

Chronologie des principales actions des autorités publiques françaises sur l'intelligence artificielle

Contexte // Cette chronologie a pour objectif de présenter, de manière non exhaustive, les principaux rapports, avis, stratégies et initiatives réalisés par les autorités publiques françaises sur l'intelligence artificielle.

Le rapport Stratégie France I.A., pour le développement des technologies d'intelligence artificielle a été présenté le 21 mars 2017 à la Cité des Sciences et de l'Industrie, en présence du Président de la République, du ministre de l'Économie et des Finances, du secrétaire d'État chargé de l'Enseignement supérieur et de la Recherche, et du secrétaire d'État chargé de l'Industrie, du Numérique et de l'Innovation ([lire le rapport](#)).

Du 20 janvier au 14 mars 2017, 17 groupes de travail se sont réunis à l'initiative du Gouvernement, pour proposer des actions de politique publique visant à développer l'intelligence artificielle en France. Une cinquantaine de propositions ont été remontées par les différents groupes de travail, ainsi que lors des Journées de l'Intelligence Artificielle organisées dans l'ensemble du pays, autour des thèmes suivants :

- En matière de recherche, affirmer l'excellence de la recherche française en préservant sa diversité, en développant l'interdisciplinarité, et en valorisant les talents
- Sur la formation, améliorer l'environnement et l'appropriation des questions d'IA dans la société, pour permettre la compréhension des enjeux de l'IA et créer des vocations, susciter l'envie d'innover, et pour former à des métiers transformés en continu.
- Permettre un transfert efficace de la recherche vers les applications industrielles et économiques
- Développer une stratégie industrielle fondée sur les écosystèmes d'innovation et une "verticalisation" de l'IA. dans chaque filière
- Poursuivre le débat public sur l'IA. et développer des outils en facilitant la compréhension.

À la suite des travaux menés par les groupes de travail, plusieurs actions ont été menées tout au long de l'année 2017 :

- Mise en place du comité stratégique #FranceIA rassemblant les sphères académique, scientifique, économique et la société civile, chargé de mettre en œuvre les recommandations des groupes de travail
- Coordination par la France d'une candidature à un "projet phare de technologie émergente" ("FET flagship") sur l'IA., co-financé par l'Union européenne (1 milliard d'euros)
- Lancement d'un nouveau programme mobilisant les institutions de recherche pour identifier, attirer et retenir les meilleurs talents en IA, dans le cadre de l'action Programmes prioritaires de recherche du P.I.A.3
- Financement d'une infrastructure mutualisée pour la recherche
- Constitution d'un consortium public-privé en vue de l'identification ou de la création d'un centre interdisciplinaire pour l'intelligence artificielle
- Inclusion systématique, d'ici fin 2017, de l'IA. dans les priorités de l'ensemble des dispositifs publics de soutien à l'innovation
- Mobilisation des ressources publiques (Bpifrance, P.I.A.) et privées pour atteindre l'objectif d'ici 5 ans d'investir dans 10 start-ups françaises pour plus de 25 millions d'euros chacune
- Mobilisation des filières automobile, relation client, finances, santé et transport ferroviaire pour que chaque filière définisse une stratégie sectorielle I.A. d'ici fin 2017
- Lancement d'un appel à projets pour des plateformes sectorielles de partage de données pour 3 à 6 secteurs, d'ici la fin 2017
- Conclusion du débat éthique sur les "algorithmes" animé par la CNIL

- Lancement d'ici l'été d'une concertation de France Stratégie sur la question
- des effets de l'intelligence artificielle sur l'emploi

Décembre 2017

Résumé du rapport de l'OPECST sur l'intelligence artificielle

L'Office parlementaire d'évaluation des choix scientifiques et technologiques (OPECST) a adopté en décembre 2017 et à l'unanimité, sur la saisine de la commission des Affaires économiques du Sénat, un rapport intitulé *"Pour une intelligence artificielle maîtrisée, utile et démystifiée"* dont les rapporteurs sont M. Claude de GANAY, député (Groupe Les Républicains, Loiret) et Mme Dominique GILLOT, sénatrice (Groupe socialiste et républicain, Val d'Oise).

