



Intelligence artificielle et G/RSS

**Note consultative de l'Équipe internationale
de conseil au secteur de la sécurité (ISSAT)**

Juin 2023

Acronymes

IA

G/RSS

Sommaire exécutif

La décennie précédente, tout particulièrement la période de la COVID-19, a été marquée par une montée en puissance de l'utilisation des capacités numériques dans le domaine de la sécurité et de la justice, à travers le recours aux pratiques de gouvernance électronique, aux technologies de surveillance, ainsi qu'aux pratiques liées aux Big Data et au renforcement des mesures de la cybersécurité.

L'usage croissant de l'intelligence artificielle (IA) s'inscrit dans la lignée de cette tendance et revêt une importance majeure dans la gouvernance et la réforme du secteur de la sécurité (G/RSS).

Les systèmes d'IA contribuent déjà à améliorer l'efficacité de la prestation des services publics. Ils ont également permis de générer des prévisions basées sur des données et de simuler des scénarios complexes pour appuyer le processus décisionnel.

Cependant, ces technologies peuvent également être utilisées à mauvais escient et risquent, par conséquent, de réduire la confiance du public dans le secteur de la sécurité et de la justice en raison des lacunes en matière d'accès et de sensibilisation ainsi que des problèmes de confidentialité et de partialité.

Face à ces défis, les objectifs de la gouvernance et la réforme du secteur de la sécurité (G/RSS) demeurent tout à fait pertinents dans un monde où l'IA façonne les contours des secteurs de la sécurité et de la justice.

Une bonne gouvernance du secteur de la sécurité contribue à façonner des secteurs de sécurité axés sur les citoyens, respectueux des droits humains efficaces, et redevables.

La sensibilisation et le renforcement de l'engagement du public, le renforcement des capacités, la garantie de l'inclusion et le respect des droits humains ainsi que le soutien de la mise en place de mécanismes de supervision et de régulation sont autant de prérequis nécessaires pour ce domaine en plein essor.

Aujourd'hui, de nombreux États utilisent déjà des solutions axées sur la technologie dans leurs secteurs de sécurité et de justice, notamment dans les domaines de l'accès à la justice, de la cybersécurité, de l'élaboration des stratégies de réforme de la défense, du développement des cadres juridiques, de la gouvernance électronique et des initiatives de données ouvertes.

Ces occurrences offrent des enseignements utiles sur l'orientation que pourrait emprunter ce domaine.

Dans la présente note, l'équipe consultative du secteur de la sécurité internationale (ISSAT) met en lumière les opportunités que l'intelligence artificielle (IA) offre pour les réformes, ainsi que les défis qu'elle soulève.

Elle propose une structure clarifiant les liens entre l'IA et la gouvernance/réforme du secteur de la sécurité (G/RSS), tout en présentant certaines leçons émergentes et bonnes pratiques issues des programmes sur le terrain et des initiatives des donateurs.

Définition de l'IA

L'IA est un logiciel informatique qui imite le mode de pensée humain afin de réaliser des tâches complexes, tels que le raisonnement et l'analyse. Diverses approches permettent d'atteindre cette intelligence, notamment à travers l'apprentissage-machine (ML)¹ et le traitement automatique du langage naturel (NLP)².

L'IA peut jouer un rôle significatif aussi bien dans l'amélioration de la prestation des services de sécurité et de justice que dans leur gouvernance. À ce titre, elle peut fournir les données nécessaires à la prise de décision basée sur des preuves, à l'allocation des ressources et à l'amélioration de l'accès à la justice.

¹ Aujourd'hui, l'Apprentissage-machine (ML) s'est imposé en tant qu'un sous-ensemble de l'IA, qui le plus grand impact. Généralement, il s'agit d'une approche qui permet aux ordinateurs d'apprendre sans être explicitement programmés pour cette fin. Cela peut se faire à travers « l'initiation » d'un modèle informatique à de grandes quantités de données qu'il utilise pour s'entraîner à identifier des tendances.

² Le NLP est un logiciel qui permet aux ordinateurs de décrypter les langues « naturelles » (par exemple l'espagnol, le mandarin) et de procéder à l'interprétation de la communication humaine.

Les défis de l'IA lancés à la Justice et à la Sécurité

Le concept de l'IA a suscité depuis longtemps la peur et la fascination, et les histoires d'IA au service de la sécurité et de la défense datent en effet depuis des millénaires.³

Que notre vision de l'IA soit davantage orientée vers la destruction ou la protection, l'IA pose de nombreux défis de taille pour la justice et la sécurité. Parmi les plus débattus de ces défis aujourd'hui, figurent :

- **La pérennisation des préjugés et de la discrimination** : Les systèmes d'IA reposent principalement sur des données et sont aussi bons que les données qu'ils utilisent. Si ces données reproduisent des discriminations systémiques, la technologie les reproduira également, même de manière non intentionnelle.
- **Le jeu du pouvoir** : Tant qu'elle est dépourvue de mécanismes de supervision et de régulation, l'AI peut être utilisée comme moyen d'oppression politique, de provocation, de persécution et de manipulation, notamment à travers la surveillance, la permutation des visages (Deep fakes), et autres formes de désinformation, et l'imposition d'agendas politiques.
- **La réglementation à la poursuite de l'innovation** : Dans de nombreux contextes, l'utilisation des technologies dans le secteur de la sécurité et de la justice, intervient bien avant que des cadres juridiques adéquats ne soient développés. Cela peut porter préjudice aux droits civiques, aux droits des consommateurs et aux droits humains. L'élaboration de nouvelles normes, politiques et lois s'avère difficile en raison de la complexité et de l'évolution rapide de la technologie.

