

RAPPORT

Le service public à l'épreuve de l'intelligence artificielle

_Émilie Agnoux et Johan Theuret (coord.)
_Daniel Agacinski _Gilles Babinet
_Marius Bertolucci _Gabrielle du Boucher
_Anca Caruntu _Diana Filippova
_Yann Ferguson _Matthieu Girier
_Corinne Narassiguin _Soizic Pénicaud
_Jacques Priol _Thaima Samman



Sens du service public

Fondation
Jean Jaurès
ÉDITIONS

Introduction

Garder la main sur nos destins

– Émilie Agnoux

Cofondatrice du Sens du service public, experte associée à la Fondation Jean-Jaurès, autrice de *Puissance publique. Contre les démolisseurs d'État*, l'Aube/Fondation Jean-Jaurès, 2025

– Johan Theuret

Cofondateur du Sens du service public et directeur général adjoint de la Ville et Métropole de Rennes

Changement de monde ?

Actuellement, impossible d'échapper à toutes les études annonçant régulièrement de nouveaux records en matière d'intelligence artificielle (IA). Les États ne sont pas en reste et engagent des milliards de fonds publics, comme au dernier Sommet de l'IA à Paris en février 2025 pour répondre aux enjeux de l'innovation.

Pourtant, notre société prend progressivement conscience que l'usage des technologies numériques n'a rien de neutre, tant sur un plan géopolitique (l'entité qui détient et maîtrise la technologie détient le pouvoir), de santé publique (on peut se référer aux débats sur la nocivité des écrans et des réseaux sociaux), que de cohésion sociale (les « smombies¹ » courbés sur leurs smartphones).

L'IA n'est pas un sujet nouveau, pas plus dans le secteur public que dans le reste de la société. De nombreuses entités publiques mobilisent des technologies d'IA depuis plusieurs années (recrutement, gestion de l'eau, de l'énergie, des déplacements ou de l'espace public, agents conversationnels, trafic routier...). Au-delà des dispositifs déployés, des acteurs

publics organisent des conventions citoyennes sur l'IA et élaborent des stratégies politiques permettant de développer une vision stratégique et de poser un cadre de régulation.

Un récent numéro de *La Gazette des communes*² met en valeur les initiatives portées par des entités publiques. L'Observatoire Data Publica recense également plusieurs centaines d'expérimentations. Contrairement à certaines idées reçues, le secteur public n'est pas en retard par rapport au secteur privé. Mais l'IA ne peut s'y déployer de la même manière tant leur objet politique diffère. C'est un modèle d'IA d'intérêt général qu'il convient de bâtir³. Des cas d'usage existent et méritent d'être évalués avec une vue à 360 degrés pour déterminer les conditions optimales et souhaitables de déploiement de ce type de technologies.

La percée fulgurante de l'IA générative dans nos vies quotidiennes a porté un nouvel éclairage sur la mobilisation de ces outils dans le pilotage, la conception et la mise en œuvre de l'action publique. Toutes les échelles (locale, nationale) et tous les services publics sont concernés.

L'IA est pleine de promesses, mais aussi d'interrogations, voire de risques. Elle suscite aussi bien de

1. Néologisme pour désigner les personnes absorbées par leur smartphone dans l'espace public. Voir Hubert Beroche, *Smombies. La ville à l'épreuve des écrans*, La Tour d'Aigues, l'Aube, 2025.

2. « Intelligence artificielle », *La Gazette des communes*, n°25/2771, semaine du 16 au 22 juin 2025.

3. Émilie Agnoux, *Puissance publique. Contre les démolisseurs d'État*, La Tour-d'Aigues/Paris, l'Aube/Fondation Jean-Jaurès, 2025.

l'enthousiasme, de l'inquiétude que de l'indifférence. C'est pourquoi ce sujet mérite d'être abordé avec nuances, dans une approche pluridisciplinaire, d'autant qu'un certain nombre de ses effets ne seront perceptibles qu'à moyen terme.

Garder la maîtrise publique : un enjeu démocratique et de souveraineté

L'impasse budgétaire dans laquelle nous nous trouvons pourrait conduire à renforcer certains biais de surconfiance dans le technosolutionnisme et la tentation de n'aborder l'IA que sous l'angle des potentiels gains de productivité¹. Cette vision n'intègre pas, d'une part, le coût réel du déploiement et de la maintenance de ce type de technologies et, d'autre part, n'apprécie pas le risque de perte de maîtrise publique dans la conduite des affaires d'intérêt général, *a fortiori* si cette tendance devait se traduire par la disparition de savoir-faire internes et d'expertises.

S'il est indispensable d'abord de comprendre et de maîtriser la technique, l'IA interroge davantage le projet politique et le modèle de société que nous voulons fonder. Face aux enjeux d'efficacité du secteur public, elle érige une problématique démocratique. C'est notamment le sujet de la transparence, plus ou moins poussée (information du traitement par une IA, accès au code source, connaissance des algorithmes...) et de l'appropriation citoyenne des usages de l'IA qui est ainsi débattu dans ce rapport, au travers de points de vue personnels complémentaires.

Le cadre juridique érigé ces dernières années au niveau européen et national pour accompagner ces grandes transformations fait encore l'objet de débats, dont ce rapport est le reflet, car l'IA interroge le rapport à la puissance publique, les droits des citoyens et la place des acteurs privés fournisseurs de solutions numériques. L'IA reconfigure les rapports de

force et mérite d'être appréhendée comme une dimension essentielle de notre souveraineté.

L'IA crée de nouvelles dépendances : aux pays dotés en matières premières, à ceux qui fabriquent les matériaux, à ceux qui disposent de moyens en recherche et en innovation, aux fournisseurs de solutions ou de centres de données... Elle peut aussi créer de nouvelles dépendances à la machine elle-même, voire accroître la vulnérabilité de nos organisations publiques en démultipliant les points d'entrée propices aux cyberattaques. Enfin, son coût écologique n'est pas anodin et doit être considéré dans l'équation, quand bien même elle serait au service de la transition écologique.

L'IA nous enjoint donc à renoncer à la facilité et à la naïveté. L'automatisation de l'administration nous oblige à refonder une doctrine de l'action publique et à actualiser le cadre conceptuel.

Faut-il interdire, généraliser, accompagner, réguler ou bien encore hybrider ? Autant de questions plus politiques et démocratiques que techniques.

L'IA est, au final, une opportunité pour débattre et trancher de sujets longtemps mis en sommeil, à commencer par la redevabilité des acteurs publics ou encore la compréhension des mécanismes de fonctionnement et de décision institutionnels. Qui prend la décision ? Sur quels fondements ? Avec quels recours possibles ?

Clarifier le projet politique : l'IA au service de l'humain, et pas l'inverse

L'IA interroge les modalités d'intervention de la puissance publique dans l'espace social.

Nous commençons tout juste à prendre la mesure des conséquences politiques et pratiques des démarches de dématérialisation des dernières décennies qui,

1. Johan Theuret, « Intelligence artificielle dans le service public : attention aux fausses illusions ! », *Le Nouvel Obs*, 27 mai 2025.

si elles ont simplifié les démarches d'un grand nombre de Françaises et de Français, ont aussi renforcé des formes d'éloignement de l'État et des difficultés d'accès aux dispositifs¹.

À ce stade, l'IA pourrait tout aussi bien résoudre ces problématiques (par exemple à travers « l'administration proactive » qui vise à anticiper les besoins et à ouvrir automatiquement des droits aux bénéficiaires) que les aggraver (par exemple par l'introduction ou le renforcement de certains biais dans l'accès aux dispositifs publics ou le non-accès effectif à un recours humain).

L'acceptabilité sociale conditionne le succès de toute politique publique. Pour autant, le débat ne peut se limiter aux quelques exemples régulièrement cités (Parcoursup, l'algorithme de la CAF, la vidéosurveillance algorithmique...) pour embrasser la complexité de l'objet.

Il apparaît primordial de placer la question de la qualité de la relation humaine au centre de la réflexion, en renouant au passage avec une préoccupation philosophique ancienne quant à la place du progrès technologique dans nos vies, qui façonne un certain rapport au monde, aux autres, à soi-même et, en l'occurrence, aux institutions publiques.

D'après le baromètre Data Publica, 80 % des sondés préfèrent avoir accès à un humain, même si cela ne

peut se faire qu'aux horaires d'ouverture des guichets, contre 20 % qui préfèrent avoir un accès à une messagerie instantanée avec une IA qui peut répondre à n'importe quelle heure².

Naturellement, les métiers et compétences publics se trouvent mis en tension, comme dans tout secteur professionnel, l'enjeu étant d'anticiper, de décider de la trajectoire et d'accompagner les changements sur le temps long, en reclarifiant la place de nos agents publics, qui ne peuvent être simplement appréhendus comme des superviseurs, des médiateurs ou des recours en dernière instance. Il nous faudra pour cela définir, enfin, ce que la société attend de ses services publics.

Régulièrement mise à distance dans le cadre du dialogue social, l'utilisation des technologies pour concevoir et mettre en œuvre le service public doit désormais y occuper une place centrale, au risque que les évolutions échappent également aux principaux concernés.

Les échéances électorales des deux ans à venir au niveau local et national seront des occasions de clarifier les différents projets politiques en matière d'IA. Mais n'attendons pas pour en faire dès à présent un objet de débat et de formation à l'intérieur et à l'extérieur de nos organisations publiques.

1. « Un tiers des adultes ont renoncé à effectuer une démarche administrative », *Insee Études*, n°267, 11 mai 2022.

2. *Les Français et leur confiance dans la gestion des données par les acteurs publics*, Ipsos pour Civiteo, Data Publica, juillet 2024.

Quinze propositions

La volonté d'optimiser et l'attente d'efficacité nous conduisent à aller vite, toujours plus vite. Si les services publics ne peuvent rester en dehors de cette tendance, car l'intelligence artificielle (IA) est dans nos vies, ils doivent en déployer des usages avec vision, méthode, clairvoyance. Rappelons au préalable que tout ceci nécessite des moyens humains, financiers, techniques et logistiques, à toutes les échelles d'action (locale, nationale, européenne), qui doivent être précisément évalués et garantis pour garder la maîtrise publique sur les évolutions en cours et à venir. Si le cadre juridique mérite encore d'évoluer pour mieux encadrer les usages, garantir de nouveaux droits aux citoyens et tenir compte des évolutions

rapides des technologies, des actions très concrètes peuvent d'ores et déjà être entreprises.

Le déploiement de l'IA ne doit pas être motivé par la recherche d'économies budgétaires ou de remplacement des agents publics. Si elle peut permettre d'augmenter leur productivité, elle doit libérer du temps pour qu'ils se concentrent sur des tâches à plus forte valeur ajoutée tout en améliorant la qualité des services aux usagers et les conditions de travail des agents. Ce rapport en appelle à un déploiement réfléchi de l'IA afin de ne pas reproduire certains errements de la dématérialisation des démarches administratives.

1.

Obliger à l'adoption de stratégies politiques et de doctrines claires de mobilisation de l'IA dans l'action publique

Faire en sorte que toutes les administrations locales et nationales se dotent d'une feuille de route stratégique et éthique de l'IA, rendue publique et facilement accessible.

2.

Désigner des entités référentes pour piloter la stratégie IA

Faire de l'IA une politique publique transversale incarnée par des organismes et des personnalités en capacité de défendre une vision stratégique et de la décliner opérationnellement (nouvelle agence ou extension de prérogatives d'entités existantes, *chief AI officer*, élu référent en collectivité...).

3.

Renforcer la recherche publique pour une IA d'intérêt général

Cibler une part significative du budget sur la recherche en IA d'intérêt général, soutenir les initiatives et les réseaux et encourager la création de chaires IA et service public dans les grandes écoles et universités.

4.

Garantir le caractère inclusif des IA utilisées dans les services publics

Privilégier des IA entraînées de manière inclusive et accessible à toutes et tous, à commencer par les personnes âgées, les personnes en situation de handicap ou peu à l'aise avec le numérique.

5.

Former systématiquement, et en préalable à la mobilisation des outils d'IA, les élus et les agents publics à l'usage, à la compréhension et aux limites de l'IA

Déployer rapidement un plan national de formation pour tous les élus et agents publics concernés, afin qu'ils puissent utiliser, superviser ou critiquer les outils d'IA de manière éclairée.

6.

Rendre effective la supervision humaine dans tous les processus assistés par IA

Instaurer une garantie de validation préalable par une personne compétente avant toute décision individuelle fondée sur un traitement algorithmique, en particulier dans les domaines sensibles (santé, justice, allocations). Il faut donc aller plus loin que le droit à l'information et à la transparence en instaurant une validation préalable par un agent humain avant que la décision ne prenne effet.

7.

Rendre effective l'obligation d'information des citoyens en cas de recours à des processus algorithmiques et expliciter leurs finalités

Obliger la publication des algorithmes et des critères de traitement lorsqu'ils influencent les décisions administratives en veillant à l'application effective de cette transparence. Obliger les administrations à fournir aux citoyens une explication compréhensible lorsqu'une décision les concernant est prise par un algorithme. Imposer des audits indépendants réguliers sur les biais, les erreurs ou les dérives potentielles des algorithmes utilisés.

8.

Favoriser l'usage de modèles d'IA frugaux et de petite taille dans les services publics

Privilégier les modèles légers, spécialisés et économes en énergie plutôt que les grands modèles généralistes coûteux et énergivores. Consommant moins d'énergie car ils nécessitent moins de puissance de calcul et des serveurs moins performants, ces modèles peuvent être très efficaces pour des tâches ciblées (classer des documents, répondre à des questions simples, automatiser des démarches administratives).

9.

Investir dans la création de modèles d'IA souverains et ouverts adaptés aux besoins publics

Soutenir le développement de modèles d'IA français ou européens, *open source* et entraînés sur des données publiques non sensibles.

10.

Protéger strictement les données publiques utilisées dans l'entraînement ou le fonctionnement des IA

Construire les conditions opérationnelles d'une maîtrise des usages des données publiques et s'assurer que les données soient utilisées dans le cadre d'algorithmes conformes à l'intérêt général. Encadrer tout usage commercial des données produites ou détenu par les services publics (transparence sur les usages des données d'entraînement des IA).

11.

Imposer des études d'impacts environnementaux, éthiques et sociaux, avant tout déploiement d'IA

Évaluer systématiquement l'ensemble des risques (discrimination, exclusion, emploi ou consommation énergétique) liés à chaque usage.

12.

Systématiser des dispositifs pérennes de participation citoyenne autour des usages de l'IA

Généraliser les dispositifs de type conventions citoyennes, conseils citoyens, forums délibératifs ou plateformes de consultation sur les usages de l'IA dans les services publics.

13.

Créer un observatoire public indépendant de l'usage de l'IA dans les administrations

Mettre en place une instance nationale qui recense, audite et analyse les usages de l'IA dans le secteur public, en lien avec les usagers et les agents, sur le modèle de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (OBVIA).

14.

Inclure des clauses environnementales et sociales dans tous les marchés publics liés à l'IA

Obliger les fournisseurs à justifier l'empreinte énergétique, la traçabilité et la réversibilité des données ainsi que l'impact social de leurs solutions. Exiger que les prestataires d'IA démontrent une consommation énergétique maîtrisée et imposer des critères d'inclusion dans la conception des systèmes.

15.

Identifier et protéger les compétences humaines stratégiques menacées par l'automatisation

Cartographier les métiers et savoir-faire exposés aux effets de substitution de l'IA et construire des plans de sauvegarde ou de transformation des compétences.

Connaître et former :

de l'utilité de maîtriser

au préalable de quoi on parle

Gérer les risques et accompagner les usagers : l'IA comme opportunité pour l'action publique

– Anca Caruntu

Directrice Affaires européennes, cabinet d'avocats Samman

– Thaima Samman

Associée fondatrice, cabinet d'avocats Samman

Comme la prose de Monsieur Jourdain, nous utilisons déjà tous l'intelligence artificielle (IA), dernier développement de l'ère numérique, dans notre vie quotidienne, pour nos transports, nos achats en ligne, nos rendez-vous médicaux...

Son usage donne lieu à un cadre législatif et réglementaire qui peut être divisé en deux catégories : les problématiques transversales, mais pas forcément propres, au numérique – protection des données, cybersécurité, propriété intellectuelle... – et celles de nature catégorielle – secteur d'activité, par exemple. Le Règlement sur l'intelligence artificielle (RIA) qui encadre depuis le 1^{er} août 2024 la mise sur le marché, la mise en service et l'usage des systèmes d'IA dans l'Union européenne n'est, par conséquent, qu'une des composantes de ce cadre.

Si pour beaucoup l'IA, comme le numérique en général, paraît inaccessible en raison de sa nature technologique et d'un cadre légal complexe, l'erreur serait de se laisser impressionner, comme le malade imaginaire par les propos en latin et en grec de ses médecins. Finalement, point n'est besoin d'être ingénieur et juriste pour maîtriser les concepts permettant aux acteurs publics d'en tirer les principaux

avantages, pour autant qu'on prenne le temps de s'y intéresser et de s'y former.

De quoi l'IA est-elle faite ?

Derrière le terme générique d'IA, se cache une technologie reposant sur une combinaison de ressources – données, algorithmes et infrastructures de calcul –, dont la qualité, l'accès et le contrôle conditionnent à la fois ses performances et ses effets.