Le concept d'intelligence artificielle renvoie à des technologies multiples, nées dans la seconde moitié du XX^e siècle, qui reposent sur l'utilisation d'algorithmes. Leur essor récent reste soumis à une contrainte d'acceptabilité sociale assez forte sous l'effet de représentations souvent catastrophistes. Ces technologies, dont les combinaisons sont en évolution constante, représentent un bouleversement de nature à transformer profondément nos sociétés et nos économies. Les applications sectorielles présentes ou futures sont considérables et les évolutions peuvent être rapides, que l'on pense par exemple à l'éducation, à l'environnement, à l'énergie, aux transports, à l'aéronautique, à l'agriculture, au commerce, à la finance, à la défense, à la sécurité, à la sécurité informatique, à la communication, aux loisirs, à la santé, à la dépendance ou, encore, au handicap. Le rapport *"Pour une intelligence artificielle maîtrisée, utile et démystifiée"* fournit un **état de la recherche** en intelligence artificielle et fait le point sur de nombreux autres rapports parus récemment sur le sujet en France et dans le monde. Il présente aussi les **enjeux éthiques, juridiques, économiques, sociaux et scientifiques** de ces technologies, parmi lesquels la place prépondérante de la recherche privée, dominée par les entreprises américaines et, potentiellement, chinoises, l'accélération du passage à une économie globalisée dominée par des "plateformes", les transformations du marché du travail, les régimes de responsabilité, les biais et les problèmes posés par les données et les algorithmes, le phénomène de "boîtes noires" des algorithmes et la question des "bulles d'information". Il évoque, par ailleurs, certains sujets d'interrogation liés à la "singularité", à la "convergence NBIC" et au "transhumanisme" ainsi que la nécessité d'une prise en compte grandissante de **règles éthiques**.

Les rapporteurs invitent à conduire la réflexion sur l'intelligence artificielle de manière sereine et rationnelle, afin de mettre en avant ses **opportunités** tout autant que ses **risques**, de rassurer le public et de démystifier les représentations biaisées de l'intelligence artificielle. Les progrès en intelligence artificielle posent des questions auxquelles toute la société doit être sensibilisée. Ces progrès sont d'abord et avant tout bénéfiques. Ils comportent aussi des risques, qui peuvent et doivent être identifiés, anticipés et maîtrisés. L'avènement d'une super-intelligence ne fait pas partie de ces risques à court et moyen terme. À long terme, la réalité de cette menace n'est pas certaine. Quant à son imminence à court ou moyen terme, elle relève du fantasme. Ni quête vaine ni projet de remplacement de l'homme par la machine, **l'intelligence artificielle représente une chance à saisir** pour nos sociétés et nos économies. La France doit relever ce défi. Plutôt qu'une hypothétique confrontation dans le futur entre les hommes et les machines, qui relève d'une forme de science-fiction catastrophiste, les rapporteurs sont convaincus du bel avenir de la complémentarité homme-machine. Nous allons bien plus vers une intelligence humaine augmentée que vers une intelligence artificielle concurrençant l'homme. **Le rapport, à travers ses 15 propositions, se prononce pour une intelligence artificielle maîtrisée, utile et démystifiée** : maîtrisée, parce que ces technologies devront être les plus sûres, les plus transparentes et les plus justes possibles ; utile parce qu'elles doivent, dans le respect des valeurs humanistes, profiter à tous au terme d'un large débat public ; démystifiée, enfin, parce que les difficultés d'acceptabilité sociale de

l'intelligence artificielle résultent largement de visions catastrophistes sans fondement.

Décembre 2017

Résumé du rapport de la CNIL sur les enjeux éthiques des algorithmes et de l'IA

En décembre 2017, la CNIL a publié le rapport de synthèse du débat public qu'elle a animé sur les enjeux éthiques des algorithmes et de l'intelligence artificielle. Isabelle FALQUE-PIERROTIN, présidente de la CNIL, le présente aujourd'hui en présence de Mounir MAHJOUBI, Secrétaire d'Etat chargé du Numérique et de Cédric VILLANI, Député, chargé par le gouvernement d'une mission sur l'intelligence artificielle ([lire le rapport](#)).

Vers l'affirmation de deux principes fondateurs : loyauté et vigilance

Les débats ont permis de dégager deux principes fondateurs pour une intelligence artificielle au service de l'homme. Ces principes pourraient s'inscrire dans une nouvelle génération de garanties et de droits fondamentaux à l'ère numérique, des « droits-système » organisant la gouvernance mondiale de notre univers numérique :