L'IA et la G/RSS se croisent particulièrement dans des domaines programmatiques clés : **La cybersécurité, l'État du droit et gouvernance et la réforme inclusive par la technologie** (ou l'utilisation des technologies pour promouvoir la supervision, l'inclusion et la redevabilité).

Une bonne gouvernance du secteur de la sécurité contribue de manière substantielle à atténuer les risques que l'IA fait peser sur les droits humains, la sécurité et la sûreté. Cela se fait ainsi par le biais de réformes en matière de gouvernance et d'État du droit, en créant des mécanismes de surveillance et de redevabilité, en sécurisant les plateformes de données, ainsi qu'en utilisant les technologies en faveur de réformes plus inclusives.

Les systèmes d'IA nécessitent la mise en place de cadres législatifs et réglementaires appropriés pour sécuriser les données privées, contrôler les activités du gouvernement et du secteur privé et interdire une mise en œuvre déficiente susceptible de porter atteinte aux droits fondamentaux.

Cinq façons dont l'IA peut soutenir la G/RSS

L'IA peut améliorer l'efficacité de la prestation des services publics, en augmentant la productivité tout en réduisant les coûts opérationnels. Les solutions numériques offrent non seulement des processus plus accessibles et plus efficaces mais elles permettent également une meilleure gestion de l'information grâce à la centralisation de l'acquisition et de la gestion des données.

La croissance exponentielle de l'utilisation de l'IA est en train d'opérer une transformation dans le secteur de la sécurité et de la justice. L'IA soutient la disposition et la résilience en période de crise, comme une pandémie. Les pays ayant mis en place des solutions technologiques robustes avant la COVID-19 ont généralement fait preuve d'une plus grande résilience pendant la pandémie.

L'IA peut améliorer la transparence dans le secteur de la sécurité et de la justice. A titre d'illustration, le suivi des affaires de corruption peut s'avérer une œuvre de longue haleine en présence de données volumineuses et difficiles à gérer. Les outils d'IA peuvent aider à donner un sens au flux d'informations recueillies.

Les partenariats public-privé offrent de nombreuses opportunités. La coopération entre les acteurs publics et privés constitue et continuera de constituer le fondement de l'expansion de l'IA dans les secteurs de la sécurité et de la justice. Les technologies fournies par le secteur privé seront appliquées aux données traditionnellement gérées par le secteur public. Pour garantir la légitimité de ce processus, il est essentiel d'impliquer les institutions de l'État de droit, la société civile, le milieu académique et les citoyens dans les décisions concernant les types de données à collecter ainsi que les mécanismes de protection des données.

³ Alex Shashkevich, "Stanford researcher examines earliest concepts of artificial intelligence, robots in ancient myths," 28 February 2019, <https://news.stanford.edu/2019/02/28/ancient-myths-reveal-early-fantasies-artificial-life/>

L'IA permet aussi de favoriser des réformes basées sur des données en améliorant la collecte et le stockage des données, ce qui permet de générer davantage d'informations basées sur les données en ce qui concerne les besoins, l'accessibilité et les modèles d'utilisation. Ce système favorise, à moyen et à long terme, l'instauration d'une gouvernance basée sur les données.

Que pourrait faire mieux la Communauté Internationale ?

- 1. Soutenir le développement de cadres nationaux pour l'Etat de droit**, lesquels peuvent être adaptés aux nouvelles réalités de l'IA. L'IA peut conférer des pouvoirs étendus au secteur de la sécurité, et les pays disposant de cadres faibles pour l'état de droit sont exposés aux risques les plus élevés d'excès et d'abus de la part du secteur de la sécurité, en raison des lacunes dans les mécanismes de régulation.
- 2. Fournir les meilleures pratiques internationales** sur la manière d'intégrer la technologie dans la prestation et la supervision des services de sécurité et de justice. Dans de nombreux cas, les applications et les systèmes d'IA sont fournis aux gouvernements par le secteur privé ou grâce à l'aide des donateurs. Les autorités nationales pourraient ne pas disposer de capacités suffisantes pour encadrer les aspects éthiques, juridiques et de communication liée à la gestion des données.
- 3. Élaborer des directives opérationnelles** sur les modalités de collecte et de gestion des données à caractère personnel pour répondre aux besoins définis par le gouvernement.
La transparence des activités gouvernementales peut être grandement améliorée au moyen des technologies numériques. Toutefois, il est crucial de communiquer à la communauté que ces données personnelles peuvent devenir accessibles et utilisables par le secteur de la sécurité.
- 4. Renforcer les capacités des acteurs étatiques** en matière d'acquisition, d'utilisation, de maintenance et de durabilité des applications d'IA. L'utilisation de nouvelles technologies dans les institutions publiques crée inévitablement un écart de compétences.
Le personnel actuel sera dans l'obligation de se familiariser avec les nouvelles pratiques et méthodologies de travail afin de garantir que l'utilisation des technologies soit efficacement optimisée et durable.
- 5. Faciliter le développement de partenariats public-privé** axés sur le développement de la technologie et des systèmes d'IA. Les partenaires internationaux ont un rôle-clé à jouer dans le développement des partenariats public-privé, en s'appuyant sur les ressources dont dispose le secteur privé, tout en veillant à ce que la technologie soit développée conformément aux normes publiques.
- 6. Encourager et orienter les autorités nationales vers des approches pangouvernementales et pansectorielles** pour renforcer la coordination et la collaboration numérique dans le secteur public. L'application des solutions d'IA doivent s'inscrire dans une démarche de cohérence intergouvernementale afin d'éviter toute répercussion négative sur la collaboration interinstitutionnelle. Il s'agit ainsi d'une pratique fort utile pour encourager la transformation sectorielle.
- 7. Fournir des conseils techniques sur la manière de joindre les personnes souvent « transparentes » pour le système.** Ceux qui ont été marginalisés par les systèmes traditionnels sont souvent moins instruits, souffrent de discrimination fondée sur l'identité, d'inégalités socioéconomiques ou d'un manque de conscience ou de volonté de s'impliquer dans le système.

