



Commission d'Oxford sur l'IA
et la bonne gouvernance



L'IA dans le service public: Des principes à la pratique

Commission d'Oxford sur l'IA et la bonne gouvernance

DÉCEMBRE 2021

L'IA dans le service public: Des principes à la pratique

Commission d'Oxford sur l'IA et la bonne gouvernance

DÉCEMBRE 2021



Commission d'Oxford sur l'IA
et la bonne gouvernance



Sommaire

Résumé	1
Avant-propos	2
1. Introduction	3
2. Défis	4
3. Principes	7
4. Recommandations	8
Qui doit porter conseil sur l'IA du service public ?	8
Organe consultatif scientifique sur les répercussions sur le public	8
Organe arbitral indépendant et accessible	9
Travailler ensemble	10
Recommandation 1	10
Comment renforcer les capacités du service public en matière d'IA pour favoriser la bonne gouvernance ?	11
Recommandation 2	11
Comment garantir la fiabilité et la crédibilité de l'IA du service public ?	12
Recommandation 3	12
5. Conclusion : Idées pour un plan d'action immédiat	13
Récapitulatif des objectifs	14
6. Recherche de l'OxCAIGG	15
À propos de la Commission	15
Commissaires	16
Remerciements	18



Résumé

Les premières tentatives d'introduction de l'intelligence artificielle dans le service public indiquent que les États du monde entier auront de lourds défis à relever pour s'engager dans cette voie.

C'est pourtant une occasion unique de s'appuyer sur la coopération pour que le recours aux systèmes d'IA soit propice à une bonne gouvernance et nous permette de résoudre certains des problèmes les plus urgents et les plus insolubles du secteur public.

Au terme de travaux de recherche et de délibérations, la Commission d'Oxford sur l'IA et la bonne gouvernance a étudié les principaux défis à relever afin d'exploiter efficacement les systèmes d'IA pour le service public. Ces efforts ont débouché sur trois *questions* essentielles :

- Qui doit porter conseil sur l'IA du service public ?
- Comment renforcer les capacités du service public en matière d'IA pour favoriser la bonne gouvernance ?
- Comment garantir la fiabilité et la crédibilité de l'IA du service public ?

Nous en avons conclu *qu'il convient d'atteindre plusieurs objectifs* pour utiliser l'IA dans l'intérêt public

- La conception de cette IA doit se faire selon une approche inclusive.
.....
- Toute acquisition d'IA doit se faire sur les conseils d'une agence publique bien éclairée.
.....
- Le déploiement de l'IA dans le service public doit être réfléchi.
.....
- Les systèmes d'IA doivent être constamment exploités dans le respect d'obligations redditionnelles vis-à-vis des parties prenantes.
.....

Pour atteindre ces objectifs, nous formulons trois recommandations :

- I Les États, les industries et la société civile doivent collaborer à l'échelle internationale pour mettre en place et habiliter a) un **organe scientifique** international afin de faire avancer la recherche sur les applications à IA du service public et ; b) un **organe d'arbitrage** à même de trancher les litiges susceptibles de survenir entre les parties prenantes impliquées dans les systèmes d'IA du service public.
.....
- 2 Les États d'envergure nationale, profitant du soutien de ces deux nouvelles organisations internationales, doivent a) **renforcer les capacités** de leur service public respectif en faveur d'engagements profonds en conception, acquisition, déploiement et en obligations redditionnelles vis-à-vis des systèmes d'IA du service public et b) proposer des **boîtes à outils au personnel** des agences publiques pour qu'il puisse accomplir sa mission de surveillance.
.....
- 3 La confiance accordée au recours à l'IA doit être renforcée au moyen de campagnes de sensibilisation du public sur des applications quotidiennes et de véritables cas pratiques qui se profilent à l'horizon, les impacts et risques connexes ; a) en demandant aux États de **communiquer en toute franchise** sur l'usage prévu des technologies d'IA dans le service public et ; b) en mettant en place une agence multisectorielle qui **établit un système de certification de base** permettant d'approuver continuellement les applications et d'instiller un sentiment de confiance dans l'esprit du public au fil du temps.
.....

L'atteinte de ces objectifs peut être favorisée par l'adoption immédiate de trois mesures :

- I Réaliser des études de faisabilité pour les agences scientifiques et d'arbitrage.
- II Consulter les agences nationales et multilatérales existantes sur les possibilités à envisager par ces organes en matière d'élargissement et de perfectionnement des capacités et bonnes pratiques actuelles.
- III Planifier l'engagement à l'occasion de plusieurs des événements marquants sur l'innovation technologique figurant déjà à l'agenda international des 36 prochains mois.



Avant-propos

Si l'intelligence artificielle semble prometteuse pour résoudre certains des défis les plus urgents en matière de politique publique et de changement sociétal, les États ne sont peut-être pas prêts à la déployer en faveur d'une bonne gouvernance.

Ces dernières années, bon nombre d'États ont cherché à tirer parti de nouvelles technologies comme l'apprentissage automatique, le big data et d'autres outils algorithmiques, afin d'élaborer et de mettre en œuvre de meilleurs programmes et politiques pour leurs administrés. Bien que l'innovation dans le secteur public doive faire l'objet de mesures incitatives, nombre de projets de cette nature ont connu des débuts difficiles, d'autant plus que certaines collaborations avec les industries ont échoué. Par ailleurs, des données de qualité médiocre et des technologies défectueuses ont produit des résultats inattendus et souvent biaisés.

Ces expériences ont soulevé une question cruciale : comment exploiter au mieux l'IA pour le service public ?

De 2020 à 2021, nous avons participé à un projet, soutenu par l'université d'Oxford, pour répondre à cette même question, à savoir, la Commission d'Oxford sur l'IA et la gouvernance (OxCAIGG). Notre groupe était constitué de responsables, de cadres, d'universitaires, d'avocats indépendants et de conseillers gouvernementaux représentant un éventail de postes et d'intérêts, en provenance du monde entier. Sur une période de dix-huit mois, l'OxCAIGG a organisé des briefings d'experts, consulté des conseillers techniques de l'université d'Oxford, produit des recherches exclusives sur des sujets liés à l'IA et à la bonne gouvernance, et organisé plusieurs journées d'ateliers.

Selon nos points de vue en tant que commissaires, les recherches effectuées à ce jour et nos propres expériences, nous pensons qu'il est temps que des organes indépendants et internationaux assument la responsabilité et la mission de concevoir et de soutenir la mise en œuvre de normes et de bonnes pratiques pour une application efficace de l'IA dans le secteur public.

Au fur et à mesure que chacun des États du monde réalise des progrès par rapport à l'adoption de ces technologies, ceux-ci pourraient tirer parti d'agences de facilitation susceptibles de leur fournir des conseils sur les échanges avec l'industrie et des moyens d'éviter les résultats politiques ou programmatiques biaisés ou de piètre qualité, ainsi que des services d'arbitrage ou de règlement des différends. De telles agences se révéleraient également utiles à de nombreux États dans le monde dont les capacités institutionnelles ou réglementaires sont limitées en matière d'adoption de technologies.

En tant que commission, nous avons élaboré une série de documents de travail sur l'évaluation des opportunités et risques intrinsèques au recours à l'IA dans le service public dont nous avons tiré des enseignements pour élaborer nos recommandations. Notre but consiste à réduire les résultats contre-productifs du recours à l'IA et à soutenir le développement d'une gouvernance mondiale plus normalisée pour cette technologie émergente.

Nous avons également pour but de formuler des recommandations de nature structurelle, d'ordre général dans leurs descriptions, et basées sur un ensemble de principes s'inscrivant dans le droit fil de nos recherches. Cette approche apportera des marges de manœuvre pour leur mise en œuvre avec souplesse et innovation en tenant compte des observations des parties prenantes concernées.

En tant que commission, nous avons cherché à aller au-delà de l'élaboration de principes d'utilisation de l'IA (dont il existe déjà de nombreuses listes) afin de mettre en place des directives politiques et des cadres organisationnels concrets que les acteurs nationaux et internationaux pertinents puissent prendre en considération.