- Un principe de *loyauté* appliqué à tous les algorithmes et intégrant les impacts collectifs, et pas seulement personnels, de ces derniers. Tout algorithme, qu'il traite ou non des données personnelles, doit être loyal envers ses utilisateurs, non pas seulement en tant que consommateurs, mais également en tant que citoyens, voire envers des communautés ou de grands intérêts collectifs dont l'existence pourrait être directement affectée. L'intérêt des utilisateurs doit primer. Par exemple, un tel principe pourrait avoir vocation à s'appliquer à l'impact potentiel des réseaux sociaux sur la structure du débat public dans nos démocraties (segmentation du corps politique par le ciblage de l'information) ou à celui d'algorithmes de police prédictive sur des communautés ou quartiers entiers.
- Un principe de *vigilance/réflexivité* : il s'agit d'organiser une forme de questionnement régulier, méthodique et délibératif à l'égard de ces objets mouvants. Ce principe constitue une réponse directe aux exigences qu'imposent ces objets technologiques du fait de leur nature imprévisible (inhérente au *machine learning*), du caractère très compartimenté des chaînes algorithmiques au sein desquels ils s'insèrent et, enfin, de la confiance excessive à laquelle ils donnent souvent lieu. C'est l'ensemble des maillons de la chaîne algorithmique (concepteurs, entreprises, citoyens) qui doivent être mobilisés pour donner corps à ce principe, au moyen de procédures concrètes (par exemple, des comités d'éthique assurant un dialogue systématique et continu entre les différentes parties-prenantes).

Ces principes fondateurs sont complétés par des principes organisationnels ayant traits à l'intelligibilité et à la responsabilité des systèmes algorithmiques ainsi qu'à la nature de l'intervention humaine dans la prise de décision algorithmique.

Les 6 recommandations opérationnelles

Ces principes font l'objet d'une déclinaison sous la forme de 6 recommandations opérationnelles à destination tant des pouvoirs publics que des diverses composantes de la société civile (entreprises, grand public, etc.) :

- Former à l'éthique tous les acteurs-maillons de la « chaîne algorithmique » (concepteurs, professionnels, citoyens) : l'alphabétisation au numérique doit permettre à chaque humain de comprendre les ressorts de la machine ;
- Rendre les systèmes algorithmiques compréhensibles en renforçant les droits existants et en organisant la médiation avec les utilisateurs ;
- Travailler le *design* des systèmes algorithmiques au service de la liberté humaine, pour contrer l'effet « boîtes noires » ;
- Constituer une plateforme nationale d'audit des algorithmes ;
- Encourager la recherche sur l'IA éthique et lancer une grande cause nationale participative autour d'un projet de recherche d'intérêt général ;
- Renforcer la fonction éthique au sein des entreprises (par exemple, l'élaboration de comités d'éthique, la diffusion de bonnes pratiques sectorielles ou la révision de chartes de déontologie peuvent être envisagées).

Les problématiques soulevées par les algorithmes et l'IA

Pour formuler ces principes fondateurs et les recommandations, le rapport s'est basé sur 6 problématiques essentielles remontées lors des débats afin de relever le défi éthique posé par ces technologies :

L'autonomie humaine au défi de l'autonomie des machines : Algorithmes et intelligence artificielle permettent la délégation croissante de tâches, de raisonnements ou de décisions de plus en plus critiques à des machines. Souvent jugées infaillibles et « neutres », celles-ci n'ouvrent-elles pas la voie à une confiance excessive et à la tentation pour chacun de ne pas exercer pleinement ses responsabilités ? Comment faire face aux formes nouvelles de dilution de la responsabilité qu'impliquent des systèmes algorithmiques complexes et très segmentés ?

Biais, discrimination et exclusion : Biais, discrimination et exclusion constituent des effets clairement identifiés des algorithmes et de l'intelligence artificielle. Parfois volontaires, ces phénomènes sont, à l'heure du déploiement des algorithmes de *machine learning*, le plus souvent inconscients et difficilement repérables. Comment appréhender cette nouvelle problématique ?

Fragmentation algorithmique : la personnalisation contre les logiques collectives :

La personnalisation constitue l'une des grandes promesses de l'algorithme tel qu'il s'est construit avec le Web. Bien que des bénéfices certains pour l'individu émanent de la segmentation et du profilage, comment peuvent-ils tout autant affecter vigoureusement des logiques collectives essentielles à la vie de nos sociétés (pluralisme démocratique et culturel, mutualisation du risque) ?

Entre limitation des mégafichiers et développement de l'intelligence artificielle : un équilibre à réinventer : L'intelligence artificielle nécessite d'importantes quantités de données à des fins d'apprentissage. Elle suscite également de nombreuses promesses. Comment exploiter celles-ci tout en maintenant l'impératif de protéger les libertés

individuelles de chacun, intrinsèque à la législation européenne de protection des données personnelles ?

Qualité, quantité, pertinence : l'enjeu des données fournies à l'IA : L'entraînement d'un algorithme d'apprentissage implique de l'alimenter avec des données sélectionnées avec soin, pertinentes au regard de l'objectif poursuivi et en quantité suffisante. Comment parvenir à une telle attitude critique en dépit de la tendance à une confiance excessive dans la machine ?