La technologie pourrait aliéner davantage ces groupes. Les partenaires internationaux interviennent dans ce cas pour soutenir le financement et le déploiement d'études telles que : les évaluations des besoins communautaires, l'analyse des parties prenantes, le diagnostic des insuffisances ainsi que l'analyse de genre et des conflits.

Chacune de ces études devrait offrir une meilleure compréhension de ceux qui ont besoin de sécurité et de justice et de comment combler au mieux leur écart d'accessibilité au système.

Appui international à la stratégie nationale et à l'Etat de droit

Les stratégies nationales et cadres juridiques devraient orienter et promouvoir le développement des systèmes d'IA, en renforçant la gouvernance globale et les mécanismes de redevabilité, et coordonner les partenariats public-privé dans ce domaine.



Stratégie nationale de l'IA

Une stratégie nationale en matière d'IA est un cadre politique canalisant les ressources nationales d'un pays vers le développement de ses capacités en IA, en s'engageant clairement à mettre à jour et à déployer des mécanismes de supervision et de régulation correspondants.

Une telle stratégie doit avoir un objectif clair et réaliste sur la manière dont les avantages des systèmes d'IA peuvent être exploités, tout en veillant à promouvoir une culture de redevabilité.

Elle identifie ce qui doit changer et pourquoi ce changement est nécessaire. Cette vision est nécessaire pour élaborer une feuille de route politique claire et efficace, avec des cibles précises pour le changement.

Une stratégie nationale d'IA doit également évaluer le capital et l'infrastructure disponibles pour le déploiement de la vision. Dans ce cadre, une approche sectorielle pourrait bien être utile pour explorer le potentiel de l'IA dans chaque secteur concerné.

Elle doit inclure un échéancier permettant une mise en œuvre progressive ainsi qu'un mécanisme de suivi qui assure des mises à jour régulières des progrès réalisés, évalués par rapport à des résultats concrets et mesurables.

Pays disposant de stratégie nationale globale (ou en voie de l'être) en 2021

Pays de l'OCDE	Europe	MENA	Asie	Amérique Latine	Afrique
Autriche	Russie	Emirats Arabes Unis	Chine	Chili	Kenya
États-Unis	Estonie	Oman	Singapour	Brésil	Maurice
Canada	Lettonie	Arabie Saoudite	Taïwan	Uruguay	
Royaume-Uni	Serbie	Tunisie	Malaisie	Argentine	
Norvège	Roumanie	Egypte	Indonésie	Colombie	
France	Slovénie	Israël	Vietnam	Mexique	
Australie	République Tchèque	Chypre	Inde		
Nouvelle-Zélande	Slovaquie				

Danemark	Hongrie
Pays-Bas	Croatie
Allemagne	Pologne
Irlande	Bulgarie
Belgique	Ukraine
Italie	Lituanie
Portugal	Malte
Finlande	
Grèce	
Espagne	
Japon	
Corée du Sud	

Dimensions-clés d'une stratégie nationale en matière d'IA

Mettre en place un ensemble de lois standardisées sur la protection des données et faire face aux préoccupations éthiques

Les données sont au cœur des mécanismes d'IA. Raison pour laquelle, les lois sur la protection des données sont nécessaires pour répondre au mieux aux préoccupations éthiques et réglementer l'accès aux données et leur utilisation.

Ces lois devraient définir avec clarté les relations entre les personnes concernées par les données et les parties prenantes qui les traitent, en expliquant la manière dont ces données sont collectées, stockées, traitées, partagées et supprimées.

Dans le secteur de la sécurité, les lois relatives à la protection des données devraient accorder la priorité aux engagements nationaux et internationaux conclus par le pays en matière de droits humains.

Créer un environnement de recherche solide et promouvoir l'intégration public-privé

Les stratégies nationales doivent permettre et définir un espace pour les partenariats public-privé orientés vers la recherche et le développement des systèmes d'IA. Bien que le secteur public détienne diverses banques de données, comme il collecte naturellement des informations pour guider la prise de décision, il n'en demeure pas moins qu'il ne détient pas toujours la capacité de développer et de mettre en œuvre, les technologies d'IA de manière autonome. Cependant, ce partenariat comporte des risques et doit être guidé par des cadres éthiques de gouvernance.