De la donnée

Les données sont le carburant de l'IA dont le volume et la qualité conditionnent la performance des systèmes. Il n'existe pas d'IA fiable sans données fiables. Le traitement des données fait depuis plusieurs années l'objet d'un encadrement législatif et réglementaire par l'Union européenne : le Règlement général sur la protection de données (RGPD, 2016)¹ qui en restreint l'accès, la directive européenne « Police-Justice² » (2016), mais aussi des textes comme le

1. Règlement (UE) 2016/679 du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE.

2. Directive (UE) 2016/680 du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales et à la libre circulation de ces données, et abrogeant la décision-cadre 2008/977/JAI du Conseil.

Règlement sur le libre flux des données non personnelles (2018)¹, le Règlement sur la gouvernance des données (2018)², la Directive sur l'ouverture des données du service public³ (2019) ou le Règlement sur les données, sur l'utilisation, la réutilisation, l'accès ou le partage des données, personnelles ou non personnelles⁴ (2023). L'enjeu désormais est d'intégrer cet ensemble de manière cohérente, afin de permettre le déploiement de l'innovation technologique sans discrimination, dans le respect des droits fondamentaux.

De l'algorithme

Les outils d'IA se distinguent par leur capacité d'inférence : à partir d'une base de données d'entraînement, ils peuvent générer du contenu, des recommandations, ou dans certains cas prendre des décisions. Cette capacité repose sur des techniques dites d'apprentissage automatique, elles-mêmes fondées sur des algorithmes – une série d'instructions déterminant une tâche spécifique.

Parmi les principaux acteurs de l'IA, le RIA identifie les fournisseurs de modèles d'IA à usage général et/ou des systèmes d'IA (par exemple ChatGPT, Gemini, Le Chat Mistral, Cohere Command, Claude...), mais aussi les déployeurs de ces systèmes, dont les acteurs publics, c'est-à-dire ses utilisateurs en capacité de modifier/ajuster les modèles ou les systèmes d'IA à leurs usages et environnements spécifiques, avec leurs propres données et fonctions.

Ces mécanismes soulèvent des enjeux de transparence, de responsabilité et de supervision que le RIA vient précisément encadrer. Le législateur européen a fait à la fois le choix d'une approche en fonction des usages et des risques inhérents à ces derniers, et transversale pour l'IA générative (IAG).

Il a notamment interdit certaines pratiques, par exemple la notation sociale, et définit des règles

proportionnelles aux risques par nature des usages – en matière de santé, de sécurité, de droits fondamentaux, de la démocratie et de l'État de droit... Plus les risques sont socialement élevés, plus les obligations sont strictes en matière de qualité des données, de supervision humaine, de traçabilité, de gestion des risques.

De la puissance de calcul

Les unités de traitement graphique (GPU), rendues célèbres par l'industrie du jeu vidéo, sont devenues des composants essentiels pour entraîner et faire fonctionner les grands modèles d'IA. Au-delà de ces composants, l'IA repose sur toute une infrastructure de calcul sous-jacente – centres de données, services de cloud, réseaux de distribution – dans laquelle les acteurs américains et asiatiques ont investi depuis des décennies, leur offrant un avantage net.

La Commission européenne prépare actuellement un futur Acte pour le développement du *cloud* et de l'IA destiné à combler les lacunes dans les capacités européennes.

L'IA et les libertés individuelles : le cas de la reconnaissance faciale

Le choix de l'approche par les risques algorithmiques a été nourri par des débats médiatiques, notamment de l'usage controversé de l'IA dans la sphère publique comme la notation sociale en Chine. Désormais explicitement prohibée dans l'Union européenne, la notation sociale a illustré les dérives possibles d'une technologie dans les mains d'un gouvernement autoritaire.

1. Règlement (UE) 2018/1807 du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne.
2. Règlement (UE) 2022/868 sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724.
3. Directive (UE) 2019/1024 du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public.
4. Règlement (UE) 2023/2854 du 13 décembre 2023 concernant des règles harmonisées portant sur l'équité de l'accès aux données et de l'utilisation des données et modifiant le règlement (UE) 2017/2394 et la directive (UE) 2020/1828.

Mais au-delà de ce cas dystopique, d'autres usages plus ambivalents, comme la reconnaissance faciale, ont fait l'objet de nombreux débats. Pour les autorités chargées du maintien de l'ordre, les outils d'IA présentent des avantages majeurs face aux menaces complexes. Les méthodes traditionnelles d'enquête ne suffisent plus à elles seules alors que les données générées en ligne ou par les équipements connectés augmentent de manière exponentielle. Dans ce contexte, l'Europol prône, par exemple, le recours à l'IA pour analyser des millions d'images de caméras vidéo en temps réel et notifier des activités suspectes, telles que des véhicules près de lieux sensibles, des objets non surveillés comme un sac oublié dans une gare ou des entrées non autorisées¹.

Mais l'utilisation de l'IA dans les espaces publics soulève des questionnements sur le risque de fichage généralisé – du fait de la collecte et du traitement de données biométriques sensibles – ou de stigmatisation de certaines populations.

Le RIA a tranché par une interdiction de principe des systèmes d'IA utilisés aux fins d'identification biométrique à distance, en temps réel, dans des espaces publics à des fins répressives, sauf exceptions strictement encadrées, telles que la recherche ciblée de victimes d'enlèvement ou la prévention d'une menace spécifique.

Ce nouveau cadre européen vient conforter des principes déjà ancrés dans le droit français et européen : la loi « informatique et libertés » autorise certains traitements biométriques par décret, après avis de la Commission nationale de l'informatique et des libertés (Cnil).

À l'occasion de l'organisation des Jeux olympiques (JO) de 2024, une expérimentation de caméras augmentées a été autorisée par la loi². Si elle n'incluait pas l'autorisation de la reconnaissance faciale, elle autorisait l'analyse automatisée d'images en temps réel permettant aux autorités compétentes et aux opérateurs de certaines infrastructures de transport

de détecter des situations inhabituelles et de prendre des décisions d'ordre public en conséquence. Arrivée à son terme le 31 mars dernier, l'expérimentation devrait être pérennisée dans le cadre d'un projet de loi actuellement à l'étude, relatif à l'organisation des Jeux olympiques et paralympiques (JOP) d'hiver 2030.

La loi « Jeux olympiques » illustre une approche finalement assez pragmatique, consistant à tester des dispositifs, tout en restant vigilants sur leur encadrement. Elle montre l'importance d'une mise en œuvre réfléchie et encadrée afin de minimiser les risques et de garantir une utilisation responsable des technologies.

Au-delà des enjeux de sécurité, quels usages de l'IA dans la sphère publique ?

Le RIA prévoit que la prise des décisions affectant la trajectoire d'un individu ou son rapport à l'administration soit accompagnée de garanties renforcées. Ainsi, les systèmes d'IA ayant un impact sur l'accès des citoyens aux services publics et prestations sociales essentielles, ceux utilisés dans le domaine de l'éducation et de la formation professionnelle ou bien pour le recrutement et l'évaluation de la performance des agents, pourraient être qualifiés d'IA à haut risque.

Sans être interdit, leur usage impose aux déploieurs des obligations spécifiques, en fonction du niveau de contrôle qu'ils exercent sur l'outil et notamment :

- d'acculturer les personnes concernées aux enjeux et aux usages des systèmes d'IA ;
- de tenir un journal d'événements ;
- de mettre en place des mesures techniques et organisationnelles appropriées ;
- de réaliser des analyses d'impact sur la protection

1. *AI and policing. The benefits and challenges of artificial intelligence for law enforcement*, Europol, 2024.

2. Loi n° 2023-380 du 19 mai 2023 relative aux Jeux olympiques et paralympiques de 2024 et portant diverses autres dispositions ; décret n° 2023-828 du 28 août 2023 relatif aux modalités de mise en œuvre des traitements algorithmiques sur les images collectées au moyen de systèmes de vidéoprotection et de caméras installées sur des aéronefs, pris en application de l'article 10 de la loi n° 2023-380 du 19 mai 2023 relative aux Jeux olympiques et paralympiques de 2024 et portant diverses autres dispositions.

de données à caractère personnel, conformément au RGPD ou à la directive européenne « Police-Justice » et sur les droits fondamentaux, conformément au RIA...

Documentation, transparence, analyse d'impact... Si une culture de la conformité est déjà à l'œuvre au sein d'une organisation, l'adoption des systèmes d'IA à haut risque peut s'opérer dans la continuité des méthodes existantes, de façon structurée.

L'enjeu principal n'est peut-être pas tant juridique qu'opérationnel : intégrer des outils d'IA de manière à accompagner le travail humain, développer les expertises internes, est un enjeu et une opportunité pour l'action publique.

Pas de « droit à la paresse » : la technologie évolue, les organisations aussi

Si l'exigence était déjà induite par la technologie et son encadrement, le RIA introduit de façon encore plus explicite l'obligation de se former et de former (*AI literacy*), pour toute organisation utilisant l'IA.

Pour accompagner cet effort, la Commission européenne¹ a publié un répertoire de cas pratiques publics et privés, d'utilisateurs, fournisseurs ou déployeurs en Europe, comme celui de l'entreprise publique espagnole Ineco qui a choisi de d'abord former l'ensemble des équipes aux outils d'IA, avant d'approfondir et de spécialiser celles et ceux qui développent ou supervisent des projets spécifiques.

Tout comme pour l'avènement de la messagerie électronique ou de l'internet, l'adoption de l'IA appelle à une acculturation progressive, avec une mise à jour des compétences, une réflexion sur les usages et une révision des pratiques de conformité, dans l'intérêt des usagers et du bon fonctionnement des institutions.

1. « Répertoire vivant pour favoriser l'apprentissage et l'échange sur l'alphabétisation en matière d'IA », Commission européenne, 4 février 2025.

Donner un sens au travail avec l'intelligence artificielle

– Yann Ferguson

Sociologue à l'Inria, directeur scientifique du LaborIA

La France, « start-up nation », s'est rapidement positionnée sur l'intelligence artificielle (IA) en étant, dès 2018, un des premiers pays à se doter d'une stratégie. La mission, confiée au mathématicien Cédric Villani, a produit un rapport remarqué pour son équilibre scientifique, technique et social, intitulé *Donner un sens à l'intelligence artificielle*¹, soit :

- donner un cap industriel en priorisant quatre secteurs – la santé, l'écologie, les transports-mobilités et la défense-sécurité – considérés comme étant au service de l'intérêt général et des grands défis de notre époque ;
- donner une signification sociale : promouvoir 1) une complémentarité « capacitante » entre IA et humaine et 2) une IA inclusive, qui ne contribue pas à accroître les inégalités sociales et humaines ;
- donner une explication : démystifier l'IA auprès du grand public et financer la recherche sur l'explicabilité de l'IA neuronale en vue de la rendre plus transparente et interprétable.

« C'est de manière plus générale une réflexion collective qui doit être menée sur ces technologies : la constante accélération des rythmes de déploiement ne doit pas empêcher une discussion politique sur les objectifs que nous poursuivons et leur bien-fondé² », recommande le rapport, affectant à l'action de « donner un sens » un volet délibératif et citoyen. Concernant le travail, le ton est d'abord grave : « Le

monde du travail est à l'aube de grandes transformations et n'y est encore que peu préparé³ ». Mais la mission considère aussi que « l'automatisation des tâches et des métiers peut constituer une chance historique de désautomatisation du travail humain : elle permet de développer des capacités proprement humaines (créativité, dextérité manuelle, pensée abstraite, résolution de problèmes)⁴ ».

De sa lettre de mission reçue du Premier ministre Édouard Philippe, le député Villani a bien retenu et même dépassé le premier objectif, « permettre l'émergence d'un écosystème performant et de champions ». En revanche, le second, « décrire les meilleures pratiques internationales d'application de l'IA au service de la transformation et de l'amélioration des politiques publiques, et identifier les applications prioritaires à déployer à l'intérieur de la sphère publique en France » est nettement moins développé. En effet, le sujet est traité dans la partie consacrée à l'inclusion qui enjoint la puissance publique de veiller à ce que l'IA ne pénalise pas certains publics, notamment les plus précaires, pour, au contraire, la mobiliser « pour absorber la complexité administrative, fluidifier et personnaliser l'expérience d'utilisation des services publics⁵ ». Si le rapport donne un sens à l'IA, celui-ci ne concerne donc pas la sphère publique. Le Conseil d'État va s'en charger, quatre ans plus tard.

1. Cédric Villani et al., *Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne*, rapport remis au Premier ministre, 28 mars 2018.

2. *Ibid.*, p. 13.

3. *Ibid.*, p. 17.

4. *Ibid.*, p. 105.

5. Lettre de mission du Premier ministre Édouard Philippe au député Cédric Villani, 8 septembre 2017.

L'IA dans l'action publique : servir dans la performance dans un cadre de confiance

Le rapport du Conseil d'État de mars 2022, « Intelligence artificielle et action publique : construire la confiance, servir la performance¹ » prend le parti assumé d'« aborder le sujet par les opportunités avant d'en exposer les risques ». Il affirme ainsi la nécessité d'une démarche proactive, « le secteur public ne doit pas attendre le moment, mais le créer² », plutôt que de devoir réagir à une pression citoyenne, budgétaire voire judiciaire, les rapporteurs anticipant l'émergence d'un « droit à l'IA » qui verrait l'acteur public condamné pour ses carences. En ce sens, une grande partie de l'étude se consacre à décrire l'état des lieux européen pour cerner le potentiel de l'IA. Celle-ci apparaît comme une technologie d'application générale adressant tous les domaines de l'action publique : gestion des territoires, défense, sécurité, justice, emploi, éducation, protection sociale, santé. S'appuyant sur un recensement opéré par AI Watch qui a répertorié 376 cas d'usage en Europe, dont la France représente 8 % en septembre 2021, le rapport observe que « la France n'accuse donc pas de retard au regard des pratiques constatées dans les administrations des autres États membres de l'Union européenne, mais elle ne bénéficie pas non plus d'une quelconque avance³ ».

Sept principes pour une « IA publique de confiance » visant à renforcer la confiance dans ces systèmes, à maîtriser les risques qu'ils présentent, notamment au regard des droits fondamentaux et des libertés publiques, et à assurer leur acceptabilité sociale sont posés : la primauté humaine (les systèmes d'IA doivent rester des outils au service de l'humain, de l'intérêt général et garantir une supervision humaine effective), la performance (les systèmes d'intelligence

artificielle (SIA) doivent être exacts, robustes, durables et améliorer concrètement le service public), l'équité et la non-discrimination (les systèmes doivent prévenir toute discrimination, consciente ou involontaire, et garantir l'accessibilité universelle), la transparence (les usagers doivent être informés, les systèmes documentés et les décisions explicables), la sûreté (les SIA doivent intégrer des protections contre les cyberattaques et anticiper les vulnérabilités propres à l'IA), la soutenabilité environnementale (l'impact écologique des systèmes doit être mesuré, limité et justifié au regard de leur utilité) et enfin l'autonomie stratégique (les choix technologiques doivent préserver la souveraineté nationale, notamment par la maîtrise des dépendances critiques, des données sensibles et par le développement de capacités européennes). Ces principes garantiraient l'alignement entre les usages de l'IA, la mission de service public et l'intérêt général. Enfin, le rapport préconise des investissements précoces et durables, tant techniques qu'humains, une stratégie RH étant indispensable pour éviter « un archipel de l'IA publique composé d'une poignée de services en pointe mais qui n'auront ni la capacité d'entraînement, ni même la volonté, pour faire monter en compétence les autres entités publiques⁴ ».

Dans son travail sur l'IA dans les politiques publiques (2024), la Cour des comptes s'appuie sur l'exemple du déploiement ambitieux de l'IA dans le ministère de l'Économie et des Finances⁵. Elle a notamment évalué la maturité de la confiance des cas d'usage d'IA, une grosse trentaine, à partir d'une grille légèrement différente de celle du Conseil d'État (n'intégrant pas notamment l'autonomie stratégique). La maturité globale moyenne est estimée à un niveau 2 sur 5. Cela signifie que si la stratégie de confiance commence à être gérée au-delà des seules équipes projet, elle n'est pas encore standardisée à l'échelle ministérielle, ni pilotée par les risques, et l'ouverture à des tiers pour des évaluations régulières reste rare.

1. « Intelligence artificielle et action publique : construire la confiance, servir la performance », Conseil d'État, 31 août 2022.

2. *Ibid.*, p. 84.

3. *Ibid.*, p. 81.

4. *Ibid.*, p. 87.

5. *L'intelligence artificielle dans les politiques publiques : l'exemple du ministère de l'Économie et des Finances, exercices 2018-2023*, Cour des comptes, 18 juillet 2024.

En revanche, l'application « Signaux faibles », développée par la Direction générale des entreprises (DGE), dont le but est de calculer le risque de défaillance à dix-huit mois pour les entreprises de plus de dix salariés, se distingue par un fort niveau de confiance (4 sur 5), grâce à une gouvernance formalisée, une forte explicabilité, un pilotage continu, une sécurité certifiée, une conformité rigoureuse au Règlement général sur la protection des données (RGPD) et au Référentiel général de sécurité (RGS), ainsi qu'une répartition claire des responsabilités. Si l'alignement des SIA sur les exigences d'un service public performant au service de l'intérêt général constitue un véritable défi, il n'est donc pas insurmontable dès lors qu'on en fait la priorité. Pour autant, cet alignement, indispensable, n'est pas suffisant pour garantir le sens du travail des agents publics.