L'identité humaine au défi de l'intelligence artificielle : L'idée d'une spécificité humaine irréductible se voit questionnée par l'autonomie grandissante des machines ainsi que par l'apparition de formes d'hybridation entre humains et machines. Faut-il et est-il possible de parler au sens propre d'« éthique des algorithmes » ? Comment appréhender cette nouvelle classe d'objets que sont les robots humanoïdes, susceptibles d'engendrer chez l'homme des formes d'affect ?

Mars 2018

Résumé du rapport "Donner un sens à l'intelligence artificielle"

En septembre 2017, M. Cédric Villani, député, a été chargé de conduire une mission sur la mise en œuvre d'une stratégie française et européenne en intelligence artificielle (IA). Le rapport, présenté en mars 2018, aborde successivement différentes facettes de l'IA : politique économique, recherche, emploi, éthique, cohésion sociale. Cinq annexes insistent par ailleurs sur des domaines d'intérêt particulier : éducation, santé, agriculture, transport, défense et sécurité. ([consulter le site aiforhumanity](#)).

L'intelligence artificielle est actuellement dominée par cinq pays : les États-Unis, la Chine et, dans une moindre mesure, le Royaume-Uni, le Canada et Israël. Selon le rapport, "la France dispose de tous les atouts pour exister pleinement sur la scène internationale", mais il est nécessaire de mettre en place une véritable stratégie de développement. Cet effort devrait être concentré sur quatre secteurs stratégiques :

- la santé ;
- l'environnement ;
- les transports ;
- la défense.

Dans ces secteurs, le développement de l'IA va passer par la mise en place de plateformes de mutualisation alliant le public et le privé. Grâce à ces plateformes, des données seront accessibles, et à partir de ces données, des nouvelles fonctionnalités seront élaborées.

Le développement de l'IA passe par une politique de la donnée : création de "communs" de la donnée (ressources dont l'usage et la gouvernance sont définies par une communauté), accélération du mouvement d'ouverture des données publiques (open data), définition de modalités d'accès aux données privées, protection du transfert de données hors de l'Union européenne, etc.

L'État doit être un "puissant moteur" des transformations. Pour cela, l'intelligence artificielle devrait être intégrée à la conduite des politiques publiques et un coordinateur interministériel devrait être installé.

Quelle politique pour la recherche ?

La mission déplore une certaine fuite des cerveaux français vers de grandes entreprises américaines. Pour retenir les chercheurs français en France et inciter des chercheurs étrangers à s'installer en France, le rapport insiste sur la nécessité de créer un environnement de travail attractif. Cela passe par de meilleurs salaires mais aussi par "l'allègement drastique" des démarches administratives. Les "Instituts interdisciplinaires d'intelligence artificielle" (3IA) que le rapport appelle de ses vœux devront être conçus comme des "zones franches de l'IA".

Pour rivaliser avec les moyens des grands acteurs privés tels que Facebook ou Google, un supercalculateur conçu pour les applications d'intelligence artificielle devrait être mis en place à la disposition des instituts 3IA.

Quels enjeux sociaux et éthiques ?

Les développements de l'intelligence artificielle, de la robotisation et de l'automatisation laissent présager des transformations radicales du monde du travail. Cette transition doit être pensée et accompagnée.

Le développement de l'IA ne doit pas s'opérer sans tenir compte de sa soutenabilité. L'IA doit être "nativement écologique". Par exemple, les besoins en espace de stockage de données à l'horizon 2040 risquent d'être supérieurs aux capacités d'extraction du silicium. Les impacts de l'intelligence artificielle doivent être anticipés et il faut rechercher des voies pour une IA moins consommatrice d'énergie.

Enfin, sur le plan éthique, le rapport pointe notamment les problèmes posés par les "boîtes noires" (algorithmes dont on ne parvient pas à expliquer le fonctionnement interne). Plus largement, les considérations éthiques devraient accompagner tout développement d'algorithme d'IA. Il faut "obliger les développeurs à se poser les bonnes questions". La mission recommande ainsi la création d'un comité d'éthique de l'IA.

Les 6 parties du rapport sont les suivantes :

- 01 - Pour une politique offensive de la donnée
- 02 - Miser sur 4 secteurs stratégiques
- 03 - Libérer les potentiels de la recherche française
- 04 - Anticiper l'impact de l'IA sur le travail et expérimenter
- 05 - Pour une IA écologique
- 06 - Ouvrir les boîtes noires de l'IA
- 07 - Pour une intelligence artificielle inclusive et diverse

Depuis 2017 mais surtout à la suite du rapport dit Villani, le Gouvernement français a lancé une réflexion autour du développement de l'intelligence artificielle. Celle-ci prend la forme d'une stratégie nationale, divisée en deux phases entre 2018 et 2025 ([pour en savoir plus](#)).