Adapter le personnel à la technologie de l'IA

Tout développement de la technologie de l'IA doit anticiper l'impact qu'il pourrait produire sur les effectifs en place ainsi que le besoin nécessaire en matière de compétences et d'infrastructures futures qui devraient accompagner la transition vers les systèmes d'IA.

La technologie de l'IA requiert en effet de nouvelles compétences et infrastructures et va, par ricochet, supprimer certains types d'emplois, via l'automatisation de tâches qui étaient auparavant exécutées par l'homme.

Un plan de renforcement des capacités ainsi que des évaluations fonctionnelles nécessaires, doivent être intégrés dans la stratégie. Le personnel de sécurité et de justice devrait être capables de comprendre les modes de fonctionnement des systèmes d'IA et de surveiller de manière adéquate leurs performances.

S'engager dans une collaboration internationale

Les technologies de l'IA font appel à des compétences et à des capacités inégalement réparties dans le monde. Compte tenu du rythme et de l'ampleur de la recherche mondiale sur l'IA, les gouvernements ne peuvent pas travailler de manière isolée.

L'échange d'informations et l'instauration de partenariats internationaux devraient favoriser le déploiement des systèmes d'IA dans les pays disposant de capacités de développement des technologies de l'information et de la communication (TIC) plus faibles. La collaboration internationale est également essentielle pour renforcer la gouvernance de l'IA conformément aux meilleures pratiques internationales.

Lois et réglementation éthique sur la protection des données

La réglementation éthique de l'IA définit un système de principes moraux destinés à orienter le développement et l'utilisation de la technologie de l'IA. Ces principes-clés tournent généralement autour des concepts « d'autonomie », « de non-nuisance », de « bénéfice » et de « justice ». Ces concepts sont essentiels pour établir des cadres d'audit adéquats, qui superviseront la responsabilité de la technologie de l'IA.



Autonomie

La mise en œuvre des applications d'IA implique de céder le pouvoir de décision humain aux solutions technologiques. La question-clé qui se pose à la suite de cette délégation étant : 'A quel point' ?

L'équilibre entre le pouvoir de décision détenu par l'homme et sa délégation à la machine se situe au cœur même du principe « d'autonomie ». L'objectif est d'éviter qu'une autonomie artificielle accrue n'entrave l'épanouissement de l'autonomie humaine.

A cette fin, les systèmes d'IA doivent être développés et déployés avec l'idée selon laquelle l'autonomie humaine doit être favorisée tandis que l'autonomie des machines doit rester restreinte et réversible, au cas où il serait nécessaire de rétablir le pouvoir décisionnel humain.

Principes connexes au concept « d'autonomie »

Pouvoir de décider	Contrôle humain	Surveillance humaine	Transparence
Interprétabilité	Explicabilité	Liberté	Ouverture
Droits fondamentaux	Valeurs humaines	Vie privée	Protection de la vie privée

Bénéfice

Le concept de « bénéfice » dans l'éthique de l'IA exprime bien la nécessité « prioriser le bien-être humain en tant que résultat, dans tous les processus de conception ». ⁴ La technologie de l'IA doit garantir les conditions préalables de base pour la vie sur la planète, la prospérité continue de l'humanité et la préservation de l'environnement pour les générations futures.

Principes connexes au concept de « bénéfice »

Promouvoir le bien-être	Bénéfice à la société	Générer des bénéfices nets
Durabilité	Impact	Efficacité
Explicabilité	Conception axée sur l'utilisateur	Approche centrée sur les citoyens

Non -Nuisance

Le concept de « Non Nuisance » met en avant la nécessité d'instituer des garanties de protection des droits fondamentaux des individus contre l'usage excessif ou abusif des technologies de l'IA. La prévention des atteintes à la vie privée est la dimension essentielle sous-tendant ce concept.

Il comporte également d'autres dispositions se rapportant à d'autres menaces, telles qu'une éventuelle course aux armements de l'IA ou l'auto-amélioration récursive des systèmes d'IA. Le concept de « Non Nuisance » couvre à la fois les personnes chargées de développer l'IA et la technologie elle-même.

Principes connexes au concept de Non Nuisance

Contrôle des risques	Sûreté	Sécurité
Capacité de prudence	Protection des données	Vie privée (pour éviter le préjudice)
Explicabilité	Transparence (pour éviter les préjudices)	Reproductibilité
Précision	Fiabilité	Déploiement responsable

Justice

Le concept de « Justice » traite des conséquences des disparités au niveau des capacités de prise de décision au sein d'un pays. Il veille à ce que l'IA contribue à la justice mondiale, à un accès équitable aux bénéfices, à une prospérité partagée et à l'élimination de toutes les formes de discrimination. Il vise à atténuer le risque de biais dans les ensembles de données, fortement influencés par les normes systémiques.

