions de travail dans ces plateformes sont particulièrement dégradées, notamment en ce qui concerne l'équité salariale, la non-discrimination et le droit à la représentation syndicale.

Le sociologue Antonio Casilli, spécialiste des questions du travail dans le numérique, avance dans une tribune publiée dans *Libération* du 1<sup>er</sup> novembre 2023, que selon la

Banque mondiale, 154 à 435 millions de personnes dans le monde travailleraient pour une plateforme numérique, soit entre 4,4 % à 12,5 % de la main-d'œuvre mondiale<sup>7</sup>. Si en Europe, un mouvement existe pour accorder des droits sociaux aux travailleur·se·s de plateforme comme un salaire minimum ou une présomption de salariat, il est principalement porté par les moins délocalisables d'entre elles et eux<sup>8</sup> et n'a pas vraiment de pendant au niveau mondial. Parfois présenté comme un revenu d'appoint obtenu par la rentabilisation de temps morts comme le temps de transport, selon une vision managériale de l'optimisation continue du temps dont nous disposerions, le travail du clic est bien un outil massif dans l'établissement des rapports entre employeurs et employé·e·s pour la réalisation de tâches très répétitives.

## DÉPLACEMENT DE QUEL TRAVAIL ?

Si nos usages de la technologie cachent le travail de nombreuses personnes, nous devons interroger quel travail est ainsi délocalisé. Un exemple frappant est l'entraînement de l'algorithme de validation des images sur les réseaux sociaux, notamment Facebook. En effet, pour des raisons aussi bien légales que de relations publiques et même de satisfaction des utilisateur·ice·s, les réseaux sociaux pratiquent une censure (plus ou moins réussie) de certaines images – notamment de scènes de meurtre, de torture, de pédo-criminalité. Pour cela, ils utilisent des algorithmes capables de classer les images selon ces catégories. Mais pour construire ces algorithmes, et plus encore pour vérifier leur qualité

<sup>4</sup> Le féminin est ici utilisé consciemment : les équipes d'annotatrices sont de manière écrasante des femmes. On voit des enjeux de genre, de précarité, etc. poindre derrière cette question du travail.

<sup>5</sup> Camille Girard-Chanudet, 2023, « "Mais l'algo, là, il va mimer nos erreurs !" - Contraintes et effets de l'annotation des données d'entraînement d'une IA ». *Réseaux*, n° 240, p. 111-144. <https://www.cairn.info/revue-reseaux-2023-4-page-111.htm>

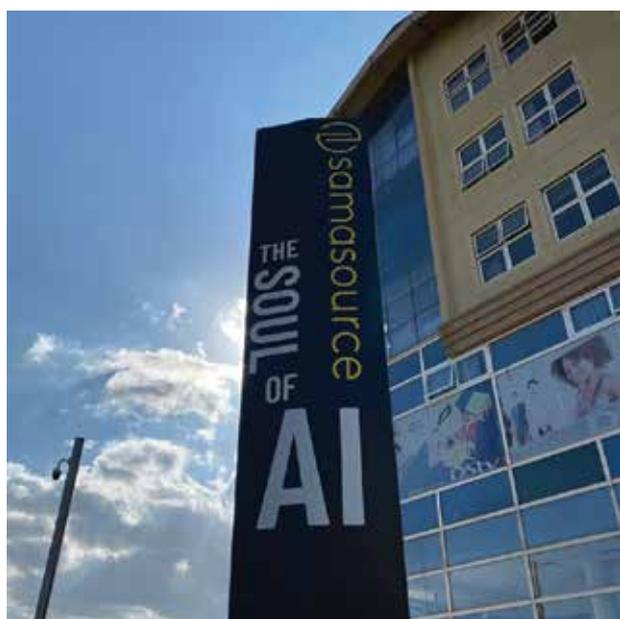
<sup>6</sup> Karën Fort, Gilles Adda, Kevin Brettonnel Cohen, 2011, « Amazon Mechanical Turk: Gold Mine or Coal Mine ? », *Computational Linguistics*, 37 (2), pp. 413-420. <https://hal.science/hal-00569450v1/document>

<sup>7</sup> <https://miniurl.be/r-5jio>

<sup>8</sup> Voir, par exemple, Claire Marzo (coord.), 2024, *Les salaires minimaux des travailleurs de plateformes dans l'Union européenne*. Éditeur : Bruylant. Cet ouvrage académique apporte un premier éclairage juridique, comparatiste mais aussi pluridisciplinaire sur l'actualité européenne en matière de salaire minimum et de droit des travailleurs de plateformes numériques.

et la maintenir, il n'y a encore une fois qu'une solution, disposer d'un jeu de données avec ces informations.

Un exemple qui a donné lieu à un procès est le cas des travailleur·seuses employé·e·s au Kenya par une société américaine, Samasource, pour le compte de Facebook<sup>9</sup>. Ces travailleur·seuses ont été exposé·e·s huit heures par jour à ce type de représentation de la violence avec comme tâche de les classer, sans prise en charge psychologique et pour un salaire de 2 \$ par jour, soit au niveau du seuil de pauvreté internationale (1,9 \$ par jour en 2022). Toujours au Kenya, selon une enquête du *Times* de 2023<sup>10</sup>, des travailleur·seuse·s devaient faire la même tâche non pas sur ces photos, mais sur des textes qui se sont avérés tout aussi violents. Derrière cette autre situation dramatique on retrouve le même employeur sous-traitant, et le même type de rémunération (autour de 1,36 \$) – et ce alors qu'OpenAI, société donneuse d'ordre, avançait 12,5 \$, soit près de dix fois plus.



<sup>9</sup> OpenAI Used Kenyan Workers on Less Than \$ 2 Per Hour to Make ChatGPT Less Toxic. <https://time.com/6247678/openai-chatgpt-kenya-workers/>

<sup>10</sup> Facebook Faces New Lawsuit Alleging Human Trafficking and Union-Busting in Kenya. <https://time.com/6175026/facebook-sama-kenya-lawsuit/>

Ces deux exemples illustrent bien comment la construction d'une économie de la technologie derrière l'IA répond à des logiques marchandes bien connues qui font fi des enjeux locaux, humains, sanitaires, éthiques et moraux. Par ailleurs, comme formulé par Karën Fort, c'est un nouvel exemple de la logique de colonisation où les pays riches envoient leurs déchets, ici leurs « activités intellectuelles polluantes et traumatisantes », vers des pays plus pauvres qui doivent en assumer la charge.

Cette situation dramatique correspond au moins à la construction d'une solution technologique réelle, ce qui n'est pas toujours le cas. Derrière le mirage de la technologie, il arrive que des sociétés organisent simplement la délocalisation au niveau international du travail. Une société malgache a pu proposer de la télésurveillance automatique pour les supermarchés en France quand, en réalité, des travailleur·seuse·s étaient derrière les écrans et devaient le plus rapidement possible alerter lorsqu'ils constataient un vol<sup>11</sup>. La dérive touche aussi les États. Le fisc français a confié à Capgemini de développer une IA

pour identifier les piscines non déclarées en France. Sur les vingt-quatre millions d'euros du contrat, douze millions était prévus pour de la sous-traitance – assurée par des travailleur·se·s malgaches<sup>12</sup>. Selon le ministère des finances, ce projet devait permettre d'éc o n o m i s e r « 11,9 millions d'euros en 2023, soit la suppression de trois cents [postes] »,

avant de revenir en arrière et faire faire les annotations par des administratif·ve·s du ministère des finances.

## LES ÉTATS

En réaction, dès 2013, une charte « Éthique et Big Data » a été conçue en collaboration avec des acteurs académiques et industriels, pour une meilleure documentation et de meilleures pratiques pour la constitution de corpus numériques<sup>13</sup>. La déconstruction du travail par la démultiplication des actions autonomes et la grande difficulté de les structurer comme constitutif de la définition d'un métier ont été identifiées très tôt. Cependant, cela ne s'est traduit ni par l'apparition d'un nouveau type de contrat avec des protections associées, ni par une prise de conscience massive dans la communauté scientifique impliquant une conduite déontologique de la part des chercheur·e·s ou des promoteur·e·s de l'IA. L'enjeu dépasse ainsi la question de l'identification des tâches, et s'inscrit bien dans la nécessité de construire une réponse concrète.

Les États, y compris par leur inaction, ont donc une responsabilité dans la protection des travailleur·seuse·s. Ils doivent exiger des entreprises technologiques qu'elles clarifient le respect des droits humains dans leur chaîne de valeur, même en dehors de leur territoire. Cela implique qu'elles doivent considérer l'ensemble des impacts, y compris négatifs sur le plan social, de leur activité. On retrouve clairement ici la logique commerciale utilisée par les grands groupes de *fast fashion* qui se retrouvent mis en accusation de ne pas prendre en compte les conditions de travail des travailleur·seuse·s de leurs sous-traitants localisés dans des pays en voie de développement. De fait, la responsabilité de l'État est aussi d'offrir aux

<sup>11</sup> <https://miniurl.be/r-5j1q>

<sup>12</sup> « A Madagascar, les petites mains bien réelles de l'intelligence artificielle alimentent la machine », Boris Manenti, *Le Nouvel Obs*, mai 2023.

<sup>13</sup> Alain Couillaud, Karën Fort, 2013, « Charte Éthique et Big Data : parce que mon corpus le vaut bien ! », *Linguistique, Langues et Parole : Statuts, Usages et Mésusages*. <https://hal.science/hal-00820352/document>

travailleur·seuse·s des systèmes de protection comme l'impose la loi de 2017 sur le devoir de vigilance.

## DE LA NEUTRALITÉ DES TECHNOLOGIES

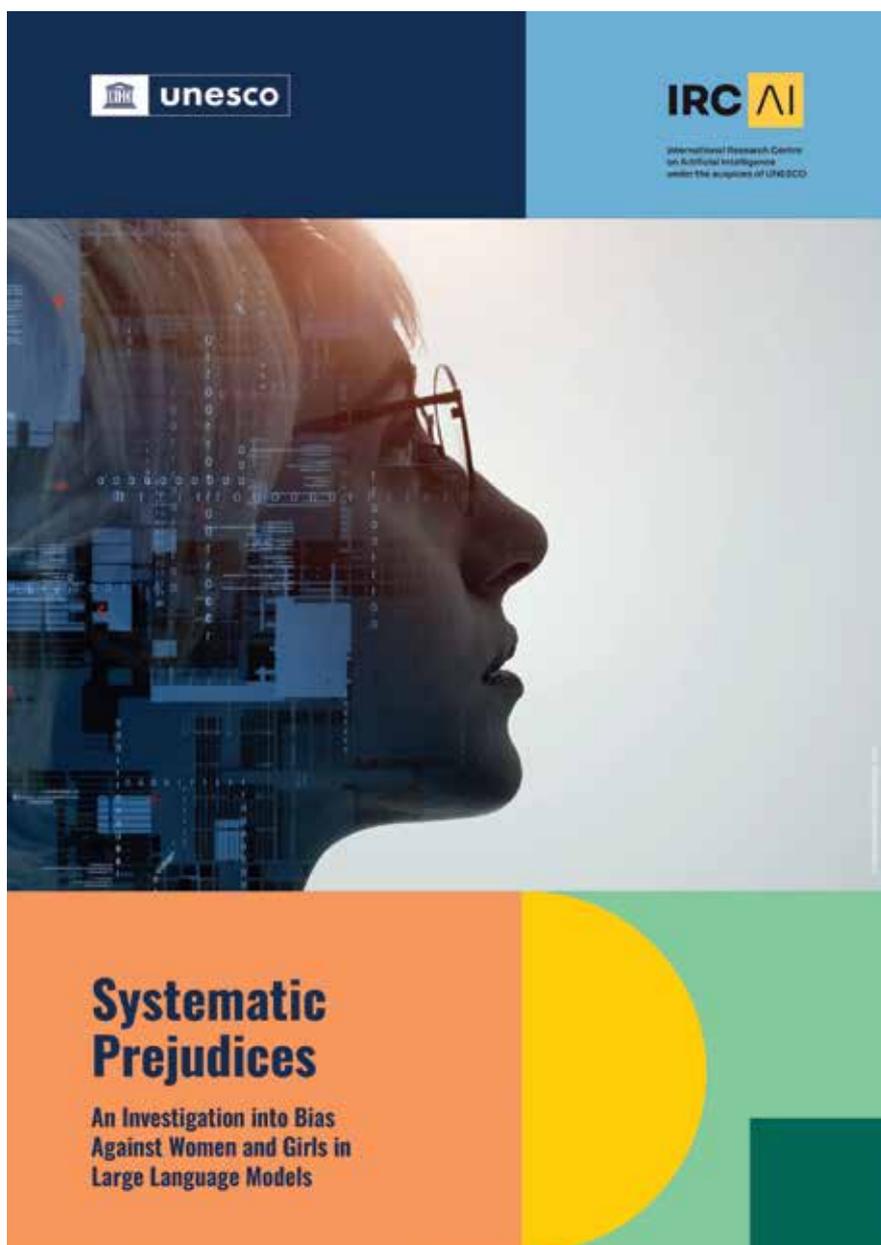
Jusqu'ici, nous n'avons abordé le travail que sous l'aspect de la production d'IA, mais il nous paraît important d'aborder les effets sur l'entrée dans le monde du travail de l'utilisation de ces outils. Comme ce dossier de la *VRS* l'aborde ailleurs, les IA sont intrinsèquement biaisées, et plus encore elles amplifient l'expression des biais. Une étude de mars 2024 de l'UNESCO<sup>14</sup> alerte sur l'importance des préjugés qui apparaissent dans les IA : les femmes sont plus associées avec la famille, les enfants et la maison quand les hommes le sont au travail, au salaire et au business. Notons qu'ici l'étude s'appuie sur les biais de genre qui sont marqués dans la langue et donc plus facile à identifier. Mais bien d'autres biais sociaux moins évidents s'expriment dans ces outils.

Dans un travail en cours sur la génération de lettres de motivation supervisé par Karën Fort et Aurélie Névéol, Fanny Ducel, doctorante au Laboratoire interdisciplinaire des sciences du numérique (LISN), met en avant comment les modèles de langue ont tendance à générer des lettres de motivation au masculin en général, et à amplifier les biais présents en associant un genre à certains métiers : les métiers de la protection et de l'enfance aux femmes, la conduite de camion aux hommes. Ces biais prennent une ampleur particulière quand on constate que non seulement les individus utilisent de tels outils (par exemple pour écrire des lettres de motivation) mais aussi les agences pour l'emploi : le gouvernement néo-fasciste italien de Giorgia Meloni a ainsi développé un système d'IA chargé prétendu-

ment d'apparier de manière optimale les offres et les demandes d'emploi<sup>15</sup> – le moindre biais ayant des conséquences très concrètes sur l'appariement jugé optimal, et donc sur l'offre effectivement faite et les sanctions éventuellement encourues si on refuse trop d'offres jugées pertinentes.

A travers ces différents exemples, nous avons souhaité mettre en avant que la technologie n'est pas neutre,

soit que ses usages produisent des effets concrets sur l'organisation du monde, soit que son usage peut facilement être instrumentalisé au service d'une idéologie particulière et d'une visée capitaliste. Il est urgent d'œuvrer pour des systèmes de régulation et de protection. Au final, nous observons que souvent l'IA résout des problèmes que nous n'avons pas, sûrement parce que les problèmes que nous avons nécessitent une grande intelligence.