Faisant suite à la vision apportée par la mission Villani, la stratégie nationale pour l'intelligence artificielle a jeté les bases d'une structuration de long terme de l'écosystème d'IA, à tous les stades du développement technologique : recherche, développements et innovations, applicatifs, mise sur le marché et diffusion intersectorielle, soutien et encadrement du déploiement.

Renforcer les capacités de recherche (2018-2022)

Initialement dotée de près de 1,5 milliard d'euros sur la période 2018-2022, la première phase de la stratégie nationale pour l'IA tend à positionner la France comme l'un des leaders mondiaux de cet ensemble de disciplines scientifiques et de technologies-clés du traitement de l'information.

Comment ? En favorisant la création et le développement d'un réseau d'instituts interdisciplinaires d'intelligence artificielle, le soutien à des chaires d'excellence en IA, le financement de programmes doctoraux et l'investissement dans les capacités de calcul de la recherche publique (ex : le supercalculateur Jean Zay). Depuis 2018, cette phase s'est ainsi largement concentrée sur le renforcement des capacités de recherche.

81

Nombre de laboratoires d'IA en France en 2021. C'est le plus grand nombre parmi les pays européens.

502

Nombre de startups spécialisées en IA en 2021, soit une hausse de 11 % par rapport à 2020.

13 459

Nombre de personnes qui travaillent dans les start-ups de l'IA en 2021 (pour 70000 emplois indirects générés). 9000 personnes devraient être recrutées sur l'année 2022 dans ces start-ups selon le recensement de France Digitale.

Former et attirer les meilleurs talents en IA (2021-2025)

Le Gouvernement a lancé le 8 novembre 2021 la deuxième phase de la stratégie nationale pour l'IA, afin d'accroître le nombre de talents formés dans ce domaine et d'accélérer le potentiel de recherche et développement en succès économiques.

Un appel à projets sur l'intelligence artificielle embarquée

Bpi France a lancé le 8 novembre 2021 un appel à projet (AAP) visant à soutenir les innovations matérielles et logicielles liées aux applications d'IA embarquée, depuis leur conception jusqu'à la démonstration de leur efficacité en situation réelle. L'appel à projets est ouvert jusqu'au 5 octobre 2022. En savoir plus.

Pour cette deuxième phase, il est prévu de consacrer au total 2,22 milliards d'euros à l'IA pour les cinq ans qui viennent, dont 1,5 milliard d'euros de financements publics et 506 millions d'euros de cofinancements privés.

Dans le détail, la formation représente plus de la moitié de l'effort (50 %). 40 % des investissements sont également destinés aux mesures de soutien à l'innovation et aux mesures économiques. Enfin, une part de près de 10 % du budget est destinée à financer les mesures en faveur de la recherche scientifique et de transfert.

Juin 2020

Présentation du Partenariat mondial sur l'IA

Le Partenariat mondial sur l'intelligence artificielle (PMIA) a été lancé en juin 2020. Il s'agit d'une initiative internationale et multipartite visant à guider le développement et l'utilisation responsables de l'IA, dans le respect des droits de la personne, des libertés fondamentales et de nos valeurs démocratiques communes, conformément à la Recommandation de l'OCDE sur l'IA.

Les membres fondateurs du PMIA comprennent l'Allemagne, l'Australie, le Canada, la République de Corée, les États-Unis d'Amérique, la France, l'Inde, l'Italie, le Japon, le Mexique, la Nouvelle-Zélande, le Royaume-Uni, Singapour, la Slovaquie et l'Union européenne ([plus d'informations](#)).

Conscients de la nécessité d'une coopération à l'échelle internationale pour exploiter le plein potentiel de l'intelligence artificielle (IA) et en faire bénéficier l'ensemble des citoyens, le Canada et la France lancent conjointement avec l'Allemagne, l'Australie, la République de Corée, les États-Unis d'Amérique, l'Italie, l'Inde, le Japon, le Mexique, la Nouvelle-Zélande, le Royaume-Uni, Singapour, la Slovaquie et l'Union européenne le Partenariat Mondial sur l'intelligence artificielle (PMIA) qui encouragera et guidera le développement responsable d'une intelligence artificielle fondée sur les droits de l'homme, l'inclusion, la diversité, l'innovation et la croissance économique.

Le PMIA est une initiative internationale et multipartite visant à guider le développement et l'utilisation responsables de l'IA, dans un esprit de respect des droits de la personne, d'inclusion, de diversité, d'innovation et de croissance économique. Afin d'atteindre cet objectif, les pays membres s'emploieront à jeter des ponts entre la théorie et la pratique et soutiendront des activités de recherche de pointe ainsi que des activités de mise en application liées aux priorités en matière d'IA.