Ce rapport, *L'IA dans le service public*, n'est pas un recueil bibliographique en soi ; les références figureront dans des résultats de recherche particuliers produits au cours de l'existence de la Commission. Il s'agit d'une déclaration succincte de nos délibérations et consultations, visant à récapituler ce que nous considérons comme les questions critiques et les objectifs vitaux, nos recommandations, et les prochaines étapes du programme nécessaires pour mettre l'IA au service de la bonne gouvernance.

Les commissaires de la Commission d'Oxford sur l'IA et la bonne gouvernance



1 Introduction

Les expériences et résultats précoces indiquent qu'il y aura de sérieux défis à relever lorsque les États tenteront de mettre en pratique les outils de l'intelligence artificielle dans le service public. Ces balbutiements sont une occasion unique durant laquelle la coopération et les normes consensuelles pourraient avoir un impact significatif pour atténuer ces défis tout en maximisant les opportunités.

La Commission d'Oxford sur l'IA et la bonne gouvernance (OxCAIGG) a étudié plusieurs défis concrets autour de l'IA dans le service public, en mettant un accent particulier sur utilisation, par les démocraties du monde entier, des outils d'IA à des fins de programmes publics. L'objectif de ce rapport final de la Commission est d'évaluer ces défis et de s'appuyer sur les principes fondamentaux de l'utilisation de l'IA pour la bonne gouvernance afin de proposer des recommandations spécifiques permettant de surmonter ces défis.

Depuis son lancement en 2020 jusqu'à la production de ce rapport final en 2021, l'OxCAIGG a :

- Financé et produit sept documents de travail de recherche exclusifs, sept articles de blog et deux éditoriaux d'opinion sur les questions contemporaines concernant l'IA et la bonne gouvernance.
- Organisé huit réunions d'experts auprès de parties prenantes de l'industrie, de la société civile et du secteur public.
- Organisé quatre tables rondes avec les commissaires et les conseillers techniques de l'OxCAIGG.
- Présenté son travail à des décideurs de haut niveau d'agences publiques nationales, supranationales et multilatérales.
- Présenté des éléments tangibles en réponse aux réunions de témoignage d'experts issus d'agences publiques nationales et multilatérales.

Ces recherches exclusives, conjuguées à nos briefings d'experts et à notre engagement auprès des responsables publics, ont documenté nos délibérations et les recommandations présentées dans le présent document. Dans ce rapport, nous soulignons les principaux défis, proposons un ensemble de principes fondamentaux et identifions les mesures à prendre dès maintenant.

Les mesures proposées sont récapitulées sous forme de recommandations exploitables, fondées sur des données probantes, qui permettront aux États et aux administrations publiques d'exploiter les bienfaits de l'IA. Dans le respect des principes fondamentaux de l'OxCAIGG, nous nous engageons à :

- Agir avec agilité et détermination pour comprendre l'impact des outils de l'IA actuellement utilisés dans la gouvernance et, dans l'administration publique, pour garantir le bien-être social ainsi que pour fournir des biens et services publics.
- Aider les décideurs et les entrepreneurs à répondre aux questions politiques et aux applications du secteur public, et à concevoir des idées qui placent l'IA et l'apprentissage automatique au cœur du service public.
- Identifier des solutions et orienter les processus d'élaboration des politiques de manière à renforcer la voie de développement d'intégration de l'IA pour résoudre les problèmes sociaux plutôt que les compliquer ainsi que pour instaurer la confiance dans l'esprit du public quant à l'utilisation de ces outils.



2 Les défis

L'adoption des outils d'IA par le secteur public conjugue les risques systémiques inhérents aux technologies de l'IA aux questions pratiques liées à l'élaboration de politiques et programmes innovants pour le secteur public.

En se servant de ces outils, les États n'auront d'autre choix que de faire évoluer leurs pratiques d'achat, de développement et de mise en œuvre des programmes. Ils seront confrontés aux questions relatives à la collecte de données de formation, à l'évaluation de systèmes technologiques complexes, à la formation du personnel et à l'embauche de nouveaux experts, entre autres.

Les recherches menées par l'OxCAIGG ont passé en revue un grand nombre des usages les plus courants de l'IA en matière de gouvernance et de service public à ce jour, et se sont également penchées sur l'avenir, en étudiant certains défis à relever prochainement. Nos chercheurs ont mis en évidence une série de défis qui entravent l'adoption de l'IA au sein du service public. Nos recommandations visent à atténuer les difficultés des défis actuels et à proposer la souplesse nécessaire pour s'adapter aux défis de demain. Nous présenterons ici des récapitulatifs des recherches entreprises par l'OxCAIGG avant d'évoquer les principaux défis identifiés dans le cadre de ces travaux.

Les réactions dans le monde face à l'émergence de l'IA, de l'apprentissage automatique et de la prise de décision automatisée^[1]

Dans leur document de travail s'appuyant sur des données d'enquête issues d'un échantillon de 154 195 personnes interrogées dans 142 pays, recueillies dans le cadre de l'enquête sur les risques mondiaux de la Fondation Lloyd's Register, Lisa-Maria Neudert, Aleks Knuutila et Philip Howard analysent les principaux indicateurs des perceptions du public concernant les inconvénients et opportunités potentiels de l'intégration de l'IA à nos affaires personnelles et à notre vie publique. Ces experts ont constaté que la perception de l'IA par le public varie fortement selon les régions et les groupes socio-économiques considérés. Il existe des divisions Est-Ouest notables, les préoccupations du public concernant l'IA étant les plus élevées en Europe (43 %), en Amérique latine (49 %) et en Amérique du Nord (47 %), tandis qu'en Asie du Sud-Est (25 %) et en Asie de l'Est (11 %), des proportions relativement faibles de personnes pensent que l'IA serait nuisible. Dans les différentes professions, cadres d'entreprise et responsables publics (47 %) et les autres professionnels (44 %) sont les plus enthousiastes à l'égard de l'IA, tandis que les travailleurs de l'industrie manufacturière (35 %) et des services (35 %) sont moins confiants.

Enseignements pratiques pour projets d'IA d'État : Témoignages de quatre initiatives de villes intelligentes^[2]

À l'occasion de ces travaux, Godofredo Ramizo Jr. étudie l'exploitation par les États de l'intelligence artificielle dans leurs prestations de services publics en se concentrant sur les projets de villes intelligentes axés sur l'IA. En s'appuyant sur une analyse documentaire approfondie et des entretiens exclusifs réalisés auprès de hauts fonctionnaires de Hong Kong, de Malaisie et de Singapour qui ont travaillé sur des projets de villes intelligentes et des projets similaires axés sur l'IA, Ramizo démontre la diversité des projets étatiques d'IA et identifie des principes pratiques pour favoriser la protection de l'intérêt public. Son document de travail révèle que les États rencontrent des difficultés en matière d'acquisition, de mise en œuvre et d'évaluation de l'impact de l'IA, et qu'ils peinent à déterminer la viabilité des projets sur les plans financier, technique et politique. En particulier, lorsque les entreprises technologiques disposent de ressources et d'une influence vertigineuses, le pouvoir de négociation des États est mis à mal.

L'intelligence artificielle dans les administrations locales^[3]

Dans ce document de travail, Thomas Vogl se penche sur le recours à l'intelligence artificielle par les autorités locales britanniques. Bien que l'on compte un certain nombre de projets aboutis liés à l'automatisation du back-office, à l'analyse prédictive pour l'aide à la décision et à l'utilisation de chatbots pour échanger avec les résidents, M. Vogl constate que les États sont confrontés à des défis pratiques majeurs pour réussir leur adoption de l'IA. Il démontre que les autorités locales doivent améliorer leurs capacités de collecte et d'analyse des données, définir plus clairement les problèmes avant de rechercher des solutions à base d'IA, et transmettre aux fournisseurs des connaissances contextuelles sur les autorités locales et leurs processus.



Old Cracks, New Tech : Intelligence artificielle, droits de l'homme et bonne gouvernance dans des sociétés hautement fragmentées et socialement stratifiées.