De l'IA de confiance au travail à la confiance dans l'IA au travail

Dans l'approche des sciences du travail (ergonomie, sociologie, psychologie du travail...), le travail est une activité par laquelle les humains transforment le monde naturel et social et se transforment eux-mêmes. Plus précisément, Thomas Coutrot et Coralie Pérez ont établi trois critères pour définir le sens du travail : l'utilité sociale (« mon travail est utile aux autres »), la cohérence éthique (« je ressens la fierté du travail bien fait ») et la capacité de développement (« dans mon travail je peux développer mes compétences »¹¹). Ainsi, un travail doté de sens permet de se sentir utile, de se reconnaître dans sa réalisation qui respecte les règles du métier et l'éthique commune et de développer ses habiletés et son expérience. Dans cette perspective, les enquêtes Conditions de travail menées régulièrement par la

Dares montrent que ceux qui travaillent en contact avec le public dans le secteur public ou associatif, portés par un fort sentiment d'utilité, mettent particulièrement en avant le sens du travail. À cet égard, la Commission IA générative, chargée de réactualiser la stratégie nationale, dresse, dans son rapport de mars 2024, un bilan nuancé de la transformation numérique des services publics². Outre sa superficialité, s'arrêtant à la dématérialisation sans transformer en profondeur la circulation de l'information ou le traitement des demandes, la commission souligne un sentiment de déshumanisation et d'éloignement du service public. En réponse, l'IA, en particulier générative, constituerait une opportunité de « libérer les agents de tâches répétitives, ou chronophages, tout en améliorant la qualité de service » et de « libérer la créativité des agents en permettant d'expérimenter la technologie à leur niveau sans avoir toujours besoin d'un système spécifique³ ».

Les travaux du LaborIA, programme de recherche sur l'IA et le travail initié fin 2021 par le ministère du Travail et de l'Emploi et l'Institut national de recherche en sciences et technologies du numérique (Inria), évite une lecture déterministe qui ferait de l'IA une dynamique destructrice du sens du travail pour les uns ou libératrices pour les autres. Ses travaux proposent une lecture moins technocentrée et davantage « anthropocentrée » de l'impact de l'IA sur le travail qui résulte moins des propriétés techniques de l'IA que de la manière dont son intégration se pense avec les professionnels, les futurs utilisateurs. Or, dans la plupart des environnements de travail, publics comme privés, le LaborIA constate, dans son étude publiée en 2024, un « conflit de rationalité⁴ » entre des solutions qui colportent une logique gestionnaire de l'activité et les situations réellement vécues par les travailleurs, totalement ou partiellement exclus de l'expression des besoins comme de la réalisation des projets. Non seulement l'efficacité, l'utilité et la performance des SIA sont relatives, mais leur utilisation nécessite des interventions de préparation des données et de vérification des résultats,

1. Thomas Coutrot et Coralie Pérez, *Redonner du sens au travail. Une aspiration révolutionnaire*, Paris, La République des idées/Seuil, 2022.

2. *IA : notre ambition pour la France*, Commission de l'intelligence artificielle, mars 2024.

3. *Ibid.*, p. 77.

4. « Synthèse générale : Étude des impacts de l'IA sur le travail », LaborIA, 17 mai 2024.

une activité fastidieuse et peu stimulante. En somme, l'intégration des SIA sans associer les utilisateurs détricote l'activité, supprimant parfois des tâches gratifiantes, en générant de nouvelles orientées vers le SIA, plutôt que libérer des tâches répétitives et la créativité des agents. S'occuper de la machine plutôt que des usagers, de la mission proprement dite : certains agents dénoncent une industrialisation des procédés et une centralisation massive des procédures où les indicateurs remplacent les relations. Dans ce contexte de dépérissement du travail vivant, notion qui désigne l'activité productive d'individus mettant en jeu leurs ressources subjectives, physiques, intellectuelles et émotionnelles, comment se sentir utile, faire du bon travail et s'épanouir ? Dans l'autre sens, l'utilisation spontanée et discrète de nombreux agents de solutions grand public, à leurs frais et sans l'assentiment de l'administration, démontre leur intérêt dès lors qu'ils peuvent conserver leur autonomie. Pour autant, ces pratiques ne sont pas soutenables à terme. Non seulement les solutions privilégiées, par ailleurs non souveraines, sont portées par des valeurs divergentes, voire antagonistes à l'intérêt général, mais ces utilisations informelles et individuelles peuvent provoquer une atomisation des collectifs de travail et des pratiques professionnelles qui fragiliseraient la cohérence du service public et de la notion même d'intérêt général.

À partir des travaux du LaborIA, Data Publica et Civiteo ont conçu un dispositif dont l'ambition principale est de questionner et d'anticiper collectivement les effets de l'IA sur les agents des collectivités (leur engagement, la qualité de vie et des conditions de travail), les managers (leur rôle et leurs responsabilités) et l'organisation dans son ensemble. Pour chaque projet, elle permet de faire émerger des valeurs positives à viser et des points de vigilance, autour des critères clés qui nourrissent le sens du travail : l'autonomie, la responsabilité, le savoir-faire, la reconnaissance, les relations sociales, la santé mentale, le partage du travail, la coopération entre agents, la prise de décision. Ces questions sont certes pertinentes pour tous les projets de transformation, mais l'IA comporte aussi des enjeux spécifiques, voire inédits : un imaginaire très riche et anxiogène alimenté par la fiction, mais également par les leaders actuels de l'IA, des modalités d'interaction humain-machine en langage naturel, des apprentissages croisés entre machine et travailleur et, bien sûr, des enjeux éthiques sensibles. En matière d'IA, le secteur public aurait, selon le Conseil d'État, « un devoir particulier d'exemplarité¹ ». Ce devoir prévaut autant pour l'alignement du développement des SIA sur un cadre de confiance que pour le sens du travail de ses agents. Car il ne peut y avoir un cadre de confiance pour l'IA de l'action publique qui serait détaché de l'expérience sensible de l'IA au travail par ses agents.

1. « Intelligence artificielle et action publique : construire la confiance, servir la performance », Conseil d'État, *op. cit.*, 2022, p. 94.

IA et éducation : faire avec ou faire sans ?

– Diana Filippova¹

Romancière et essayiste, autrice de *Technopouvoir. Dépolitiser pour mieux régner* (Les liens qui libèrent, 2019) et de *Rien n'est plus grand que la mère des hommes* (Albin Michel, 2025)

En quoi l'intelligence artificielle est-elle un enjeu pour l'éducation ?

Je formulerais les choses dans l'ordre inverse : en matière d'intelligence artificielle (IA), l'impact sur l'éducation est le sujet le plus important. La faute que nous avons collectivement commise au moment de l'arrivée de l'Internet, des réseaux sociaux ou du smartphone a été de penser uniquement en termes d'adaptation des compétences des adultes, d'un côté, et de nouveaux usages des enfants au sein du système éducatif, de l'autre côté, ce qui s'est essentiellement matérialisé par une logique d'équipement et d'adaptation des écoles aux nouveaux usages sans réflexion critique sur les transformations profondes – voire anthropologiques – introduites par ces nouvelles technologies.

Avant même de parler de l'IA, notons que les effets d'un usage même contrôlé des réseaux sociaux sur l'attention, la santé mentale, la sécurité en ligne, le rapport au monde, les sociabilités ou encore sur la façon d'être sont tels que nous sommes face à une bombe d'une puissance inégalée. De ce point de vue, l'analogie avec les addictions est édifiante : imagine-t-on une proportion écrasante d'une classe d'âge accéder librement à une drogue avec la bénédiction des parents et la relative indifférence des pouvoirs publics ? Cela est d'autant plus problématique que ces usages se déploient à un âge où les enfants sont censés développer leur libre arbitre et leur capacité à dire non. Mais croire qu'à l'âge de l'entrée au col-

lège, les enfants ont les ressources intellectuelles, mentales et émotionnelles de réguler leurs propres usages de techniques *addictive by design*, cela relève de la pensée magique, de l'inconsistance ou de l'irresponsabilité.

Comme d'habitude, les entreprises technologiques et, dans leur sillage, les pouvoirs publics ont tendance à formuler le problème en matière de responsabilité individuelle. Cela veut dire : éduquons les parents qui, à leur tour, éduqueront leurs enfants. Il n'est guère surprenant que l'individualisation du problème est vouée à l'échec. C'est pour cela qu'il est crucial de formuler ce problème, comme il est, je crois, de plus en plus la norme aujourd'hui, en termes de santé publique et de responsabilité collective. Mais cela, il me semble, ne suffit pas à saisir l'ampleur des bouleversements qu'introduit la superposition de l'IA sur cet impensé et incontrôlé qu'est l'Internet.

Je ne vais pas revenir sur une définition spéculative de ce qu'est l'IA. Ce qui m'intéresse, c'est ce que représente l'IA dans le contexte singulier de l'enfance, de l'adolescence, des apprentissages et des différents niveaux de socialisation qui se déploient à ces périodes déterminantes pour l'avenir et le devenir de notre humanité (je n'exagère nullement en affirmant cela).

Il y a, dans ce contexte, trois enjeux cruciaux que je voudrais soulever dans le cadre de cet entretien.

D'abord, l'IA bouleverse entièrement notre rapport au langage parlé et écrit, plus particulièrement en ce

1. Propos recueillis par Émilie Agnoux le 11 juin 2025.

que l'accès à une quasi-infinité de sources primaires en désacralise la littérarité. Il est très possible que ce que le journaliste américain spécialiste des techniques Jeff Jarvis appelle la « parenthèse Gutenberg¹ » soit en passe d'être refermée. Qu'est-ce que cela implique ? Non pas qu'on lise de moins en moins de textes ou qu'on ait accès à moins d'informations, mais plutôt que les sources en soient de moins en moins à la fois des sources écrites (imprimées) et des sources primaires. Pour le dire simplement, au lieu de lire *Madame Bovary*, un élève va demander à une IA de la résumer puis de lui lire le résumé à voix haute, peut-être assorti d'un commentaire critique. Sans s'égarer dans les méandres de ce que cela implique pour son développement cognitif et son rapport au monde (et cela implique beaucoup de choses), il suffit de constater que cela va à l'encontre des fondements mêmes de l'éducation fondée sur la *literacy*, pour reprendre un terme anglais difficilement traduisible, et sur la pensée critique qui se déploie tant à l'écrit qu'à l'oral dans une communauté humaine.

Il y a bien sûr un deuxième problème, intimement lié à ce que je viens d'expliquer : la vérifiabilité des sources et, plus fondamentalement, le rapport à la vérité, d'une part, et à la création, de l'autre. Pour le dire très simplement, le recours à un grand modèle de langage (LLM) ne développe en soi ni l'esprit critique, ni l'effort, ni la diligence. Dans un rapport récent², le MIT Media Lab étudie l'effet comparatif sur le cerveau du recours aux LLM *versus* un effort cognitif autonome : bien qu'il nous faille les prendre avec toutes les précautions, les résultats sont éloquentes. Les utilisateurs des LLM accumulent une sorte de « dette cognitive », avec une activité, une mémoire et des performances significativement plus basses que les individus qui n'ont recours à aucun outil numérique et, de manière notable, à ceux qui s'appuient sur des moteurs de recherche. Et là, on parle d'adultes ! Que dire des enfants, dont l'apprentissage consiste principalement à développer leur

capacité à fournir une attention et des efforts continus, soutenus, réguliers, sans gratification immédiate ?

Enfin, troisième enjeu, et pas des moindres : la sécurité. Je ne parle pas seulement de la sécurité au sens premier du terme, même s'il n'échappe à personne que l'IA permet désormais de faire passer à peu près n'importe quelle image et information pour véridique, depuis les campagnes de haine conçues de toutes pièces jusqu'aux vidéos d'enfants exploités sur le *dark web* par les réseaux pédocriminels. Je parle de sécurité affective, de sécurité mentale, bref, de cette capacité de se mouvoir dans un monde perçu comme relativement stable, rassurant et authentique. Avec l'IA, la désintermédiation commencée par les réseaux sociaux introduit une rupture. À quoi bon faire l'effort d'apprendre dans un monde où tous les problèmes peuvent être réglés avec le bon *prompt*³ ? Comment apprendre à faire confiance, respecter l'autorité, nouer des liens avec les pairs et les adultes de référence alors qu'on est dès l'enfance en présence d'un simulacre de pensée ? Quel intérêt à se former à des métiers qui, demain, auront été balayés par un nouveau modèle d'IA ? Attention, je ne suis pas en train d'affirmer que ces activités deviennent obsolètes ou disparaîtront de la surface de la Terre, bien au contraire : je crois profondément à leur valeur en tant qu'activité en soi, pour soi et pour les autres. Mais la dégradation de leur valeur sociale, économique et politique suffit à ternir un attribut précieux de toutes les jeunesses de toutes les époques : l'espoir et la conviction qu'elles ont prise sur le monde et qu'elles peuvent le changer pour le mieux.

Quelles en sont les implications pour l'Éducation nationale ?

La première chose à faire, c'est de tirer les enseignements du passé, c'est-à-dire de l'arrivée des ordinateurs personnels, de l'Internet, des réseaux sociaux, de l'hyperconnexion non seulement dans la vie des enfants, mais au cœur de notre système sociopolitique.

1. Jeff Jarvis, *The Gutenberg Parenthesis. The Age of Print and Its Lessons for the Age of the Internet*, Bloomsbury Academic, juin 2023.

2. Nataliya Kosmyna, Xian-Hao Liao, Eugene Hauptmann, Ashly Vivian Beresnitzky, Ye Tong Yuan, Iris Braunstein, Jessica Situ et Pattie Maes, *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*, preprint, MIT, juin 2025.

3. Un *prompt* est une instruction ou des données adressées à un système d'IA permettant de préciser ce qui est attendu de la machine en termes de résultats et de livrables.

Je crois que les politiques, en particulier les politiques de gauche, séduits par les promesses progressistes qu'ils croyaient déceler dans la vision du monde portée par les entrepreneurs de la Silicon Valley, ont largement échoué à jouer leur rôle de régulateur, voire de frein. Nous avons collectivement perdu la capacité à dire « non ». Je voudrais reprendre ici un concept de Michel Foucault qui m'est particulièrement cher, celui de biopolitique et de biopouvoir¹. Cette idée que le pouvoir se décentralise, se fragmente, s'insinue dans le moindre repli de nos existences, presque jusque dans notre chair, a été fondatrice pour la réactualisation que j'ai tenté d'en faire dans mon essai *Technopouvoir*².

La thèse centrale en est simple : toutes les technologies sont toujours en même temps des technologies de pouvoir ; toutes les manières de gouverner et de contrôler s'appuient toujours sur des technologies qui, par définition, ne sont pas neutres. Il me semble que ce cadre d'analyse n'a rien perdu de sa puissance d'explication en matière d'IA : il permet ainsi de comprendre que l'enjeu, pour l'Éducation nationale, n'est pas d'« intégrer un outil dans ses programmes », comme on équiperait une classe de tablettes ou proposerait des séances de familiarisation avec le code dès la maternelle. De la même manière, la question de la fraude aux examens et de la triche à l'occasion des devoirs à la maison me semble relativement secondaire : il suffit d'avoir des bons outils de détection de LLM et d'avoir un minimum de jugeote, par exemple si un élève qui écrit d'habitude très mal se met à user d'un langage extrêmement soutenu.

Je formulerais l'enjeu de la manière suivante : comment préparer les futurs adultes à habiter et à façonner un monde incertain qui leur échappe ? Comment leur permettre de garder le contrôle quant aux choix individuels et collectifs qu'ils seront amenés à faire ? Comment tirer avantage de nouvelles techniques en conservant un rapport avec le réel, avec l'autre, avec le vivant, avec le monde, avec soi ?

C'est somme toute un vieux casse-tête qui a animé des générations successives de professeurs, de penseurs, de décideurs. Je trouve cela, *in fine*, assez

rassurant et même porteur d'espoir : l'urgence de vivre et d'enseigner avec une technologie aussi pervasive, sinieuse, évolutive et peu fiable nous rappelle à l'urgence de refonder notre système public et privé d'éducation.

La difficulté vient sans doute du fait que notre service public de l'éducation n'est pas en grande forme, et cela depuis bien avant les technologies dont nous parlons dans cet entretien. Que ce soit en matière de production des inégalités, de système à plusieurs vitesses, de difficulté croissante à prendre en charge les jeunes en échec scolaire doublée d'une compétitivité proprement nuisible à la confiance et au bien-être mental, ou encore de recrutement et de valorisation des professeurs : le mal est ancien. Cette faiblesse devenue structurelle n'a, je crois, rien à avoir avec son échec à remplir ses missions, mais plutôt avec le sous-investissement chronique, le mépris et l'abandon de plus faibles à leur sort. Il faut bien le comprendre.

Prenons Parcoursup : nous n'avons fait que passer un échelon supplémentaire dans le sentiment d'abandon de l'ambition méritocratique, avec une désintermédiation entre les jeunes, d'un côté, et ceux qui sont censés organiser les conditions de leur réussite, de l'autre côté, et un algorithme qui enferme socialement et territorialement. Sans parler du coup psychique et physique de se retrouver, pour des milliers de jeunes, sans aucune affectation après des mois de préparation épuisante. C'est dans ces conditions, alors que beaucoup de ceux qui en ont les moyens se désolidarisent de l'effort national, que l'Éducation nationale doit accueillir la rupture majeure que représente la pervasivité de l'IA.

Comment reprendre le contrôle ?

J'aimerais pouvoir y répondre en quelques lignes ! Hélas, tout cela est très compliqué et je crains qu'une bonne partie des bouleversements dont nous sommes les spectateurs nous ont définitivement échappé. Nous avons cédé nos données aux géants de la tech, nous avons accepté les serveurs et les technologies

1. Michel Foucault, *Naissance de la biopolitique. Cours au Collège de France 1978-1979*, Paris, Gallimard, 2004 ; et Michel Foucault, *Du gouvernement des vivants : cours au Collège de France, 1979-1980*, Paris, Seuil, 2012.