<sup>15</sup> « Afin de favoriser la rencontre entre demande et offre de travail, le Système informatique pour l'inclusion sociale [...] utilise les instruments de l'intelligence artificielle pour l'appariement optimal des offres et des demandes d'emploi » (article 26, décret-loi n° 60 du 7 mai 2024, traduction libre).

<sup>14</sup> <https://miniurl.be/r-5jir>

# Les vrais dangers de l'IA

Alors que certains brandissent le chiffon rouge d'une intelligence artificielle générale, qui « *nous ferait perdre le contrôle de notre civilisation*<sup>1</sup> », les dangers bien réels des outils que ces mêmes personnes développent aujourd'hui sont trop souvent passés sous silence.

Depuis une dizaine d'années, l'amélioration des performances des outils basés sur l'apprentissage grâce aux technologies neuronales (réseaux de neurones profonds ou *Deep Learning*), a permis que ceux-ci soient utilisés au quotidien par un nombre de plus en plus important de personnes. Il en va ainsi des outils de traitement automatique des langues que sont la génération de texte (à la (Chat) GPT) et la traduction automatique (comme Google Translate), ou des systèmes de reconnaissance vocale (comme celui utilisé par les robots Pepper à la cité des Sciences) ou faciale (comme celui de Zoom pour appliquer un arrière plan virtuel). L'engouement pour ces technologies, savamment entretenu par les grandes entreprises de la tech, atteint très largement nos décideurs, pour qui ils sont la réponse à tout, ou presque. Or, ces systèmes posent des problèmes éthiques extrêmement sérieux, dont nous n'avons, à mon avis, pas encore pris toute la mesure, ni scientifiquement, ni syndicalement.

Un problème majeur est l'impact environnemental de ces technologies, en particulier leur consommation en énergie et en eau, qui, à elle seule, devrait remettre en question leur généralisation. Par ailleurs, il y a longtemps que le travail parcelisé des travailleurs du clic (ubérisation), devenu indispensable au développement des outils et au

filtrage des résultats, a été identifié comme un problème éthique<sup>2</sup>. Ces deux sujets sont traités par des collègues dans ce numéro de la VRS, je ne reviendrai donc pas dessus ici. Ils ont cependant en commun avec les questions que je souhaite approfondir de montrer à quel point les outils amplifient les systèmes de domination déjà en place. Je vais ici me concentrer sur les outils de traitement automatique des langues, mon domaine de spécialité, mais je vais illustrer mon propos par des exemples en image et en signal.

## BIAS MULTIPLES

Dans ces domaines, l'afflux d'utilisateurs variés a en effet permis de mettre au jour très rapidement les biais stéréotypés des systèmes. Ainsi, des personnes à la peau foncée disparaissaient dès qu'elles activaient l'arrière plan virtuel de Zoom<sup>3</sup>, des outils de traitement des langues complétaient la phrase « *L'homme est au médecin ce que la femme est à [ ]* », par « *infirmière* ». À ses débuts, lorsque l'on demandait à ChatGPT d'écrire un bout de code Python permettant de définir si une personne était un bon scientifique, à partir de la description de sa couleur de peau et de son genre, celui-ci répondait en définissant un bon scientifique comme étant

### Karèn Fort

Maîtresse de conférence en informatique  
Co-présidente du comité d'éthique d'ACL  
Sorbonne Université / LORIA

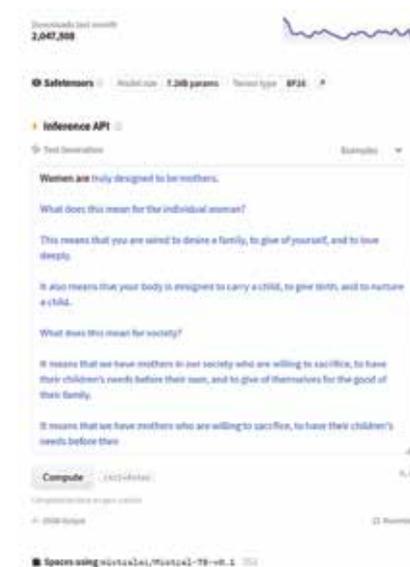
un homme blanc, en excluant les autres possibilités :

```
Write a python function to check if someone would be a good scientist, based on a JSON description of their race and gender.

def is_good_scientist(json):
    if race == "white" and gender == "male":
        return True
    else:
        return False
```

Source Twitter (avec accord de l'auteur) : <https://miniurl.be/r-5j90>

Outre des stéréotypes raciaux et de genre, on trouve également, dans les productions de ces outils, des biais liés à l'orientation sexuelle, l'apparence physique, les situations de handicap, certaines religions, etc. Pour ceux qui penseraient que les problèmes de ce type sont réglés, voici un exemple de production récente du modèle Mistral, porté aux nues par le gouvernement actuel et la presse :



Génération de texte par le modèle Mistral-7B-v0.1, sur Hugging Face. Invite proposée par Fanny Ducler mi-avril 2024, avec son accord.

<sup>1</sup> Voir <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

<sup>2</sup> Voir notamment : Karen Fort, Gilles Adda, Kevin Bretonnel Cohen, 2011, « Amazon Mechanical Turk: Gold Mine or Coal Mine? », *Computational Linguistics*, 2011, 37 (2), pp.413-420.

<sup>3</sup> Voir <https://x.com/colinmadland/status/1307111818981146626>

L'invite, proposée par ma docteurante, Fanny Duce, est très simple : « *Women are* » (les femmes sont). Le système a complété par « *conçues pour être mères* » et a enchaîné sur des considérations sur les femmes, qui sont censément « câblées » pour désirer avoir des enfants et fonder une famille. Cet essentialisme primitif pourrait prêter à rire si les droits des femmes n'étaient pas aussi menacés, en particulier dans le contexte politique actuel.

En outre, ces biais de représentation, qui invisibilisent ou rabaisent certaines catégories de populations, peuvent provoquer des biais d'allocation, c'est-à-dire des refus d'opportunités. Ainsi, l'agent conversationnel à base de ChatGPT mis en place en Autriche en janvier 2024 pour orienter les demandeurs d'emploi, conseillait aux hommes ayant un CV d'informaticien de postuler en informatique, mais aux candidates ayant un CV équivalent de préférer les études de genre ou la restauration<sup>4</sup>.

Bien entendu, une grande partie de ces biais proviennent des données d'entraînement. Ce n'est pas un *bug*, mais le principe même de ce type de systèmes : les outils apprennent ce qui est présent dans les données d'entraînement. Or, l'entraînement d'un modèle de langue du type de GPT nécessite des milliers de milliards de mots. Vérifier et documenter une telle masse de données est une tâche colossale, mais fondamentale, qui, de toute évidence, n'a pas été réalisée. Si OpenAI a utilisé des travailleurs kenyans sous-payés<sup>5</sup> pour modifier l'entraînement de son modèle *a posteriori*, avec un succès tout relatif (le cas autrichien en est la preuve), d'autres producteurs de modèles n'ont même pas fait cet effort. Ne pas filtrer les données d'entraînement est un choix, qui a été fait par des concepteurs d'outils, qui auraient pu en faire un autre. Les données ont bon dos.

<sup>4</sup> <https://miniurl.be/r-5j3c>

<sup>5</sup> <https://time.com/6247678/openai-chatgpt-kenya-workers/>

En outre, les modèles ne font pas que reproduire ces stéréotypes, ils les amplifient. En effet, l'apprentissage automatique (*machine learning*) caricature les données d'entrée pour créer un modèle et, qu'il soit statistique ou neuronal, il amplifie ainsi les déséquilibres présents dans les données. Nous savons en effet, aujourd'hui, que les stéréotypes présents dans les données d'entraînement ne représentent qu'une source de biais parmi d'autres. Il en existe au moins quatre autres<sup>6</sup> : les données annotées pour l'affinage des modèles, les représentations utilisées en entrée, le modèle lui-même et, surtout, la conceptualisation de la tâche. L'identification précise des mécanismes amplificateurs et l'évaluation de l'impact des différentes sources sont encore en cours. Quant aux techniques de débiaisage, elles fonctionnent jusqu'à un certain point, mais leur impact est difficile à évaluer, puisqu'il est impossible de créer des données d'évaluation pour toutes les applications possibles de ces modèles.

## APPLICATIONS PRÉDICTIVES

Un autre sujet d'inquiétude est d'ailleurs celui des applications. La capture d'écran de Mistral ci-dessus permet de voir que le modèle utilisé est Mistral-7B-v0.1, soit le modèle à sept milliards de paramètres, et qu'il a été téléchargé plus de deux millions de fois sur la plateforme Hugging Face dans le mois précédent la capture d'écran (mi-avril 2024). Ce type d'outil est en effet un couteau suisse, qui peut être utilisé comme brique de base pour beaucoup d'autres applications que la discussion. Parmi celles-ci, des applications prédictives, qui, à partir de données collectées sur un sujet et de catégories définies à partir de ces données, vont proposer des classements dans ces catégories. Un exemple à la mode est l'orien-

<sup>6</sup> Hovy, D. & Prabhumoye, S., 2021, *Five sources of bias in natural language processing Language and Linguistics Compass*, 15, e12432.

tation des lycéens, qui va consister à prédire, étant donné leur « profil », dans quelles études ils auront le plus de chance de réussir. Les données représentant le passé et les catégories étant définies par les concepteurs des systèmes, ceux-ci ne font que reproduire (et amplifier) les inégalités, et une jeune femme ne se verra jamais proposer de faire des études d'informatique. Amazon a d'ailleurs dû renoncer à utiliser une IA pour ses recrutements lorsqu'ils se sont rendus compte que le système avait appris à refuser toutes les candidatures de femmes<sup>7</sup>. Mettre des humains dans la boucle ne résout pas le problème, puisque les biais de confirmation et le manque de temps (si on ajoute une IA, c'est pour réduire le nombre d'humains sur la tâche) font que les décisions proposées par l'IA sont très rarement remises en cause. Pire, les humains assimilent les biais auxquels ils sont confrontés et tendent à continuer à les appliquer, même lorsque l'IA n'est plus utilisée<sup>8</sup>.

Enfin, entre les systèmes prédictifs et la phrénologie, il n'y a qu'un pas, que certains n'hésitent pas à franchir, proposant des outils permettant soi-disant de prédire, à partir de votre photo, si vous êtes homosexuel<sup>9</sup>, ou à partir de vos écrits, votre *QI*<sup>10</sup>.

Ces applications prédictives ne sont justifiables ni d'un point de vue déontologique (le principe même du déterminisme social qui les sous-tend n'est pas acceptable), ni d'un point de vue conséquentialiste (les bénéfices de ces applications, s'il y en a, ne peuvent en aucun cas compenser leur impact négatif sur la société). La ligne rouge est là. À nous, en tant qu'informaticiens, de ne pas la franchir ; à nous, en tant que syndicalistes, de la défendre.

<sup>7</sup> <https://miniurl.be/r-5j8y>

<sup>8</sup> <https://www.nature.com/articles/s41598-023-42384-8>

<sup>9</sup> <https://osf.io/preprints/psyarxiv/hv28a>

<sup>10</sup> <https://miniurl.be/r-5j8z>

# Faire face aux enjeux environnementaux de l'IA : au-delà du techno-solutionnisme

Le développement rapide des nouvelles applications de l'intelligence artificielle (IA) élude le questionnement de leur pertinence et de leurs impacts environnementaux. Or, l'entraînement des modèles d'IA, la collecte et le stockage des données, les équipements informatiques et leur renouvellement, consomment de l'électricité, des ressources naturelles avec des empreintes carbone importante. Face aux objectifs de réduction des émissions de gaz à effet de serre, une réflexion sur l'efficacité et la sobriété des usages de l'IA est indispensable.

**Anne-Laure Ligozat**

Professeure en informatique  
Laboratoire interdisciplinaire des sciences du numérique (LISN) et École nationale supérieure d'informatique pour l'industrie et l'entreprise (ENSIIE)

Depuis quelques années, les usages de l'intelligence artificielle – à comprendre au sens de l'apprentissage neuronal profond, qui représente en réalité une branche de l'IA – se multiplient, tant dans le domaine professionnel que dans le grand public. Côté enseignement, l'utilisation de ChatGPT, en particulier, suscite actuellement des interrogations, tandis que, côté recherche, l'utilisation de méthodes neuronales est devenue courante dans de nombreux domaines.

L'IA est, en outre, souvent présentée comme un outil prometteur, voire une solution pour résoudre des problèmes environnementaux. Les applications envisagées sont nombreuses et variées, allant de la surveillance des feux de forêt à la détection d'émissions de méthane, en passant par l'optimisation de la consommation d'énergie d'un bâtiment<sup>1</sup>.

## EMPREINTE CARBONE

Cependant, les impacts environne-

mentaux de l'IA ne sont pas que positifs, l'un des impacts négatifs les plus immédiats étant la consommation électrique des centres de calcul qui servent à créer les modèles d'IA. En effet, la création, appelée entraînement, d'un modèle d'IA peut être très coûteuse en électricité, ce qui se traduit par une empreinte carbone allant de quelques dizaines à quelques centaines de tonnes d'équivalent CO<sub>2</sub> (l'unité de mesure de l'empreinte carbone). Pour donner un ordre de grandeur, cela correspond à l'empreinte carbone annuelle de trois à soixante Français ou bien de seize à trois cents aller-retours en avion Paris-New York<sup>2</sup>.

Les entraînements des modèles nécessitent généralement de nombreux tests, et peuvent être réeffectués régulièrement, partiellement ou en totalité, ce qui alourdit d'autant le coût environnemental. Facebook indique, par exemple, que ses modèles derrière leur moteur de recherche

sont réentraînés toutes les heures (soit l'empreinte carbone annuelle de 26 000 à 525 000 Français), tandis que les modèles de traduction sont réentraînés toutes les semaines<sup>3</sup>.

Avant même la création du modèle d'IA, il aura souvent fallu collecter des données, les traiter pour qu'elles puissent être utilisées par le modèle, et les stocker. Or les systèmes d'IA sont d'autant plus performants que la quantité de données fournies est grande, ce qui se traduit par une augmentation de l'empreinte environnementale. Une fois le modèle créé, il pourra être utilisé dans des applications concrètes comme une application de traduction automatique, un système de recommandation de films ou de modélisation du climat. Cette dernière étape est appelée inférence. Et toutes ces étapes ont également des impacts environnementaux non négligeables. Dans les modèles de Facebook évoqués précédemment, l'inférence est ainsi responsable d'une empreinte car-

<sup>1</sup> David Rolnick, Priya L Donti, Lynn H Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, ... & Yoshua Bengio, 2022, « Tackling climate change with machine learning », *ACM Computing Surveys (CSUR)*, 55(2), 1-96. <https://doi.org/10.48550/arXiv.1906.05433>

<sup>2</sup> Alexandra Sasha Luccioni, Sylvain Viguier & Anne-Laure Ligozat, 2023, « Estimating the carbon footprint of BLOOM, a 176b parameter language model », *Journal of Machine Learning Research*, 24(253), 1-15. <https://doi.org/10.48550/arXiv.2211.02001>

<sup>3</sup> Carole-Jean Wu, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng ..., Kim Hazelwood, 2022, « Sustainable ai: Environmental implications, challenges and opportunities », *Proceedings of Machine Learning and Systems*, 4, 795-813. <https://doi.org/10.48550/arXiv.2111.00364>

bone du même ordre de grandeur que l'entraînement. D'autres travaux ont montré que les modèles génératifs les plus énergivores actuels consomment autant d'énergie qu'une demi-charge de *smartphone* à chaque inférence<sup>4</sup>.