Le cas du Kenya^[4]

Dans le cadre de son travail, Nanjala Nyabola enquête sur la politique du gouvernement kényan en matière d'IA et de technologie blockchain, et évalue son succès. En effectuant une revue de la documentation connexe et une analyse de documents politiques et de travail, elle démontre qu'au Kenya, des principales mises en pratique de cette technologie sont axées sur des soins de santé abordables, la sécurité alimentaire, la fabrication, le logement, la cybersécurité et la constitution de titres fonciers. Mme Nyabola constate que le déploiement de l'IA dans des sociétés très fragmentées comme le Kenya risque de creuser les clivages existants, notamment ceux entre classes sociales et identités, et que l'éthique liée à l'utilisation de l'IA en milieu industriel diffère dans ses implications et son impact sociétal comparativement au secteur public. À l'instar de nombreux pays en développement, le Kenya commence tout juste à fixer des cadres juridiques pour régir l'exploitation des technologies.

La surveillance en tant que service : le marché européen de la surveillance de masse assistée par IA^[5]

Dans son document de travail, Yung Au examine le marché européen de la production et de l'exportation de systèmes de surveillance assistés par IA vers des États du monde entier. Cette commission étudie ce qu'elle appelle la surveillance en tant que service : à savoir des services et des logiciels fournis à des fins de surveillance, et qui consistent en des systèmes complexes proposés avec des interfaces intuitives, ainsi que des services continus de maintenance, de mise à jour et d'assistance dépannage. Son analyse se concentre sur trois exemples de services de ce type qui se sont imposés comme des priorités d'une réglementation majeure ces derniers temps : la reconnaissance et les analyses faciales ; la reconnaissance et l'analyse de la parole ; et les systèmes d'analyse comportementale et de nudging. La frontière devient poreuse entre technologies d'IA et applications de surveillance de masse, d'autant plus que leur potentiel de nuisance ne cesse de progresser. En l'absence de réglementation suffisante, ce marché pourrait être à l'origine de préjudices de longue durée.

Harmonisation de l'intelligence artificielle : le rôle des normes dans le règlement de l'UE sur l'IA^[6]

Dans ce document de travail, Mark McFadden, Kate Jones, Emily Taylor et Georgia Osborn étudient le rôle des normes technologiques en faveur du développement sûr, équitable et innovant de l'intelligence artificielle, défini dans le projet de règlement européen sur l'IA. Ce document de travail met en exergue la complexité d'une normalisation dans ce contexte, et que le lien entre les normes et les objectifs de la Commission européenne repose sur une rencontre difficile entre des parties prenantes, des intérêts économiques et des organisations de

développement de normes de renom. S'appuyant sur de vastes recherches et la consultation des parties prenantes, ce projet de règlement définit un cadre exhaustif pour la gouvernance et les normes de l'IA. Ce document se concentre sur le rôle accordé par le projet de règlement aux normes en matière d'IA. Plus précisément, la conformité aux normes harmonisées instituera une présomption de conformité pour les applications et services d'IA à haut risque, offrant ainsi un niveau de confiance quant à leur conformité aux exigences onéreuses et complexes du règlement proposé tout en créant de fortes incitations pour que l'industrie se plie aux normes européennes.

Les principaux défis

Ces rapports et documents d'information présentent des recherches exclusives, ou de manière singulière, un large éventail de recherches de pointe. Cet ensemble de travaux s'appuie sur le volume grandissant de la documentation liée aux sciences sociales et politiques sur l'usage de l'IA dans le service public, et retient des expériences et des études de cas du monde entier. Ces travaux identifient les principaux défis à relever pour mettre l'IA au service de l'intérêt public. Les exemples positifs et constructifs peuvent inspirer l'espoir, tandis que les mauvaises expériences méritent d'être critiquées et doivent éclairer les décisions de demain.

Tout d'abord, plusieurs contributions soulignent les complexités institutionnelles et structurelles de la mise en place de normes pour l'IA. Bien que les avantages d'un certain niveau de normalisation soient évidents, une feuille de route scientifiquement étayée pour la normalisation et des capacités dédiées à la normalisation de l'IA s'imposent. Actuellement, il n'existe pas de mécanismes efficaces de règlement des différends.

Deuxièmement, l'adoption de l'IA pour le service public se heurte à des défis très concrets et pratiques liés à l'acquisition de l'IA ainsi qu'à la collecte et à l'analyse des données de formation. Les fonctionnaires manquent non seulement d'expertise et de compétences, mais également de boîtes à outils pratiques pour prendre de bonnes décisions. De toute évidence, les puissantes entreprises technologiques ont un pouvoir de négociation et une expertise supérieurs à ceux des États et des administrateurs publics. Les fonctionnaires ont besoin de capacités techniques et pratiques pour l'adoption d'une IA propice à la bonne gouvernance.

Troisièmement, l'un des principaux défis auxquels se heurte le recours à l'IA dans le secteur public est la confiance des citoyens. Il est toujours primordial que les administrés aient confiance dans les services publics, et toute idée selon laquelle la mise en pratique de l'apprentissage automatique serait coûteuse, porterait préjudice aux principes d'équité ou serait source de nouveaux problèmes, ne fera que compliquer la tâche.

Enfin, même un usage bien intentionné de l'IA par les États pourrait renforcer les biais et les inégalités existants. L'insuffisance de cadres juridiques ou pratiques pour fonctionnaires pose



des problèmes à l'adoption de ces nouvelles technologies. Les technologies d'IA utilisées pour la gouvernance peuvent présenter des biais systémiques, des conséquences imprévues, et même des risques systématiques pour les droits de l'homme. Qui plus est, un marché sous-réglementé peut créer des normes techniques de facto qui ont un impact sur la vie publique en entretenant ces biais et ces inégalités, car aucune intervention d'intérêt public sur ces marchés ne capte ni ne corrige les biais et les inégalités à rectifier.

Regroupement de recherches

À partir des défis mis en évidence par ces recherches, et sur la base de nos réunions d'experts et de nos tables rondes internes, nous avons identifié trois lacunes fondamentales au niveau de la réglementation sur l'IA dans les services publics.

Qui doit porter conseil sur l'IA du service public ? Dans leur grande majorité, les États reconnaissent le potentiel de l'IA et souhaitent l'exploiter à des fins de bonne gouvernance. Toutefois, l'adoption de technologies d'IA s'accompagne de défis pratiques et normatifs très spécifiques que les États ne sont pas en mesure de relever seuls.

Comment renforcer les capacités du service public en matière d'IA pour favoriser la bonne gouvernance ? Les fonctionnaires jouent un rôle essentiel dans l'adoption de l'IA dans le service public. Le déploiement éclairé de l'IA nécessite des compétences technologiques et une sensibilisation auxquelles les agences publiques ont rarement accès.

Comment garantir la fiabilité et la crédibilité de l'IA du service public ? Nos recherches indiquent que le public se méfie de l'IA. Nous devons veiller à ce que les systèmes d'IA soient à la fois fiables et dignes de confiance pour qu'ils puissent être utilisés avec succès.



3 Les principes

Le paysage actuel de l'IA regorge de codes éthiques et de directives normatives. Les experts de renom des plus grandes institutions au monde ont proposé des principes pour régir de l'intelligence artificielle dans des domaines comme la justice pénale, les soins de santé et le développement durable.