Principes connexes au concept de justice

Équité	Droits fondamentaux	Égalité	Non-discrimination
Éviter les biais	Inclusivité	Diversité	Neutralité des données
Données représentatives	Prospérité partagée	Impacts sociaux et économiques	Éviter les disparités
Atténuer la désagrégation du tissu social	Préserver la solidarité	Accessibilité	Explicabilité
Transparence (aux fins de redevabilité)	Ouverture (aux fins de) redevabilité	Redevabilité	Auditabilité
Responsabilité	Transparence judiciaire	Gouvernance ouverte	Conformité réglementaire et légale

⁴ The IEEE Initiative on Ethics of Autonomous and Intelligent Systems (2017). Ethically Aligned Design, v2. <https://ethicsinaction.ieee.org>

Niveaux de la législation d'IA

Règles relatives aux données

Les dispositions relatives aux données conditionnent l'accès à la matière première nécessaire au développement de produits d'IA et constituent ainsi le premier niveau de conformité pour système d'IA.

Elles régissent la manière dont les données peuvent être collectées, traitées et transférées. Les dispositions relatives aux données sont souvent liées à des cadres nationaux. Bien qu'il existe des exemples de législation supranationale sur les données, telles que le RGPD de l'UE, ces réglementations sont souvent complétées par des règles nationales supplémentaires en matière de protection des données.

Règles spécifiques aux applications d'IA

Les règles spécifiques aux applications d'IA représentent le deuxième niveau de conformité. Ces dispositions ciblent de manière explicite des applications ou des domaines spécifiques d'IA. Bien que ces règles soient encore rares, elles sont susceptibles d'augmenter, notamment en ce qui concerne les normes techniques des produits.

Ces règles peuvent aller d'une interdiction pure et simple de technologies spécifiques à des normes techniques détaillées régissant les systèmes d'IA.

Elles combinent des législations traditionnelles et des règles constitutionnelles assorties de sanctions, en plus de normes techniques nouvellement développées par des acteurs plus amorphes et non gouvernementaux, tels que les organisations de standardisation. Quel que soit leur origine, ces règles seraient toujours contraignantes, *de facto et de jure*.

Règles générales d'IA

Troisièmement, les règles générales d'IA couvrent les fonctions d'IA, telles que la prise de décision automatisée. A titre d'exemple, le RGPD de l'UE a prescrit une interdiction partielle de « la prise de décision basée uniquement sur le traitement automatisé de données personnelles. »

Le règlement européen considère que le traitement automatisé ne peut être effectué que si seulement la décision s'avère nécessaire à la conclusion ou à l'exécution d'un contrat, si elle est autorisée par la loi ou si la personne concernée a donné son consentement explicite.

Il y a fort à parier que les règles générales d'IA évolueront en termes de profondeur et d'étendue, car les futurs dispositifs de procédures judiciaires vont contraindre le système judiciaire à s'adapter et à mieux interagir avec les technologies de l'IA.

Règles spécifiques aux applications ne relevant pas de l'IA

Quatrièmement, ces règles s'appliquent à des applications spécifiques sans considérer l'IA en tant que telle. L'utilisation de l'IA ne dispense pas des réglementations appliquées à des secteurs spécifiques. Tant qu'aucune exemption spécifique n'est en place, les systèmes d'IA sont tenus de respecter la réglementation sectorielle spécifique. A titre d'illustration, les réglementations maritimes stipulent qu'un navire doit être dirigé par un capitaine et un équipage remplissant des fonctions humaines bien spécifiques. En tant que telle, cette disposition serait considérée comme un obstacle à l'automatisation complète d'un cargo de charge transportant des marchandises sur plusieurs continents.

Dispositions générales non liées à l'IA

Enfin, le 5^e niveau de conformité, les dispositions générales non liées à l'IA présentent un cadre juridique général pour le contrôle des comportements. Généralement, les lois contre la discrimination peuvent être classées comme réglementations générales non liées à l'IA, mais essentielles au développement de l'IA.

Par exemple, ces règles peuvent influencer la conception des algorithmes, le type de données utilisables, les paramètres décisionnels éligibles et le cadre des décisions applicables.

IA et secteur de la sécurité et de la justice

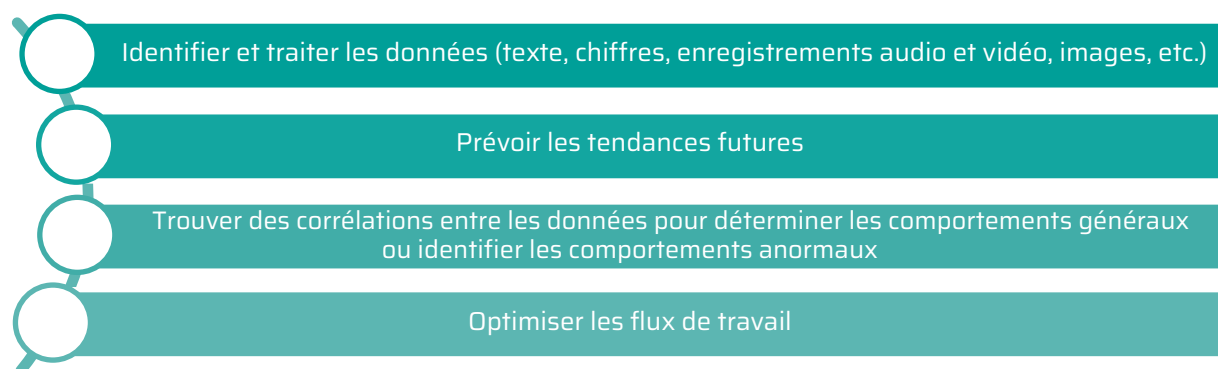
L'application de l'IA dans le secteur de la sécurité et de la justice en est dans ses premiers balbutiements ; son utilité ainsi que ses risques potentiels sont encore en cours d'évaluation. Le Parlement européen a résumé cela en 2021⁵ :