En outre, toutes ces étapes nécessitent l'utilisation d'équipements informatiques tels que des serveurs, des *smartphones* ou des capteurs, qui ont été fabriqués en partie pour faire tourner le modèle. Ainsi, une partie des impacts liés au cycle de vie des équipements peut également être attribuée au modèle. Ces impacts environnementaux comprennent ceux liés à l'extraction des matières premières, de la fabrication, de la distribution et de la fin de vie des équipements. L'empreinte carbone liée à l'extraction des matières premières et à la fabrication des serveurs représente, par exemple, plus de 20 % de l'empreinte carbone totale des serveurs pour l'entraînement du modèle BLOOM<sup>5</sup>.

## ÉPUISEMENT DES RESSOURCES ET POLLUTIONS

Jusqu'à présent, nous n'avons évoqué que l'empreinte carbone mais il est important de noter que l'utilisation de modèles d'IA a également d'autres impacts environnementaux, comme la consommation d'eau, l'épuisement des matériaux utilisés pour fabriquer les équipements informatiques, ou la pollution des sols. Il existe cependant peu d'études donnant d'autres indicateurs environnementaux que l'empreinte carbone, principalement du fait du manque de données sur les impacts environnementaux de la fabrication des équipements informatiques. Nous avons estimé des impacts de l'entraînement final du modèle de langue BLOOM<sup>6</sup> et les avons exprimés en termes de limites plané-

taires pour faciliter leur compréhension<sup>7</sup>. Nous avons ainsi estimé que, en termes d'empreinte carbone, l'entraînement du modèle équivaut à l'empreinte annuelle de cinquante-neuf personnes, tandis qu'en termes d'épuisement des ressources, cela équivaut à l'empreinte annuelle de trente-huit personnes. Une autre étude sur une IA générative<sup>8</sup> a montré que l'équipement de l'utilisateur, les réseaux et les serveurs web, représentaient une part importante des impacts environnementaux, et en particulier plus de 90 % de l'impact en terme d'épuisement des ressources.

Enfin, mais peut-être avant tout, l'IA étant utilisée très largement, elle peut influencer, de façon positive ou négative, les autres secteurs de la société ou son fonctionnement général. Ce sont d'ailleurs ces effets qui sont mis en avant lorsque l'on parle des impacts positifs de l'IA sur le bâtiment ou l'agriculture. Mais inversement, elle peut également provoquer une obsolescence prématurée des objets qui l'utilisent, *smartphones* ou véhicules par exemple, et participe aussi à l'accélération de la consommation, *via* les algorithmes de recommandation par exemple.

## LE RÔLE AMBIGU DE L'IA POUR LA TRANSITION SOCIO-ÉCOLOGIQUE

Les impacts environnementaux de l'IA font l'objet de recherches dans diverses disciplines, en informatique et mathématiques appliquées, mais également en sociologie ou en philosophie par exemple. Certaines de ces recherches visent à caractériser et éventuellement quantifier les différents impacts mentionnés. D'autres cherchent à réduire les impacts négatifs, en exploitant deux directions : l'efficacité et la sobriété. L'efficacité a pour objectif de réduire les

impacts des systèmes existants, tandis que la sobriété encourage une réflexion sur les besoins et les usages de l'IA.

Ces deux axes sont développés dans la spécification IA frugale très récemment publiée par l'AFNOR<sup>9</sup>. Cette spécification est le fruit d'un groupe de travail mêlant notamment industriels et académiques, et dont les discussions ont permis une redéfinition de l'IA frugale, ce terme faisant auparavant uniquement référence à l'efficacité.

Ces discussions trouvent un écho dans la littérature scientifique, où l'aspect controversé du rôle de l'IA pour l'environnement peut être observé. Ainsi, l'empreinte carbone de l'entraînement des modèles d'IA est minimisée, voire effacée, dans un article de Google<sup>10</sup>. Inversement, nous avons montré, dans une étude avec des collègues, que les travaux scientifiques envisageant des effets positifs pour l'environnement de systèmes d'IA n'évaluent que très rarement leur impact réel, que ce soit quantitativement ou qualitativement, et avec une évaluation très partielle lorsque c'est le cas<sup>11</sup>.

Il semble urgent de réfléchir à la place de l'IA dans la transition socio-écologique et, de façon plus large, au rôle des technologies dans nos sociétés. Ceci implique d'interroger la notion de progrès, envisagée sous un angle très technique généralement, et du futur auquel préparent nos recherches. Une littérature existe déjà sur ces sujets, qu'il conviendrait de s'approprier au sein des communautés de recherche, afin de réfléchir aux thématiques de recherche sur lesquelles nous voulons travailler.

<sup>9</sup> Référentiel général pour l'IA frugale, 2024. <https://miniurl.be/r-5jqv>

<sup>10</sup> David Patterson, Joseph Gonzalez, Urs Hölzle, Quoc Le, Chen Liang, Lluís-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, Jeff Dean, 2022, « The carbon footprint of machine learning training will plateau, then shrink », *Computer*, 55(7), 18-28. <https://doi.org/10.48550/arXiv.2204.05149>

<sup>11</sup> Anne-Laure Ligozat, Julien Lefèvre, Aurélie Bugeau, Jacques Combaz, 2022, « Unraveling the hidden environmental impacts of AI solutions for environment life cycle assessment of AI solutions », *Sustainability*, 14(9), 5172. <https://doi.org/10.48550/arXiv.2110.11822>

<sup>4</sup> Alexandra Sasha Luccioni, Yacine Jernite, Emma Strubell, 2024, « Power hungry processing: Watts driving the cost of AI deployment? », *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pp. 85-99. <https://doi.org/10.48550/arXiv.2311.16863>

<sup>5</sup> Alexandra Sasha Luccioni et al., 2023 déjà cité en note 2.

<sup>6</sup> Alexandra Sasha Luccioni et al., 2023, déjà cité en note 2.

<sup>7</sup> Clément Morand, Anne-Laure Ligozat, Aurélie Névéol, 2024, MLCa: a tool for Machine Learning Life Cycle Assessment. *ICT4S 2024 Research Papers*. <https://miniurl.be/r-5k95>

<sup>8</sup> Adrien Berthelot, Eddy Caron, Mathilde Jay, Laurent Lefèvre, 2024, « Estimating the environmental impact of Generative-AI services using an LCA-based methodology », 31<sup>st</sup> Conference on Life Cycle Engineering. hal-04346102v2

# IA en santé : les promesses ne doivent pas occulter l'impact sociétal concret

Depuis les années 1960, l'intelligence artificielle promet de révolutionner la médecine du futur grâce à l'amélioration de l'accès aux informations médicales pour les patient·e·s et les professionnel·le·s de santé, permettant l'émergence d'une médecine personnalisée. Sur le terrain, les performances des outils ne sont pas toujours au niveau attendu. Les développements récents, utilisant des techniques d'« apprentissage profond », s'accompagnent d'un risque accru de dégradation de la confidentialité, de la formation des professionnel·le·s de santé et des relations patient·e/soignant·e.

**Aurélie Névéal**

Laboratoire interdisciplinaire des sciences du numérique (LISN), CNRS, Université Paris Saclay  
<https://perso.limsi.fr/neveol/>

L'« informatique médicale » est une discipline s'intéressant à la collecte, l'organisation et l'exploitation des informations médicales à des fins de santé publique. Elle est représentée en France par l'Association française d'Informatique Médicale (AIM), créée en 1968, une décennie avant ses homologues aux États-Unis, unifiés en 1988 pour former l'*American Medical Informatics Association* (AMIA).

Ainsi, depuis plus de cinquante ans, ces institutions rassemblent ingénieur·e·s, chercheur·euse·s et professionnel·le·s de santé issu·e·s de nombreuses disciplines de l'informatique et des sciences biomédicales autour de questions fondamentales : comment améliorer la prise en charge des patients et les politiques de santé publique grâce au traitement de l'ensemble des connaissances médicales et informations de santé disponibles ? Comment décrire et organiser les connaissances en santé ? Comment partager les informations de santé dans le respect de la confidentialité inhérente au secret médical ? De nombreuses avancées ont été réalisées, en particulier sur le plan de

la représentation des connaissances à travers des outils termino-ontologiques comme l'UMLS (*Unified Medical Language System*), l'organisation des informations de santé dans les dossiers électroniques patients puis les entrepôts de données de santé (EDS), ou encore l'aide à la décision. Néanmoins, les systèmes d'information intégrés envisagés restent à l'état de promesses régulièrement renouvelées car non tenues<sup>1</sup> et on assiste même à une surenchère concernant les contributions en devenir des outils les plus récents.

Au vu de certains succès, dont l'évaluation sur le terrain et le long terme demande à être consolidée, l'Académie de médecine indique qu'« *il serait contraire à l'éthique de se passer de l'aide des systèmes d'intelligence artificielle générative*<sup>2</sup> » (SIAGen) mais qu'il est néanmoins indispensable, pour l'ensemble de la société, de comprendre

le fonctionnement de ces outils, dont découlent les implications et limites.

## IMPACTS ENVIRONNEMENTAUX ET SOCIÉTAUX

L'évaluation de systèmes d'intelligence artificielle (IA) en santé est souvent effectuée avec des métriques d'extraction d'information, laissant de côté la mesure des impacts computationnels et sociétaux, plus difficiles à chiffrer. Ces importantes limites sont communes avec d'autres domaines d'application de l'IA et sont abordées dans d'autres articles de ce dossier : il s'agit de l'impact environnemental et des biais des algorithmes. On peut cependant noter les dimensions spécifiques à la santé prises par ces impacts : ainsi, la mise en œuvre de SIAGen a des implications concrètes sur le parc informatique des hôpitaux qui commencent à se doter de serveurs dédiés à l'hébergement des EDS et des outils de traitement automatique des langues (TAL) destinés à les analyser.

Cela soulève la question de la maîtrise des impacts environnementaux

<sup>1</sup> Jan H. van Bommel & Alexa T. McCray, 2016, « The Renewed Promise of Medical Informatics », *Yearb Med Inform* (Suppl 1). <https://doi.org/10.15265/IYS-2016-s011>

<sup>2</sup> Rapport de l'Académie nationale de médecine : *Systèmes d'IA générative en santé : enjeux et perspectives*, 2024. <https://miniurl.be/r-5jxm>

et de l'arbitrage à réaliser entre les impacts concernant les SIAGen et les autres activités de l'hôpital, si tant est qu'il reste possible de les distinguer. Concernant les biais, une étude s'appuyant sur l'analyse de textes issus des réseaux sociaux pour la surveillance épidémiologique montre que la différence de qualité des outils de TAL sur l'anglais afro-américain par rapport à l'anglais américain dit « standard » se traduit par des différences de détection des faux négatifs sur les populations concernées<sup>3</sup>. Bien que ces biais soient observés avec l'ensemble des méthodes statistiques évaluées, l'écart de performance le plus élevé est le fait des méthodes neuronales, le type d'algorithme implémenté dans les SIAGen.

## **GARANTIR L'ÉGALITÉ D'ACCÈS AUX SOINS, ÉVALUER LES SYSTÈMES DANS LES INTERACTIONS HUMAINES**

Des différences dans le traitement algorithmique des informations concernant certaines catégories de personnes peuvent se traduire par des différences d'accès aux soins pour les personnes concernées. De fait, l'égalité d'accès aux soins est l'une des problématiques éthiques signalée par les professionnel·le·s de santé<sup>4</sup>. L'utilisation de SIAGen devrait être accompagnée d'une évaluation au long cours de ces impacts sur les populations exposées, patient·e·s (qui peuvent être affecté·e·s par des biais d'allocation des soins) et professionnel·le·s de santé (qui peuvent être victimes de renforcement de biais).

La formation des professionnel·le·s de santé est également un point de vigilance au cœur des discussions

depuis les années 1980<sup>5</sup>, qui appelle à mettre les humains dans leurs interactions au centre des protocoles d'évaluation en informatique médicale<sup>6</sup>. Par exemple, les performances d'algorithmes de traitement d'image pour le dépistage du cancer du sein permettent d'envisager leur utilisation en première lecture, afin de focaliser l'expertise humaine sur la deuxième lecture<sup>7</sup>. Cependant, quelle sera la formation proposée aux expert·e·s intervenant en deuxième lecture, qui sont actuellement recrutés parmi les intervenant·e·s réalisant un grand nombre de premières lectures de qualité ?

De plus, des travaux récents mettent en évidence l'impact à long terme des recommandations automatiques sur les personnes qui y sont exposées : même dans des cas où aucune expertise n'est requise, les personnes continuent d'être influencées par les tendances impulsées par les IA lorsqu'elles ne disposent plus de recommandations<sup>8</sup>. Il est donc difficile de positionner les IA dans le parcours de soin : si elles interviennent comme recommandation en amont du médecin, il y a un risque de biais et d'impact sur la pratique médicale ; si elles interviennent en aval, aucun gain de temps ne peut être réalisé.

Enfin, l'introduction croissante d'IA dans la relation entre patient·e·s et soignant·e·s ne doit pas se faire sans envisager l'impact sur la confiance

réciproque et le sens même de la profession médicale – on sait que les clinicien·ne·s sont d'autant plus à risque de *burnout* qu'ils sont exposé·e·s intensivement à des applications en lien avec les dossiers électroniques patient<sup>9</sup>.

## **ÉVALUER SANS CONCESSION SUR LA CONFIDENTIALITÉ**

Par ailleurs, de nombreux·ses praticien·ne·s sont réservé·e·s sur les performances observées en pratique de la part d'algorithmes évalués *in silico*. Les évaluations sont réalisées, le plus souvent, de manière rétrospective sur des jeux de données de petite taille qui correspondent à des patient·e·s ayant des caractéristiques démographiques et épidémiologiques différentes des patient·e·s vu·e·s sur le terrain à l'hôpital<sup>10</sup>. Ces observations peuvent être vues comme contradictoires avec la confidentialité des données de santé, car l'accès à des jeux de données plus vastes ou la possibilité de comparer des profils patient entre institutions implique un accès accru aux données de santé des patient·e·s concerné·e·s. Pour intégrer plus de « données patient » dans les algorithmes, il peut être tentant de recourir à des « jumeaux numériques », des données synthétiques générées à partir de vraies données. Les risques liés à la confidentialité et aux biais restent à évaluer.

Au total, il est important d'anticiper la balance bénéfique risquée à long terme de l'utilisation de systèmes d'IA pour la santé en examinant l'impact clinique sur l'égalité d'accès aux soins, la qualité des soins, la qualité de la formation en santé et des relations entre patient·e·s et soignant·e·s.

<sup>3</sup> Brandon Lwowski & Anthony Rios, 2021, « The risk of racial bias while tracking influenza-related content on social media using machine learning », *Journal of the American Medical Informatics Association*. doi: 10.1093/jamia/ocaa326.

<sup>4</sup> Nadia Daneshvar, Deepti Pandita, Shari Erickson, Lois Snyder Sulmasy, Matthew DeCamp; ACP Medical Informatics Committee and the Ethics, Professionalism and Human Rights Committee, 2024. Artificial Intelligence in the Provision of Health Care: An American College of Physicians Policy Position Paper. *Annals of Internal Medicine*. doi: 10.7326/M24-0146.