En collaboration avec des partenaires et des organisations internationales, le PMIA rassemblera des experts issus de l'industrie, de la société civile, des gouvernements et du milieu universitaire. Ces experts se réuniront au sein de Groupes de travail axés sur quatre thèmes : 1) l'utilisation responsable de l'IA ; 2) la gouvernance des données ; 3) l'avenir du travail ; et 4) l'innovation et la commercialisation. À court terme, les experts participants se pencheront également sur l'apport possible de l'IA comme moyen de répondre à la pandémie de COVID-19 et de la surmonter.

Le PMIA sera appuyé par un secrétariat, hébergé par l'OCDE à Paris, ainsi que par deux Centres d'expertise, l'un à Montréal, l'autre à Paris. La collaboration avec l'OCDE donnera lieu à de fortes synergies entre les travaux scientifiques et techniques du PMIA et le leadership international exercé par l'OCDE en matière de politiques liées à l'IA, ce qui approfondira la base de données probantes sous-tendant les politiques sur l'utilisation responsable de l'IA. Les Centres d'expertise fourniront un soutien administratif ainsi qu'un soutien à la recherche au titre de projets pratiques menés ou évalués par les experts des différents Groupes de travail issus de divers secteurs et disciplines. Les centres organiseront également les séances plénières annuelles du groupe d'experts multipartite du PMIA. Le Canada sera l'hôte de la première de ces séances en décembre 2020.

Le Centre d'expertise de Paris, piloté par Inria, interviendra en soutien des deux groupes d'experts sur 3) l'avenir du travail et 4) l'innovation et la commercialisation. Il sera en lien avec le centre d'expertise de Montréal qui gèrera les autres groupes.

Ce lancement couronne deux années de travail de la diplomatie numérique française et de ses partenaires du Canada visant à mettre en œuvre l'appel lancé par la France et le Canada dans la déclaration franco-canadienne sur l'intelligence artificielle de juin 2018.

juin 2020

Synthèse de l'appel du Défenseur des droits et de la CNIL sur les biais algorithmiques

En juin 2020, le Défenseur des droits a appelé, avec la CNIL, à une mobilisation collective sur le sujet des algorithmes et des discriminations ([en savoir plus](#)).

Pour prévenir les biais discriminatoires des algorithmes, les corriger et sanctionner leurs auteurs, le Défenseur des droits appelle à une prise de conscience collective et engage les pouvoirs publics et les acteurs concernés à prendre des mesures pour éviter que les discriminations soient reproduites et amplifiées par ces technologies. Le Défenseur des droits, en partenariat avec la CNIL, [propose les orientations suivantes](#) :

- Former et sensibiliser les professionnels des métiers techniques et d'ingénierie informatique aux risques que les algorithmes font peser sur les droits fondamentaux ;
- Soutenir la recherche pour développer les études de mesure et de prévention des biais, et approfondir la notion de « *fair learning* » – c'est-à-dire la conception d'algorithmes répondant à des objectifs d'égalité et de compréhension, et non seulement de performance ;
- Renforcer les obligations légales en matière d'information, de transparence et d'explicabilité des algorithmes à l'égard des usagers et personnes concernées, mais également des tiers et des professionnels utilisateurs de ces systèmes, au nom de l'intérêt général, comme en témoignent par exemple les questions suscitées par Parcoursup ;
- Réaliser des études d'impact pour anticiper les effets discriminatoires des algorithmes et contrôler leurs effets après leur déploiement.

En juin 2021, le Défenseur des droits a aussi collaboré à une formation en ligne du Conseil de l'Europe sur l'intelligence artificielle et les discriminations. Cette formation présente les réglementations et les normes européennes en vigueur et à venir, des études de cas, des bonnes pratiques, et permet d'engager une discussion sur les objectifs et les méthodes de coopération entre les différents régulateurs dans ce domaine ([pour en savoir plus](#)).

Présentation du règlement dit législation sur l'IA

En avril 2021, la Commission européenne a dévoilé le premier cadre juridique sur l'intelligence artificielle (IA) au sein de l'Union européenne. L'objectif est d'asseoir une vision européenne de l'IA basée sur l'éthique en prévenant les risques inhérents à ces technologies par un règlement commun permettant d'éviter certaines dérives.

Avec la proposition de nouvelle réglementation sur l'IA présentée le 21 avril 2021, la Commission européenne réaffirme l'importance stratégique de l'IA pour l'Europe et la nécessité d'encadrer son usage dans les différents secteurs d'application.