« les technologies numériques en général et, plus particulièrement, la prolifération du traitement et de l'analyse des données grâce à l'intelligence artificielle (IA) sont sources de grandes promesses et d'immenses risques; que de grands progrès ont été accomplis ces dernières années dans le développement de l'IA, en faisant de celle-ci l'une des technologies stratégiques du 21^e siècle; qu'elle est susceptible de générer des avantages substantiels en matière d'efficacité, de précision et de facilité, et ainsi d'apporter un changement positif à l'économie et à la société européennes, mais qu'elle fait également peser de lourds risques sur les droits fondamentaux et les démocraties fondées sur l'État de droit; que l'IA ne doit pas être considérée comme une fin en soi, mais comme un outil pour servir les personnes, dans le but ultime d'accroître le bien-être humain, les capacités humaines et la sécurité. »

Compte tenu du pouvoir considérable de l'IA, les risques et les opportunités véhiculés par son utilisation doivent être traités avec un surcroît de prudence. Dans de nombreux pays fragiles et touchés par des conflits, les institutions de sécurité et de justice ont du mal à s'acquitter de leurs tâches, y compris la couverture de l'intégralité de l'espace géographique ainsi que la réponse aux besoins de sécurité humaine de la population. Cela a été une source de préoccupation pour la paix et la sécurité globales.

Le manque de capacité à offrir la réponse appropriée aux défis de sécurité au moment opportun et avec manière la plus adéquate a sapé la légitimité de l'État et la crédibilité des systèmes de gouvernance en place, remettant ainsi en cause le contrat social dans ces pays.

L'IA pourrait offrir une nouvelle opportunité de réforme plus inclusive, efficace et efficiente. Les institutions de sécurité et de justice utilisent l'IA dans les quatre scénarios ci-dessous :



IA et secteur de la défense

Dans un contexte de défense, l'IA offre une série d'avantages :

- Une prise de décision opérationnelle plus rapide grâce à un accès rapide aux données et à leur capacité à identifier **les tendances sous-jacentes** qui ne sont pas facilement repérables par les humains, permettant ainsi une meilleure **évaluation des menaces et des réponses optimales**.
- Le gain du temps pour accéder et traiter les données grâce à l'IA permet de disposer d'une grande latitude d'action au niveau de la planification et la conduite des opérations.
- Identification des **facteurs de risque liés aux environnements** et aux conditions de travail et proposition de mesures de protection pour atténuer les impacts négatifs.
- Exécuter des **simulations** pour renforcer la capacité de formation du personnel de service.
- En association avec la robotique, aider le personnel de service à maintenir une distance de sécurité lorsqu'il est question de traiter avec des environnements contaminés.

⁵ European Parliament, "Texts Adopted - Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial Authorities in Criminal Matters - Wednesday, 6 October 2021," October 6, 2021, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html.

- Effectuer des **tâches auxiliaires et répétitives** qui sont chronophages pour le personnel. En allégeant le fardeau de ces charges, l'IA permet au personnel en service de consacrer plus de temps aux tâches à forte valeur ajoutée, telles que la réflexion stratégique et la prise de décision.
- Développer des **modèles de prévision** permettant d'appuyer l'optimisation des réseaux logistiques, la gestion et la maintenance des équipements, les engagements financiers et le recrutement.

IA et police

Dans le contexte des services de police, l'IA offre une panoplie d'applications bien établies :

- **La reconnaissance faciale** consiste à capturer des images du public, à effectuer une recherche dans une base de données préexistante de suspects, voire créer une nouvelle base de données.
Cette pratique est menée en direct ou à partir d'images enregistrées. Les forces de police ont également accès à la reconnaissance faciale rétrospective grâce aux bases de données des forces de l'ordre qui regorgent des millions d'images faciales.
- **Traitement d'images** permettant de lire sur les lèvres, d'analyser les styles d'écriture, d'identifier les voitures volées et de détecter les comportements des voleurs à l'étalage.
- **Analyse financière au moyen de l'IA** permet d'identifier des activités criminelles telles que la fraude et le blanchiment d'argent.
- **La police prédictive** est une pratique qui permet d'identifier les endroits ou les individus présentant un risque plus élevé de commettre une activité criminelle.
Les points chauds et les schémas de criminalité sont cartographiés dans une base de données historique. Cette analyse alimente les stratégies de prévention de la criminalité, par exemple, l'organisation des patrouilles de police et autres allocations de ressources.
- **Évaluation de la solvabilité** L'IA peut évaluer la probabilité qu'une affaire soit résolue. Dans ce cas, les algorithmes peuvent recommander une ventilation des ressources entre certaines affaires.

IA et secteur de la justice

Les moteurs de recherche avancés sur la jurisprudence utilisent l'IA pour faciliter la recherche d'informations relatives à des affaires de justice. L'IA permet la création **d'un système de recherche juridique intelligent** qui analyse les antécédents judiciaires et fournit une analyse statistique automatisée ainsi qu'un sommaire des concepts juridiques utilisés dans des affaires similaires. La technologie de l'IA peut également servir de **mécanisme alternatif de résolution des litiges**.