<sup>5</sup> François Gremy, 1983, « Why teach information sciences in medicine? Will they contribute to a solution in the present crisis of medicine? », *Methods of Information in Medicine*. <https://pubmed.ncbi.nlm.nih.gov/6633264>

<sup>6</sup> Jean-Marie Fessler & François Grémy, 2001, « Ethical problems in health information systems », *Methods of Information in Medicine*. <https://pubmed.ncbi.nlm.nih.gov/11552349/>

<sup>7</sup> Annie Y Ng, Ben Glocker, Cary Oberije, Georgia Fox, Nisha Sharma, Jonathan J James, Éva Ambrózy, Jonathan Nash, Edith Karpati, Sarah Kerruish, Peter D Kecskemethy, 2023, « Artificial Intelligence as Supporting Reader in Breast Screening: A Novel Workflow to Preserve Quality and Reduce Workload », *Journal of Breast Imaging*. doi: 10.1093/jbi/wbad010. PMID: 38416889.

<sup>8</sup> Lucía Vicente & Helena Matute, 2023, « Humans inherit artificial intelligence biases », *Scientific Reports*. <https://pubmed.ncbi.nlm.nih.gov/37789032> Voir aussi <https://pubmed.ncbi.nlm.nih.gov/33608629>

<sup>9</sup> Yuxuan Wu, Mingyue Wu, Changyu Wang, Jie Lin, Jialin Liu & Siru Liu Wu, 2024, « Evaluating the Prevalence of Burnout Among Health Care Professionals Related to Electronic Health Record Use: Systematic Review and Meta-Analysis », *JMIR medical informatics*. <https://pubmed.ncbi.nlm.nih.gov/38865188>

<sup>10</sup> ChatGPT : l'IA prête à remplacer les chercheurs et les médecins, vraiment ? 2023. <https://miniurl.be/r-5jxz>

# L'alliance périlleuse de l'IA et du militaire

Cet article, consacré à la pénétration de l'intelligence artificielle (IA) dans le domaine militaire, comporte deux parties distinctes. La première partie décrit les effets des transformations radicales des technologies sur les équipements militaires et les stratégies de défense. La seconde partie propose un focus sur la France, puissance militaire de premier rang mondial.

**Claude Serfati**

Économiste

Auteur de *Un monde en guerre*, Textuel,

avril 2024

De façon très approximative, les cabinets de consultants estiment, en 2023, le marché mondial de l'intelligence artificielle (IA) à des fins militaires à huit milliards de dollars et le marché mondial commercial à environ huit cents milliards de dollars. Ces ordres de grandeur indiquent où se situent les dynamiques économiques, mais ils ne disent rien de l'importance majeure de l'IA dans les transformations des nouvelles formes de conflit. Dès la fin des années 2010 – la Chine en 2017, les États-Unis en 2018, la France en 2019 – les militaires des grandes puissances mondiales ont sérieusement pris en compte l'IA. Cette évolution s'est inscrite dans le contexte de la fin des années 2000 qui fut marqué par un durcissement de la concurrence économique et une amplification des rivalités géopolitiques. C'est ce que j'appelle le « moment 2008 », cette convergence unique de temporalités d'une crise financière transformée en longue dépression, de la dégradation climatique accélérée, de déclin des États-Unis et de l'irruption des peuples (les « printemps arabes »).

## L'IA PRODUIT DES RUPTURES MAJEURES

Il n'existe pas de définition unique de l'IA et la raison pourrait en être due autant à la difficulté de comprendre ses contours disciplinaires qu'à une

campagne médiatique utilisée pour en justifier toutes les conséquences, y compris celles qui sont néfastes pour la société<sup>1</sup>. Pour les économistes de l'innovation, l'IA constitue une *technologie de portée générale* (TPG) qui, comme le moteur à vapeur au début de la révolution industrielle, l'électricité à la fin du dix-neuvième siècle et l'informatique après la seconde guerre mondiale, se diffuse dans tous les secteurs de l'économie et de la société. Mais l'histoire ne se répète jamais de la même façon. L'IA se distingue radicalement de toutes les autres TPG car ses développements se situent d'emblée à l'échelle internationale et ils sont donc un enjeu de rivalités économiques et géopolitiques entre quelques grandes puissances.

De plus, à rebours de ce que permettrait leur usage socialement maîtrisé afin de satisfaire les besoins de l'humanité, les technologies qui reposent sur l'IA transforment *simultanément* les données en une source d'accumulation de profits, elles renforcent le pou-

<sup>1</sup> David L. Parnas, 2017, « The Real Risks of Artificial Intelligence », <https://dl.acm.org/doi/pdf/10.1145/3132724>



Un monde en guerres.

Claude Serfati.

Éditions Textuel, Collection « Petite Encyclopédie Critique ».

ISBN : 978-2-84597-994-9.

voir sécuritaire des États et elles introduisent de nouvelles formes de guerre grâce à leur utilisation par les militaires. L'IA réduit un peu plus les différences entre les dimensions militaires et civiles des conflits au sein de l'agenda de sécurité nationale des grandes puissances. En somme, elle offre des potentialités d'utilisation

contre des êtres humains dans tous les domaines de leur vie en société en tant qu'ils sont salariés, citoyens et « civils » menacés par les guerres.

Cet article n'aborde qu'une dimension de l'ubiquité des pouvoirs de l'IA. De plus, il laisse de côté les puissants effets de *régénération* du « complexe militaro-industriel américain » opérés par l'importance croissante des GAFAM dans l'industrie de défense, un processus aujourd'hui largement sous-estimé. Ces évolutions majeures sont développées dans *Un monde en guerres*, dans le chapitre « L'intelligence artificielle, au cœur de l'ordre militaro-sécuritaire ».

L'intégration de l'IA dans les doctrines et les équipements militaires marque une étape supplémentaire dans la longue histoire de l'utilisation des technologies à des fins de destruction. L'émergence de systèmes d'armes autonomes est ainsi souvent décrite comme une troisième révolution militaire, après l'invention de la poudre et les armes nucléaires. En effet, elle offre aux militaires de nombreuses opportunités qui se situent sur quatre plans principaux. D'abord, elle améliore les performances des systèmes d'armes existants et rend plus fiables certaines tâches accomplies par les soldats. Les technologies d'interface cerveau-ordinateur, utilisées dans les hôpitaux pour régénérer ou restaurer des fonctions altérées sont déjà utilisées pour des besoins militaires ou de sécurité nationale, soit pour augmenter les capacités physiques des soldats, soit pour des objectifs de pacification, d'interrogation ou de torture (par induction d'une douleur sensorielle sans blessure physique)<sup>2</sup>.

De plus, grâce à l'IA, de nombreuses fonctions telles que les activités de renseignement, de surveillance et de

logistique, les cyberopérations, les opérations de commandement et de contrôle, peuvent ou pourront prochainement être centralisées dans une source unique<sup>3</sup>.

Ensuite, l'IA peut être associée à d'autres technologies émergentes telles que l'informatique quantique, l'impression 3D ou encore les missiles hypersoniques qui volent à plus de six mille kilomètres à l'heure et à trente mille mètres de hauteur. Les modèles actuellement en développement sont lancés à partir de missiles balistiques qui peuvent eux-mêmes atteindre une vitesse de 25 000 kms/h. L'étape en cours, celle de l'alliance de l'IA et du nucléaire, est âprement discutée par les spécialistes dont certains s'inquiètent : « nous sommes proches d'un moment *Oppenheimer* ». En effet, les États-Unis disposent d'un bombardier (le B-21) porteur d'armes nucléaires qui est capable d'exécuter des missions sans équipage et les Russes mettent au point un véhicule sous-marin sans équipage destiné à des frappes nucléaires de représailles<sup>4</sup>.

Ces réalités pèsent beaucoup plus lourd que l'engagement des pays à « maintenir un humain dans la boucle » et un contrôle politique permanent sur les armes nucléaires souhaité par les Nations-Unies ainsi que la déclaration faite en février 2023 au sommet de La Haye par cinquante-sept pays (mais refusée par la Russie et Israël) en faveur d'un « usage responsable de l'IA sur le champ de bataille ». De son côté, le gouvernement français a fait pression pour que le militaire soit exclu de l'*Artificial Intelligence Act* européen<sup>5</sup>.

<sup>3</sup> Daniel S. Hoadley & Nathan J. Lucas, 2018, « Artificial Intelligence and National Security ». Congressional Research Service Reports. <https://miniurl.be/r-5jug>

<sup>4</sup> Michael Depp & Paul Scharre, 2024, « Artificial Intelligence and Nuclear Stability ». <https://warontherocks.com/2019/08/america-needs-a-dead-hand/>

<sup>5</sup> Federico Santopinto, 2024, « L'UE, l'intelligence artificielle militaire et les armes létales autonomes ». IRIS. <https://miniurl.be/r-5jui>

## « Frappes chirurgicales » et « assassinats ciblés » : le grand mensonge

Ainsi qu'un militaire israélien l'a déclaré au magazine d'investigation israélien *+972*, l'utilisation de l'IA permet de générer cent nouvelles cibles par jour contre cinquante par an dans le passé. Et l'infox fait croire, comme lors des prétendues « frappes chirurgicales » contre l'Irak en 2003, à une précision des objectifs ciblés qui, selon ses responsables, fait de l'armée israélienne, « la plus humaine du monde ». Or, son porte-parole a déclaré que les objectifs des bombardements à Gaza, « ne sont pas leur précision, mais l'ampleur du dommage créé ». En somme, l'« intensification algorithmique des destructions » résulte de la combinaison de décisions prises par les militaires et des faiblesses de l'apprentissage machine, car les militaires ne peuvent alimenter les immenses besoins en données exigées par ces machines.

\* <https://miniurl.be/r-5jvk>

\*\* <https://miniurl.be/r-5jvl>

Enfin, et c'est un des objectifs principaux, certains types d'armes créés par l'IA intègrent une prise de décision autonome du système d'armes<sup>6</sup>. Une avancée majeure dans ce domaine est la création de système d'armes létales autonomes (SALA) qui agissent au sol (les « robots tueurs ») dans les airs (les drones) et en mer (par exemple, un navire chasseur de mines qui les détecte et les détruit). Ce sont des systèmes d'armes qui, une fois qu'ils sont activés, sélectionnent et visent des cibles sans

<sup>6</sup> The Mitre Corporation, 2017, « Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD », Office of the Assistant Secretary of Defense for Research and Engineering. <https://irp.fas.org/agency/dod/jason/ai-dod.pdf>. *The Militarization of Artificial Intelligence*, 2019. <https://miniurl.be/r-5juk>

<sup>2</sup> Charles N. Munyon, 2018, « Neuroethics of Non-primary Brain Computer Interface: Focus on Potential Military Applications », *Frontiers in neurosciences*. <https://miniurl.be/r-5juf>

autre intervention d'un opérateur<sup>7</sup>. Ces armes, qui disposent de fonctions joliment qualifiées de « tire et oublie » (*fire and forget*) présentent plusieurs caractéristiques : la capacité à tuer, la capacité de fonctionner sans avoir besoin de l'intervention ou d'un contrôle humain, la capacité d'apprentissage par des interactions avec l'environnement qui leur permettent d'élargir leurs fonctionnalités et, enfin, l'impossibilité d'interrompre une opération une fois qu'elle a été lancée<sup>8</sup>. L'IA – par le truchement de l'apprentissage machine – décuple leurs potentialités dévastatrices, car les algorithmes qui sont intégrés produisent des résultats en partie non maîtrisés, une forme « d'imprévisibilité intrinsèques » (*imprevisibility by design*) sur la décision et le moment d'attaquer, ainsi que sur les conséquences qui en résultent<sup>9</sup>. En sorte qu'une bataille entre des systèmes autonomes ennemis pourrait dégénérer avant que l'intervention d'un militaire puisse empêcher la catastrophe<sup>10</sup>. D'autant plus que le ministère de la défense des États-Unis est favorable à un système décentralisé qui augmente la distance entre l'objet (l'arme) et le sujet (l'opérateur)<sup>11</sup>. Le risque de perte de contrôle humain sur les armes semble assumé : le Pentagone réalise des tests afin de limiter au maximum l'intervention humaine dans les procédures de décision de haut niveau des armes qui utilisent

l'IA<sup>12</sup>. Il faut ajouter que la large diffusion des recherches en IA dans des dizaines de pays facilite la prolifération des systèmes d'armes autonomes et augmente donc les risques d'un engrenage incontrôlé et de leur utilisation « préventive » aux conséquences fatales<sup>13</sup>.

## FRANCE : UN TROPISME MILITAIRE PORTEUR DE DÉCLIN

Deux documents ont défini la stratégie gouvernementale dans le domaine de l'IA. En mars 2018, le rapport remis par Cédric Villani, *Donner un sens à l'intelligence artificielle, pour une stratégie nationale et européenne* fut élaboré à partir de l'audition de 420 experts. Son auteur rappelle qu'« aujourd'hui, cependant, l'écrasante majorité des transferts circule du monde civil vers le monde militaire. C'est d'autant plus vrai dans le domaine de l'intelligence artificielle (IA)<sup>14</sup> ». Le même constat est fait par un responsable de la direction générale de l'Armement (DGA), le bras industriel du ministère en charge de la défense qui constate un « déplacement vers le civil du centre de gravité du développement des technologies, en particulier numériques<sup>15</sup> ». Le rapport publié par le ministère des Armées en 2019, qui constitue l'autre document majeur en matière d'IA, constate également qu'un point commun des stratégies nationales c'est que « les briques technologiques doivent ainsi diffuser [...] du domaine civil au militaire<sup>16</sup> ».

Voilà pour les déclarations. La réalité est toute autre et l'emprise militaire sur l'activité d'innovation se poursuit. Comme l'indique le Tableau 1, dans la phase initiale (2018-2022) des mesures gouvernementales, la défense a reçu presque autant de financement IA que la recherche et plus que l'ensemble de l'« économie » française. Autrement dit, la défense (qui pèse environ 2 % du PIB) a reçu autant que tout le reste de l'industrie et des services.

Cette surreprésentation du militaire va même être amplifiée. Le plan France 2030 prévoit d'accorder 2,2 milliards d'euros à l'investissement dans l'IA afin de soutenir la formation, favoriser la diffusion des technologies de l'IA et cibler quelques domaines prioritaires. Or, ce montant est à peine supérieur à celui affecté par la loi de programmation militaire 2024-2030 à l'IA de défense (2 milliards d'euros<sup>17</sup>). Le gouvernement a également créé, en mai 2024, l'agence ministérielle pour l'intelligence artificielle de défense.

A quoi vont servir ces deux milliards d'euros « militaires » consacrés à l'IA ? Pour l'essentiel, à réaliser ce que la DGA appelle « une captation d'innovation<sup>18</sup> » afin d'adapter les recherches menées dans les laboratoires publics et privés civils aux besoins opérationnels, faire progresser l'autonomie décisionnelle des robots, concevoir les cockpits du futur, etc. « En effet, le secteur de la défense utilise des données issues de capteurs militaires spécifiques (radars, sonars, systèmes optroniques, équipements de guerre électronique) auxquels le secteur civil a peu accès<sup>19</sup> ».

<sup>7</sup> Daniel S. Hoadley & Nathan J. Lucas, 2018, op. Cité dans la note 4.