L'intention de l'OxCAIGG n'est pas d'étoffer l'ensemble grandissant d'impératifs de haut vol pour l'IA. Au contraire, compte tenu de l'expérience et de l'expertise de ce groupe, nous avons cherché à élaborer un ensemble net et concis de recommandations exploitables pouvant orienter l'utilisation de l'IA par le secteur public. Pourtant, dès le lancement de la Commission, nous avons souhaité exprimer sans équivoque notre mission, à savoir faire progresser l'utilisation de l'IA précisément dans le service public et faire progresser la bonne gouvernance. Ainsi, nos principes équilibrent le besoin de cadres suffisamment flexibles pour permettre une mise en pratique globale, mais relativement précise pour orienter des mesures très spécifiques prises par les fonctionnaires et les politiciens pour le déploiement de ces nouveaux outils.^[7]

Le travail de l'OxCAIGG s'appuie sur les quatre principes clés suivants, qui ont constitué la base de notre travail en tant que commission et les recommandations présentées dans ce document :

- I La conception de l'IA doit être inclusive :**
Les outils et programmes d'IA doivent s'inspirer de l'expérience des experts du secteur public pour surmonter les problèmes de discrimination et de partialité liés à l'utilisation d'ensembles de données inadéquats, à l'exclusion des minorités et des groupes sous-représentés, ainsi qu'au manque de diversité dans la conception.
- II L'acquisition de l'IA doit se faire en toute transparence :**
Cela permettra de surmonter les défis liés à l'acquisition et au développement, à la conception et à la facilité d'utilisation des outils d'IA. Le processus d'acquisition doit inclure une évaluation des risques et des avantages du déploiement des outils d'IA.
- III Le déploiement de l'IA doit se faire en connaissance de cause :**
Les fonctionnaires doivent être formés aux questions d'interopérabilité, d'explicabilité, de partialité et d'intégration aux processus décisionnels.
- IV L'IA doit être exploitée dans le respect d'obligations redditionnelles :**
Les décisions prises par un système d'IA doivent être prises en toute transparence et éviter les résultats de type « boîte noire ». Cela implique la mise en place de processus pour surveiller et auditer le système d'IA.



4 Les recommandations

QUI DOIT PORTER CONSEIL SUR L'IA DU SERVICE PUBLIC ?

La Commission a discuté de questions fondamentales à propos du développement, de l'acquisition et de l'utilisation de l'intelligence artificielle par le secteur public. Au cours de nos délibérations, nous avons identifié deux nécessités réglementaires. La première est qu'un organe scientifique international spécialisé fasse progresser la recherche sur les audits algorithmiques, les impacts sociaux, les exemples d'utilisation et les bonnes pratiques politiques, et qu'il diffuse ces recherches afin d'inspirer et de coordonner l'exploitation de nouveaux systèmes d'IA pour résoudre les problèmes nécessitant une mobilisation collective. La seconde porte sur un organe d'arbitrage capable de régler rapidement et efficacement les différends entre développeurs, régulateurs et personnes concernées des systèmes d'IA à usage public. Ces organisations seraient indépendantes, mais complémentaires : l'organe scientifique fournirait des preuves impartiales à l'organe d'arbitrage, tandis que ce dernier signalerait la nécessité de mener des recherches sur les questions portées à sa connaissance.

Organe consultatif scientifique sur les incidences publiques

Le premier organe mondial, *un organe consultatif scientifique*, serait consacré aux discussions scientifiques, techniques et d'ingénierie sur l'impact de l'IA, de l'apprentissage automatique et d'autres systèmes algorithmiques de pointe sur les problèmes du secteur public.

S'inspirant du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC), une telle agence coordonnerait les discussions scientifiques, faciliterait l'examen par les pairs des recherches sur les développements techniques et les systèmes d'audit, et identifierait les points de consensus sur l'envergure du caractère inclusif des processus de conception, l'efficacité des processus d'acquisition, le fonctionnement des systèmes de mise en œuvre et l'impact des systèmes d'IA sur le public qu'ils sont censés servir. Toutefois, le plus important est que cet organe exerce non seulement un rôle technocratique mais également politique. Son équipe de direction doit être en mesure de produire des preuves de résultats économiques, culturels, et politiques des décisions techniques et d'ingénierie.

Cet organe scientifique doit disposer de l'expertise technique nécessaire pour comprendre les interactions entre intelligence artificielle et systèmes sociaux. Il s'agit d'un projet intrinsèquement multidisciplinaire qui cherche à expliquer l'impact concret de l'IA sur la vie culturelle, économique et politique dans le monde, ainsi qu'à recueillir et à évaluer les preuves de l'impact de l'IA sur les droits de l'homme des individus. L'opérationnalisation d'une conception inclusive, d'une acquisition éclairée, d'une mise en œuvre ciblée et d'une responsabilité persistante nécessitera l'attention d'informaticiens et de spécialistes des sciences sociales.



Il s'agira d'un organe essentiel pour de nombreux États dans le monde dépourvus d'une communauté d'ingénieurs experts et de chercheurs sociaux, propre à leur territoire. Un organe de recherche doté d'un personnel adéquat aurait la responsabilité d'évaluer les voies politiques empruntées par les États membres, d'organiser des discussions poussées sur les défis à relever à mesure des évolutions technologiques, d'évaluer les signes d'évolution de l'égalité sociale et de prévoir un forum pour recueillir des idées et avis de l'industrie, de la société civile et de la myriade d'autres groupes d'intérêt public touchés par l'IA dans le service public.

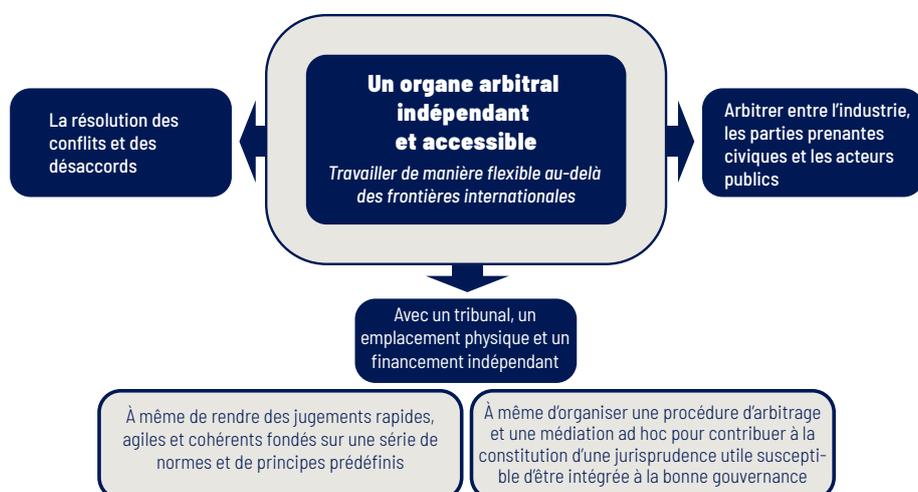
Avant tout, cet organe scientifique devra produire des éléments tangibles à prendre en considération pour la mise en place de systèmes de normes et de certification ainsi que de boîtes à outils pour décideurs politiques. Les travaux d'audit d'algorithmes, d'interprétation des impacts et d'évaluation des bonnes pratiques devront être menés par des chercheurs en informatique, en sciences sociales et en politique pour être crédibles et avoir un impact.

Il existe plusieurs types d'associations universitaires internationales, d'organes de recherche multilatéraux et d'agences intergouvernementales qui disposent déjà de services de recherche qui se penchent sur certaines de ces questions. Néanmoins, la recherche d'un consensus scientifique coordonné autour des preuves de l'impact global de l'IA du service public n'est pas organisée de façon méthodique. Heureusement, le GIEC a prévu un modèle de coordination de l'apprentissage scientifique. Il existe assurément des recherches illustrant la mécanique de fossilisation des principaux domaines d'enquête sociale au fil du temps, mais nous en savons suffisamment sur les incitations à la diversité et à la créativité en enquête sociale pour atténuer ce phénomène. En outre, les résultats d'un tel organe scientifique seraient essentiels pour les recommandations à venir : un consensus scientifique et des preuves solides seront à exploiter par les autres organes et processus que nous recommandons ici.

Un organe arbitral indépendant et accessible

Le deuxième organe, *un organe d'arbitrage mondial*, permettrait de résoudre les conflits et les désaccords, en jouant le rôle d'arbitre entre l'industrie, les parties prenantes civiques et les acteurs publics en cas de problèmes critiques. En tant que commission, nous avons observé que le rythme de l'innovation est d'une rapidité incroyable. Il est difficile d'essayer d'anticiper l'usage des nouveaux outils d'IA dans le secteur public. Dans le cadre de leur part de marché, les acteurs industriels peuvent mettre en place des normes techniques de facto, alors que la législation prendra du temps pour se mettre au diapason. De toute évidence, des désaccords et des applications mal avisées se produiront, et des erreurs techniques et des défauts de conception compliqueront l'utilisation de l'apprentissage automatique dans le secteur public. Officiellement, certains pays disposent de voies légales qui permettraient d'engager toute action civile. Ceci étant, beaucoup n'y ont pas recours, et bien de ces désaccords se posent au niveau des frontières internationales, alors que la grande majorité des tribunaux d'arbitrage nationaux ont une compétence technique limitée pour trancher les litiges. Qui plus est, le rythme de l'innovation s'est avéré prendre de vitesse la législation ayant a priori été pensée pour l'avenir.