2. Diana Filippova, *Technopouvoir. Dépolitiser pour mieux régner*, Paris Les liens qui libèrent, 2019.

non françaises et non européennes qui sont administrées par des entreprises américaines appliquant la règle de l'extraterritorialité, nous avons renoncé à la souveraineté depuis trop longtemps pour avoir une chance de rattraper tout notre retard. Je crois cependant qu'il n'est jamais trop tard, et il est temps plus que jamais de miser sur une IA souveraine, versée aux communs, intégrant dans son dessein et dans sa vision les valeurs intellectuelles, politiques, culturelles et sociales qui sont les nôtres. De ce point de vue, la France a plutôt fait preuve de volontarisme.

Mais il me semble hasardeux de mettre tous nos œufs dans le panier de ce projet, tant il dépend peu de nous en tant que citoyens et de la France en tant que puissance souveraine. Je propose dans l'ordre quelques idées qui pourraient participer à une appropriation de l'IA et à une mise au service d'un service public de l'éducation tel que j'en ai décrit les enjeux un peu plus tôt.

D'abord, parlons des réseaux sociaux. Je suis tout à fait favorable à une interdiction des réseaux jusqu'à un certain âge. Je serais même favorable pour mettre sur la table des décideurs une interdiction des smartphones, dont les réseaux sociaux sont une excroissance, jusqu'au lycée. Jonathan Haidt, le grand spécialiste américain des ravages que les réseaux font parmi les jeunes, réhabilite par exemple l'idée d'un téléphone fixe¹ : avec une dose de volontarisme et d'organisation parentale, cela fonctionne. À ceux qui font des gros yeux, je réponds ceci : imaginez-vous il y a même cinq ans que votre ado de 15 ans en crise demanderait à ChatGPT de lui prodiguer des conseils psy au lieu d'en parler à vous, à ses amis ou à un professionnel ? Rien n'est impossible si on se donne les moyens d'y parvenir. Par le passé, nous avons été capables d'interdire la vente aux jeunes des substances toxiques, comme le tabac. Nous devons aujourd'hui pareillement reconnaître le caractère nocif d'outils numériques pour nos jeunes, d'autant que tous les enfants, tous les milieux et toutes les familles ne sont pas égaux face à la régulation des usages.

Il est urgent de former les professeurs. Car ceux-ci sont aujourd'hui, au mieux, laissés seuls face aux

usages de l'IA et, au pire, soumis à des injonctions contradictoires. La dernière mouture du « Cadre IA en éducation » proposé par le ministère de l'Éducation nationale édicte des principes et se contente d'indiquer les grandes directions sans donner d'outil concret, à charge, semble-t-il, de chaque établissement pour en organiser le mode opératoire. Ce n'est pas suffisant. Enseigner aux élèves la bonne utilisation d'un *prompt*, les procédés pour croiser les données et vérifier leur efficacité, comprendre le fonctionnement des IA génératives, leurs possibilités et leurs limites, tout cela doit faire à mon sens l'objet d'une priorité nationale, élaborée de manière centralisée et collaborative en y associant des professionnels et des chercheurs de tous horizons.

Une telle institution *ad hoc* pourrait, par exemple, prendre la forme d'une autorité administrative indépendante (AAI), ou, plus simplement, être associée à une institution existante – je pense par exemple à l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse (ARCEP). Il est très important qu'une telle autorité, quelle qu'en soit la forme, associe – outre les chercheurs tant en IA qu'en littérature ou en médecine – la société civile. Et, plus encore, qu'elle soit dotée d'un statut et d'instances lui permettant de peser dans le rapport de force avec les acteurs de l'IA.

L'une de ses missions devrait être d'élaborer un programme, des méthodes, une organisation qui s'applique à toute la France et dont la mise en œuvre, en revanche, devrait être pensée de concert avec les acteurs de l'Éducation nationale à tous les niveaux. Face à l'ampleur des bouleversements – ou cataclysmes, selon notre capacité à s'en emparer – qu'apporte l'IA, on ne peut plus se cantonner de sensibiliser, d'avertir, d'amorcer : il faut prendre la question à bras-le-corps.

Je conclurai sur une note plus personnelle, celle d'une écrivaine qui a grandi et doit son salut intellectuel aux livres. À l'heure où ceux-ci perdent des milliers de lecteurs chaque année, où le smartphone et les loisirs numériques ont remplacé cet objet dans les mains des enfants dès le plus jeune âge, nous pourrions être tentés de conclure à la fermeture de cette

1. Jonathan Haidt, *Génération anxieuse. Comment les réseaux sociaux menacent la santé des jeunes*, Paris, Les Arènes, 2025.

fameuse parenthèse Gutenberg. Les conséquences ne se limiteraient pas à un passage de l'ère de l'écrit à une nouvelle ère de l'oralité et de la conversation. Je suis intimement convaincue que nos démocraties n'y survivraient pas. Car la désintermédiation, « l'immédiateté du capitalisme tardif¹ », pour reprendre la thèse développée par la chercheuse en littérature américaine Anna Kornbluh, la méfiance face aux faits et la relative crédulité face aux sources tant qu'elles sont aisément accessibles rompent quelque chose de bien plus fondamental que notre relation à l'écrit ou à l'information.

Je pense d'abord à l'existence de principes universels irréductibles à la somme des connaissances ou des informations disponibles, une raison qui nous guident dans nos choix tant intellectuels que pratiques. Or l'IA n'est jamais qu'une supersynthèse, la somme de tout ce qui est disponible, et une somme très souvent erronée. De façon liée, il s'agit de l'idée que cette raison est indissociable du corps, c'est-à-dire qu'elle est incorporée, ce que l'IA n'est pas et, je le pense,

ne sera jamais. Dans la mesure où cette incorporation se joue à la fois dans chaque corps pris séparément et dans les relations infinies qui se forment entre les individus d'une société, sur un plan intersubjectif, elle fonde l'idée même de l'altérité comme principe fondateur de nos sociétés.

Enfin, l'IA est incapable de créer quelque chose qui n'existe pas, ni une œuvre, ni une idée, ni même le cours le plus simple. Or la capacité de briser les codes établis, d'enfreindre un ordre, de créer une contre-esthétique ou d'agir envers et contre toutes les règles sociales, c'est le fond même de la liberté. C'est grâce à cette capacité de rompre avec ce qui est, d'imaginer d'autres horizons, d'autres règles, d'autres œuvres, que l'humanité a su très souvent d'une trajectoire parfois tragique renouer avec un horizon désirable. D'un tel changement de perspective, nos enfants ont aujourd'hui plus besoin que jamais. Gageons que ce n'est pas le nez dans un *prompt* qu'ils y parviendront.

1. Anna Kornbluh, *Immediacy, or The Style of Too Late Capitalism*, Londres, Verso Books, 2024.

Encadrer : pour une IA
au service de la relation,
de la qualité des services publics
et des droits des citoyens

La transparence, un prérequis essentiel pour l'intelligence artificielle dans le secteur public

– Soizic Pénicaud

Chercheuse et consultante indépendante en politiques publiques de l'intelligence artificielle, cofondatrice de l'Observatoire des algorithmes publics, enseignante à Sciences Po Paris¹

L'intelligence artificielle dans le secteur public, une déferlante opaque

Depuis l'irruption de l'intelligence artificielle (IA) générative, l'IA a le vent en poupe dans le secteur public. En février 2025, le gouvernement annonçait vouloir « généraliser le recours à l'IA pour améliorer la qualité et l'efficacité de l'action publique² ». Selon le Baromètre de l'Observatoire Data Publica³, plusieurs centaines de collectivités de toutes tailles ont mené des projets d'IA en 2024.

En réalité, les administrations n'ont pas attendu l'IA générative pour utiliser des algorithmes dans leurs missions de service public. La numérisation du calcul des impôts remonte déjà à plus de cinquante ans et plusieurs algorithmes publics ont créé le débat ces dernières années, de Parcoursup⁴ à l'algorithme de *scoring*⁵ utilisé par les caisses d'allocations familiales pour contrôler les allocataires, en passant par la vidéosurveillance algorithmique (VSA)⁶ lors des Jeux olympiques et paralympiques (JOP). D'autres

systèmes sont moins connus du grand public, mais tout aussi importants, tels que ceux de Météo France⁷ pour les prédictions météorologiques. Ces différents systèmes emploient des technologies plus ou moins complexes, ont des conséquences plus ou moins importantes et sont utilisés dans des contextes divers. Cependant, ils s'inscrivent tous dans une dynamique d'automatisation croissante de l'administration.

Problème : ces développements sont souvent réalisés dans l'opacité. Bien qu'ils aient des conséquences importantes sur le service public et les citoyennes et citoyens, il reste très difficile de savoir quels systèmes algorithmiques et d'IA sont utilisés par les administrations, et dans quelles circonstances. L'administration ne réalise pas, à ce jour, de panorama exhaustif de ses usages de l'IA.

Face à cette opacité, l'Observatoire des algorithmes publics (Odap) alimente depuis 2024 un inventaire, disponible en ligne, compilant les informations disponibles en ligne sur les algorithmes utilisés par les ministères et institutions nationales françaises. Les résultats de l'association montrent que même lorsque l'on connaît l'existence des systèmes,

1. Auparavant, elle a également travaillé sur la transparence des algorithmes publics à Etalab (Direction interministérielle du numérique).

2. *Faire de la France une puissance de l'IA*, Comité interministériel de l'intelligence artificielle, 6 février 2025.

3. *Le Baromètre de l'Observatoire Data Publica 2024 : les collectivités territoriales et la donnée*, Observatoire Data Publica, 2024.

4. Sébastien Gavois, « Comment fonctionnent les algorithmes de Parcoursup », Next, 8 mars 2025.

5. Gabriel Geiger, Soizic Pénicaud, Manon Romain et Adrien Sénécat, « Profilage et discriminations : enquête sur les dérives de l'algorithme des caisses d'allocations familiales », *Le Monde*, 4 décembre 2023.

6. Florian Reynaud et Martin Untersinger, « Paris 2024 : la vidéosurveillance algorithmique à l'épreuve des Jeux olympiques », *Le Monde*, 23 juillet 2024.

7. « Les prévisions météo plus fiables grâce à l'intelligence artificielle », France Info, 22 mai 2025.

les informations à leur sujet sont rares. Sur les 72 systèmes recensés par l'association, seulement 3 disposent d'un budget public et 4 % d'évaluations publiques sur leurs résultats ou leurs enjeux pour les droits fondamentaux et la protection des données personnelles¹.

La transparence, un prérequis démocratique

Or, mieux connaître ces systèmes et leurs conséquences est un enjeu démocratique : les administrations sont responsables et redevables de leurs politiques publiques, y compris de celles qu'elles mettent en œuvre *via* l'IA. L'article 15 de la Déclaration des droits de l'homme et du citoyen de 1789 prévoit que « la société a le droit de demander compte à tout agent public de son administration ». La transparence est un prérequis de cette redevabilité, car elle permet à toute personne de comprendre les systèmes utilisés et en débattre, voire les contester ou les refuser. Il s'agit pour les citoyens et citoyennes de comprendre quels algorithmes sont utilisés pour prendre des décisions qui les concernent, mais aussi pour les agents publics, notamment *via* les organisations syndicales, d'avoir une voix dans les outils qu'elles et ils vont utiliser.

La transparence est de surcroît un garde-fou contre les dérives des systèmes d'IA. Le manque d'évaluations publiques masque le fait qu'énormément de systèmes algorithmiques ne remplissent pas leurs promesses d'amélioration du service public. À New York, un *chatbot* d'IA générative² mis en place en 2024 par la municipalité pour accompagner les employés leur donnait des conseils contraires au droit

du travail. D'après une étude du Data Justice Lab³ de l'université de Cardiff, l'abandon de projets d'IA est, dans la majorité des cas, dû à un manque d'efficacité ou d'efficience des systèmes. En France, alors que le comité indépendant chargé de l'évaluation de la VSA pendant les JOP a rendu un rapport en demi-teinte⁴ sur l'efficacité de l'expérimentation, celle-ci a manqué de peu⁵ d'être reconduite et pourrait désormais l'être dans le cadre du projet de loi JOP 2030.

On sait également que l'IA peut coûter cher, à la fois dans ses développements et dans ses conséquences. En Australie, l'algorithme Robodebt, utilisé de 2016 à 2019 pour récupérer des trop-perçus chez les allocataires de prestations sociales, a coûté près de 2 milliards d'euros⁶ à l'administration du fait de ses dysfonctionnements.

Bien sûr, la transparence des systèmes d'IA ne mène pas automatiquement⁷ à une plus grande exemplarité. Mais le fait de devoir produire des informations et les publier peut pousser les administrations à se questionner davantage sur les systèmes qu'elles vont mettre en place.

Que rendre transparent ? Au-delà de la transparence technique

Comme toute politique publique, un système algorithmique est affaire de choix. De manière instinctive, on pense aux choix techniques, par exemple dans les paramètres des systèmes. Les travaux de Vincent Tiberj à l'IEP de Bordeaux⁸ en sont une très bonne illustration : dans le cadre de Parcoursup, l'IEP a développé un algorithme alternatif à celui mis à

1. Voir la liste complète des algorithmes sur le site de l'Odap : <https://odap.fr/inventaire/synthese/>.

2. Colin Lecher, « NYC AI chatbot touted by Adams Tells Businesses to break the law », *The City*, 29 mars 2024.

3. « Automating public services: learning from cancelled systems », Data Justice Lab, 2022.

4. Yannick Falt, « Vidéosurveillance algorithmique : le comité d'évaluation émet des doutes sur l'efficacité du dispositif mis en place pour les Jeux de Paris », France Info, 15 janvier 2025.

5. « Comment le gouvernement a voulu prolonger la vidéosurveillance algorithmique », Amnesty International, 24 mars 2025.

6. « Automating public services: learning from cancelled systems », *op. cit.*, 2022.

7. « Making algorithm registers Work for meaningful transparency », IA Ciudadana, 13 mars 2025.

8. *Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ? Points de vigilance et recommandations*, Défenseur des droits, 13 novembre 2024.

disposition par le ministère pour classer les candidatures. Les deux algorithmes mènent à des résultats drastiquement différents : en 2023, seulement 3 % des élèves ayant bénéficié du programme « Cordée de la réussite » étaient considérés comme admissibles par l'algorithme du ministère, contre 19 % par celui de l'IEP de Bordeaux.

Mais les choix vont au-delà des paramètres du système ou des données utilisées. Par exemple, confie-t-on le développement d'un système d'IA à un prestataire privé ou à une équipe en interne ? Sur quelles infrastructures sont stockées les données ? Quel est le coût environnemental du système ? Et, à l'origine, pourquoi décide-t-on de résoudre tel ou tel problème de politique publique avec un système d'IA, et aurait-on pu faire différemment ? Même quand la complexité technique fait d'un système d'IA une prétendue « boîte noire », ou que le code source d'un algorithme ne peut pas être publié, beaucoup d'informations peuvent quand même être partagées. Tous ces choix sont importants, car ils sont, *in fine*, politiques.

Une transparence possible

La transparence des systèmes d'IA dans l'administration est pourtant possible. Dans certains cas, c'est même une obligation légale. En France, la loi pour une République numérique de 2016 a introduit des obligations¹ pour les administrations qui utilisent des algorithmes pour prendre des décisions. Ainsi, elles doivent notifier les personnes concernées de l'utilisation d'un algorithme, expliquer les critères de prise de décision et publier en ligne un inventaire des algorithmes et de leurs règles.

Le Règlement européen sur l'intelligence artificielle (*AI Act*) prévoit quant à lui la création, d'ici à août 2026, d'une base de données des systèmes d'IA au niveau européen. Cette base devrait recenser les systèmes d'IA à haut risque, y compris dans les services publics, la justice, l'éducation, la police et l'immigration.

Ces deux cadres réglementaires ne sont pas sans limites, et beaucoup de systèmes algorithmiques critiques vont rester invisibles. En France, les algorithmes couverts par certains secrets sont exemptés de ces obligations, dont les algorithmes de lutte contre la fraude. La Caisse nationale d'allocations familiales a par exemple invoqué cette exception pour refuser de publier les paramètres de son algorithme de contrôle, qui attribue un score à 13 millions de foyers en France². Par ailleurs, seuls les algorithmes de prise de décision sont concernés, ce qui exclut un certain nombre de systèmes, y compris les agents conversationnels (*chatbots*) utilisés dans les relations entre l'administration et les citoyennes et citoyens. Quant au règlement européen sur l'IA, il prévoit notamment que les systèmes utilisés par les forces de l'ordre et dans les domaines liés à l'asile et l'immigration seront recensés dans une partie « non publique » (comprendre : privée) de la base de données. Cela veut dire que les systèmes de vidéosurveillance algorithmique (utilisés lors des JOP 2024) ou bien les algorithmes utilisés par l'Union européenne dans le cadre du contrôle aux frontières³ vont rester cachés.

Cependant, ces cadres peuvent donner de l'élan aux administrations et rien ne les empêche d'aller plus loin. En France, quelques collectivités se sont conformées à l'obligation d'inventaire créée par la loi pour une République numérique citée plus haut. La ville d'Antibes⁴, le département d'Ille-et-Vilaine⁵ et la ville de Paris⁶ ont tous publié sur leur site une page listant les algorithmes qu'ils utilisent et leurs règles.

1. « Les algorithmes publics : enjeux et obligations », Data.gouv.fr.

2. « Rapport d'activité 2022-2023 », Commission d'accès aux documents administratifs, 2024, p. 43.