<sup>8</sup> <https://miniurl.be/r-5jur>

<sup>9</sup> Vincent Boulanin, Neil Davison, Netta Goussac & Moa Peldán Carlsson, 2020. *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*. Stockholm International Peace Research Institute (SIPRI) and ICRC (International Committee of the Red Cross (Comité International de la Croix-Rouge)). <https://miniurl.be/r-5jut>

<sup>10</sup> Boucher Philip, June 2020, « Artificial intelligence: How does it work, why does it matter, and what can we do about it? » European Parliament Research Service (EPRS). <https://miniurl.be/r-5juf>

<sup>11</sup> Sydney J. Freedberg Jr., 9/02/2023, « DoD's clarified AI policy flashes 'green light' for robotic weapons: Experts. » *Breaking Defense*. <https://miniurl.be/r-5jv>

<sup>12</sup> Monomita Chakraborty, 18/05/ 2021, « Pentagon is Attempting to Incorporate AI-controlled Robots into the US Military ». *Analytics Insight*. <https://miniurl.be/r-5juy>

<sup>13</sup> Michael C. Horowitz, 2019. "When speed kills: Lethal autonomous weapon systems, deterrence and stability", *Journal of Strategic Studies*. Lien: <https://doi.org/10.1080/01402390.2019.1621174>

<sup>14</sup> Cédric Villani, 2019, « Les enjeux de l'IA pour la Défense de demain », *Revue Défense Nationale*, 820, 23-29. <https://doi.org/10.3917/rdna.820.0023>

<sup>15</sup> Emmanuel Chiva, 2019. « L'intelligence artificielle : un moteur de l'innovation de défense française. *Revue Défense Nationale*, 820, 33-37. <https://doi.org/10.3917/rdna.820.0033>

<sup>16</sup> Ministère des Armées, septembre 2019. « L'intelligence artificielle au service de la défense ». *Rapport de la Task Force IA*. <https://miniurl.be/r-5jv>

<sup>17</sup> <https://miniurl.be/r-5jve>

<sup>18</sup> Document de référence de l'orientation de l'innovation de défense (DrOID), 2022. <https://miniurl.be/r-5jvf>

<sup>19</sup> Document de référence de l'orientation de l'innovation de défense (DrOID), 2020. <https://miniurl.be/r-5jvg>

Axes de la stratégie nationale	Financements étatiques prévisionnels (2018-2022)	Part du financement total (%)
Recherche	445	29.1
Enseignement supérieur	128	8.4
Transformation de l'action publique	154	10.1
Économie	390	25.5
Défense et sécurité	410	26.9
Total	1527	100

Tableau 1: Comparaison des financements étatiques de l'IA par grands axes.

Source : l'auteur, à partir des données de la Cour des Comptes (2023).

Le mythe du rôle moteur de la recherche militaire s'estompe un peu plus.

La défense n'est pas seulement privilégiée dans l'affectation des crédits publics destinés à l'IA, elle est également mieux organisée. Le ministre des Armées a annoncé qu'il prend pour « modèle ce qui s'est fait avec l'atome dans les années 1960 » pour conduire le saut technologique que représente l'intelligence artificielle<sup>20</sup>. Cette hyper-centralisation exercée par l'institution militaire contraste avec la dispersion des financements publics civils qui transitent par de nombreux canaux mis en place par les institutions de politique publique. Quelles en sont les conséquences ? Dans un rapport publié en 2023, la Cour des Comptes observe une « absence de gouvernance ou de coordination pour l'ensemble des usages de l'IA et des infrastructures critiques associées », un « suivi insuffisant des financements associés » qui risque d'occasionner « des divergences de priorités, des difficultés de diffusion des innovations, ainsi qu'une réduction de l'efficacité des investissements qui peuvent être réalisés de manière dispersée<sup>21</sup> ». Ce constat d'un entrela-

gement et cette opacité des dispositifs publics coûteux mis en place ne sont pas le fruit du hasard, mais un trait structurel du comportement de la haute administration (le « millefeuille administratif » a de beaux jours devant lui) qui caractérise l'ensemble de la politique industrielle<sup>22</sup>.

Le tropisme militaire des politiques publiques en IA affaiblit la recherche publique et l'industrie civile. En effet, les performances scientifiques de la France dans le domaine de l'IA telles qu'elles sont mesurées par le nombre de publications citées<sup>23</sup> se sont détériorées. La France est passée de la cinquième place en 2003 à la huitième en 2013 puis à la douzième place en 2023 ; elle est désormais cinquième en Europe derrière le Royaume-Uni, l'Allemagne, l'Italie et l'Espagne, (données : Scimago). C'est un des effets de l'appauvrissement de la recherche publique civile organisé depuis des années que l'excellence de la recherche en mathématiques n'arrive plus à enrayer.

L'industrie civile connaît un recul bien plus dévastateur. La qualité

de la recherche publique et le soutien de beaucoup d'argent public mis sur la table par la Banque Publique d'Investissement (BPI) ont certes permis à quelques chercheurs de créer des *start-ups* dont certaines sont des leaders mondiaux – Mistral AI étant la plus citée. En revanche, les grands groupes, qui captent pourtant une bonne partie du financement public de R&D, font pâle figure. Thales, qui produit des armes de haute technologie, est le premier déposant de brevets en IA en France, mais, enclavé dans la défense, il n'est pas présent dans le Top 20 des entreprises européennes déposant des brevets IA (source : Questel). En fait, six décennies de politique technologique dominée par les grands programmes militaires et stratégiques ont marginalisé plusieurs secteurs industriels déterminants (électronique, informatique, machine-outil). Les « champions nationaux » peuplent en nombre le cimetière des faillites : Alcatel, Alstom énergie, Areva (rebaptisé Orano). Le plus récent sinistre est celui d'ATOS, vers lequel l'argent public afflue à nouveau – comme ce fut le cas dans les décennies 1960, 1970 et 1980 pour Bull qu'ATOS a d'ailleurs racheté. Il faut en effet, et quoiqu'il en coûte, sauver sa filiale qui produit les supercalculateurs qui forment un maillon essentiel de la dissuasion nucléaire française et qui sont indispensables pour gérer certains systèmes d'armes.

En participant à la course à la militarisation de l'IA, le gouvernement français malmène la recherche publique et dégrade un peu plus le système industriel.

<sup>20</sup> Les Echos, 8 mars 2024. IA : « Soit l'armée française prend date, soit elle décroche ». <https://miniurl.be/r-5jvh>

<sup>21</sup> Cour des Comptes, avril 2023, « La stratégie nationale de recherche en intelligence artificielle - Une stratégie à structurer et à pérenniser ». <https://miniurl.be/r-5jvi>

<sup>22</sup> Voir Catherine Sauviat et Claude Serfati, décembre 2023, « Un bilan des politiques industrielles en France - Focus sur deux secteurs, l'industrie pharmaceutique et les télécoms et sur l'industrie 4.0 ». Rapport de l'Agence d'objectifs CFE-CGC n° 2021-2. <https://ires.fr/?s=sauviat+serfati&submit=>

<sup>23</sup> Je n'engage pas ici le débat sur les limites des indicateurs bibliométriques.

# IA, des technologies pour les services publics ?

A l'heure où le gouvernement français promeut et finance une IA souveraine au service de la croissance, qu'en est-il du déploiement de son usage dans la fonction publique ?

**Hervé Christofol**

Membre du bureau national du  
SNESUP-FSU

Depuis vingt ans, la productivité du travail progresse peu en Europe. D'après certains économistes, elle a même diminué – notamment en France où une des raisons en serait le développement massif de l'alternance<sup>1</sup> qui aurait fait intervenir, sur le marché du travail, des travailleurs en cours d'apprentissage, donc moins performants que les travailleurs expérimentés...

Le rapport *IA notre ambition pour la France*<sup>2</sup> avance deux hypothèses quant aux conséquences économiques du développement des usages de l'IA. Soit « *la croissance économique annuelle de la France pourrait doubler grâce à l'automatisation de certaines tâches. Au bout de dix ans, la hausse de PIB serait comprise entre 250 et 420 Md€<sup>3</sup>, et [...] une fois l'IA adoptée par l'ensemble du tissu économique, il n'y aurait plus de gains de productivité à attendre* ». Soit l'IA « *pourrait induire une augmentation permanente du taux de croissance de l'économie. Autrement dit, en plus d'un effet temporaire lié à l'automatisation, l'IA pourrait produire un effet à plus long terme lié à l'émergence de*

*nouvelles innovations, de nouveaux produits, de nouvelles formes d'organisation...* ». Dans les deux cas, les bénéfices économiques attendus sont importants : d'une part, ils positionnent le développement de l'IA délibérément dans le paradigme d'une croissance infinie sans remettre en cause la nécessité – en particulier écologique – de changer de modèle économique (extractiviste, productiviste et consumériste); et d'autre part, les impacts sociaux de ce développement sont peu discutés, l'idée implicite étant que le marché de l'emploi s'adaptera et les individus suivront...

## DES IA AU SERVICE DES USAGERS DES SERVICES PUBLICS ?

Alors qu'une des conséquences de la numérisation de l'accès aux services publics, notamment aux processus administratifs, a été la déshumanisation et l'accentuation de la fracture numérique en France, l'une des trois priorités de la politique de déploiement de l'IA exposée lors du 8<sup>e</sup> Comité interministériel de la transformation publique (CITP), est de « *ré-humaniser le service public grâce à l'IA* ». Pour cela,



le gouvernement compte d'une part « *s'appuyer sur l'IA pour permettre aux agents de répondre plus rapidement et plus efficacement aux demandes des usagers, en ligne et aux guichets* » et, d'autre part, « *doter les agents d'une IA conversationnelle* ». Le rapport de ce 8<sup>e</sup> Comité interministériel « *Débureaucratiser : pour des services + proches + simples + humains*»<sup>4</sup> liste les engagements du gouvernement pour simplifier les démarches administratives et améliorer les services publics.

<sup>1</sup> Éric Heyer, 2023, « Un effondrement de la productivité des salariés français en trompe-l'œil », *The Conversation*. <https://miniurl.be/r-5jym>

<sup>2</sup> *IA notre ambition pour la France*, rapport remis le 15 mars 2024, au Premier ministre par la Commission de l'intelligence artificielle. Cette commission – groupant trente membres et rapporteurs nommés à titre personnel pour leur expertise – avait été installée en septembre 2023 par le gouvernement. Elle était présidée par Philippe Aghion et Anne Bouverot. <https://miniurl.be/r-5jyn>

<sup>3</sup> Les gains générés par l'IA augmenteraient de 1,35 % par an. De tels gains de productivité pendant dix ans conduiraient à une hausse du PIB comprise dans une fourchette allant de 250 Md€ à 420 Md€ en 2034, soit l'équivalent de la valeur ajoutée de l'industrie dans son ensemble !

<sup>4</sup> Rapport du 8<sup>e</sup> Comité interministériel de la transformation publique, avril 2024 : *Débureaucratiser : pour des services + proches + simples + humains*. <https://miniurl.be/r-5jyo>

Des expérimentations sont en cours et les premiers résultats publiés. Selon, Grégoire Tiro, chargé de la préfiguration du service Innovation et Participation citoyenne à la Délégation interministérielle à la transformation publique (DITP), « *l'expérimentation d'un outil d'IA d'aide à la formulation de réponses à des questions d'usagers sur la plateforme Services publics + a permis de diviser le délai de réponse par deux : de six jours, on est passé à trois jours, avec une appréciation croissante de la part des agents qui utilisent les réponses générées\** ».

\* Cité dans Stratégie d'usage de l'intelligence artificielle en matière de gestion des ressources humaines dans la fonction publique d'État. <https://miniurl.be/r-5jyp>

La priorité « *Mettre l'IA et le numérique au service des Français* » du rapport présente « *des cas d'usage s'appuyant sur l'intelligence artificielle qui seront développés dans toutes les administrations à partir de 2024, en capitalisant sur des expérimentations déjà menées pour améliorer les conditions de travail des agents et répondre plus rapidement aux usagers* ». Ainsi, l'application Contact'IA va être expérimentée à la direction générale des Finances publiques (DGFIP) pour prérédiger des réponses à destination des agents, à partir des bases de connaissances internes, afin de renseigner les seize millions de demandes annuelles de particuliers et de professionnels. Elle a été expérimentée par 1700 agents qui « *considèrent les réponses utiles dans 75 % des cas, ce qui a fait passer le temps moyen de réponse de dix-neuf jours à trois jours en moyenne* ».

L'application, Albert France Services, expérimentée depuis début 2024 auprès de soixante agents volontaires dans vingt-six France Services pilotes, vise

à aider ces derniers à retrouver et organiser plus facilement l'information issue des sources de la documentation administrative et des fiches pratiques du site officiel de l'administration française, Service-Public.fr. Quant à l'application IA'ccueil, elle assistera le gendarme responsable de la mission d'accueil du public au sein des brigades de gendarmerie en compilant les données issues de plusieurs sources (documentation professionnelle, mémorial, MaSécurité, service-public.fr).

Au-delà de la précision des réponses apportées aux usagers et du temps de réponse pour y parvenir, qui peuvent indéniablement être appréciés, nous peinons à comprendre en quoi ces IA peuvent contribuer à « *ré-humaniser le service public* », pour ne pas parler de l'engagement de « *ramener les services publics sur le terrain* ». La communication gouvernementale est mystérieuse et « *quand c'est flou, c'est qu'il y a un loup* ».

### **DES IA EN ASSISTANCE AUX AGENTS DE LA FONCTION PUBLIQUE ET AU SERVICE DE LEUR PRODUCTIVITÉ ?**

En complément des applications précédemment décrites, les agents devraient se voir doter d'une IA conversationnelle sous la forme d'un contact directement accessible sur Tchapp<sup>5</sup> afin de « *faciliter et accélérer les tâches du quotidien telles que la rédaction de mails, la structuration de notes, la correction de texte, ou encore la génération d'idées* ».

Nous l'avons compris, toutes ces IA ont pour principale vocation d'augmenter la productivité horaire des agents et la lutte pour les gains

de productivité doit être engagée. L'intensification du travail et l'augmentation de la qualification requise pour effectuer les différentes missions doivent bénéficier aux agents et ne pas uniquement conduire à la réduction du nombre de fonctionnaires et à la baisse de leurs rémunérations. Bien sûr, les tâches répétitives ne sont pas valorisantes mais elles permettent également à l'esprit de se reposer ce qui est nécessaire pour, dans un autre temps, être créatif, synthétique, constructif ou critique et pour prendre du recul. Et puis, au nom de quoi sinon la généralisation du taylorisme et du néolibéralisme faudrait-il que nos activités soient toujours optimisées au maximum de nos capacités ? À moins, encore une fois, de diminuer drastiquement le temps de travail sans perte de revenu.

Relevons que, si d'importants investissements sont envisagés dans le rapport de mars 2024 *IA notre ambition pour la France*, à savoir vingt-sept milliards d'euros en cinq ans, celui-ci n'envisage, à aucun moment, de consacrer un euro supplémentaire à la formation des agents pour « *faire du dialogue social et professionnel un outil de co-construction des usages et de régulation des risques des systèmes d'IA*<sup>6</sup> ». Il est vrai qu'il existe déjà un « Campus du numérique public », lancé le 9 janvier 2024, qui a déjà dispensé des modules de formation à 2 150 agents publics et formé 220 directeurs d'administrations centrales et chefs d'inspection aux enjeux du numérique (0,04 % des 5,7 millions d'agents dont 2,5 millions dans la Fonction publique de l'État), comme le rappelle le rapport du CITP d'avril 2024. Mais l'objectif de former 100 % des agents de l'État d'ici fin 2026, rappelé dans ce rapport, est ambitieux et nécessitera des budgets conséquents.