Un organe d'arbitrage dédié doté d'un secrétariat aux capacités énormes, d'un emplacement physique, d'un financement indépendant et à même de proposer des services d'arbitrage et de médiation, contribuerait à instituer des précédents commodes susceptibles d'enrichir en partie la bonne gouvernance. La mission du tribunal arbitral serait de rendre des sentences rapides, agiles et cohérentes fondées sur une série de normes et de principes qui seraient, vraisemblablement, suggérés par l'organe scientifique. La rapidité des avancées technologiques fait qu'il est aujourd'hui quasiment impossible d'invoquer des lois nationales au périmètre très vaste et parfois contradictoires. Globalement, un règlement d'arbitrage flexible s'impose au sujet de questions posées en temps réel au fur et à mesure des évolutions technologiques.



Bien évidemment, les spécificités de fonctionnement de l'organe d'arbitrage doivent être négociées, peut-être comme l'une des préoccupations initiales de l'organe scientifique. Le panel d'arbitres doit être diversifié. Il doit disposer d'un siège ou d'un emplacement officiel dans un territoire stable, politiquement neutre et scrupuleusement respectueux de la dimension législative des sentences arbitrales. Il lui faudra un secrétariat pour l'aider à recueillir les preuves et les dépôts de pièces des parties au litige. Il conviendra également de prévoir un mécanisme de financement, dans l'idéal, un cocktail d'aides publiques et de soutien de l'industrie, assorti de systèmes pare-feu appropriés. Il sera crucial que l'organe arbitral soit en mesure de recevoir des observations d'un large éventail de parties prenantes, à l'instar du rôle que jouent les mémoires d'amicus curiae dans certains systèmes judiciaires nationaux. Cela permettra aux parties prenantes étrangères à un désaccord de contribuer au processus d'arbitrage en proposant des informations, une expertise et des connaissances ayant des incidences sur les questions à trancher.

Le tribunal statuera sur les différends ou les désaccords entre et parmi les acteurs des secteurs public, privé et civil. Ces conflits pourront se poser, par exemple, à propos de droits de propriété intellectuelle, de la portée de missions d'ordre public en matière de prestation de services, voire de répercussions découlant de l'utilisation de l'IA par les agences publiques. Au fil de nos discussions, la Commission a conclu que le groupe d'arbitrage devrait être doté d'un pouvoir, par le biais d'un choix (« opt-in ») contractuel ou déclaratoire, des acteurs des secteurs public et privé, qui apprécieraient les avantages de normes transparentes, d'évaluations crédibles, de réponses rapides et de décisions cohérentes à l'échelle mondiale. Il sera essentiel que les deux secteurs acceptent d'être liés par les sentences du tribunal arbitral. Autrement dit, les États devront renoncer aux demandes d'immunité souveraine et convenir d'être liés par la réglementation et les décisions relevant du périmètre des lois nationales. De même, les acteurs privés devraient accepter de se soumettre aux sentences de l'organe arbitral, ni demander toute autre réparation ou décision différente en formant un recours devant d'autres organes moins compétents ou pertinents.

L'avantage d'un tribunal arbitral est qu'il rend activement des sentences qui font office de jurisprudence utile, souvent de manière flexible en réponse à des questions d'actualité, et qu'il peut faire avancer toute situation plus rapidement que des procédures qui progressent au rythme d'une législation dense. D'autant plus qu'un tribunal arbitral dédié à l'utilisation de l'IA dans le secteur public apporterait une profondeur d'engagement et une capacité technique particulières dans le règlement des désaccords.

Travailler ensemble

L'IA est devenue prolifique, mais son utilisation à des fins de bonne gouvernance manque de mécanismes d'inclusion, de collaboration et de participation entre le gouvernement, l'industrie, les parties prenantes civiles et les chercheurs. Une collaboration plus transnationale entre les organes scientifiques et d'arbitrage pourrait favoriser la confiance à l'égard de la réglementation de l'IA, accroître le consensus et, en fin de compte, renforcer le soutien

civique vis-à-vis du déploiement de l'IA et de son utilisation en matière de gouvernance. Exiger « plus de collaboration » n'est guère suffisant. La collaboration doit se révéler utile, en engageant les parties prenantes internationales de manière indépendante et en privilégiant le dialogue et le consensus, et ce, de manière inclusive et ciblée.

Les initiatives et objectifs politiques communs au niveau européen ou transatlantique, la collaboration avec l'industrie, les conférences scientifiques et le financement de projets de recherche et de plaidoyer axés sur l'IA fixent un point de départ pour faciliter ce type de mise en réseau et d'échange de connaissances. Néanmoins, les affrontements géopolitiques et les pressions concurrentielles ont entravé la collaboration sans lien de dépendance entre les divers groupes de parties prenantes et les frontières nationales. Les systèmes d'arbitrage indépendants et les organes de concertation scientifique peuvent contribuer à faire progresser la collaboration (ou à tout le moins le dialogue) entre le gouvernement, l'industrie et la société civile, sur les plans national et international.

Somme toute, cet ensemble d'organisations, s'il est doté de ressources suffisantes, pourrait prodiguer des conseils exhaustifs au fil de l'apparition d'innovations au cours des années à venir. Ces organisations formeraient un duo de choc, l'une se consacrant à la production de preuves grâce aux dernières méthodes scientifiques et l'autre à l'évaluation des conséquences sociales en s'appuyant sur ces preuves. Il existe de bons précédents dans le système international portant sur ce genre d'organisation. Les sentences arbitrales sont exécutées par le biais de la Convention de New York sur l'arbitrage de 1958 (Convention pour la reconnaissance et l'exécution des sentences arbitrales étrangères), dont 168 pays sont signataires. Le tribunal arbitral du sport jouit d'une solide réputation internationale pour la rapidité de ses décisions et ses prouesses techniques, et a également acquis une grande crédibilité pour ses travaux. Le Conseil international pour l'autorégulation de la publicité et la Chambre de commerce international sont deux organisations bien établies qui fixent des normes, proposent des procédures d'arbitrage contraignantes en droit privé et ont un statut consultatif auprès des Nations Unies. Il conviendrait certainement d'approfondir les recherches pour établir le cadre juridique de collaboration entre l'industrie, le gouvernement et la société civile, probablement par le biais du droit privé. Cela étant, les modèles de mise en œuvre d'une telle collaboration se veulent encourageants.

RECOMMANDATION N° 1

Les États, les industries et la société civile doivent collaborer à l'échelle internationale pour mettre en place et habiller a) un organe scientifique international afin de faire avancer la recherche sur les applications à IA du service public et b) un organe d'arbitrage à même de trancher les litiges susceptibles de survenir entre les parties prenantes impliquées dans les systèmes d'IA du service public.



COMMENT RENFORCER LES CAPACITÉS DU SERVICE PUBLIC EN MATIÈRE D'IA POUR FAVORISER LA BONNE GOUVERNANCE ?

L'IA, l'apprentissage automatique et d'autres systèmes algorithmiques de pointe ont évolué extrêmement rapidement, et il faut s'attendre à ce qu'ils poursuivent dans cette voie. L'IA englobe pratiquement tous les domaines de la vie publique, elle est donc apparue comme un défi multidisciplinaire en matière de réglementation. Nous avons proposé que des organes soient chargés d'élaborer des principes et des normes et d'en contrôler le respect. Toutefois, les questions relatives à l'utilisation quotidienne de l'IA dans le service public sont souvent beaucoup plus banales et de nature pratique. Quelles ressources faut-il fournir aux fonctionnaires afin de réguler efficacement l'IA dans le secteur public ? Nous proposons ici une conception au sens large de la réglementation, prévoyant l'élaboration de règles administratives dans le cadre du mandat de principes ou de normes, tels que ceux élaborés par l'organe scientifique proposé.