Pour que ces nouvelles règles entrent en vigueur, le Parlement européen et les États membres doivent néanmoins encore les adopter.

Priorité à l'évaluation des risques

L'approche de la Commission européenne en matière d'intelligence artificielle se fonde d'abord sur une identification et un encadrement des risques selon les domaines d'applications – publics ou privés – concernés :

- les risques inacceptables dans le champ de la sécurité, des moyens de subsistance et des droits des personnes ;
- les risques élevés dans les domaines touchant les infrastructures critiques (énergie, transports...), l'éducation, la formation professionnelle, l'emploi, les ressources humaines, les services privés et publics essentiels, le maintien de l'ordre, la justice, les processus démocratiques...

Ainsi, les systèmes d'identification biométrique à distance fondés sur l'IA, sont considérés à haut risque et la Commission rappelle l'interdiction de leur utilisation dans l'espace public et en temps réel "*aux fins du maintien de l'ordre*" en dehors de cas spécifiques encadrés judiciairement.

Dans les secteurs où les risques sont limités, la Commission recommande une transparence d'information sur la présence de l'IA.

Dans d'autres domaines d'application comme les jeux vidéos ou les filtres antispam, les risques liés à la présence de l'IA sont considérés comme minimes et ne nécessitent pas d'encadrement spécifique.

Affirmer le modèle européen

La nouvelle réglementation s'accompagne d'un plan coordonné destiné à accélérer les investissements dans l'IA et à stimuler la mise en œuvre de stratégies nationales. Son financement est assuré par des programmes existants comme le fonds pour une Europe numérique ou Horizon Europe et une dotation issue du plan de relance économique européen décidé dans le contexte du Covid-19.

Ce plan européen prévoit de :

- créer les conditions propices au développement de l'IA ;

- favoriser l'excellence en matière d'IA (partenariat public-privé) ;
- former de nouvelles compétences dans les technologies de l'IA ;
- établir un *leadership* technologique de l'Europe en matière d'IA dans les secteurs vertueux (environnement durable, santé, agriculture..).

Mai et novembre 2021

Résumé des avis du CNPEN sur les véhicules autonomes et les agents conversationnels

Le Comité national pilote d'éthique du numérique, institution créée en 2019 à la suite du rapport Villani, a rendu en 2021 deux avis sur les agents conversationnels et les véhicules autonomes.

Sur les véhicules autonomes (avril 2021)

Dans sa lettre du 15 juillet 2019 donnant mission au président du CCNE de mettre en œuvre une démarche pilote concernant les questions d'éthique des sciences, technologies, usages et innovations du numérique et de l'intelligence artificielle¹, le Premier ministre a souhaité que les travaux conduits dans cette phase pilote concernent en particulier le « véhicule autonome » en coordination avec la mission concernant la stratégie nationale de développement de la mobilité routière automatisée confiée à Mme Anne-Marie Idrac².

Le terme « véhicule autonome » est généralement utilisé pour désigner des véhicules routiers équipés de fonctions de conduite automatisées pouvant se substituer au conducteur, qu'ils soient individuels ou partagés, destinés au transport de passagers ou de biens. Le déploiement à grande échelle de tels véhicules rendrait nécessaire une réglementation internationale. Au préalable, les questions éthiques qu'ils soulèvent doivent être étudiées.

Dans cet avis, le CNPEN met en lumière les enjeux liés :

- à la terminologie utilisée à propos du véhicule et des fonctions liées à son automatisation ;
- à la sécurité routière et à la sûreté de fonctionnement ;
- au contrôle humain du véhicule ;
- aux libertés des personnes ;
- aux impacts sociaux et environnementaux.

La pertinence des motivations invoquées pour l'automatisation des fonctions de conduite y est également discutée : enjeux de sécurité routière, bénéfice social, impact environnemental et écologie ou encore motifs économiques.

[Lire l'avis](#)

Sur les agents conversationnels

Les agents conversationnels (aussi appelés « chatbots ») sont des systèmes numériques capables d'interagir avec leurs utilisateurs en langage naturel, à l'écrit comme à l'oral. Aujourd'hui, on en trouve dans de nombreux objets de notre quotidien (smartphones, enceintes connectées, voitures, sites internet) : ils apportent des conseils de santé, facilitent le recrutement, gèrent des services après-vente, forment à distance, aident les personnes vulnérables, etc.

En France comme au niveau européen, les agents conversationnels, parce qu'ils s'appuient sur des données personnelles, commencent à attirer l'attention des autorités, des régulateurs et du public. Au-delà de ces points réglementaires, se posent de nombreuses questions éthiques. Quels comportements et croyances avons-nous par rapport aux agents conversationnels ? Un chatbot peut-il mentir à son utilisateur ? Les erreurs des agents conversationnels seront-elles plus acceptables que celles d'un être humain ? Un chatbot doit-il imiter systématiquement l'être humain ?