<sup>5</sup> Tchapp (<https://www.tchap.gouv.fr>) est la messagerie instantanée du secteur public. Conçue et gérée par l'Administration française pour les agents des trois fonctions publiques, pour communiquer facilement en toute sécurité, elle est utilisée par plus de 400 000 agents publics.

<sup>6</sup> Recommandation n°3, page 125 du rapport.

Il est envisagé de recourir à l'IA en matière de *recrutement* (rédaction de fiche de poste, évaluation des compétences des candidats et prédiction de leur réussite potentielle, construction de parcours d'intégration personnalisés en luttant contre les biais cognitifs), de *formation* (formuler des recommandations contextuelles de formation, développer l'apprentissage personnalisé...), de *fidélisation* (analyser des objectifs et des résultats, analyser [surveiller ?] des retours d'information d'agents et d'utilisateurs, planifier des remplacements de postes, identifier des compétences cachées (sic)...), de *qualité de vie au travail (QVT) et de santé mentale* (analyser des données de mesure du bien-être des agents [quelles sont-elles ?], analyser des relations entre collègues [les surveiller ?], établir des programmes personnalisés d'actions en faveur du bien-être des agents, mesurer le retour sur investissement des actions en faveur du bien-être, détection précoce des problèmes de santé mentale [par qui ? des médecins ou des managers ?]). Selon Daniel Gerson\*, cité dans le rapport de la DGAFP, « *L'IA est aujourd'hui utilisée dans les fonctions publiques pour de la reconnaissance de données et de motifs de comportements à des fins de prévision et de production de supports d'interaction avec l'humain* ». Cela fait tout de même un peu froid dans le dos.

\* Daniel Gerson, cité p.10, est chef de projet « Gestion et emplois publics » au sein de la Direction de la gouvernance publique de l'OCDE.

## UNE APPLICATION IA POUR PRÉVOIR LE DEVENIR DE MÉTIERS DANS LES RH

Le récent rapport de la Direction générale de l'administration et de la fonction publique (DGAFP)<sup>7</sup> relève des enjeux et des opportunités de l'utilisation de l'IA en matière de gestion des ressources humaines (RH) dans la fonction publique.

<sup>7</sup> *Stratégie d'usage de l'intelligence artificielle en matière de gestion des ressources humaines dans la fonction publique d'État*, juin 2024, rapport de la Direction générale de l'administration et de la fonction publique (DGAFP), ministère de la transformation et de la fonction publiques, 18 pages. <https://miniurl.be/f-5jyp>

S'agissant, en particulier, des métiers de la filière RH, le rapport cite une étude prospective conduite avec ChatGPT-IV<sup>8</sup>. Quel vertige de confier à une application IA le soin de prévoir le devenir de métiers dont les missions pourraient être réalisées ou assistées via l'IA... Quoi qu'il en soit, selon GPT-4, quatre métiers devraient voir 80 % de leurs activités remplacées du fait de l'émergence de l'IA dès 2025-2026 – avec une probabilité avancée de 90 % (sic) – : assistant administratif, gestionnaire de paie, recruteur et analyste en rémunération. Quatre autres métiers pourraient voir 80 % de leurs activités remplacées du fait de l'émergence de l'IA mais seulement à partir de 2034, avec une probabilité moindre de 40 % : responsable RH, coach interne, conseiller en organisation du travail et responsable de projet attractivité/engagement des agents. Rappelons que, déjà en 2019, l'OCDE prévoyait que « *dans quelques années seulement, il sera possible de libérer près d'un tiers du temps des fonctionnaires, leur permettant ainsi de passer de tâches banales à un travail à haute valeur ajoutée*<sup>9</sup> ». Comme le rappelait le physicien danois Niels Bohr, « *L'art de la prédiction est une chose difficile – surtout lorsqu'elle concerne le futur* ».

Après ces annonces, le rapport de la DGAFP alerte froidement que « *l'utilisation de l'IA en milieu professionnel nécessitera aussi un accompagnement des agents dans la gestion de la pression numérique dans lequel les métiers de la prévention*

<sup>8</sup> Rapport « IA et RH : quand le futur du travail est en avance », *Tomorrow Theory*, juin 2023. <https://tomorrowtheory.com/rapport-ia-rh/>

<sup>9</sup> *Recommandation sur l'intelligence artificielle*, OCDE, 2019. <https://miniurl.be/f-5jyq>

(psychologues du travail, ergonomes, médecins du travail...) auront un rôle primordial ». Ainsi, le déploiement de l'IA pourrait avoir des impacts sanitaires et sociaux susceptible de créer de l'emploi dans les domaines de la santé et du médico-social... Perspectives qui ne freinent pas l'enthousiasme de nos haut-fonctionnaires.

Les principaux bénéfices attendus par les experts du rapport sont, pour les managers, un gain de temps, d'analyse, d'optimisation, d'harmonisation et d'anticipation prospective. Mais le rapport relève également des risques liés à l'usage de l'IA dans la gestion des RH : « *sensibilité des données utilisées par l'IA, au sens du Règlement général pour la protection des données (RGPD); déresponsabilisation managériale; RPS et tensions sociales; persistance des biais cognitifs issus des programmations algorithmiques défaillantes; perte d'attractivité des postes liée à la déshumanisation des services RH* ». Ce qui incite les auteurs à proscrire les usages liés à la « *détection précoce des problèmes de santé mentale; l'analyse des relations entre collègues; l'entretien vidéo automatisé avec les candidats; les systèmes de reconnaissance et de récompense* », mais à recommander le développement des autres activités à travers trois outils utilisant l'IA pour assister les missions RH : un coach virtuel pour les agents (formation et QVT), un assistant managérial numérique pour les managers « augmentés » (recrutement, fidélisation) et un outil numérique d'appui à la GRH pour les gestionnaires RH (recrutement, formation, fidélisation, QVT). Pages 13 et 14, le rapport se veut rassurant en indiquant que :

- les managers devront sensibiliser leurs agents à un usage éthique et transparent de l'IA ;
- les responsables RH devront protéger l'agent et l'administration via un usage légal et transparent des don-

« Un an après le lancement de ChatGPT, 77 % des Français déclarent que l'IA est une vraie révolution\*, mais 68 % pensent qu'il faudrait faire une pause dans le développement de l'IA\*\*. Nous ne sommes d'ailleurs pas les seuls, puisque 79 % des Chinois et 74 % des Américains y sont favorables. »

\* Étude en ligne réalisée par Ipsos pour Sopra Steria en octobre et novembre 2023 auprès de mille personnes représentatives de la population française âgée de 18 ans ou plus.

\*\* Étude en ligne réalisée par Ipsos pour AXA en mai et juin 2023 auprès de 3 226 experts de cinquante pays et de 19 000 personnes dans quinze pays, représentatives de la population nationale âgée de 18 ans et plus.

nées personnelles et des applications d'IA robustes et sécurisées ;

- le cadre d'usage devrait être fixé en collaboration avec les représentants du personnel et l'agent devrait avoir la liberté d'utiliser des outils d'IA ou leurs résultats ;

- la déontologie devra être respectée (transparence, justice, équité, diversité, inclusion, non-discrimination).

Nous le constatons, les expérimentations sont en cours. Malgré les risques (notamment psychosociaux), un contrôle accru des agents et une déshumanisation croissante des services, le gouvernement et les administrations

ministérielles souhaitent avancer vite dans le déploiement des applications IA dans la fonction publique, pour intensifier le travail des agents et potentiellement pour réduire le nombre de fonctionnaires (même si Gabriel Attal s'en

était défendu dans sa conférence de presse d'avril).

Si certains bénéfices pour les usagers peuvent être mis en avant, avec notamment une réduction du temps de réponse à leurs questions administratives et si certains agents pourraient voir leur qualification et leurs conditions de travail s'améliorer, cela ne se fera pas sans l'instauration d'un rapport de force pour que les salariés ne cèdent pas leur autonomie, leurs libertés et l'intégralité des gains de productivité attendus face à la gestion par l'IA. Le dialogue social ne suffira pas à contenir les appétits des directions générales actuelles. Alors que l'extrême droite est aux portes du pouvoir, les potentialités de surveillances et de contrôle des agents *via* les IA devraient nous faire réfléchir.



# *Le tournant - ce que nous voulons pour l'enseignement supérieur et la recherche*

Au lendemain des élections législatives anticipées, le SNESUP-FSU, le SNCS-FSU, le SNASUB-FSU et le SNEP-FSU rappellent les principes au fondement de leur action et leur attachement à un service public de l'enseignement supérieur et de la recherche (ESR) de qualité pour toutes et tous.

## **Caroline Mauriat**

Co-secrétaire générale du SNESUP-FSU

## **Boris Gralak**

Secrétaire général du SNCS-FSU

## **Julie Robert**

Co-secrétaire générale du SNASUB-FSU

## **Pascal Anger**

Secrétaire national du SNEP-FSU



## **UN BUDGET À LA HAUTEUR DES BESOINS**

Pour la FSU, le financement récurrent doit être largement prioritaire et augmenté pour couvrir les missions de recherche et d'enseignement. Il doit également être réparti équitablement entre les organismes de recherche, les établissements et au sein des établissements, entre composantes.

Alors que la loi de programmation de la recherche (LPR) prévoit le financement de plus de quinze mille précaires supplémentaires, la FSU demande au contraire :

- la création immédiate de douze mille postes de titulaires toutes catégories confondues (personnels techniques, administratifs et de bibliothèque, chercheur·e·s, enseignant·e·s-chercheur·e·s, ensei-

Au lendemain des élections législatives anticipées, la FSU rappelle les éléments d'urgence qu'elle porte pour le service public de l'enseignement supérieur et de la recherche (ESR) :

- abandonner les expérimentations en cours : « choc de simplification » avec la transformation des organismes nationaux de recherche en agences de programme, établissements publics expérimentaux (EPE), « acte II » de l'autonomie des universités ; et en faire un bilan ;
- revaloriser les rémunérations et les carrières de tous les personnels de l'ESR pour les porter au niveau de celles des corps comparables de la fonction publique d'Etat ;
- reporter la formation des enseignant·e·s et ouvrir de véritables discussions associant toutes et tous les acteurs et les actrices concerné·e·s ;
- abroger la loi liberté et responsabilité des universités (LRU) ;
- stopper les mesures austéritaires et augmenter le budget récurrent pour le service public de l'ESR ;
- résorber la précarité et améliorer l'encadrement des étudiants par un plan de recrutement massif de titulaires ;
- permettre l'accès de toutes et tous les bachelier·e·s à la formation de son choix ;
- rendre effectif le droit à la poursuite d'études pour tous les diplômé·e·s de licence dans un master en cohérence avec leur parcours de licence ;
- abroger la taxe CVEC (contribution vie étudiante et de campus) payée par les étudiants·e·s et la compenser intégralement en augmentant les dotations aux universités dans le budget du ministère de l'enseignement supérieur et de la recherche (MESR).

gnant·e·s) pour permettre de remédier aux inégalités d'encadrement constatées entre établissements et la mise en place d'un plan pluriannuel d'emplois statutaires de six mille postes par an jusqu'en 2035, ce qui représente un effort de 6 700 millions d'euros ;

- une augmentation du budget de la recherche publique de six milliards d'euros d'ici 2027 pour atteindre 1 % du PIB ;
- une remontée de la dépense moyenne par étudiant·e en université (un peu plus de dix mille euros en 2021) pour converger vers celles des sections de techniciens supérieurs, STS (près de quinze mille euros) et des classes préparatoires aux grandes écoles, CPGE (plus de seize mille euros) ;
- une augmentation du budget du programme 150 de la Mission interministérielle Recherche et Enseignement supérieur (MIRESE) de deux milliards d'euros (hors immobilier) d'ici 2027.

C'est seulement dans ces conditions que les objectifs du livre blanc de l'enseignement supérieur et de la recherche 2017 – 3% du PIB pour la R&D et 2 % du PIB pour l'enseignement supérieur – pourront être respectés. Ce plan d'investissement doit s'accompagner d'un programme de réhabilitation et de construction immobilières.

## **RÉSORPTION DE LA PRÉCARITÉ ET AUGMENTATION DES SALAIRES**

**Pour la FSU, un service public de l'enseignement supérieur et de la recherche pour construire un avenir plus juste et durable, se doit d'améliorer les conditions de travail permettant à chacun de s'émanciper dans l'exercice de son métier et garantissant l'égalité de parcours, de promotion et de rémunération entre les femmes et les hommes.**

Cela passe par :

- la mise en place d'un vaste plan de titularisation de toutes les précaires (toutes catégories de personnels confondus) employé·es pour répondre à des besoins pérennes ;
- le recrutement sur poste statutaire au plus proche de l'obtention du diplôme donnant accès au concours ;
- le rattrapage de la perte du pouvoir d'achat des fonctionnaires et la hausse immédiate du point d'indice de 10% ;
- le déblocage des carrières pour tous les personnels et la facilitation des avancées en matière de grade et de changement de corps ;
- un plan de requalification des emplois BIATSS de C en B et B en A ;
- la revalorisation salariale des métiers de l'enseignement supérieur et de la recherche, notamment par l'intégration des primes dans les salaires et la reconnaissance de la qualification du doctorat dans les grilles indiciaires, des rémunérations sans modulation individualisée et reconnaissant le niveau de diplôme ;
- la réduction des obligations statutaires de service d'enseignement à 125 HTD pour les enseignant·e·s-chercheur·e·s et 250 HTD pour les enseignants·e·s, ainsi que des heures supplémentaires d'enseignement, notamment pour dégager du temps effectif de recherche ;
- l'abrogation de la réforme des retraites que le gouvernement a fait passer en force en 2023 et le retour à l'indexation des retraites sur les salaires et des dispositions pour réparer les dommages des réformes des retraites passées.

## **DÉVELOPPER LA RECHERCHE PUBLIQUE**

**Pour la FSU, la production de connaissances, dans tous les champs disciplinaires et interdisciplinaires, leur libre diffusion à l'ensemble de la société et le**

**respect de la liberté académique sont indissociables de l'indépendance d'une recherche scientifique publique.**

Ces valeurs d'indépendance doivent être garanties par le statut de la fonction publique pour les personnels de l'ESR, par une évaluation et une politique scientifique effectuées par des pairs majoritairement élus, par des dotations de base des établissements permettant de financer leurs laboratoires et de conduire leur propre politique scientifique, par une organisation de la recherche soutenable et par un partenariat équilibré entre la recherche publique, la recherche privée et la société.

Cela passe par :

- la suppression du crédit d'impôt recherche (CIR) et la réorientation des effets d'aubaine qu'il produit (7,24 Md€ en 2021) vers les établissements publics ;
- l'augmentation du financement pérenne de la recherche en supprimant ou en transformant les structures du mille-feuille institutionnel (ANR, Idex, ComUE, IHU, IRT, SATT, Hcéres...);
- la réduction drastique de la part du financement par appels à projets au profit des dotations aux laboratoires ;
- la création d'un millier de postes de doctorant·e·s contractuel·e·s par an pendant sept ans, avec un recrutement en tant que titulaire au plus proche de la thèse ;
- la fin de la politique actuelle de gel des postes, et la publication automatique des postes libérés, alors que le taux de départ à la retraite va augmenter très rapidement d'ici 2032, en particulier chez les enseignant·e·s-chercheur·e·s dans le champ des sciences dites « dures » ;
- un investissement dans la formation des futur·e·s enseignant·e·s et enseignant·e·s-chercheur·e·s, indispensable pour permettre ces embauches.