Tout d'abord, nous recommandons d'appliquer les principes évoqués dans la section 3, et nous recommandons une réglementation à propos de la *conception* des systèmes d'IA ; des processus d'*acquisition* par lesquels les agences publiques acquièrent ou accordent des licences pour les systèmes d'apprentissage automatique ; des processus de *déploiement*, y compris les consultations publiques, les campagnes de sensibilisation et l'accès aux informations ; et des processus d'*obligations redditionnelles* à long terme pour rassembler les observations du public sur les conséquences à long terme de l'apprentissage automatique appliqué aux problèmes d'action collective.

De nombreux pays ont déjà adopté avec succès l'IA dans le service public, avec ou sans réglementation concernant la conception, l'acquisition, le déploiement et la reddition de comptes. Néanmoins, nos recherches soulignent que les fonctionnaires manquent encore largement des capacités nécessaires à l'acquisition, à la conception, à l'évaluation et au déploiement de l'IA. L'étape suivante consiste à offrir aux fonctionnaires une expertise fondamentale afin de leur permettre de prendre de bonnes décisions en matière d'IA. Il est certain que la sensibilisation des fonctionnaires et la mise à niveau des agences de réglementation sont au cœur de ce défi, mais ces processus sont gourmands en ressources et chronophages. Par ailleurs, des boîtes à outils pratiques préconisant des bonnes pratiques et des matrices de décision simples peuvent résoudre certains des défis les plus urgents. Il est vital que les États élaborent une base de connaissances centrale et accessible, afin que les stratégies et l'expertise avérées puissent circuler d'un département à l'autre.

Pour ce faire, il est capital de doter les régulateurs des compétences indispensables et de leur conférer la mission nécessaire pour mettre au point les instruments consultatifs et les

mécanismes de conformité dont ils ont besoin. Sans nul doute, certains aspects de réglementation concernant la conception, l'acquisition, le déploiement et la reddition de comptes relatifs à l'IA pourraient être externalisés. Les États n'ont pas à se charger de tout. Il est probable que les agences publiques doivent diriger les trois premières disciplines, les travaux de reddition de comptes et de certification étant effectués par des tiers selon un certain nombre de directives publiques.

Enfin, s'il est utile d'identifier ce qu'il convient de réglementer, il est primordial de souligner en l'occurrence que le processus de réglementation (et tout programme réglementaire qui l'accompagne) doit disposer de ressources et de capacités suffisantes pour apporter un soutien aux équipes de personnel intergouvernemental pouvant se spécialiser dans les questions liées à l'IA dans le secteur public. Il sera vital d'entretenir dûment le niveau de formation des administrateurs des programmes publics, mais il devrait également être possible d'élargir le champ de compétences technologiques à l'ensemble des agences publiques de manière à doter les États plus largement de capacités d'interaction avec d'autres parties prenantes de l'industrie et de la société civile.

RECOMMANDATION N° 2

Les États d'envergure nationale, profitant du soutien de ces deux nouveaux organes internationaux, doivent a) renforcer les capacités de leur service public respectif en faveur d'engagements profonds en conception, acquisition, déploiement et en obligations redditionnelles vis-à-vis des systèmes d'IA du service public et b) proposer des boîtes à outils au personnel des agences publiques pour qu'il puisse accomplir sa mission de surveillance.



COMMENT GARANTIR LA FIABILITÉ ET LA CRÉDIBILITÉ DE L'IA DU SERVICE PUBLIC ?

Un manque de confiance pourrait devenir un obstacle crucial au déploiement réussi et opportun de l'IA dans le service public : si le public ne fait pas confiance à l'IA, les prises de décisions politiques basées sur les systèmes d'IA se heurteront forcément à une opposition véhémente. Il est indubitable que la confiance du public envers l'IA n'est envisageable que lorsque les systèmes techniques sont effectivement dignes de confiance. Par conséquent, l'OxCAIGG soutient que le renforcement du soutien du public tout comme les évaluations de l'IA sont essentiels au déploiement de l'IA à des fins de bonne gouvernance. Nous recommandons aux États et aux agences publiques de répondre aux inquiétudes des citoyens concernant l'utilisation de l'IA par des mesures visant la sensibilisation, les compétences et la certification.

Tout d'abord, pour renforcer la confiance du public vis-à-vis de l'IA du service public, nous recommandons aux États de lancer des campagnes publiques de communication et d'apprentissage qui mettent en évidence les usages bienveillants de l'IA pour des questions d'intérêt public dans des domaines que leurs citoyens connaissent déjà. Les campagnes de communication doivent également prendre en compte les points de préoccupation légitimes concernant les défauts systémiques et l'utilisation à mauvais escient des systèmes d'IA par les États. Nous recommandons aux États d'allouer des budgets de communication substantiels à des campagnes et programmes publicitaires opportuns et intermédiaires, visant à atteindre divers groupes démographiques. Étant donné que les expériences négatives et les fausses informations sur l'IA sont devenues légion, la mise en lumière du recours à l'IA dans des domaines tels que les diagnostics précoces, la gestion du trafic et les technologies climatiques pourrait favoriser l'engagement des citoyens par rapport aux technologies à base d'IA.

Ensuite, pour orienter les travaux futurs dans ce domaine à plus long terme, les gouvernements devraient s'efforcer de mieux communiquer sur les circonstances et modalités d'exploitation courante ou prévue de l'IA. Dévoiler les lieux et modalités d'utilisation de l'IA dans le service public permettrait d'accroître la transparence et, en fin de compte, la confiance vis-à-vis de l'IA. Les notes d'information sur les pages Web des États et les bases de données centrales, ainsi que les listes accessibles au public des fournisseurs d'IA et des systèmes d'IA mis au point par les États, pourraient servir de point de départ à des systèmes de communication exhaustifs. Il y a un avantage à créer une base de données collective et internationale, puisqu'elle pourrait être proposée par l'organe scientifique que nous avons recommandé.

Enfin, nous proposons un système de certification mondial pour le recours à l'IA dans le service public. Cette certification de l'IA devrait prendre en compte des indicateurs de sécurité et de qualité, et tenir un registre public sur l'utilisation par tout État des systèmes d'IA qui aurait été jugée insuffisante. Le débat intellectuel sur l'opérationnalisation de l'IA bienveillante et digne de confiance est bien avancé, et les chercheurs ont proposé des cadres éthiques normatifs et des indicateurs mesurables afin d'évaluer la transparence, l'interprétabilité et la reddition de comptes des systèmes d'IA. Par conséquent, les évaluations de la qualité, des risques et de l'impact, les tests rigoureux des processus de conception, l'évaluation des données de formation et la maintenance des systèmes ont été plus largement adoptés dans l'industrie. Néanmoins il y a encore lieu de fixer de bonnes pratiques ou normes exhaustives. Il y a de bonnes raisons de ne pas redouter que la certification décourage les efforts d'innovation. Dans le monde entier, on relève des pratiques d'audit et de certification efficaces pour les logiciels de transactions financières, les machines de jeux d'argent vidéo et le traitement des données. La conception de systèmes de certification complets et de boîtes à outils pour les politiques pourrait être l'une des premières missions de l'organe scientifique proposé, bien qu'une agence internationale de normalisation ou une agence basée sur l'industrie puisse finalement être mieux placée pour assurer la certification.

RECOMMANDATION N° 3

La confiance accordée au recours à l'IA doit être renforcée au moyen de campagnes de sensibilisation du public sur des applications quotidiennes et de véritables cas pratiques qui se profilent à l'horizon, les impacts et risques connexes, a) en demandant aux États de communiquer en toute franchise sur l'usage prévu des technologies d'IA dans le service public et b) en mettant en place une agence multisectorielle qui établisse un système de certification de base permettant d'approuver continuellement les applications et d'instiller un sentiment de confiance dans l'esprit du public au fil du temps.

5 Conclusion : des idées pour un plan d'action immédiat

L'identification de grands principes est cruciale pour instaurer un cadre permettant de comprendre un problème d'ordre public. Cela nous a permis d'évaluer les défis et d'identifier des recommandations concrètes pour un plan d'action.