3. « Automated Suspicion: the EU's new travel surveillance initiatives », Statewatch, 2020 et « Automating Authority: Artificial Intelligence in European Police and Border Regimes », Statewatch, 2025.

4. Alexandre Léchenet, « Antibes publie l'inventaire de ses algorithmes », *La Gazette des communes*, 11 février 2021.

5. « Recensement des traitements algorithmiques », Data Ille-et-Vilaine.

6. « Consultation des algorithmes publics de la ville de Paris », Open data Paris.

Dans le monde, d'autres États ont déjà mis en place des inventaires à un niveau national. C'est le cas des Pays-Bas¹, dont l'inventaire compte à ce jour près d'un millier d'algorithmes et est régulièrement utilisé par la Commission nationale de l'informatique et des libertés (Cnil) et le Défenseur des droits néerlandais pour identifier les algorithmes à risques. En France, la Direction interministérielle de la transformation publique mène également des travaux intéressants sur les explications individuelles², dans le cadre du programme « Administration proactive ».

Mais force est de constater que, jusqu'ici, peu d'administrations se sont emparées du sujet. On peut

espérer que le règlement européen sur l'IA et le regain d'intérêt autour de l'« éthique » de l'IA provoqué par l'IA générative motivent plus de services à être exemplaires dans leurs projets. De leur côté, de plus en plus d'associations³, syndicats⁴ et journalistes⁵ s'organisent pour arracher la transparence aux administrations réfractaires, *via* des enquêtes, des appels à témoignages et des demandes d'accès aux documents administratifs. Et c'est tant mieux : alors que les services publics sont de plus en plus automatisés, il est désormais indispensable de s'intéresser aux algorithmes qui les sous-tendent pour garantir qu'ils restent accessibles, justes et égalitaires.

1. « The algorithm register of the Dutch government », Overheid.

2. « Réaliser des communications à destination des usagers, en proactivité », Direction interministérielle de la transformation publique, 5 novembre 2024.

3. « France contrôle. Gestion algorithmique des populations », La Quadrature du Net.

4. Voir dossier « IA à la DGFIP », Solidaires finances publiques.

5. « La police nationale utilise illégalement un logiciel israélien de reconnaissance faciale », Disclose, 14 novembre 2023.

L'IA dans la relation à l'utilisateur : vers la déshumanisation automatisée ?

– Marius Bertolucci

Maître de conférences en sciences de gestion à l'Institut de management public et gouvernance territoriale (IMPGT) et du laboratoire CERGAM à Aix-Marseille Université, auteur de *L'Homme diminué par l'IA* (Hermann, 2023) – prix Jacques Ellul 2024

À l'heure des choix budgétaires difficiles, l'arrivée de l'intelligence artificielle (IA) dans les services publics est présentée comme une manne providentielle. Elle serait à la fois source d'économies, de gains de performance et même de « re-humanisation » du service public, selon le 8^e Comité interministériel de la transformation publique¹. Pourtant, au-delà des discours enthousiastes politiques et commerciaux, les recherches académiques et les expériences de terrain montrent une réalité bien différente. Les services publics se trouvent confrontés à un double risque : la précipitation vers des IA non fiables et l'assistance des agents par des IA provoquant déqualification et déresponsabilisation.

Diagnostic : du non-accès au service public à son algorithmisation

Nul besoin que les humains soient remplacés par des IA pour que des risques apparaissent, une simple assistance suffit. Parcourons quelques exemples pour nous en convaincre.

L'accès direct pour les usagers à l'administration est souvent difficile. C'est particulièrement visible à

travers l'accès téléphonique aux services publics, comme le montre une étude récente conduite par 60 millions de consommateurs et la Défenseure des droits : 54 % des appels vers la caisse d'allocation familiale (CAF) et 72 % vers l'Assurance maladie restent sans réponse². Si certaines collectivités expérimentent timidement des IA vocales, comme à Chartres³, la majorité reste prudente face aux IA en lien direct avec les usagers, craignant à la fois leur manque de fiabilité technique et la perte du lien humain. Notons que l'agent conversationnel IssyGPT d'Issy-les-Moulineaux se trompe pour des informations comme la pharmacie de garde⁴... L'État, quant à lui, adopte progressivement l'IA en soutien aux agents publics en interaction avec les usagers, une pratique non dénuée de risques et de perte de lien. En 2024, un couple dans l'incapacité de contacter la CAF par téléphone se résout à l'écrit. La réponse obtenue *via* la plateforme Services publics+, rédigée dans une langue impersonnelle et mécanique, précise que « cette réponse a été facilitée par l'IA à 88 %, après traitement et validation par un agent ». Cet exemple de validation dit de « l'humain dans la boucle » est censé compenser les déficiences des IA. Si tel était le cas, les nombreuses catastrophes n'auraient jamais eu lieu, comme celle du scandale des algorithmes de détection des fraudes aux allocations familiales aux Pays-Bas ou aux allocations chômage dans le Michigan,

1. « 8^e Comité interministériel de la transformation publique », Direction interministérielle de la transformation publique, 23 avril 2024.
2. Enquête « Services publics : y'a quelqu'un au bout du fil ? », 60 millions de consommateurs, 26 janvier 2023.
3. Xavier Biseul, « Chartres met de l'intelligence artificielle dans son standard téléphonique », *Acteurs publics*, 11 février 2025.
4. Émile Marzolf, « 5 exemples de mise en pratique de l'IA générative dans le secteur public », *Acteurs publics*, 2 juillet 2024.

alors que le scandale Robodebt en Australie a touché 500 000 personnes¹.

À cet égard, penchons-nous sur le cas de l'IA VioGen en Espagne. Conçu pour aider les policiers à évaluer les risques des victimes de violences conjugales, ce système algorithmique s'appuie sur une trentaine de questions standardisées. En 2018², une femme menacée de mort ainsi que ses deux filles par son ex-mari (enregistrement audio à l'appui) se rendent dans un commissariat. Le policier assisté de VioGen classe son cas comme un risque faible, décision que suivra également un juge sur cette même base. Dans un contexte où les policiers ont parfois tendance à minimiser les violences domestiques, le solutionnisme technologique accentue ces difficultés en déplaçant la responsabilité de la décision sur l'IA. Ce drame révèle comment les biais humains peuvent être exacerbés par la dépendance à la décision algorithmique. Dans 95 % des cas³, les policiers espagnols suivent aveuglément la recommandation de l'IA alors qu'ils ont la possibilité de la modifier manuellement. Le cadre posé de l'interaction par VioGen induit un interrogatoire déshumanisé plutôt qu'une écoute et l'usage du jugement professionnel.

Ces risques sont largement confirmés par la littérature scientifique. La base de risques liés à l'IA du MIT⁴ en répertorie 1 600 classés en sept catégories, notamment celle intitulée « Interaction humain-machine ». Il en ressort des études montrant la perte d'autonomie, les effets de déresponsabilisation, la sur-confiance dans les décisions de l'IA. À titre d'exemple, les consultants du Boston Consulting Group (BCG) voient leur performance baisser de 23 % lorsqu'ils sont assistés par ChatGPT, lequel les induit en erreur⁵. Dans le même ordre d'idées, l'utilisation d'IA pour évaluer les remises en liberté favorise une déresponsabilisation des juges⁶.

Ces phénomènes s'expliquent notamment par la capacité persuasive des IA génératives.

Recommandations pour une IA au service des citoyens

Face à ces constats et enjeux, nous proposons plusieurs mesures.

Du côté des usagers, il est indispensable de garantir le droit à une interaction directe avec un agent humain, que ce soit en présentiel, par téléphone ou par courriel. En cas d'interaction initiale avec un dispositif d'IA, les usagers doivent pouvoir exiger immédiatement et sans coûts supplémentaires l'intervention d'un humain. De plus, une transparence réelle doit être exigée quant à l'utilisation de l'IA dans les processus administratifs, notamment lorsqu'elle intervient en appui des décisions ou des réponses fournies aux usagers.

Du côté des acteurs publics, aucune réduction des effectifs des agents en contact direct avec les usagers ne doit être mise en œuvre sans évaluation préalable des conséquences potentielles sur la qualité et l'accès au service public. Il convient donc d'imposer systématiquement la réalisation d'audits en amont du déploiement de tout système algorithmique destiné à assister les interactions avec les usagers. Ces audits doivent évaluer les besoins et risques spécifiques aux populations concernées. De même, des audits réguliers doivent être réalisés après l'installation des dispositifs automatisés, afin d'évaluer leurs effets réels, en particulier les éventuels biais d'automatisation et les risques de déshumanisation. Ces évaluations doivent être menées par des équipes mixtes réunissant

1. Peter Whiteford, « Robodebt was a fiasco with a cost we have yet to fully appreciate », *The Conversation*, 16 novembre 2020.
2. Michele Catanzaro, « In Spain, the VioGen algorithm attempts to forecast gender violence », *Automating society report*, Algorithm Watch, Bertelsmann Stiftung, 2020.
3. Gema Galdón Clavell et Éticas Fondation, « Inteligencia artificial, ¿empobrecimiento humano? », *La economía española ante una encrucijada crítica*, 2023.
4. Consultable sur le site MIT AI Risk.
5. Fabrizio Dell'Acqua, Edward McFowland III, Ethan Mollick, Hila Lifshitz-Assaf et al., « Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality », Harvard Business School, working paper 24-013, 2023.
6. Ryad Titah, « How AI skews our sense of responsibility », *MIT Sloan Management Review*, vol. 65, n°4, 13 mai 2024, pp. 18-19.

à la fois des acteurs internes aux administrations (inspecteurs, représentants syndicaux) et des parties externes indépendantes (chercheurs, associations d'utilisateurs) afin d'assurer une véritable transparence et une pluralité des perspectives. Les principes énoncés par le Conseil d'État en 2022¹ pour une IA de confiance que sont la primauté humaine, l'équité et la non-discrimination ainsi que la transparence devront servir de guide, notamment au travers de charte de déontologie. Les IA doivent rester limitées à des fonctions d'assistance dans des tâches secondaires, sans se substituer au jugement professionnel.

Enfin, il est nécessaire de préserver les compétences relationnelles et décisionnelles des agents publics, notamment par une formation initiale et continue adaptée, qui réaffirme clairement la primauté de

l'humain dans la relation à l'utilisateur. L'objectif est ainsi de maintenir une culture du service public fondée sur l'écoute, l'empathie et la responsabilité individuelle des agents, essentielle pour la confiance des citoyens envers leur administration.

La finalité dans les services publics doit toujours précéder les moyens. Si les relations des utilisateurs avec le service public sont redéfinies sur la base d'un cadre imposé par les possibilités de la technique, alors le service public risque fort d'entrer dans une dépossesion de lui-même. Au-delà de la promesse d'économie à court terme, le remplacement progressif des interactions humaines par des interfaces impersonnelles risque d'aggraver le sentiment d'éloignement ressenti par une partie croissante des citoyens, comme l'a rappelé la crise des « gilets jaunes ».

1. « Intelligence artificielle et action publique : construire la confiance, servir la performance », Conseil d'État, 2022.

Comment garantir les droits des usagers face à l'intelligence artificielle dans les services publics ? Pour un droit à l'explication

– Daniel Agacinski

Délégué général à la médiation (Défenseur des droits)

– Gabrielle du Boucher

Chargée de mission numérique, droits et libertés (Défenseur des droits)

Les promoteurs des solutions d'intelligence artificielle (IA) savent nous éblouir, mais il importe de ne pas nous laisser aveugler. Dans son rôle de vigie, le Défenseur des droits s'attache à analyser les évolutions de l'action publique afin de s'assurer qu'elles n'affectent pas les droits garantis aux usagers des services publics. Au cours des dernières années, la numérisation des démarches administratives a déjà fortement transformé les modalités d'accès aux droits et aux services publics, entraînant d'importantes inégalités¹ ; à son tour, l'intégration de systèmes d'IA dans le fonctionnement même des administrations ne va pas sans créer des risques d'inégalité de traitement et, plus largement, d'atteinte aux droits des usagers.

Dans le rapport *Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ?*², le Défenseur des droits interroge l'effectivité des garanties qui encadrent aujourd'hui le recours à des solutions d'IA dans les administrations et formule des recommandations afin d'assurer que, même « augmenté », le service public reste fidèle à ses principes.

Des outils déjà déployés qui suscitent des questions juridiques

À toutes les échelles, des administrations mobilisent déjà des systèmes algorithmiques et d'IA dans le traitement des données, souvent massives, de leurs usagers. On utilise aujourd'hui des systèmes d'IA pour procéder à l'anonymisation des décisions de justice, pour identifier des événements « à risque » dans l'espace public via la vidéosurveillance algorithmique, pour améliorer le diagnostic et la prévention en matière médicale, pour personnaliser les apprentissages, pour lutter contre le gaspillage alimentaire dans les cantines des établissements publics, pour guider le déploiement de politiques publiques ou encore analyser des besoins territoriaux (gestion de l'eau ou de l'éclairage public)...

Le cadre juridique actuel, dans lequel s'inscrivent ces déploiements, n'a rien d'une page blanche : dès qu'il

1. Voir *Dématérialisation des services publics : trois ans après, où en est-on ?*, Défenseur des droits, 16 février 2022 et *L'administration numérique pour les étrangers en France (ANEF) : une dématérialisation à l'origine d'atteintes massives aux droits des usagers*, Défenseur des droits, 11 décembre 2024.
2. *Les droits des usagers des services publics face aux algorithmes et aux systèmes d'IA : points de vigilance et recommandations de la Défenseure des droits*, Défenseur des droits, 13 novembre 2024.

Il y a traitement de données à caractère personnel, les dispositions de la loi « informatique et libertés¹ » s'appliquent ; dès qu'un algorithme est utilisé par une administration pour préparer ou prendre une décision individuelle, le code des relations entre le public et l'administration (CRPA) prévoit des obligations², auxquelles s'ajoutent d'autres dispositions propres à certains secteurs comme la santé ou l'enseignement supérieur. En outre, s'appliquent progressivement le règlement sur l'IA adopté par l'Union européenne en 2024³, ainsi que la convention-cadre sur l'IA adoptée par le Conseil de l'Europe⁴ la même année.

Le premier problème dans l'application de ce cadre tient à la nature même des systèmes dont on parle : afin de s'assurer que les systèmes mis en œuvre respectent le cadre juridique applicable, il faut que ceux qui les utilisent et ceux qui y sont soumis puissent comprendre leur fonctionnement, leur intérêt et leurs limites. Or, la caractéristique d'un système d'IA, par rapport à un algorithme « déterministe », qui applique la règle qu'on lui a donnée au départ, réside dans son caractère « auto-apprenant », c'est-à-dire dans sa capacité à élaborer lui-même les règles lui permettant d'atteindre le résultat demandé – au point que même les concepteurs de ces algorithmes « perdent la capacité à comprendre la logique des résultats produits », selon les termes de la Commission nationale de l'informatique et des libertés (Cnil)⁵.

Sur cette base, le Défenseur des droits a interrogé deux enjeux majeurs au regard de ses missions : d'une part, le dispositif juridique qui encadre aujourd'hui les décisions automatisées, en prévoyant notamment la nécessité d'une intervention humaine, est-il satisfaisant et quelle est concrètement la substance de la maîtrise humaine de ces systèmes ? D'autre part, les garanties de transparence et d'explicabilité actuellement prévues sont-elles suffisantes pour assurer le respect des droits des usagers ?

Automatiser les décisions : oui, mais jusqu'où ?

Le contrôle humain des systèmes automatisés, pré-requis à leur utilisation, s'opère, de façon schématique, à deux niveaux : celui du « système » lui-même, qui concerne le fonctionnement global de l'algorithme ou du système d'IA ; celui de la prise de décision « individuelle », quand le système est utilisé pour ce faire.

Le contrôle « système » vise la conception comme l'exploitation du système et implique des tests et audits réguliers, afin de s'assurer de la conformité globale du système au droit et de vérifier qu'il reflète les choix opérés. Par exemple, le code, quand il y en a un, respecte-t-il les règles de priorité fixées par certains textes ? Le risque de biais discriminatoire, que le Défenseur des droits avait souligné dans un rapport dédié⁶, est-il suffisamment étudié et évité ? Ce contrôle doit aussi permettre d'éviter qu'un système fonctionne sur la base d'impensés : des choix qui n'auraient pas été faits au niveau politique ou administratif et qui seraient codés ou opérés au niveau informatique, sans réelle réflexion et décision. Enfin, ce contrôle implique de pouvoir arrêter le système ou de le revoir s'il produit des suggestions inexplicables, des résultats dangereux, ou si son utilité n'est pas attestée. Il s'agit en résumé d'assurer la supervision effective de l'outil, tout au long de son cycle.

Quand le système est utilisé par l'administration pour prendre une décision individuelle, le cadre juridique actuel envisage deux situations distinctes : soit la décision est « partiellement automatisée », c'est-à-dire qu'un agent intervient au moment de la prise de décision, le système ayant constitué une aide à la décision ; soit la décision est « entièrement automatisée » – ce qui est possible depuis la loi « informatique et libertés » de 2018⁷.

1. Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

2. Voir notamment les articles L. 311-3-1, L. 311-5, L. 312-1-3, R. 311-3-1-1 et R. 311-3-1-2.

3. Règl. (UE) 2024/1689, 13 juin 2024, établissant des règles harmonisées concernant l'intelligence artificielle.

4. Convention-cadre sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, adoptée par les ministres des Affaires étrangères des 46 pays membres du Conseil de l'Europe, le 17 mai 2024 à Strasbourg.

5. *Comment permettre à l'Homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Cnil, 2017, p. 55.

6. *Algorithmes : prévenir l'automatisation des discriminations*, Défenseur des droits, 31 mai 2020.