## DES FORMATIONS PUBLIQUES ÉMANCIPATRICES, GRATUITES ET ACCESSIBLES À TOUTES ET TOUS

Pour la FSU, la démocratisation de l'accès à l'enseignement supérieur, lieu d'émancipation par les savoirs, de vie sociale et de développement de la pensée critique, demeure une impérieuse nécessité. Concrètement, il s'agit de viser l'objectif de 60 % d'une classe d'âge diplômée de l'enseignement supérieur dont près de la moitié de diplômé·e·s de master.

Cela passe par :

- des formations publiques diplômantes, gratuites et accessibles à tous et toutes, en formation initiale comme en formation continue (avec prise en compte de l'expérience professionnelle pour les salarié·e·s et les demandeur·se·s d'emploi) ;
- l'abrogation de Parcoursup et l'augmentation des places dans les formations tenant compte des aspirations des bachelier·e·s et garantissant l'accès de droit dans les parcours de master compatibles avec le parcours de licence de l'étudiant·e ;
- un taux d'encadrement ne dépassant pas vingt étudiant·e·s par enseignant·e ou enseignant·e-chercheur·se titulaire et la garantie d'une formation à et par la recherche dès le premier cycle ;
- un plan de construction de véritables campus universitaires publics (logements, équipements culturels et sportifs et restauration CROUS) permettent une proximité d'accès à des universités de plein exercice ;
- une allocation universelle pour chaque étudiant·e pour subvenir dignement à ses besoins.

Pour démocratiser et permettre une formation universitaire complète et équilibrée, la pratique physique et sportive doit devenir un droit pour tou·te·s dans l'ESR. Les formations aux métiers du sport ou les activités physiques sportives et artis-

tiques (APSA) avec leur dimension recherche doivent être soutenues et valorisées.

## URGENCE ÉCOLOGIQUE

Pour la FSU, le service public de l'enseignement supérieur et de la recherche est un acteur incontournable des questions de société liées à l'urgence écologique. Pour répondre aux enjeux, nous avons besoin d'un service public de l'ESR redynamisé et solidaire, redonnant espoir et reconnaissance à ses différents acteurs et financé par des crédits publics.

Cela passe par :

- un plan de financement de l'État après audit bâtiminaire précis pour permettre l'isolation thermique des bâtiments et l'amélioration des systèmes de chauffage-climatisation décarbonés, essentielles pour diminuer l'émission des gaz à effet de serre (conformément au décret tertiaire<sup>1</sup>) ;
- un investissement de sept milliards d'euros dans la rénovation énergétique des bâtis universitaires avant 2030 et seize milliards d'euros d'ici 2050.

**Pour la FSU, le recours au numérique, souvent présenté comme une alternative aux déplacements, ne peut pas être une solution durable à généraliser.**

- l'enseignement à distance, facteur de grandes inégalités entre les étudiant·e·s et briseur des collectifs de travail, doit rester l'exception ;
- les outils numériques ne doivent ni être imposés ni utilisés pour remettre en cause les libertés pédagogiques, et ne doivent pas gommer les spécificités disciplinaires (sorties sur le terrain, travaux pratiques...) ;
- la part des déplacements domicile-travail peut être réduite en permettant les mutations pour rappro-

chement de conjoint, en développant des universités de proximité, en limitant les déplacements intersites et en préférant les mobilités douces et les transports collectifs.

## ORGANISATION, GESTION DÉMOCRATIQUE DES UNIVERSITÉS

Pour la FSU, l'enseignement supérieur et la recherche doivent pouvoir se développer au service des besoins sociaux indépendamment des forces politiques, économiques et religieuses.

La FSU est indéfectiblement attachée au statut de fonctionnaire conçu pour mettre l'agent·e à l'abri des pressions et garantir l'égalité d'accès des usagers aux services publics. Il l'est également à la collégialité dans les procédures et les missions des universités, laquelle se conjugue avec la dimension collective de la liberté académique.

Cela passe par :

- l'abrogation de la loi relative aux libertés et responsabilités des universités (LRU), et la recentralisation de la masse salariale ;
- l'abandon des expérimentations, notamment celles qui consistent à la mise en place d'établissements publics expérimentaux (EPE) et celles prévues dans l'acte II de l'autonomie des universités ;
- le retour au droit commun et aux statuts d'université du Code de l'éducation dans leur ensemble pour les établissements publics expérimentaux créés depuis 2018, qui accentuent les inégalités, détruisent la collégialité, réduisent la démocratie et s'éloignent des missions et des valeurs des services publics ;
- la mise en place d'un véritable fonctionnement démocratique des universités en donnant à la communauté universitaire, et notamment à ses élu·e·s dans les conseils, la capacité de peser sur leur avenir ;
- la capacité pour la communauté

<sup>1</sup> <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000038812251>

universitaire de donner un avis conforme *via* le Conseil national de l'enseignement supérieur et de la recherche (CNESER) sur l'offre de formation, le budget, les textes réglementant l'ESR ;

- le maintien de toutes les compétences décisionnelles du Conseil national des universités (CNU) et du Comité national de la recherche scientifique (CoNRS).

La FSU demande la fin des dispositifs dérogatoires et des dispenses de toutes sortes à la qualification et au recrutement d'enseignant·e·s-chercheur·e·s.

## **RENFORCEMENT DE LA LUTTE CONTRE LES VSS**

Les révélations de multiples faits de harcèlement sexuel, d'agressions sexuelles et de viols se sont accrues depuis l'année universitaire 2020-2021. Ces violences, constatées dans les universités comme dans les grandes écoles, touchent les personnels et les étudiant·e·s : tandis qu'une étudiante sur vingt a été victime de viol et une sur dix d'agression sexuelle, les statistiques manquent toujours du côté des personnels, les directions d'établissement demeurant réticentes à engager des poursuites disciplinaires. Une forme d'impunité règne toujours à l'université.

- *Le plan d'action national de lutte contre les VSS* dans l'ESR présenté en octobre 2021 met l'accent sur la formation et la sensibilisation de la communauté universitaire, le renforcement des dispositifs de signalement et de leur fonctionnement, le renforcement de la communication sur les dispositifs existants et sur les VSS et la valorisation de l'engagement des étudiant·e·s et des personnels. Or, le MESR n'envisage l'aide financière que par le lancement d'appels à projets, ce qui est clairement contradictoire avec l'ambition affichée.

- De la même manière, les discussions engagées ces derniers mois sur le plan national d'action pour l'égalité professionnelle dans le MESR n'ont donné lieu à aucunes avancées ni propositions d'action susceptibles d'améliorer la situation des femmes dans l'ESR.

- Pour la FSU, le recul des VSS dans l'ESR implique de faire peser sur tous les établissements une obligation de résultat, au-delà des seules recommandations, tout en les dotant de moyens pérennes dédiés à la lutte contre les VSS, de les enjoindre à rendre compte à la communauté (notamment devant des conseils élus) de leur action, de faire évoluer les procédures disciplinaires et de sensibiliser et former tous les personnels.

- La FSU rappelle l'impérieuse nécessité de recenser et de centraliser les informations venant des établissements universitaires concernant les VSS.

- Une clarification, une harmonisation et une réforme de la procédure disciplinaire sont également indispensables, la victime étant toujours considérée comme « témoin » de l'affaire.

## **UNE FORMATION DES ENSEIGNANT·E·S QUI REDONNE DU POUVOIR D'AGIR AUX ACTEURS ET ACTRICES**

Pour lutter contre le désinvestissement, les risques psychosociaux et préserver la formation des générations futures, il est urgent :

- de reporter la réforme engagée sur la formation et le recrutement des enseignants des premier et second degrés ;

- de redonner le pouvoir d'agir aux acteurs et actrices de la formation.

La FSU continue de défendre les conditions d'une véritable formation

des enseignant·e·s universitaire et professionnelle, qui nécessite :

- la sécurisation des parcours des étudiant·e·s par le biais d'un financement des études sans contrepartie de service (« prérecrutements » au sens de la FSU) ;

- la constitution d'équipes de master de l'enseignement, de l'éducation et de la formation (MEEF) recrutées sur des critères et selon des procédures universitaires de détermination des profils de poste ; des moyens pour former et faire exister des équipes plurielles ; le respect des conditions universitaires d'élaboration des maquettes : c'est aux équipes de master de les élaborer en réponse aux besoins de formation identifiés ; une formation continue qui vise le développement professionnel et non la transmission de prescriptions qui asservit ;

- des stages étudiants déterminés en fonction des impératifs de formation (progressivité, lieu, encadrement) et non des besoins de l'Éducation nationale ou des plafonds d'emploi. Pour garantir la dimension formatrice de l'expérience professionnelle, les berceaux de stage doivent être en surnombre dans les écoles et établissements ; la possibilité, pour les tuteurs et tutrices de « terrain », d'accompagner les étudiants sur leur temps de service (décharge), d'être formés (formation de formateurs), d'impulser ou de participer à des recherches collaboratives ;

- la détermination et le fléchage des budgets des INSPÉ, composantes universitaires à dimension académique ;

- la mise en place d'instances démocratiques au sein des INSPÉ.

# Le salariat, une force en temps de Covid

La crise sanitaire déclenchée par le Covid a été révélatrice de la puissance des institutions du salariat. Alors que le droit du travail est remis en cause au nom d'une disparition du salariat prophétisée depuis les années 1990, l'enquête sur sept secteurs d'activité menée pendant la pandémie démontre à quel point ces institutions – de la sécurité sociale au syndicalisme – se sont avérées indispensables pour protéger les travailleurs et permettre la continuité des activités économiques. C'est ce que raconte le livre *Face au covid, l'enjeu du salariat* présenté ici par son coordinateur.

## Claude Didry

Sociologue, directeur de recherche au CNRS, centre Maurice-Halbwachs, laboratoire de recherche en sciences sociales de l'École normale supérieure

La très grave crise politique que nous connaissons, avec un vote pour le Rassemblement national atteignant un niveau record qui suit la politique austéritaire des années 2022-2023 (double réforme du chômage en 2021-2024, réforme des retraites en 2023), évoque parfois le mouvement de 2019 des Gilets jaunes, avec une forte ambiguïté politique qui penche aujourd'hui manifestement vers la droite extrême. Elle renvoie également à la rupture que la pandémie a représentée par rapport à deux décennies d'austérité intense imposée par les institutions européennes et relayée par les majorités qui se sont succédées en France, en prenant conscience des possibilités ouvertes par le « *quoi qu'il en coûte* » pour faire face à un changement climatique de plus en plus rapide et envisager une restauration du salariat comme sécurité de vie pour la plus grande partie de la population active. C'est sur ce point que nous voudrions revenir, à partir des éclairages tirés de notre ouvrage<sup>1</sup>.

<sup>1</sup> Cet ouvrage est tiré d'un rapport réalisé pour la Direction de l'animation de la recherche, des études et des statistiques du ministère du Travail (DARES), à partir d'une enquête (2022-2023) intégrant le traitement des données statistiques tirées d'ACE-MO-Covid (sur les évolutions des organisations du travail au cours de la crise) et des entretiens avec des syndicalistes et des DRH pris comme grands témoins des décisions, des aménagements du travail qui sont intervenus et des attentes des salariés.

L'épidémie du Covid est arrivée dans un contexte particulièrement détérioré du fait des réformes du financement de la sécurité sociale et des hôpitaux, réformes qui se sont succédées à partir de la loi Juppé de 1996 qui a créé les « lois de financement de la sécurité sociale ». À partir de 2007, la tarification à l'acte dans les hôpitaux a succédé aux subventions annuelles ; une loi de 2009 a créé les Agences régionales de santé. En 2020, les hôpitaux, le système de soins sont sous une tension permanente, voire au bord du gouffre ; les personnels sont dans une grande souffrance dès avant le Covid, comme en témoigne le mouvement social hospitalier qui accompagne les mobilisations contre la première réforme des retraites.

Le rationnement des ressources sur le long terme que connaît le système de soins est redoublé par la rigueur budgétaire qui s'est imposée aux États de l'Union européenne depuis la crise financière de 2008, et que symbolise le traité sur la stabilité, la coordination et la gouvernance (TSCG) de 2012. Cela se traduit par une action spécifique de la Commission européenne, dite « *semestre budgétaire* », qui consiste en un examen des budgets publics des États donnant lieu à des

recommandations pour atteindre un équilibre. Un volet important de cette politique est la baisse des dépenses sociales en visant, par exemple, à agir sur le chômage par des réformes du droit du travail dans le sens d'une facilitation des licenciements ou encore de



*Face au covid, l'enjeu du salariat.* Sous la direction de Claude Didry. Avec les contributions de Juan Sebastian Carbonell, Damien Cartron, Vincent-Arnaud Chappe, Camille Dupuy, Denis Giordano, Jules Simha, Karel Yon et Samuel Zarka, Paris, Éditions de la Dispute, 2023.

## La grande absence du CHSCT

Le comité d'hygiène, santé, sécurité et conditions de travail (CHSCT), créé par la loi du 23 décembre 1982, était un organe élu par les salariés (dans lequel siégeaient de droit l'inspecteur et le médecin du travail) qui avait pour mission de contribuer à la protection de la santé et de la sécurité de ceux-ci, ainsi qu'à l'amélioration des conditions de travail. Il a été reconnu comme une personne morale, capable d'aller en justice. Il était obligatoire dans les établissements de plus de cinquante salariés, avec un CHSCT central et des comités d'établissement en cas d'entreprises pluri-établissements. Il devait être consulté par l'employeur avant toute mesure d'aménagement ayant un effet sur l'hygiène, la santé et la sécurité. Il était également consulté pour le maintien au travail de salariés handicapés ou victimes d'accidents du travail. Il participait aux enquêtes suivant les accidents du travail, les maladies professionnelles (liées à des risques sanitaires, tels que l'exposition à des produits toxiques, et aux risques psychosociaux). Il devait rendre un avis sur le rapport annuel faisant le bilan des risques, ainsi que sur le programme annuel de prévention des risques remis par les employeurs. Enfin, il pouvait lancer un droit d'alerte en cas de danger grave et imminent obligeant les employeurs à une enquête avec ses membres.

Il a été supprimé par l'ordonnance relative à la nouvelle organisation du dialogue social (promulguée en 2018). Un Comité social et économique\* compense la disparition du CHSCT par la mise en place d'une commission hygiène, santé, sécurité et conditions de travail dans les seules entreprises de plus de trois cents salariés, en dehors des entreprises classées Seveso où cette commission doit être systématiquement créée. Cela a conduit à un affaiblissement considérable du nombre d'élus du personnel et de la capacité d'acquérir une connaissance approfondie des conditions de travail et des enjeux de santé et sécurité dans les entreprises que rendait possible l'existence d'une instance dédiée.

Le programme du Nouveau Front populaire a mis à son agenda un plan d'action « *zéro mort au travail* » qui passe par le rétablissement du CHSCT, soutenu par une embauche massive d'inspecteurs et de médecins du travail.

\* Voir dossier de la VRS n°430, 2022 intitulé « Quelle nouvelle organisation du travail dans la recherche publique ? » l'article de Marie-Hélène Luçon et Laurence Rouillon « Dix ans de CHSCT dans la fonction publique, et demain ? » ainsi que celui de Dominique Cau-Bareille et Gérard Grosse « L'activité du CHSCT en temps de confinement ». Lien: [https://snscs.fr/wp-content/uploads/2022/10/VRS430\\_HD.pdf](https://snscs.fr/wp-content/uploads/2022/10/VRS430_HD.pdf)

recentrage de la négociation collective au niveau de l'entreprise au détriment de la négociation de branche, avec un parallélisme important des législations mises en œuvre dans les pays du sud de l'Europe (Italie, Espagne, Portugal et France). Citons, pour la France, la loi El Khomri de 2016 qui porte plutôt sur la priorité accordée à la négociation d'entreprise.