Le fait d'avoir élaboré les trois recommandations de ce rapport nous permet à présent d'indiquer quelles devraient être les prochaines étapes. Compte tenu de ce que nous savons des défis liés au déploiement de l'IA dans le service public, et de la série de recommandations dont nous bénéficions pour soutenir la bonne gouvernance grâce aux systèmes d'IA, quelles devraient être les prochaines étapes ?

I. UNE ÉTUDE DE FAISABILITÉ ORGANISATIONNELLE



Une étude de faisabilité nous permettrait d'estimer les coûts de mise en place de l'organe d'arbitrage et de l'organe scientifique nécessaires. De toute évidence, ces agences devront être « sises » dans des territoires légaux et des climats politiques propices à leur mission respective. Qui dirigerait de telles organisations, et quelles capacités organisationnelles spécifiques seront nécessaires pour mener à bien leurs tâches essentielles ? La première étape à franchir consiste à élaborer une étude de faisabilité permettant de remplir les scénarios logistiques et opérationnels afin de consolider nos capacités mondiales à intégrer l'IA au service du public.

II. CONSULTATION AUPRÈS D'AGENCES NATIONALES ET MULTILATÉRALES EXISTANTES



Il existe un certain nombre d'institutions multilatérales aux capacités énormes qui contribuent considérablement aux discussions mondiales sur l'IA dans le service public. Toutefois, aucune de ces agences n'a un champ d'action étendu ou ne considère ce travail comme sa mission centrale et dédiée. En outre, plusieurs aspects du travail à accomplir se situent entre les champs d'application des missions des agences multilatérales existantes. Les experts de l'UNESCO ont une vue d'ensemble des nombreuses tendances importantes dans les applications de

l'IA, et le Programme des Nations Unies pour le développement ainsi que le G7 et le G20 ont commencé à discuter du rôle de l'IA dans la gouvernance. Le groupe de travail inter-agences de l'ONU portant sur l'IA se réunit régulièrement. Il est dirigé par l'UNESCO et l'Union internationale des télécommunications.

Nous soutenons que les responsables politiques et les décideurs mondiaux sont ceux qui doivent ouvrir la voie à une collaboration plus prolifique. Le Partenariat mondial sur l'intelligence artificielle (GPAI) du G7 représente un effort louable en ce sens, par le regroupement d'experts en sciences, en technologies, de la société civile et en politique de manière multipartite. Toutefois, le GPAI est un groupe relativement exclusif et certains des plus importants innovateurs en matière d'IA, ainsi que le fameux Sud global, n'y sont pas représentés. La Chine, leader mondial en matière d'IA, est notamment restée éloignée du discours et des préoccupations au sens large concernant les violations des droits de l'homme liées à l'utilisation de l'IA. Sur la base de nos recherches et des discussions d'experts tenues à l'OxCAIGG, nous soulignons l'importance d'une collaboration et d'échanges d'envergure plus internationale.

III. ENGAGER DES DISCUSSIONS INCLUSIVES LORS D'ÉVÉNEMENTS INTERNATIONAUX CHARNIÈRES



Un certain nombre d'opportunités importantes se profilent à l'horizon : des événements durant lesquels des discussions inclusives peuvent être tenues sur les perspectives réalistes d'amélioration de notre capacité à garantir l'usage de l'IA à des fins de bonne gouvernance. Par exemple, en septembre 2023, l'ONU organisera le Sommet du futur, qui réunira des décideurs politiques mondiaux, en l'occurrence des chefs d'État. Cet événement peut servir de terrain de dialogue sur un territoire neutre et de point de départ pour un échange plus soutenu, y compris entre les groupes de parties prenantes qui partagent un faible consensus, le cas échéant. Un événement mondial de cette envergure peut se révéler idéal pour servir de rampe de lancement à une collaboration à long terme plus structurée.



Récapitulatif des objectifs

RECOMMANDATIONS

1

À l'échelle internationale, les gouvernements, l'industrie et la société civile doivent collaborer pour mettre en place et habilitier

- a) un organe scientifique international afin de faire avancer la recherche sur les applications de l'IA du service public et
- b) un organe arbitral afin de trancher les litiges qui pourraient survenir entre les parties prenantes impliquées dans les systèmes d'IA du service public.

2

Les gouvernements nationaux, profitant du soutien de ces deux nouvelles organisations internationales doivent

- a) renforcer la capacité de leur service public à s'impliquer profondément dans la conception, l'acquisition, le déploiement et les obligations redditionnelles des systèmes d'IA du service public et
- b) mettre à disposition des boîtes à outils pour le personnel des agences publiques pour qu'il puisse effectuer son travail d'encadrement.

3

La confiance accordée au recours à l'IA doit être renforcée au moyen de campagnes de sensibilisation du public sur

- des applications quotidiennes et de véritables cas pratiques qui se profilent à l'horizon, les impacts et risques connexes ; a) en demandant aux États de communiquer en toute franchise sur l'usage prévu des technologies d'IA dans le service public et
- b) en mettant en place une agence multisectorielle qui établisse un système de certification de base permettant d'approuver continuellement les applications et d'instiller un sentiment de confiance dans l'esprit du public au fil du temps.

PROCHAINES ÉTAPES POUR UN PLAN D'ACTION IMMÉDIAT

I

Une étude de faisabilité organisationnelle

II

Consulter les agences nationales et multilatérales existantes

III

Engager des discussions inclusives lors d'événements internationaux charnières



6. Recherche de l'OxCAIGG

- [1] Neudert, L.-M., Knuutila, A. & Howard, P. N. *Global Attitudes towards AI, Machine Learning & Automated Decision Making*. (Document de travail 2020.10, Commission d'Oxford sur l'IA et la bonne gouvernance, 2020).
- [2] Ramizo, G., Jr. *Practical Lessons for Government AI Projects*. (Document de travail 2021.1, Commission d'Oxford sur l'IA et la bonne gouvernance, 2021).
- [3] Vogl, T. *Artificial Intelligence in Local Government*. (Document de travail 2021.2, Commission d'Oxford sur l'IA et la bonne gouvernance, 2021).
- [4] Nyabola, N. *Old Cracks, New Tech: Artificial Intelligence, Human Rights and Good Governance in Highly Fragmented and Socially Stratified Societies*. Le cas du Kenya. (Document de travail 2021.3, Commission d'Oxford sur l'IA et la bonne gouvernance, 2021).
- [5] Au, Y. *Surveillance as a Service : The European AI-Assisted Mass Surveillance Marketplace*. (Document de travail 2021.4, Commission d'Oxford sur l'IA et la bonne gouvernance, 2021).
- [6] McFadden, M., Jones, K., Taylor, E. & Osborn, G. *Harmonising Artificial Intelligence : The Role of Standards in the EU AI Regulation*. (Document de travail 2021.5, Commission d'Oxford sur l'IA et la bonne gouvernance, 2021).
- [7] Neudert, L.-M. & Howard, P. N. *Four Principles for Integrating AI and Good Governance*. (Document de travail 2020.1, Commission d'Oxford sur l'IA et la bonne gouvernance, 2020).

À propos de la Commission

Le défi du recours à l'IA à des fins de bonne gouvernance et le service public est une préoccupation urgente dans les pays du monde entier. Lancé en juillet 2020, l'objectif de la Commission d'Oxford sur l'IA et la bonne gouvernance a été d'élaborer des principes et des recommandations politiques pratiques afin d'assurer l'exploitation démocratique de l'IA à des fins de bonne gouvernance.

Plus récemment, la pandémie de COVID-19 a provoqué un afflux rapide de solutions à IA. Bien qu'elles soient destinées à l'intérêt public, ces nouvelles technologies s'accompagnent de défis d'évaluation de la pertinence et de la légitimité de ces offres. La rapidité du déploiement de ces systèmes est sans précédent et démontre la nécessité de politiques à propos de ces types de produits d'IA, de leur acquisition et de leur déploiement par les États.