7. Loi n°2018-493 du 20 juin 2018 relative à la protection des données personnelles.

Pour qu'une décision administrative individuelle entièrement automatisée soit légale, il faut, d'une part, que l'organisme public maîtrise le traitement algorithmique en question et qu'il ne repose pas sur des « données sensibles » au sens de la loi « informatique et libertés » ; d'autre part, il faut que la personne concernée soit informée de cette automatisation, puisse accéder à la compréhension du fonctionnement de l'algorithme et bénéficier d'une intervention humaine en cas de recours contre la décision. Si le Conseil constitutionnel a validé ces dispositions, il a précisé que les algorithmes auto-apprenants ne pouvaient pas servir de fondement exclusif d'une décision entièrement automatisée, précisément dans la mesure où ils ne peuvent pleinement remplir la condition de compréhension, identifiée comme nécessaire.

Mais la distinction, rassurante tant pour le juriste que pour l'usager, entre décisions « partiellement » et « entièrement » automatisées, tient-elle la route ? Pour considérer qu'une décision administrative n'est que « partiellement » automatisée, il convient de s'assurer que l'intervention humaine ne se limite pas à un « simple geste symbolique » qui reviendrait uniquement à valider le résultat produit par l'algorithme. Cela suppose que l'agent concerné effectue une action positive, concrète et significative, à partir ou à côté du résultat généré par l'algorithme, par exemple une opération intellectuelle de vérification, sans quoi l'automatisation devrait être considérée comme totale.

Or, certaines décisions, aujourd'hui présentées comme « partiellement » automatisées, le sont peut-être plus qu'il n'y paraît. Deux raisons conduisent à le penser : d'une part, dans les faits, le nombre massif de certaines décisions invite à s'interroger sur la réalité de l'intervention humaine annoncée, l'idée du recours à l'algorithme étant précisément d'alléger le travail des agents – ou de réduire le nombre d'agents – et d'accélérer les décisions... ; d'autre part, l'existence largement étayée de biais de confiance¹ conduit à

relativiser la substance effective d'une intervention humaine.

Qui peut nier qu'il aurait spontanément tendance à faire confiance aux résultats de la machine (biais d'automatisation) ? Ou qu'il est plus facilement acceptable de « se tromper » en suivant le résultat prescrit par la machine qu'en ayant choisi de s'en écarter ? Ainsi les écrits de Günther Anders² ou encore de Jacques Ellul³ nous permettent de comprendre que le fonctionnement de l'objet technique tend à soumettre ou à orienter notre volonté et notre capacité de décision. On court alors le risque de voir les agents publics perdre une part de leur esprit critique face aux résultats des algorithmes avec lesquels ils travaillent et de rendre ainsi illusoire la garantie qui avait été prévue par le législateur.

Assurer la transparence : oui, mais comment ?

Quand on parle des relations entre les usagers et les services publics, la transparence est un élément fondamental. Ancrée dans notre cadre juridique depuis l'article 15 de la Déclaration de 1789 qui établit que « la société a le droit de demander compte à tout agent public de son administration », la transparence constitue, selon les termes de l'ancien vice-président du Conseil d'État Jean-Marc Sauvé, « une composante du droit à une bonne administration »⁴.

Tant au niveau de chaque usager, qui doit savoir comment est traitée sa situation, qu'au niveau collectif, auquel les services publics doivent assurer leur redevabilité démocratique, cette transparence constitue une condition nécessaire pour s'assurer que les administrations ne commettent ni abus ni discrimination. Cependant, être transparent, ce n'est pas seulement être visible ou accessible : c'est pouvoir

1. Charlotte Barot, « Contrôle humain, décisions hybrides : quels enjeux ? », Linc/Cnil, 16 octobre 2024.

2. Günther Anders, *L'obsolescence de l'homme*, Paris, L'Encyclopédie des nuisances, 2002 [1956].

3. Jacques Ellul, *Le système technicien*, Paris, Le Cherche midi, 1977.

4. Jean-Marc Sauvé, « Transparence, valeurs de l'action publique et intérêt général », exposé d'ouverture lors du colloque « Culture du secret contre transparence sans limite : quel équilibre pour garantir l'intérêt général ? », Transparence International France, Assemblée nationale, 2011.

être compris, discuté, critiqué, amendé. Cette transparence est même une condition de l'intelligibilité, par la sphère publique, de sa propre action et de ses effets.

Or, une telle compréhension, qui n'est déjà pas simple s'agissant d'algorithmes déterministes – savez-vous déchiffrer un code-source ? –, est particulièrement difficile d'accès lorsqu'on parle d'IA, c'est-à-dire d'algorithmes auto-apprenants. Du fait de cette évolution liée à l'apprentissage automatique, la pondération des critères utilisés par le système d'IA pour aboutir à ses résultats ne peut pas être connue de façon certaine et stabilisée. Si bien qu'il est douteux qu'un système d'IA utilisé en vue d'une décision administrative puisse jamais pleinement satisfaire à l'exigence réglementaire de transparence, condition de la redevabilité de l'administration qui le met en œuvre.

Pour un droit à l'explication

Partant du principe qu'un usager ne peut être certain que ses droits ont été respectés que si l'administration est en mesure de lui expliquer les motifs de la décision qui a été prise à son endroit, le Défenseur des droits a recommandé de consacrer un « droit à l'explication » pour toutes les décisions administratives individuelles partiellement et entièrement automatisées, ce qui constituerait un garde-fou indispensable, pour prévenir le déploiement de systèmes dont les résultats cesseraient d'être supervisés au regard du droit applicable.

L'effectivité d'un tel droit suppose également qu'on approfondisse la connaissance de ces systèmes, et notamment de leurs effets sur le traitement d'une grande masse de situations ; cela implique de soutenir la recherche dans ce domaine, ainsi que les projets citoyens qui visent à nourrir le débat public sur ces sujets. Là encore, la transparence est une condition de la confiance.

Penser l'intelligence artificielle au service de la relation et de la qualité du service public : l'exemple de l'hôpital public

– Matthieu Girier

Directeur d'hôpital, cofondateur du think tank Le Sens du service public,
président de l'Association pour le développement des ressources humaines
dans les établissements sanitaires et sociaux (ADRHESS)

Continuités et ruptures dans l'action publique

L'intégration de l'intelligence artificielle (IA) dans les services publics en France suscite, comme toute nouvelle technologie, autant d'attentes que de craintes. À l'instar de l'informatisation dans les années 1980 ou de l'introduction du numérique au tournant des années 2000, elle marque un tournant technologique majeur, porteur de transformations durables pour l'action publique. Cet espoir est conditionné, toutefois, par la nécessité d'être pensé avec exigence, discernement et ambition, dans le respect des valeurs démocratiques qui gouvernent notre société. L'enjeu, ici, n'est pas seulement technique : il est profondément politique, au sens noble du terme. Derrière les algorithmes, les modèles prédictifs ou les systèmes d'aide à la décision se jouent, en effet, des choix de société : sur le rôle de l'État, la place de l'humain, la finalité du service public.

Cette irruption de l'IA s'inscrit dans un contexte de transition systémique, où le modèle de service public hérité des Trente Glorieuses est confronté à des injonctions paradoxales : répondre à des besoins croissants avec des marges budgétaires contraintes ; offrir une réponse individualisée dans des cadres collectifs ; innover sans rompre avec les principes d'égalité, de solidarité et de continuité. L'IA, en tant que

technologie d'optimisation et de prédiction, semble en mesure d'offrir des réponses à certains de ces défis. Mais elle ne saurait en elle-même constituer un projet. Elle interroge notre capacité à articuler innovation technologique et valeurs républicaines, efficacité opérationnelle et justice sociale, rationalité algorithmique et intelligence sensible.

Il ne s'agit pas simplement d'adopter une nouvelle génération d'outils, mais de repenser les finalités du progrès au prisme de l'intérêt général. L'IA ne peut être qu'un vecteur, jamais un guide autonome. Elle doit être mobilisée au service d'un projet humaniste du service public, fidèle à ses principes d'universalité, de transparence, de participation et d'adaptabilité. Toute mise en œuvre de l'IA engage une certaine conception du service public : s'agit-il de réduire les coûts ? de gagner du temps ? de mieux comprendre les besoins ? de personnaliser la relation ? Ces objectifs ne sont pas neutres et les arbitrages qu'ils impliquent doivent être explicites, démocratiques et évaluable.

L'IA en appui de l'intervention humaine : des cas d'usage

Dans ce cadre, une orientation claire s'impose : l'IA doit être conçue non comme un substitut, mais

comme un appui. Appui aux agents publics, qu'elle peut aider à mieux maîtriser les situations complexes, à renforcer leur capacité d'action et à améliorer leurs conditions de travail. Appui aux usagers, citoyens et patients, en facilitant l'accès à leurs droits, en personnalisant les réponses sans les uniformiser, en rendant les démarches plus compréhensibles. En aucun cas, elle ne peut se substituer au discernement professionnel, à l'écoute active ou à la délibération publique. L'IA ne peut ni sentir, ni juger, ni prioriser selon des valeurs : elle peut seulement calculer, prédire, ordonner selon des règles qui, elles-mêmes, reflètent des choix implicites. Le risque de délégation sans contrôle, de standardisation masquée ou de fracture éthique n'est jamais totalement écarté.

C'est particulièrement vrai dans les domaines où les décisions relèvent d'un arbitrage stratégique, d'une appréciation humaine ou d'un choix politique. Par exemple, une analyse automatisée des besoins de santé d'un territoire, aussi fine soit-elle, ne saurait à elle seule trancher l'implantation d'un hôpital ou la réorganisation d'un service. Ces décisions sont éminemment politiques, et doivent être assumées comme telles. L'IA peut utilement éclairer la décision en croisant des données de santé publique, des flux de patients, des projections démographiques ou des indicateurs sociaux. Elle ne peut, en revanche, ni décider à la place du décideur, ni neutraliser les tensions éthiques, sociales et territoriales inhérentes à ces choix. L'objectivité algorithmique est souvent un leurre : tout système d'intelligence artificielle est tributaire des données sur lesquelles il a été entraîné, des biais qu'il reproduit, des finalités qu'on lui assigne.

Ce constat renforce la nécessité d'une régulation exigeante, fondée sur des principes de transparence, de loyauté, d'auditabilité et de contrôle humain. La régulation de l'IA ne peut reposer uniquement sur la conformité technique : elle doit intégrer des critères de soutenabilité sociale, d'équité territoriale, de robustesse démocratique. À cet égard, les travaux engagés au niveau européen, avec le Règlement sur l'intelligence artificielle (AI Act), ouvrent une voie structurante, à condition que les administrations publiques s'en saisissent activement et en fassent un levier d'alignement entre innovation et valeurs.

Là où l'IA trouve toute sa pertinence, c'est notamment dans la gestion d'une partie des tâches à faible

valeur ajoutée humaine : celles qui mobilisent du temps, génèrent de la fatigue organisationnelle et freinent la disponibilité pour la relation de service. L'automatisation du traitement des demandes simples, la mise à jour des plannings hospitaliers ou la reconfiguration des tâches quotidiennes sont autant d'exemples concrets où l'IA, combinée à la *robot process automation* (RPA), peut renforcer la fluidité, la fiabilité et l'efficacité des processus. Dans les hôpitaux, les caisses d'allocations familiales (CAF), les rectorats, les collectivités territoriales, des expérimentations sont déjà en cours. Elles montrent que l'IA peut contribuer à libérer du temps, à mieux sécuriser les parcours administratifs, à fiabiliser la gestion des ressources. Mais ces outils doivent être mobilisés avec proportionnalité, en fonction des objectifs poursuivis et des contextes d'usage. Il n'existe pas une IA universelle, mais des IA à concevoir sur mesure, en dialogue constant avec les agents et les usagers.

Un levier potentiel de transformation des organisations publiques à débattre et maîtriser

Ce dialogue est essentiel. Il suppose des espaces d'expérimentation et d'évaluation associant les professionnels de terrain, les usagers, les développeurs, les juristes, les chercheurs. Il suppose aussi une montée en compétences des agents publics, tant sur les usages de l'IA que sur ses limites, ses enjeux éthiques, ses impacts organisationnels. L'acculturation est un préalable à toute appropriation : sans compréhension partagée, les technologies deviennent opaques et leur légitimité s'érode. Former, outiller, accompagner les agents publics dans la compréhension de ces technologies, dans leur usage critique et dans leur articulation avec le travail réel est une condition de réussite et de justice.

Dans cette optique, l'IA peut devenir un levier de transformation positive des organisations publiques : amélioration de la qualité de vie au travail, revalorisation du temps professionnel autour du cœur de métier, développement de dynamiques collectives fondées sur la coopération homme-machine. Elle

peut aussi soutenir une logique de capacitation, en élargissant l'autonomie décisionnelle des professionnels par un meilleur accès à l'information, à l'anticipation, à la modélisation des impacts. Encore faut-il éviter deux écueils symétriques : celui d'une fascination techniciste qui survaloriserait les promesses de l'IA et celui d'un rejet de principe qui condamnerait tout usage innovant. La voie d'un humanisme technologique est étroite, mais féconde. Elle suppose une gouvernance démocratique de l'innovation, une régulation des usages, une responsabilité assumée des commanditaires publics.

À plus long terme, le défi est celui d'une IA civique. Une IA conçue, régulée et déployée selon les principes de la République, intégrant la diversité des besoins, respectant les libertés fondamentales, favorisant la cohésion sociale. Cela implique une action publique forte, capable de piloter les choix technologiques, de soutenir une filière européenne éthique de l'IA, de garantir l'indépendance des données, d'assurer la transparence des usages. Cela suppose aussi un effort soutenu de recherche publique, de normalisation éthique, de mutualisation des bonnes

pratiques. La transformation numérique du service public ne peut être laissée aux seules mains des prestataires ou des logiques de marché. Elle doit rester un projet collectif, porté par l'État, les collectivités, les établissements, les agents eux-mêmes.

Il nous faut donc adopter une approche lucide et pragmatique : l'IA ne remplace ni la compétence, ni le jugement, ni la confiance. Elle les prolonge, les enrichit, les potentialise. Ce n'est pas à elle de décider, de hiérarchiser, de dialoguer : ce rôle reste, et doit rester, l'apanage des femmes et des hommes qui font vivre le service public au quotidien. C'est à cette condition qu'elle pourra devenir un outil de renforcement du service public, et non un facteur de sa déshumanisation. À rebours d'une logique de substitution, il s'agit de faire de l'IA un levier d'émancipation pour les agents et d'accessibilité renforcée pour les usagers. Un outil d'amélioration continue et non une fin en soi. Un allié exigeant, qui oblige à repenser nos organisations, nos priorités et nos pratiques, mais qui, bien encadré, peut contribuer à redonner sens à l'action publique.

**Décider : vers une véritable
stratégie politique de l'IA
à l'échelle européenne,
nationale et locale**

Le défi de la démocratisation

– Gilles Babinet¹

Auteur², entrepreneur autodidacte, président de la mission Café IA, *Digital champion* France
auprès de la Commission européenne et membre du Comité IA,
ancien coprésident du Conseil national du numérique

Dans le rapport de la commission de l'intelligence artificielle (IA) auquel vous avez contribué, il est indiqué qu'« une mobilisation collective, massive, sans délai et au long cours est impérative ». Où en sommes-nous aujourd'hui ?

Un certain nombre d'agents utilisent aujourd'hui l'intelligence artificielle générative (IAG) sans le dire, et sans cadre au sein des administrations (*shadow AI*), ce qui présente des risques, notamment en matière de fuite de données. Il nous faut tout à la fois développer la culture de la confiance, mais également se doter de règles d'utilisation des solutions d'IA, comme nous y invitons dans notre rapport³.

Compte tenu de la situation budgétaire, des nombreuses contraintes et des multiples attentes à l'égard des institutions publiques, l'IA s'imposera et pourrait transformer en profondeur la logique verticale et cloisonnée de notre État.

Cette technologie présente à la fois des éléments de rupture et des éléments de transition par rapport à la phase de transformation numérique de l'État qui a précédé. Nous avons besoin de synchroniser ce mouvement avec celui de la société, ce qui suppose d'engager tout le système, de formaliser une doctrine stratégique de l'utilisation de l'IA et de disposer des infrastructures idoines – cloud, KYC (*Know Your Customer*)⁴, nomenclatures, modèles de langage (LLM), architecture de données fiable...

L'IA doit être au service de la qualité du service public. Elle peut avoir des apports intéressants en matière d'enseignement, de police, de justice, d'éducation pour adapter les cours aux besoins très hétérogènes des élèves, en matière d'urbanisme, de détection de la fraude fiscale, de diagnostic médical et de libération de temps médical, ou encore de *workflows*.

Pour garantir notre souveraineté, nous devons faire le choix de l'*open source*, par réallocation des budgets investis dans des logiciels propriétaires. Il faut interdire l'utilisation de certains outils comme ceux de Microsoft. Pour éviter le déclassement, il est important de passer du temps sur la technique et de définir ce qui est utile ou pas pour mener l'action publique, ce qui requiert une capacité d'évaluation des systèmes d'IA.

Nous avons, bien sûr, un enjeu de financement, évalué par la commission sur l'IA à 5 milliards d'euros par an d'investissement public pendant cinq ans⁵, notamment dans la technologie, mais nous avons aussi un besoin d'expertise internalisée au sein des administrations. Le rapport de la commission sur l'IA comporte un certain nombre de préconisations à ce sujet, en particulier pour développer l'attractivité du secteur public pour les profils numériques, ce qui pose davantage la question de la culture managériale et organisationnelle, de l'autonomie, que la question de la rémunération.