Parmi les différents enjeux d'éthique qu'identifie le comité, les principales tensions concernent le brouillage des frontières entre une machine et un être humain, l'imitation du langage et des émotions par les chatbots, ainsi que la capacité des agents conversationnels à manipuler les interlocuteurs humains.

[Lire l'avis](#)

Avril 2022

Résumé de l'avis de la CNCDH sur les impacts de l'IA sur les droits fondamentaux

Réunie en assemblée plénière le 7 avril 2022, la Commission nationale consultative des droits de l'homme (CNCDH) a adopté à l'unanimité un avis dans lequel elle formule près de 20 recommandations pour un encadrement juridique ambitieux pour l'intelligence artificielle ([lire l'avis relatif à l'impact de l'intelligence artificielle sur les droits fondamentaux](#)).

Alors que les recherches sur l'intelligence artificielle (IA) et la mise en œuvre de ses applications pratiques se développent, la réglementation actuelle demeure lacunaire pour endiguer les atteintes possiblement majeures aux droits fondamentaux. Dans le contexte de l'adoption prochaine de la proposition de règlement de l'UE sur le sujet, et des travaux en cours au sein du Conseil de l'Europe, la CNCDH invite les pouvoirs publics à promouvoir un encadrement juridique ambitieux en la matière.

Elle recommande, d'une part, d'interdire certains usages de l'IA jugés trop attentatoires aux droits fondamentaux, tels que le scoring social ou l'identification biométrique à distance des personnes dans l'espace public et les lieux accessibles au public. D'autre part, elle recommande de faire peser sur les utilisateurs d'un système d'IA des exigences en mesure de garantir le respect des droits fondamentaux : une étude d'impact, une consultation des parties prenantes, une supervision du système tout au long de son cycle de vie. La CNCDH appelle enfin à reconnaître des droits aux personnes ayant fait l'objet d'une décision impliquant un algorithme, notamment le droit à une intervention humaine dans le processus de décision, ou encore un droit au paramétrage des critères de fonctionnement du système d'IA.

Afin d'apporter un éclairage sur les enjeux de l'intelligence artificielle (IA) liés à la protection de la vie privée et d'accompagner les professionnels dans leur mise en conformité, la CNIL propose un ensemble de ressources dédiées ([plus d'informations](#)). La CNIL avait précédemment traité de l'IA notamment dans son livre blanc sur les agents conversationnels ou au sein de son laboratoire d'innovation.

Dans le cadre de ses missions d'information et de protection des droits, d'accompagnement vers la conformité et d'anticipation et d'innovation, la CNIL propose un ensemble de contenus consacrés à l'IA.

Ces ressources s'inscrivent ainsi dans une stratégie européenne visant à stimuler l'excellence dans le domaine de l'intelligence artificielle, ainsi que des règles destinées à garantir la fiabilité de ces technologies. Il s'agit en particulier d'élaborer un cadre réglementaire solide pour l'IA fondé sur les droits de l'Homme et les valeurs fondamentales et ainsi instaurer la confiance des citoyens européens.

Les éléments proposés s'adressent à trois publics distincts :

- le grand public intéressé par le fonctionnement des systèmes d'IA, leurs implications dans nos vies quotidiennes ou encore souhaitant tester leur fonctionnement ;
- les professionnels (responsables de traitement ou sous-traitants) mettant en œuvre des traitements de données personnelles reposant sur des systèmes d'IA, ou le souhaitant et qui s'interrogent sur la façon d'assurer leur conformité au RGPD ;
- les spécialistes (chercheur en IA, expert en science des données, ingénieur en apprentissage automatique, etc.) : pour les personnes manipulant l'intelligence artificielle au quotidien, curieuses des enjeux que fait peser l'intelligence artificielle sur la protection des données et qui s'intéressent à l'état de la technique sur ces questions.

Les précédents travaux de la CNIL en lien avec l'IA :

- Du fait de leur présence dans les smartphones, les voitures ou encore les réfrigérateurs, les assistants vocaux sont devenus incontournables. Afin d'en explorer les enjeux éthiques, techniques et juridiques, la CNIL a publié en septembre 2020 un livre blanc à destination des professionnels comme des utilisateurs ([plus d'informations](#)).
- Tous les articles publiés par le LINC, le laboratoire d'innovation de la CNIL sur l'intelligence artificielle sont également référencés dans un dossier sur l'IA ([plus d'informations](#)).