Généralement, les systèmes judiciaires reçoivent un nombre important de dossiers de litiges civils récurrents portant sur des dommages de très faible importance. L'IA permet aux justiciables de passer par une plateforme de diagnostic automatique du litige, ce qui aboutit à des propositions de solutions avant que l'affaire ne soit portée devant le tribunal.

La justice prédictive fait également l'objet de recherches dans le but de promouvoir des systèmes d'IA capables de prédire la teneur des décisions de justice. À titre d'exemple, ces systèmes peuvent déterminer si une personne pourrait prétendre au régime de mise en liberté conditionnelle, en déterminant si elle est susceptible de commettre d'autres délits.

Etudes de cas

Accès à la justice en Colombie

Développé par l'Université de Buenos Aires-IALAB en collaboration avec la Cour constitutionnelle colombienne et l'appui des donateurs internationaux, le programme PretorlA a été lancé en Colombie à la mi-2020, dans la perspective de promouvoir le rendement de la justice.

Ce programme permet aux citoyens de bénéficier d'une protection immédiate contre les violations des droits fondamentaux grâce au projet *Acción de Tutela* (Action constitutionnelle pour la protection des droits fondamentaux).

Le projet *Acción de Tutela* établit des précédents judiciaires en matière d'exercice des droits fondamentaux fondés sur des cas prioritaires- *key tutelas*- dont il se saisit chaque jour.

Avec plus de 3000 *tutelas* reçues chaque jour- et compte tenu du temps moyen nécessaire pour qu'une personne puisse lire, analyser et systématiser le contenu (36 minutes), il s'avère qu'il est impossible d'examiner la totalité de ces recours.

En procédant à la digitalisation des analyses nécessaires et en fournissant des prévisions de condamnation basées sur des critères prédéterminés, le programme PretorlA permet une prise de décision plus efficace et plus rapide.

L'application d'IA génère des rapports et des statistiques intuitifs qui accélèrent la procédure, permettant ainsi un déploiement plus rapide de cette protection juridique en faveur des citoyens colombiens.

L'intégration de l'IA dans l'accès à la justice en Colombie a été réalisée en partenariat avec des organisations de la société civile, ce qui a suscité des inquiétudes concernant la première version du programme PretorlA.

La Cour constitutionnelle a dû alors modifier le projet et adopter des technologies plus transparentes. Néanmoins, les défis persistants liés à la confidentialité de l'information commandent que le système d'IA adapte en permanence ses dispositions afin de refléter au mieux l'évolution du cadre juridique colombien.

Justiça para todos in Brazil (Justice pour tous au Brésil)

En 2020, le PNUD Brésil a lancé le projet *Justiça para todos* en collaboration avec le Conselho Nacional de Justiça (CNJ), une institution qui œuvre à améliorer l'action de la justice brésilienne en termes de responsabilité, d'administration et d'efficacité des processus.

Dans le cadre de l'objectif plus large de promouvoir l'innovation pour améliorer l'accès à la justice, le projet *Justiça 4.0* comprend un volet important sur l'IA. Le volume de travail des juges au Brésil étant des plus élevés au monde.

Encore dans sa phase initiale d'expérimentation, le projet a commencé par dresser une évaluation des besoins du secteur de la justice au Brésil, en matière de renforcement de l'accès à la justice par la technologie. A long terme, il envisage de développer des solutions d'IA destinées à améliorer cet accès à la justice.

La capacité des systèmes d'intelligence artificielle à analyser et à traiter de grandes quantités de données peut contribuer à identifier des modèles et des tendances, ce qui serait utile pour soutenir les juges dans le suivi des procédures judiciaires. Le soutien apporté par l'IA aux juges leur permettrait de traiter les affaires plus efficacement, ce qui, à son tour, contribuera à augmenter leur capacité à traiter les affaires.

Grâce à la capacité des juges de traiter les affaires plus rapidement, l'encombrement dans les tribunaux devrait être réduit. Cela pourrait améliorer l'accès des citoyens à la justice, dans la mesure où les retards dans le traitement des affaires seraient nécessairement réduits. En outre, en favorisant l'efficacité, le projet peut améliorer la perception de la population à l'égard des institutions judiciaires.

Au fur et à mesure que l'institution judiciaire acquiert une plus grande légitimité, la méfiance de la population serait un obstacle mineur à l'accès à la justice. L'amélioration de la prestation et de l'accès à la justice sont des éléments centraux dans les programmes de la G/RSS et l'IA a mis en évidence des perspectives prometteuses dans ce domaine.

Justiça 4.0 est un projet ambitieux visant à surmonter les obstacles structurels qui entravent l'efficacité de la justice au Brésil, en adoptant une approche globale du secteur. L'impact du projet vise à améliorer la transparence, l'accès et l'efficacité de la justice, tout comme il œuvre à renforcer les capacités du personnel de justice.

Tels qu'identifiés par l'équipe du projet à ce stade précoce, les principaux obstacles liés à la mise en œuvre dudit projet sont le manque de coordination entre les parties prenantes concernées. L'adoption d'une approche globale du secteur implique de mobiliser des acteurs aux niveaux fédéral et infranational qui ne sont pas souvent habitués à collaborer ensemble.