1. Propos recueillis par Émilie Agnoux le 5 juin 2025.

2. Il a notamment publié en mars 2024 *Green AI. L'intelligence artificielle au service du climat* (Odile Jacob, 2024).

3. *IA : notre ambition pour la France*, Commission de l'intelligence artificielle, mars 2024.

4. Processus d'identification et de vérification de l'identité.

5. *IA : notre ambition pour la France*, rapport cité, 2024.

Mais il est important de dire que le déploiement de l'IA doit d'abord prendre appui sur une vision stratégique et des choix stables dans la durée. Nous devons clarifier le projet de société à l'ère de l'IA.

Il faudrait, par exemple, un *chief technology officer* (CTO) directement rattaché au Premier ministre, qui puisse s'entretenir une fois par mois avec ce dernier afin de décliner opérationnellement les grandes orientations. Nous avons besoin avant tout de choix politiques et non pas uniquement techniques, par exemple pour décider d'adopter un *cloud* souverain. C'est un enjeu de souveraineté majeur.

Certains pays en Europe tracent une voie intéressante et peuvent servir de modèle, à l'instar du Danemark. Il est aussi intéressant de noter que ce sont des États fédéraux ou régionalisés (Espagne, Italie) qui ont fait d'importants progrès en matière numérique ces dernières années.

La ministre déléguée chargée de l'IA et du numérique, Clara Chappaz, vient de vous confier la mission de pérenniser et étendre l'initiative des cafés de l'IA recommandés par la commission de l'IA et mis en œuvre par le Conseil national du numérique. En quoi est-ce une étape importante ?

Tous les pays qui ont pris de l'avance (Taïwan, Estonie, Danemark...) ont commencé par développer une culture générale de l'IA. C'est un préalable pour développer des outils et usages de bonne qualité, mais c'est aussi un prérequis démocratique. La formation doit s'adresser à l'ensemble de la population, et commencer dès l'école. L'IA suscite à la fois des aspirations, par exemple en matière d'accès aux droits et de simplification, mais aussi des inquiétudes en matière de discriminations, d'opacité, de non-explicabilité.

Les cafés de l'IA que nous avons déjà expérimentés vont désormais pouvoir se déployer plus largement.

On comptera 2 millions de Françaises et de Français concernés d'ici à 2027.

L'objectif est de favoriser une appropriation collective, de comprendre, de débattre, de tester, de développer des compétences, de se former techniquement. Beaucoup de nos concitoyens et concitoyennes se posent des questions. Les pouvoirs publics doivent les accompagner dans ce cheminement et dans le développement d'une culture de l'IA.

Les services publics sont des lieux privilégiés pour mener ces échanges, tout comme les structures associatives, les tiers-lieux, les entreprises... Nous sommes dans une logique d'éducation populaire au numérique qui dépasse la seule logique institutionnelle. Grâce à cette généralisation des cafés de l'IA, nous pourrions à la fois nourrir et diffuser des ressources partagées.

Face au besoin d'expertise à l'intérieur des organisations publiques, mais aussi de renforcement de la confiance de la population envers l'État, vous êtes également un fervent partisan d'une mobilisation citoyenne au cœur même de l'appareil administratif. De quoi s'agit-il concrètement ?

Nous avons l'opportunité de créer une force de cyber-réserve constituée de citoyens, comme l'a fait la Suisse en matière de défense. Max Weber nous mettait en garde contre le risque de désenchantement et de perte de sens lié à la sophistication technique¹. Avec les algorithmes aujourd'hui, nous avons un risque de méta-bureaucratie, qui peut créer du ressentiment et un affaiblissement de la démocratie.

Cette force citoyenne, à mi-chemin entre agents publics et citoyens, pourrait intervenir dans cinq secteurs, en appui de l'État :

1. L'UX² de l'État et le socle numérique, afin que les outils soient pensés pour les citoyens et les agents,

1. Max Weber, *Le savant et le politique*, Paris, La Découverte, 2003 [1919].

2. L'UX, pour « user experience », désigne l'expérience utilisateur, qui se traduit par la prise en compte des besoins, des usages, des attentes de ceux à qui sont destinés les outils et les démarches afin qu'ils soient facilement accessibles et utilisables.

pour limiter les biais, améliorer les systèmes, créer de nouveaux services ;

2. une Autorité de régulation de la communication audiovisuelle et numérique (Arcom) citoyenne, sur le modèle de la nétiquette¹, afin de développer la culture du respect en ligne, l'État ne pouvant être le seul à faire la police ;
3. une Viginum² étendue pour lutter contre les ingérences étrangères et le risque de désinformation ;
4. la défense (renseignement, contre-espionnage, cyberdéfense), à l'image de ce qui se pratique en Ukraine avec des cybervolontaires ;

5. l'inclusion, pour aller chercher ceux qui sont éloignés du numérique.

Des méthodes existent pour intégrer ces ressources au fonctionnement de l'État (par exemple la qualification progressive). Le ministère des Armées a lancé une réserve opérationnelle numérique (RON) qui répond à cette exigence, avec un contrat formalisant le nombre de jours d'engagement pour chaque réserviste, et une grande diversité de missions (sécurité informatique, architecture réseau, administrateur système, *data scientist*...).

1. Ensemble de règles qui définissent un cadrage déontologique explicitant les comportements de politesse, de courtoisie, de respect à respecter sur internet.

2. Viginum est le service technique et opérationnel de l'État chargé de la vigilance et de la protection contre les ingérences numériques étrangères.

Service public local et intelligence artificielle : l'heure des premiers choix politiques

– Jacques Priol

Ancien dirigeant territorial, président de Civiteo, cofondateur de l'Observatoire Data Publica

Cela fait bientôt trois ans que nos sociétés sont entraînées dans le maëlstrom de l'intelligence artificielle (IA). Un récit majeur s'impose : cette évolution est incontournable, l'adoption des IA génératives est irréversible, la course aux grands modèles de langage devient un enjeu vital de souveraineté et d'autonomie stratégique. Malgré quelques voix dissonantes, ce récit est remarquable de puissance et de persistance.

Les débats du printemps 2023 sur la nécessité pour les uns de faire une pause et pour les autres d'accélérer la recherche face aux dérives possibles des IA génératives¹ illustrent la maîtrise de quelques acteurs majeurs sur les termes du débat public. La question pour eux n'est pas d'obtenir une décision ou une régulation, mais d'imposer les débats sur l'IA comme sujet central, qui justifie des investissements privés toujours plus massifs et des soutiens publics précipités, quand ce n'est pas inconsidérés !

Ce faisant, l'engouement des entreprises et des services publics pour l'IA (générative ou non) est incontestable. Après la sidération, la fascination et les premiers pas.

Le service public local n'est pas sous cloche

Une étude de l'Observatoire Data Publica a montré que les collectivités françaises n'ont pas tardé à s'intéresser aux systèmes d'IA (SIA). En 2021, un rapport de la Commission européenne avait identifié cinq projets locaux en France², le Conseil d'État en citait dix-neuf en août 2022³. Mi-2024, ce sont 75 % des régions, 62 % des métropoles, mais aussi 14 % des communes de moins de 10 000 habitants qui ont engagé des projets avec de l'IA. En 2025, plus d'une collectivité locale sur deux (de plus de 3 500 habitants) a commencé à utiliser des SIA pour la gestion du service public⁴ !

Cet élan montre que l'action publique locale et les élus et élus qui en ont la charge sont en prise directe avec la réalité et les débats qui irriguent notre société. Élan ne signifie pas aveuglement et se conjugue avec prudence. Des territoires freinent et vont parfois jusqu'à interdire temporairement certains usages de l'IA⁵. D'autres construisent des outils qui sont autant de garde-fous et de conditions préalables, parfois très rigoureuses, au déploiement des systèmes d'IA au sein du service public.

1. Yoshua Bengio, « Aujourd'hui, l'intelligence artificielle, c'est le Far West ! Nous devons ralentir et réguler », *Le Monde*, 28 avril 2023 ; Yann Le Cun, « L'idée même de vouloir ralentir la recherche sur l'IA s'apparente à un nouvel obscurantisme », *Le Monde*, 28 avril 2023.

2. *AI Watch, Artificial Intelligence in public services*, Commission européenne, 2020.

3. *Intelligence artificielle et action publique : construire la confiance, servir la performance*, Conseil d'État, août 2022.

4. *Les collectivités locales et la donnée*, 3^e Baromètre de l'Observatoire Data Publica, novembre 2024. Enquête menée auprès d'un échantillon de 289 collectivités et établissements publics locaux représentatif de la diversité des collectivités françaises.

5. « Montpellier interdit temporairement ChatGPT à ses agents et engage un travail éthique », *La Gazette des communes*, 20 avril 2023.

Toutes les politiques publiques sont concernées

Les 250 projets identifiés et analysés par l'Observatoire Data Publica entre septembre 2023 et décembre 2024¹ montrent que des collectivités de toutes les tailles sont concernées. Et la variété des usages possibles de l'IA est spectaculaire.

De nombreux projets concernent la vie courante du service public, avec de l'IA générative pour produire des documents internes ou même des actes juridiques, gérer les instances officielles et leurs comptes rendus, gérer des dossiers administratifs, instruire des demandes d'aide et de subvention, ou même préparer les travaux des commissions d'appels d'offres. Les directions du numérique testent des SIA pour coder, outiller leur politique de cybersécurité ou améliorer la performance de l'assistance aux utilisateurs. Les équipes financières envisagent des SIA pour du contrôle de gestion ou pour accélérer le traitement de la facturation. Les directions des ressources humaines testent des IA qui analysent et trient des candidatures.

Le recours à l'IA s'impose aussi comme un vecteur d'efficacité pour renforcer la performance de la gestion bâtementaire (réduire la consommation des fluides) ou de la gestion des espaces verts (optimiser l'arrosage, par exemple). Des IA prédictives sont embarquées dans la gestion d'automatismes industriels (dans des stations d'épuration, des centres de tri des déchets, l'éclairage urbain, etc.). En matière d'urbanisme, des IA analysent les plans d'occupation des sols, et plusieurs projets visent à accélérer l'instruction des permis de construire. L'analyse des flux de mobilité, la gestion de carrefours « intelligents » ou l'optimisation de la gestion des réseaux de transport urbain sont aussi des domaines propices au test de systèmes d'IA.

Dans des écoles, des SIA servent à prédire la fréquentation des cantines scolaires pour lutter contre le gaspillage alimentaire. Les politiques sociales ne sont pas en reste, avec des algorithmes proposés pour

lutter contre le non-recours, pour modéliser des parcours et faire de la prévention, pour trier de façon rapide les dossiers pour l'attribution de logements sociaux ou de places en crèche.

La gestion de la sécurité dans l'espace public voit apparaître une nouvelle gamme d'outils à travers les « caméras augmentées ». L'IA et la sécurité, ce sont aussi des projets destinés à améliorer la résilience des territoires : détection et modélisation de la propagation des feux de forêt, modélisation d'inondations, anticipation des risques de sécheresse.

Le recours à l'IA pour améliorer la relation aux usagers est un sujet qui intéresse de nombreux territoires, avec deux approches distinctes : l'IA au guichet avec des *chatbots* et *callbots* ou l'IA comme aide aux agents publics (renseignement et personnalisation des réponses).

Les enjeux d'un déploiement choisi et maîtrisé

Loin d'être en retard, le service public local s'est très vite engagé dans une phase de découverte des bénéfices possibles du recours à l'IA, mais aussi des limites et des risques, en engageant des démarches visant à garder le contrôle et à maîtriser l'intégration des SIA dans la production de l'action publique.

Mieux : le service public local français a très rapidement mis en exergue des risques nouveaux inhérents, non pas aux SIA eux-mêmes, mais à leur compatibilité avec des principes fondamentaux du service public comme garant de l'intérêt général.

La formation et l'acculturation des décideurs sont une étape essentielle. L'explication des mécanismes d'apprentissage des SIA est le dernier des soucis des promoteurs de solutions. Il est pourtant essentiel de comprendre comment fonctionnent ces outils probabilistes qui se nourrissent de données massives pour calculer des résultats et générer des contenus vraisemblables à des vitesses vertigineuses.

1. IA & territoires, après la découverte le temps des premiers choix, L'Observatoire Data Publica, juillet 2025.

L'Observatoire Data Publica rapporte aussi que les décideurs publics sont prudents et souhaitent vérifier, avant tout déploiement, la performance réelle et même l'utilité des systèmes d'IA qui leur sont proposés, par exemple en termes d'accélération des processus, de fiabilisation d'automatismes, d'accès à de nouvelles connaissances, de personnalisation des actions, d'économies budgétaires, d'amélioration de la performance des politiques. De nombreuses offres commerciales à l'attention du service public vantent bien plus la performance possible du système que la performance du service public : on vend des IA qui marchent, mais qui ne servent à rien.

Il est ensuite essentiel de se saisir à bras-le-corps de la question de la confiance. De nombreuses chartes ou doctrines existent. Elles exposent des principes qui sont juridiques, techniques, économiques, éthiques, mais aussi philosophiques, moraux ou politiques. On évoque la primauté humaine, la bienfaisance, la transparence, la démocratie, la souveraineté (économique, politique, technique...), l'autonomie stratégique, l'encapacitation, l'équité et la non-discrimination, la protection de la vie privée... Mais lire ces chartes, ou mieux y adhérer, n'aide guère à se faire une opinion concrète, et encore moins à piloter l'introduction d'un SIA. Ces principes doivent être traduits en repères pratiques et actionnables à l'échelle d'une collectivité publique. Ainsi, par exemple, la souveraineté et l'autonomie stratégique sont érigées en condition préalable à tout déploiement d'IA d'intérêt général par le Conseil d'État. Mais à l'échelle d'une ville de 50 000 habitants, comment s'en saisir ? On peut privilégier un modèle français, certes. Mais un modèle français entraîné sur des données inconnues n'est-il pas moins bien maîtrisé qu'un modèle américain entraîné sur des données ouvertes ? Le contrôle des biais et des erreurs et la vérification d'une absence d'atteinte à la vie privée (des usagers du service public) ne se règlent pas par une charte. Des mesures concrètes sont à prendre. Elles passent par la formation opérationnelle des agents publics, par des clauses contractuelles exigeantes vis-à-vis des fournisseurs, par des contrôles (et des compétences pour les assurer), par

l'instauration de règles managériales strictes pour mettre en place et garantir dans la durée la supervision humaine.

La question environnementale est, bien sûr, centrale. Le développement des IA génératives et leur utilisation massive font sortir de toute trajectoire carbone acceptable aussi bien les géants du numérique que les entreprises et les services publics qui les utiliseraient de façon inconsidérée¹... La question de l'empreinte des SIA utilisés est une composante essentielle de toute décision. Les décideurs locaux doivent donc se doter de moyens de mesure, et pouvoir imposer aux acteurs du numérique qu'ils livrent des chiffres vérifiables. Précisons que les collectivités françaises de plus de 50 000 habitants sont soumises depuis le 1^{er} janvier 2025 à l'obligation de se doter d'une « stratégie numérique responsable » dont l'objet est justement de limiter (ou d'infléchir) l'empreinte carbone des outils numériques du service public². Dans un contexte de dématérialisation toujours plus poussée, l'arrivée de l'IA place les décideurs face à une injonction paradoxale. Et si la France promeut dans divers cercles internationaux un référentiel pour une IA frugale³, dans bien des cas l'IA la plus frugale sera celle que l'on n'utilise pas !

Une transparence démocratique à inventer

Nos concitoyens s'interrogent et s'inquiètent des conditions d'utilisation des IA par le service public.

Plusieurs démarches récentes sont à saluer : la convention citoyenne sur l'IA organisée par Montpellier Métropole, la consultation de conseils locaux du numérique à Angers ou Rennes, la concertation « Nantes débat de l'IA ! » et bien d'autres encore. Il ressort de ces travaux des enseignements ambivalents. D'un côté, nos concitoyens s'inquiètent des risques et des dérives possibles des systèmes d'IA : empreinte environnementale non maîtrisée, atteintes

1. « Sam Altman, le patron d'OpenAI, s'inquiète du gouffre énergétique de l'IA », *Le Figaro*, 18 janvier 2024.

2. Loi n°2021-1485 du 15 novembre 2021 visant à réduire l'empreinte environnementale du numérique en France.

3. Référentiel général pour l'IA frugale, Afnor, 2024.

à la vie privée, risques de biais et d'erreurs, voire de discriminations algorithmiques, aggravation des fractures numériques, déshumanisation complète du service public...

D'un autre, ils souhaitent voir le service public local se saisir des outils les plus efficaces disponibles s'ils offrent des retombées tangibles au bénéfice du territoire et de ses habitants.

Ce sont donc les habitants qui le disent : l'enjeu est politique.

Pour être complet sur la question démocratique, il faut aussi que les élu·es et élus aient conscience des obligations de transparence qui s'imposent au service public. Au nom des libertés fondamentales de nos concitoyens, les collectivités publiques ont l'obligation d'informer les usagers en cas d'utilisation d'un algorithme pour les décisions qui les concernent. Cette obligation trouve sa source dans l'article 15 de la Déclaration des droits de l'homme et du citoyen selon lequel « la société a le droit de demander compte à tout agent public de son administration ». À l'ère de l'IA, le Conseil constitutionnel en tire une position tranchée : les exigences de transparence des décisions publiques rendent impossible l'utilisation « comme fondement exclusif d'une décision administrative individuelle, des algorithmes susceptibles de réviser eux-mêmes les règles qu'ils appliquent¹ ». En d'autres termes : l'IA ne peut pas être utilisée pour automatiser des prises de décisions individuelles. Elle ne peut être utilisée que sous le contrôle et la supervision d'agents publics, qui demeurent seuls responsables de la prise de décision. Voici une contrainte, heureuse, que ne connaissent pas les entreprises.