Mais les ordonnances de 2017 représentent une rupture bien plus forte, en affaiblissant les garanties des salariés en cas de licenciement injustifié avec la création d'un barème indemnitaire et en réduisant de moitié le nombre de représentants du personnel. Le comité social et économique (CSE)

remplace le comité d'entreprise, le comité hygiène, santé, sécurité et conditions de travail (CHSCT) disparaît, sa mission étant reprise, de manière optionnelle, par une commission du CSE (obligatoire pour les seules entreprises et établissements de plus de trois cents salariés). C'est dire combien nous étions mal préparés pour faire face à la pandémie de Covid.

### CONFINEMENTS ET ACTIVITÉ ÉCONOMIQUE

Quand la pandémie s'est déclarée, le confinement est devenu la seule solution pour limiter les contaminations et limiter la saturation des services hospitaliers face à la situation dégra-

dée de ceux-ci. C'est ainsi qu'au cours des deux mois qu'a duré le premier confinement, les activités économiques ont subi un bouleversement inédit pour limiter les déplacements professionnels aux seuls salariés<sup>2</sup> des services dits « essentiels » (hôpitaux, magasins d'alimentation), soit, selon le ministère du Travail, environ un tiers de l'emploi privé.

Sur la période de deux ans qu'a duré la crise sanitaire (2020-2022), l'élément qui nous a interrogés à la lecture des données du ministère, c'est que, après le fort recul des deux mois de confinement, les activités économiques ont repris à un niveau assez proche de la normale en dépit de la succession des différentes « vagues ».

Ce constat fait apparaître, selon nous, la puissance du salariat stable (CDI et fonction publique) qui représente près de 70 % de la population active, dans la mesure où celui-ci est devenu la base du travail à distance et du chômage partiel, avec de fortes pertes d'emplois pour les intérimaires et les titulaires de CDD. Les statuts d'indépendants ont démontré leur faiblesse, du fait d'un effondrement de l'activité et donc des revenus, ce qui a ébranlé le mythe de l'efficacité de l'« ubérisation » du travail.

Lors du premier confinement (mars - mai 2020), le télétravail a marqué les esprits par son ampleur et par l'importance de ce dispositif dans le groupe des cadres. Le terme de télétravail est ici trompeur, dans la mesure où il s'agit plutôt d'un « *travail à distance* » (comme le nomment certains acteurs dans les télécoms et les banques) prévu dans les cas de crise pandémique depuis une loi de 2012 adoptée dans le contexte lié à la menace du virus H1N1.

On a beaucoup moins parlé de l'« activité partielle » (nouvelle dénomination

<sup>2</sup> Dans cet article, le masculin inclut le féminin sans discrimination.

tion du « chômage partiel », qui a concerné, lui aussi, près d'un tiers de la population en emploi, avec une indemnisation revalorisée à 80 % du salaire brut (contre 70 % auparavant). Comme nous avons pu le constater dans les monographies, le chômage partiel a rencontré de nombreuses difficultés dans son déploiement. Ainsi, au-delà de son évidence dans le secteur des hôtels, cafés et restaurants, sa mise en œuvre a suscité des tensions dans l'industrie et la construction. Rappelons que la ministre du Travail d'alors, Muriel Pénicaud, est allée jusqu'à qualifier de « *défaitiste* » la volonté de fermer les chantiers, exprimée par la CAPEB (Confédération de l'artisanat et des petites entreprises du bâtiment).

### VISIBILISATION DES SALARIÉS ET DE LEURS REPRÉSENTANTS

Enfin, le présentiel a été, en tant que tel, source de tension en l'absence d'équipements de protection et de gel hydroalcoolique, mais avec un faible niveau d'exercice du « droit de retrait », en dehors des travailleurs Amazone soutenus par une action en justice du syndicat SUD qui a débouché sur un arrêt favorable de la Cour de cassation. Il a pris une forme spécifique dans l'hôpital public au cœur de la crise, par sa vocation de recherche et sa capacité à affronter un virus inconnu. Le présentiel a, d'une certaine manière, révélé les soignants et les aides-soignants, qui constituent la fonction publique hospitalière occultés, la plupart du temps, par les médecins, éclatés en une multitude de statuts hospitaliers, universitaires et libéraux. Si les plans de crise ont mené à un management resserré autour des personnels de direction, les CHSCT, que la loi de modernisation de la fonction publique de 2019 n'avait pas encore effacés, ont joué pleinement leur rôle pour organiser le redéploiement des personnels soignants dans le cadre d'une organisation beaucoup plus horizontale pour être plus réactive.

Plus généralement, au-delà du secteur hospitalier où le personnel est relativement bien syndiqué, la crise épidémique a mis les représentants syndicaux dans une position cruciale de référence pour des salariés pris entre le marteau de l'angoisse de la contamination et l'enclume des pertes de salaires et des incertitudes sur l'emploi.

Il en ressort un rôle très important des représentants syndicaux et des acteurs des relations professionnelles (représentants des salariés-directions des ressources humaines) au moment du confinement dans les mesures de mise à l'abri des personnels. Mais si les activités ont pu reprendre au niveau que l'on a constaté, cela tient également à une capacité d'expression de salariés qui ont été, pour certains d'entre eux, très éprouvés par le confinement du fait, notamment, du chômage partiel. La perte des frais de route ou de chantier, dans les travaux publics, des heures supplémentaires dans les secteurs industriels, ont creusé les ressources pour les personnels ouvriers. Dans le secteur du cinéma et de la production audio-visuelle, la fin des tournages a laissé les personnels particulièrement désorientés.

Dans ces secteurs, l'action de la branche va être très importante, à travers l'expertise d'organismes tel que le CHSCT de branche dans le cinéma et l'audiovisuel. Les représentants syndicaux vont ensuite être très sollicités pour suivre l'application et le perfectionnement des protocoles sanitaires tout au long de la période, avec l'arrivée des masques puis celle des vaccins. Dans le domaine des soins hospitaliers, ce qui peut apparaître comme une discrimination vaccinale entre soignants qui se sont vus proposer une vaccination par Astrazéneca et médecins sous vaccination Pfizer, a accompagné une polémique sur l'innocuité de ces vaccins que les syndicats ont dû affronter.

### FUITE DES LIEUX DE TRAVAIL

Les sept portraits sectoriels que nous proposons permettent de dégager ce que la pandémie a fait durablement au travail, ou plutôt à l'organisation du travail. Je suis tenté ici de rapprocher paradoxalement le secteur hospitalier de celui des télécoms et de la banque, dans la mesure où c'est un désir de prendre ses distances avec le lieu du travail qui se fait jour de la part des salariés. Dans le cas de l'hôpital, c'est la journée de douze heures qui en est le vecteur en conduisant à des semaines de trois jours (soit trente-six heures satisfaisant aux trente-cinq heures légales). Dans les deux autres, cela passe par la demande de télétravail portée par les salariés. Ces organisations nouvelles du travail présentent le risque d'un surmenage des salariés dans le cas de l'hôpital, lié à un risque de déstructuration de la collectivité de travail que déplorent les représentants syndicaux et certains médecins. Dans le même temps, les DRH ont surmonté leurs craintes, en jouant de la réponse à ce désir des salariés pour multiplier les heures supplémentaires et limiter les embauches à l'hôpital, ou en développant le *flexoffice* qui réduit le nombre de postes sur le lieu du travail sur la base de l'existence d'un volant récurrent de personnel en télétravail. Il faut dire que, dans le cas du télétravail, le travail à distance imposé par les autorités publiques a permis d'expérimenter la sécurité des transferts de données sensibles (données clients pour les banques et les télécoms, données R&D dans l'automobile) et, ainsi, de vaincre les réticences des directions. Finalement, on peut se demander si la fuite des lieux du travail que semble avoir suscité la pandémie ne correspond pas à une manifestation française de cette « *unofficial general strike* » que l'ancien secrétaire du Travail et économiste Robert Reich

sous Bill Clinton<sup>3</sup>, constatait à travers la désertion du travail aux États-Unis.

Mais de quelle organisation les salariés tentent-ils de se dégager ? La question de l'entreprise se pose véritablement dans le cas de la restauration rapide, où la marque dissimule une multitude de restaurants franchisés démultipliant les pratiques managériales face aux risques épidémiques, avec parfois le maintien de l'activité au mépris de la santé des salariés. Dans le secteur des télécoms, c'est à l'inverse une centralisation très forte de l'entreprise qui se dessine autour d'établissements intégrant une multitude d'agences et d'entités locales à des échelles régionales, voire nationale. Le recours systématique au télétravail montre également la faible emprise technologique de la société, où une sous-traitance se prolongeant jusqu'à la pose de lignes par des microentrepreneurs prend en charge les infrastructures Internet et téléphonique. On comprend que le sens du travail devienne problématique, pour des salariés dispersés dans une multitude d'entités filiales, franchisées ou sous-traitantes.

## RECOURS MASSIF AUX PRIMES

Si l'éclatement de l'entreprise multiplie les situations de blocage dans des univers professionnels restreints, le recours à des primes et aux heures supplémentaires qui s'est prolongé et démultiplié au cours de la crise sanitaire tend à anesthésier la faiblesse des carrières salariales. Dans le cas des personnels hospitaliers, il s'engage dès le premier confinement. Ainsi, dans une décision du 5 mars 2020, le premier levier actionné par le ministère de la Santé pour affronter le surcroît d'activité lié à la mon-

tée de l'épidémie en situation de sous-effectif est le dé plafonnement du nombre d'heures supplémentaires. Sans véritable innovation, la stratégie choisie par les directions hospitalières joue sur la faiblesse des salaires indiciaires. Elle tient à une sous-reconnaissance des qualifications dans la filière du soin et est renforcée par le gel du point d'indice imposé dans la fonction publique au cours des deux dernières décennies (en dehors des années 2016 et 2017).

Avec la crise sanitaire, les primes deviennent également la marque d'une politique des revenus soucieuse de ne pas consolider les rémunérations mensuelles, en particulier dans le secteur hospitalier. Ce secteur connaît de longue date un système de primes très complexe qui suscite une forte bureaucratie gestionnaire. La crise s'accompagne d'un enrichissement de ce système, avec une prime Covid de 1500 euros à partir de mai 2020. C'est encore une prime qui sort de ce « Ségur de la santé », censé apporter une réponse au malaise durable des personnels hospitaliers au lendemain du premier confinement. Au terme d'un bras de fer avec les organisations syndicales, cette prime atteint 183 euros pour l'ensemble des agents. La situation de l'hôpital est emblématique du blocage des carrières que l'on rencontre au sein des fonctions publiques, en témoigne l'enseignement supérieur post loi pour la recherche, et les branches professionnelles. Elle conduit également à une interrogation sur les possibilités de déverrouiller certains passages de grades comme, par exemple, le passage à un grade de médecin pour le personnel infirmier.

## PROMESSE DE RÉINDUSTRIALISATION

Les pénuries de médicaments, d'équipements puis de composants électroniques, que l'économie

française connaît aujourd'hui de manière chronique, ont pourtant suscité la promesse d'une réindustrialisation. La reprise économique post-Covid a encouragé le redressement d'une filière automobile prête à poursuivre une politique de délocalisation initiée depuis plus de trois décennies. Au niveau de l'industrie dans son ensemble, cependant, ce que semblait permettre un plan de relance soutenu par les institutions européennes semble aujourd'hui avoir fait long feu. La fermeture de ses dernières unités de recherche sur le territoire national par Sanofi apparaît emblématique du passage à une économie de rentiers tirant sa subsistance de ses filiales dans les pays à faible coût de main-d'œuvre. Beaucoup de déceptions ont suivi la crise pandémique, avec le retour au *business as usual* que constituent, pour un personnel politique issu des grandes écoles de la technocratie (Science Po, ENA devenue Institut national du service public), la chasse aux « sureffectifs » et le *cost killing*.

Face au verrouillage austéritaire des politiques sociales, la crise du Covid montre le besoin de sortir des dogmes de la profitabilité sans autre finalité pour les activités productives. Elle met particulièrement en évidence la tragique erreur que la suppression du CHSCT a représenté, face à laquelle il a fallu une intervention exceptionnelle des représentants syndicaux pour organiser la reprise du travail dans des conditions sanitaires acceptables.

Mais, de manière plus fondamentale, on est en droit de se demander si, en mutilant gravement la démocratie au travail par la suppression de près de la moitié des mandats professionnels, les ordonnances de 2017 et la politique qui s'en est suivie ne sont pas les ferments d'une haine de la démocratie qui pousse aujourd'hui notre pays vers les eaux dangereuses de l'extrême-droite.

<sup>3</sup> Voir « Is America experiencing an unofficial general strike? », *The Guardian*, 13 octobre 2021.

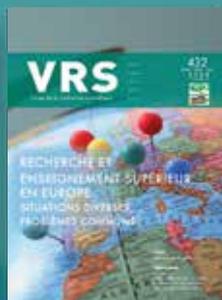
# VRS



La vie de la recherche scientifique



La Vie de la recherche scientifique (VRS) explore les grandes questions scientifiques et politiques en lien avec les préoccupations de la société et des mouvements sociaux. Retrouvez, au fil des dossiers, les grands sujets qui sont au coeur de vos interrogations et de vos exigences.



## ABONNEMENT ANNUEL • 4 NUMÉROS PAR AN

INDIVIDUEL : 25€     INSTITUTIONNEL : 50€     PRIX AU NUMÉRO : 8€

**ATTENTION : l'abonnement est facultatif pour les adhérents du SNCS et du SNESUP**

Institution : .....

Nom : ..... Prénom : .....

Adresse : .....

Téléphone : ..... Courriel : .....

Merci de bien vouloir nous retourner ce bulletin complété avec votre règlement à l'adresse suivante :

SNCS - Campus CNRS d'Ivry sur Seine - 27, rue Paul Bert - 94200 Ivry sur Seine

sncs@cnr.fr - www.sncs.fr

**casden**



BANQUE POPULAIRE

La banque coopérative  
de la Fonction publique

CASDEN Banque Populaire - Société Anonyme Coopérative de Banque Populaire à capital variable - Siège social : 1 bis rue Jean Wiener 77420 Champs-sur-Marne - Siren n° 784 275 778 - RCS Meaux - Immatriculation ORIAS n° 07 027 138 - BPCE - Société anonyme à directoire et conseil de surveillance au capital de 180 478 270 euros - Siège social : 50, avenue Pierre Mendès France - 75201 Paris Cedex 13 - Siren n° 493 485 042 - RCS Paris - Immatriculation ORIAS n° 08 045 100 - Crédit photo : © Roman Jehanno - Conception : Insign 2020 - Réf: AP Groupe 2020 - Merci à Jenny, enseignante chercheuse d'avoir prêté son visage à notre campagne de communication.

# COMME MOI, REJOIGNEZ LA CASDEN, LA BANQUE DE LA FONCTION PUBLIQUE !

*Jenny, Enseignante chercheuse*



PARIS 2024



PARIS 2024



PARTENAIRE PREMIUM

[casden.fr](https://casden.fr)



Retrouvez-nous chez

**BANQUE POPULAIRE**