L'OxCAIGG a étudié les défis en termes d'acquisition et de déploiement concernant le recours à l'IA à des fins de bonne gouvernance auxquels sont confrontées les démocraties du monde

entier. Elle a identifié les bonnes pratiques d'évaluation et de gestion des risques et des avantages, et a recommandé des stratégies dans le but de tirer pleinement parti des capacités technologiques, tout en atténuant les inconvénients potentiels des politiques publiques basées sur l'IA.

En s'appuyant sur les contributions d'experts d'un large éventail de régions géographiques et de domaines d'expertise, y compris des parties prenantes du gouvernement, de l'industrie et de la société technique et de la société civile, l'OxCAIGG a émis des recommandations appropriées et pertinentes pour l'utilisation de l'IA en vue d'une bonne gouvernance.

Nos commissaires ont mis à profit leur expérience et leurs connaissances afin de contribuer de manière réfléchie à la quête de l'OxCAIGG visant à armer et à orienter les décideurs politiques en leur prodiguant des conseils pour s'assurer que les outils liés à l'IA sont adaptés et adoptés pour une bonne gouvernance dans un avenir proche.



Les commissaires



Dr Yuichiro Anzai, commissaire de l'OxCAIGG

Conseiller principal pour la Société japonaise pour la promotion des sciences, et conseiller exécutif aux Affaires académiques à l'université Keio.

Yuichiro Anzai est chargé de conseiller, en tant que président du Conseil pour la stratégie de l'intelligence artificielle, le gouvernement japonais sur l'élaboration de politiques stratégiques. Il s'est vu décerner le titre de Personne de mérite culturel par l'État japonais pour ses travaux précurseurs sur l'intégration des sciences cognitives et de l'information.



Professeur Dame Wendy Hall (DBE, FRS, FEng), commissaire de l'OxCAIGG

Professeur Regius d'informatique, vice-présidente associée (engagement international) et directrice administrative du Web Science Institute de l'université de Southampton.

Wendy Hall est présidente de l'Institut Ada Lovelace et membre du Conseil consultatif technologique de BT. Elle a été nommée Dame Commandeur de l'ordre de l'Empire britannique dans la liste des distinctions honorifiques du Nouvel An 2009 du Royaume-Uni, et est membre de la Royal Society.



Dr Rumman Chowdhury, commissaire de l'OxCAIGG

Directrice de l'équipe META (ML Ethics, Transparency, and Accountability ou Éthique, transparence et responsabilité de l'apprentissage automatique) chez Twitter.

La passion qui anime Rumman Chowdhury réside à la jonction entre intelligence artificielle et humanité. Elle est une pionnière dans le domaine de l'éthique algorithmique appliquée, ayant élaboré des solutions socio-techniques d'avant-garde pour une IA éthique, interprétable et transparente.



Professeur Philip Howard, commissaire de l'OxCAIGG

Professeur statutaire d'études Internet, Balliol College, université d'Oxford.

Philip Howard étudie l'impact des médias numériques sur la vie politique dans le monde entier, et s'exprime fréquemment sur les médias mondiaux et les affaires politiques. Il est directeur du programme de l'université d'Oxford sur la démocratie et la technologie, qui étudie l'exploitation des algorithmes et de l'automatisation dans la vie citoyenne.



M. Tom Fletcher (CMG), commissaire de l'OxCAIGG

Principal du Hertford College, université d'Oxford, et fondateur de la Foundation for Opportunity.

Tom Fletcher a été auparavant conseiller en politique étrangère de trois premiers ministres britanniques, ambassadeur du Royaume-Uni au Liban, professeur invité à l'université de New York, conseiller de la Coalition mondiale des entreprises pour l'éducation et président du conseil international de la Fédération des industries créatives. Il est l'auteur de *The Naked Diplomat*.



Sir Julian King, commissaire de l'OxCAIGG

Ancien commissaire européen et diplomate britannique.

Julian King a précédemment occupé les fonctions d'ambassadeur du Royaume-Uni en Irlande et en France et de directeur général du Bureau pour l'Irlande du Nord (Northern Ireland Office). Il a été le dernier fonctionnaire britannique à occuper un poste et un portefeuille au sein de la Commission européenne, en tant que commissaire chargé de l'union de la sécurité.



Dr Safiya Noble, commissaire de l'OxCAIGG

Maîtresse de conférence à l'université de Californie, Los Angeles (UCLA) dans le département d'études afro-américaines et le département d'études de l'information, et co-directrice de l'UCLA Center for Critical Internet Inquiry (C2i2).

Le travail de Safiya Noble est à la fois sociologique et interdisciplinaire, soulignant l'impact et les croisements des médias numériques ont un impact et se croisent sur les questions de race, de genre, de culture et de technologie. Elle est membre du conseil d'administration de la Cyber Civil Rights Initiative, au service des personnes vulnérables au harcèlement en ligne.



M. Howard Rosen (CBE), commissaire de l'OxCAIGG

Directeur de Howard Rosen Solicitors, président des Amis suisses de l'université d'Oxford, président du Rail Working Group et administrateur du Conseil juif du Commonwealth.

Howard Rosen est spécialisé en droit des leasing, commerce international, finance, et en trusts. Il est également le fondateur et le directeur général de Rosetrust AG, une société fiduciaire sise à Zug, en Suisse, et d'Aviation Advocacy Sàrl, basée à Nyon, en Suisse.



Professeur Weixing Shen, commissaire de l'OxCAIGG

Doyen de la faculté de droit de l'université Tsinghua à Pékin, directeur de l'Institut d'études sur l'IA et le droit, et directeur de la faculté du programme LLM sur le droit informatique de l'université Tsinghua.

Weixing Shen est commissaire de l'Alliance de l'industrie de l'intelligence artificielle en Chine. Il est également membre du Conseil exécutif du Barreau de Chine et vice-président de l'Association du droit du cyber espace et de l'information du Barreau de Chine.



Baronne Joanna Shields (OBE), commissaire de l'OxCAIGG

Directrice générale, BenevolentAI

Joanna Shields est une vétérante de l'industrie technologique, créatrice d'entreprises des plus connues au monde. Les technologies bénéfiques à l'humanité sont sa grande passion. Fort d'une carrière de 30 ans, elle a essentiellement œuvré à l'exploitation de la puissance des technologies pour susciter des changements propices à l'amélioration de la connectivité, de l'humanité et de la société civile.

Remerciements

Nous adressons nos remerciements à Flora Seddon pour avoir coordonné et soutenu nos activités de l'OxCAIGG au cours des 18 derniers mois. Pour le soutien de nos consultations numériques et de nos discussions stratégiques, nous adressons nos remerciements à Bruno Selun et à l'équipe de Kumquat. Pour leurs contributions, nous remercions Hubert Au, Rutendo Chabikwa, Dr Tim Curnow, Mona Elswah, John Gilbert, Mark Healy, Dr Lucy Hennings, Mark Malbas, Dr Nahema Marchal, Sara Spinks et Niamh Walsh. Nous remercions également les experts, les chercheurs et les spécialistes du service public qui ont contribué à l'élaboration de ce rapport et aux travaux de l'OxCAIGG.

Pour avoir soutenu l'OxCAIGG, nous sommes reconnaissants envers la Fondation Adessium, Civitates, la Fondation Ford, Luminate et les Open Society Foundations. Tous les avis, constatations et conclusions ou recommandations exprimés dans le présent document sont ceux de la Commission et ne correspondent pas nécessairement aux avis de l'université d'Oxford, de nos bailleurs de fonds ou de tout commissaire. La surveillance éthique de la recherche à l'université d'Oxford est assurée par le Comité central d'éthique de la recherche de l'université. Les numéros d'approbation des résultats de recherche individuels sont indiqués dans chaque document de travail ou rapport, selon le cas.





oxcaigg.oii.ox.ac.uk

Commission d'Oxford sur l'IA et la bonne gouvernance.
L'IA dans le service public : Des principes à la pratique.
Document de travail 2021.6 Oxford, Royaume-Uni :
Commission d'Oxford sur l'IA et la bonne gouvernance. 19 pp.
Récupéré depuis : <https://oxcaigg.oii.ox.ac.uk>