Un enjeu électoral en 2026 ?

Lancer quelques projets avant la fin du mandat pour les uns, explorer des pistes nouvelles pour les autres, les candidats aux élections municipales s'intéressent à l'IA. Parmi les centaines de projets en cours, chacun peut faire ses courses. Encore faut-il avoir à l'esprit qu'il s'agit de projets... Parmi les 250 cas étudiés par l'Observatoire Data Publica entre septembre 2023 et décembre 2024, seuls 21 ont été passés en production. Et il ne s'agit pas d'une conséquence d'une faiblesse de notre service public, puisque selon Gartner le taux d'abandon des projets d'IA avant la mise en production, tous secteurs confondus, est de 80 %² !

Avant d'imaginer un usage utile, raisonnable et maîtrisé des IA au service des politiques publiques locales, beaucoup reste à faire. Malheureusement, hormis quelques financements, les élus locaux n'ont pas grand-chose à attendre de l'échelon national. Disons-le clairement ici : loin de prioriser les usages d'intérêt général, les récents gouvernements ont fait de la promotion de la *French Tech* et de l'appel aux investissements privés l'alpha et l'oméga d'une politique publique de l'IA.

D'autres réponses sont possibles, elles viendront peut-être des territoires et notamment des communes et des intercommunalités qui sauront s'organiser au printemps 2026 en désignant des élu·es et élus délégués à l'IA, en faisant des choix ciblés et maîtrisés de recours à l'IA pour des métiers précis et des priorités politiques assumées, en refusant explicitement les dérives et les offres fantaisistes qui pululent, et en informant les citoyens de ces usages tout en concertant leurs développements. Car au final, ce sont les élu·es et élus qui décident, et les citoyennes et citoyens qui choisissent.

1. Décision n°2018-765 DC du 12 juin 2018.

2. Pourquoi 80 % des projets d'IA ne voient jamais le jour, Forbes, janvier 2025.

Pour une souveraineté démocratique et sociale sur l'intelligence artificielle

– Corinne Narassiguin

Sénatrice et co-autrice du rapport parlementaire *ChatGPT et après ?*¹

L'intelligence artificielle (IA) est désormais une réalité opérationnelle.

Mais cette révolution technologique, largement conduite par quelques grandes entreprises américaines ou chinoises, échappe encore en grande partie à la puissance publique et obéit largement à des logiques commerciales. Ces outils, alors même qu'ils ont le potentiel pour faire progresser l'humanité, sont pour l'heure pensés, codés et gouvernés loin de nos choix démocratiques, de nos valeurs républicaines et de nos ambitions sociales et environnementales.

L'IA vient, en effet, percuter nos conceptions du progressisme et de l'humanisme. Ne pas repenser nos modèles de société à l'aune de cette révolution technologique serait une abdication démocratique, autant qu'une erreur historique.

L'IA est l'inverse d'un outil neutre. Elle traduit toujours des choix politiques. Quels critères sont optimisés ? Quelle transparence est garantie ? Quel contrôle démocratique est possible ? Laisser se développer des IA « boîtes noires », c'est affaiblir l'État de droit.

En 2021, le gouvernement néerlandais avait démissionné après la mise au jour de l'utilisation d'un algorithme pour la lutte contre la fraude aux allocations familiales, dont les graves biais discriminatoires avaient eu des conséquences désastreuses pour des milliers de personnes accusées à tort².

En matière de souveraineté, le constat est également préoccupant : l'Union européenne et la France dé-

pendent largement des infrastructures et modèles d'IA des Gafam. Sans investissement public massif dans les modèles de fondation (ou LLM), les infrastructures de data centers et de supercalculateurs, sans politique publique claire sur les données publiques ouvertes et l'utilisation des données privées, nous serons des clients à perpétuité de technologies conçues ailleurs.

Enfin, à l'instar des citoyennes et citoyens, l'État et les collectivités expérimentent l'IA à grande vitesse, parfois sous prétexte d'une promesse d'efficacité ou d'économies budgétaires. Mais sans cadre public clair, ces usages peuvent vite dériver : automatisation opaque de décisions administratives, biais discriminatoires, algorithmes inaccessibles aux citoyens, dépendance croissante à des fournisseurs privés étrangers, fuites de données sensibles vers des infrastructures d'entraînement d'IA commerciales...

Face à ces constats, l'IA doit devenir un objet de politique publique à part entière, et suppose donc un changement de cap dans la stratégie de l'État.

Le temps de l'encadrement éthique périphérique est déjà dépassé : ce dont nous avons besoin, c'est une politique publique offensive.

Pour servir l'intérêt public, nous avons une responsabilité : celle de défendre une IA au service de l'émancipation, de l'égalité et de la démocratie. La gauche a toujours porté la primauté du bien commun sur les logiques de marché. L'IA ne peut être l'exception.

1. Patrick Chaize et Corinne Narassiguin, *ChatGPT, et après ? Bilan et perspectives de l'intelligence artificielle*, rapport n°170, Sénat, 28 novembre 2024.

2. « Pays-Bas. Scandale des allocations familiales : un avertissement qui montre l'urgence d'interdire les algorithmes racistes », Amnesty International, 25 octobre 2021.

Une stratégie publique de l'IA doit être fondée sur :

- l'investissement public ;
- la transparence démocratique ;
- la souveraineté technologique ;
- l'éthique du service public.

Donnons à la France les moyens de choisir son avenir technologique.

Faire de l'IA un bien commun de la puissance publique française

L'État doit reprendre la main sur les grandes orientations de développement de l'IA. Cela passe par un soutien massif à la recherche publique et les partenariats public-privé, à l'image de ce que font les États-Unis avec la Defense Advanced Research Projects (DARPA) et la National Science Foundation (NSF), ou la Chine avec ses plans quinquennaux.

Appeler les entreprises privées à investir, sans fixer un cadre ambitieux de recherche et développement au service du bien commun, c'est abandonner toute ambition de mettre l'IA au service de l'intérêt général.

La recherche publique française en IA (INRIA, CNRS, universités) est historiquement de haut niveau. Mais elle est aujourd'hui sous-financée face à la puissance des investissements privés.

Le laboratoire commun entre l'INRIA et l'AP-HP développe, par exemple, des algorithmes d'aide au diagnostic éthique et validés médicalement. C'est un exemple qu'il faudrait suivre.

Nous avons besoin de renforcer les liens organiques entre le monde de la recherche et l'enseignement supérieur, par exemple en créant des chaires IA et service public dans les grandes écoles et universités.

Mais rien ne sera possible sans sortir des dogmes budgétaires actuels sur la réduction à tout va de la dépense publique. Au minimum, il faudrait doubler

le budget français de la recherche publique en IA d'ici à 2027.

Créer une véritable Agence publique de l'IA

Pour définir et mettre en œuvre une véritable action publique de l'IA, il faut un acteur institutionnel fort, capable de coordonner les stratégies nationales, de financer les projets stratégiques, de piloter l'éthique et la transparence dans les usages publics.

Créons une Agence nationale de l'IA, dotée d'une capacité d'investissement, d'un rôle d'animation territoriale et de pouvoirs de régulation. La Direction interministérielle au numérique qui existe déjà pourrait jouer ce rôle avec des pouvoirs étendus. Cette agence pourrait porter le développement de modèles ouverts et souverains, garantir la transparence algorithmique dans les administrations et soutenir les écosystèmes publics/privés ancrés localement.

D'autres pays ont déjà tracé cette voie. La Finlande a mis en place un AI Hub public rassemblant État, chercheurs et collectivités, qui a permis d'ores et déjà un déploiement maîtrisé de l'IA dans les hôpitaux et les transports¹.

Développer une IA éthique, explicable et contrôlable

Nous devons rendre obligatoires, pour tous les usages publics de l'IA, des exigences de transparence, d'explicabilité et de contrôle humain. Chaque citoyen a le droit de comprendre comment une décision le concernant est prise. Cela implique un encadrement juridique robuste, mais aussi la formation des agents publics, des élus et des citoyens.

1. Patrick Chaize et Corinne Narassiguin, *ChatGPT, et après ?*, op. cit., 2024.

Commençons par intégrer dans tous les appels d'offres publics une clause de transparence et d'explicabilité algorithmique et un cadre clair sur la maîtrise des données d'entrée et de sortie. Généralisons l'usage raisonné de l'IA comme formation obligatoire pour les agents publics.

Certaines collectivités ont pris les devants. Ainsi, Nantes métropole impose dans ses marchés publics un audit éthique des algorithmes utilisés, en partenariat avec les chercheurs du Cnam et des citoyens¹.

IA et écologie, relever le défi de l'impossible équation

Le développement de l'IA, porté par une logique capitaliste de croissance et d'innovation permanente, entre en contradiction flagrante avec les impératifs écologiques. L'entraînement des modèles d'IA nécessite des infrastructures massives, consommatrices d'énergie et de ressources naturelles, accentuant l'empreinte carbone numérique.

Mais de nombreuses applications de l'IA sont déjà très prometteuses pour faciliter la transition écologique, dans l'agriculture et dans l'industrie. Un projet progressiste et humaniste de l'IA doit viser à en faire un outil au service de la transition écologique, encadré par des objectifs collectifs et développant les énergies vertes.

Là encore, le rôle de l'investissement public et des partenariats pour créer des écosystèmes de recherche et développement public-privé est essentiel. C'est l'IA générative qui est la plus consommatrice d'énergie dans sa phase de développement. Il est souvent plus judicieux d'investir dans des IA spécialisées, peu coûteuses en énergie, pour des usages ciblés et maîtrisés.

C'est à l'État de coordonner un plan national d'usage de l'IA, avec des objectifs ambitieux : réduction drastique des usages superflus et réorientation des capa-

cités d'IA vers des secteurs utiles (agriculture durable, efficacité énergétique, santé publique).

Assurer une souveraineté européenne

L'*AI Act*², entré en vigueur en 2024, est une innovation législative importante. Elle est imparfaite, mais nécessaire. Originellement construite sur la maîtrise des risques en fonction des usages, elle introduit le bon cadre de régulation et de standardisation. Modifiée en urgence pour inclure l'avènement des IA génératives, elle montre aussi ses limites : il est difficile de réguler des technologies non expliquées et en constante évolution.

L'Union européenne devra donc accepter d'avoir à réviser régulièrement l'*AI Act*, de garder cette approche fondée sur les usages et de se concentrer essentiellement sur l'évaluation, le contrôle et la mitigation des risques et en conséquence de l'adoption de normes et de standards garantissant la qualité des produits d'IA, selon des critères éthiques permettant la confiance.

Pour cela, l'Union européenne doit se donner les moyens d'être prescripteur en matière de normes et de standards, pour peser fortement au sein de l'OCDE et ne pas se laisser dépasser par les travaux conduits aux États-Unis par le National Institute of Standards and Technology (NIST).

L'Europe investit massivement avec un plan de 200 milliards d'euros, baptisé *InvestAI*³, vise à créer des IA-gigafactories, notamment en France, pour développer des infrastructures souveraines, comme les supercalculateurs de nouvelle génération.

La startup française Mistral AI, soutenue publiquement, développe des modèles ouverts et installe un *data center* à haute performance en Île-de-France, illustrant un modèle d'innovation européenne autonome. Mais c'est un exemple trop rare.

1. « Intelligence artificielle : comment Nantes Métropole encadre et débat du recours à l'IA », Nantes.fr, 18 juin 2025.

2. Règlement européen sur l'intelligence artificielle, *Journal officiel de l'Union européenne*, 12 juillet 2024.

3. « L'UE lance l'initiative InvestAI, destinée à mobiliser 200 milliards d'euros d'investissements dans l'intelligence artificielle », Commission européenne, communiqué de presse, 11 février 2025.

Les grands modèles de fondation (ou LLM) coûtent des centaines de millions d'euros à entraîner. L'Europe doit mutualiser ses moyens pour ne pas dépendre des infrastructures étrangères.

Il est urgent d'investir dans un cloud européen de l'IA, libre et interopérable, alimenté par une énergie décarbonée.

Mais il faut voir plus loin. L'Europe doit être le lieu du prochain grand bon technologique de l'IA : nouvelles infrastructures de calcul pour l'IA frugale ; *data centers* économes en eau et en énergie ; processeurs quantiques... Qui sait aujourd'hui ce qui nous surprendra demain ?

Cela passe par des programmes de recherche et développement publics-privés européens massifs, des centres d'excellence universitaires qui travaillent en réseau pour être compétitifs internationalement avec les États-Unis.

L'Union européenne, avant d'être un système institutionnel, devrait d'abord être un système de valeurs fondées sur l'humanisme, la démocratie, la responsabilité sociale et environnementale. Assurer la souveraineté européenne en matière d'IA, c'est se donner les moyens d'agir pour le bien public.

Table des matières

01 Introduction. Garder la main sur nos destins
_Émilie Agnoux, Johan Theuret

04 15 propositions

Connaître et former : de l'utilité de maîtriser au préalable de quoi on parle

08 Gérer les risques et accompagner les usagers :
l'IA comme opportunité pour l'action publique
_Anca Caruntu, Thaima Samman

12 Donner un sens au travail avec l'intelligence artificielle
_Yann Ferguson

16 IA et éducation : faire avec ou faire sans ?
_Diana Filippova

Encadrer : pour une IA au service de la relation, de la qualité des services publics et des droits des citoyens

22 La transparence, un prérequis essentiel pour l'intelligence artificielle
dans le secteur public
_Soizic Pénicaud

26 L'IA dans la relation à l'utilisateur : vers la déshumanisation automatisée ?
_Marius Bertolucci

29 Comment garantir les droits des usagers face à l'intelligence artificielle
dans les services publics ? Pour un droit à l'explication
_Daniel Agacinski, Gabrielle du Boucher

33 Penser l'intelligence artificielle au service de la relation et de la qualité
du service public : l'exemple de l'hôpital public
_Matthieu Girier

**Décider :
vers une véritable stratégie politique de l'IA à l'échelle européenne,
nationale et locale**

- 37 Le défi de la démocratisation
 –**Gilles Babinet**
- 40 Service public local et intelligence artificielle : l'heure des premiers
 choix politiques
 –**Jacques Priol**
- 44 Pour une souveraineté démocratique et sociale sur l'intelligence artificielle
 –**Corinne Narassiguin**

Reconnue d'utilité publique dès sa création, la Fondation Jean-Jaurès est la première des fondations politiques françaises. Elle est présidée par **Jean-Marc Ayrault**.

Indépendante, européenne et sociale-démocrate, elle se veut depuis plus de trente ans un lieu de réflexion, de dialogue et d'anticipation.

La collection des « Rapports », dirigée par **Laurent Cohen** et **Jérémie Peltier**, répond à l'ambition de faire naître analyses pertinentes et propositions audacieuses, mais aussi de mettre cette production intellectuelle et politique au service de tous.

© Éditions Fondation Jean-Jaurès
12, cité Malesherbes - 75009 Paris

www.jean-jaures.org

Derniers rapports et études :

06_2025 : La « troisième gauche ». Enquête sur le tournant post-sociétal de la gauche européenne
László Andor, Arthur Borriello, Marco Enriquez-Ominami, Amaury Giraud, David Goodhart,
Justine Hermelin, Thomas Laran, Renaud Large, Marc Le Chevallier, Denis Maillard,
Lawen Redar, Jonathan Smucker, Lilith Verstrynge, Adélaïde Zulfikarpasic

06_2025 : Employés à domicile : connaître et reconnaître
Agathe Cagé, Serge Da Mariana, Édouard Lecerf, Leïla Hicheur

06_2025 : Solidarité internationale : la lucidité des Français
Hakim El Karoui, Gilles Finchelstein, Hervé Le Bras, Rémy Rioux

06_2025 : Brèves d'Olympe, et autres souvenirs de Paris 2024
Pierre Brémont

06_2025 : Un an après : les Jeux olympiques, la France et moi
Camille Andrieu, Richard Bouigue, Fabrice Février, Guénaëlle Gault, Jean-Daniel Lévy,
David Medioni, Anton Molina, Laurence de Nervaux, Jérémie Peltier, Yves-Paul Robert,
Pierre Rondeau, Adélaïde Zulfikarpasic

06_2025 : Et si l'IA était au service de la démocratie ? L'exemple du Grand débat national
Bassem Asseh, Dorian Dreuil, Antoine Jardin, Hugo Micheron, Simon-Pierre Sengayrac,
Daniel Szeftel

06_2025 : La France à table, 2^e édition. Tensions et mutations autour de notre rapport
à l'alimentation
Agnès Crozet, Guénaëlle Gault

06_2025 : Métropoles et grandes villes en 2045. Quatre concepts pour réinventer la ville
Dylan Buffinton

05_2025 : 2005-2025 : les vingt années qui ont transformé l'Europe. La méthode de l'Union
face aux crises
Renaud Bellais, Brando Benifei, Amandine Clavaud, Sylvain Kahn, Chloé Mikolajczak,
Laetitia Thissen, Théo Verdier

-  fondationjeanjaures
-  @j_jaures
-  fondation-jean-jaures
-  www.youtube.com/c/FondationJeanJaures
-  fondationjeanjaures
-  fondationjeanjaures
-  [fondationjjjaures.bsky.social](https://foundationjjjaures.bsky.social)
-  bit.ly/4g6UANC

Abonnez-vous !



www.jean-jaures.org

Fondation
Jean Jaurès
ÉDITIONS