C'est ainsi que la dotation d'une plateforme avec des processus pertinents permettra à ces parties prenantes d'exprimer leurs besoins et leur vision du projet. La mise en place de cette plateforme faciliterait une coopération plus étroite dans le cadre des activités interinstitutionnelles et accélérerait par conséquent leur mise en œuvre.

IA et gouvernance

Les services électroniques

La pandémie de COVID-19 a mené plusieurs autorités à élargir l'accès numérique aux services. Les services électroniques ou le cyber-gouvernement offrent de nombreuses possibilités d'amélioration des services de sécurité et de justice, notamment, une meilleure accessibilité aux services et une responsabilité sociale améliorée.

Au moyen d'une technologie appropriée, les dossiers numériques ont le potentiel d'accroître la transparence des activités et des services du secteur de la sécurité et de la justice, ce qui devra contribuer à démystifier l'espace de sécurité et de justice aussi bien pour les individus que pour les organisations de la société civile.

Transparence améliorée des pratiques de passation de marchés en Ukraine

Soutenu par un collectif d'acteurs de la société civile, Dozorro est un projet citoyen de technologie à but non lucratif, conçu pour détecter et prévenir l'utilisation abusive des fonds publics dans les marchés publics.

La plateforme Dozorro utilise un outil d'évaluation des risques alimenté par l'IA pour analyser la base des données générée par les activités de passation de marchés publics. Sur la base d'indicateurs de risque, le programme est capable de signaler les appels d'offres publics présentant une forte probabilité de violation.

L'appel d'offres identifié est ensuite soumis à un examen plus approfondi auprès d'un réseau de 25 OSC représentant la communauté Dozorro. Le système a été en mesure d'identifier plus de 20 000 violations et de collaborer avec plus de 900 000 utilisateurs uniques.

Dozorro facilite non seulement l'engagement de la société civile dans le contrôle externe des marchés publics mais favorise également une culture de coopération entre les différentes parties prenantes de la société civile.

Ce réseau d'OSC a développé des modes de coopération reproductibles qui devraient permettre à cette communauté de se positionner en tant qu'acteur de dialogue national sur les réformes.

La taille et le poids politique sans cesse grandissants de la communauté Dozorro devraient se traduire par un plus grand pouvoir d'action, qui, à son tour, agirait comme agent dissuasif pour quiconque tenterait de détourner des fonds publics.

Afin de créer la plateforme Dozorro, 20 experts ont évalué un ensemble de données afférents de 3500 appels d'offres pour déterminer si une offre comportait un risque, et les données obtenues ont été par la suite utilisées pour former le système d'IA. Le système d'IA évalue de manière indépendante la probabilité des risques de corruption et signale les risques potentiels aux experts.

Une fois l'expert a examiné le cas et déterminé s'il existe un risque, la réponse de l'expert, qu'elle soit positive ou négative, est enregistrée dans le système, ce qui permet à l'algorithme de continuer à apprendre et d'améliorer son exactitude.

Les premières tentatives d'automatisation du processus de vérification des appels d'offres étaient basées sur 35 indicateurs fixes tels que les signatures ou les documents manquants. Toutefois, ces indicateurs sont devenus prévisibles et donc vulnérables à ceux qui cherchent à contourner le système, à repérer des lacunes ou à éviter la détection. Le développement de Dozorro, basé sur un ensemble d'indicateurs flexibles, a considérablement réduit cette vulnérabilité.

La reproduction de ce modèle en Moldavie et en Pologne fait l'objet de discussions au sein des organisations de la société civile.

Vote électronique

En temps de la pandémie de COVID-19, de nombreuses élections ont dû être reportées ou annulées. Le vote par correspondance a souvent été préconisé comme solution alternative au vote physique.

Ce processus demeure cependant un défi logistique majeur impliquant des coûts significatifs.

Les technologies de vote électronique sont devenues donc une alternative intéressante. Les défis du vote électronique sont multiples et peuvent aller du manque de compétences techniques adéquates ; de la méfiance des électeurs envers la technologie ; des risques d'attaques et de virus informatiques jusqu'aux risques liés au vote multiple et à la complexité de l'identification correcte des électeurs.

Néanmoins, le potentiel du vote électronique devrait croître. L'écart générationnel en matière de compétences techniques se rétrécit au fur et à mesure que l'utilisation d'Internet se développe. Autant nous accumulons de l'expérience, autant la confiance en la technologie évoluera crescendo. Des solutions comme les identifiants électroniques de vote, ont déjà été utilisées avec succès.

En définitive, le vote électronique représente une opportunité précieuse pour garantir le fonctionnement des institutions publiques en cas de conjoncture de crise nationale ou mondiale.

Le vote électronique au Burkina Faso

L'agence satellitaire, SES, a apporté son soutien au Burkina Faso lors des élections de 2015 en installant 368 VSAT (Very Small Aperture Terminals) déployés dans l'ensemble du territoire du pays. Cela a permis de rendre les élections électroniques, transparentes et diffusées en direct à la télévision.

Cette initiative a connu un grand succès, et l'année précédente, ENABEL et SES ont co-signé un cadre pluriannuel prévoyant la fourniture d'une infrastructure et des services similaires par satellite dans 20 pays africains⁶.

IA et budgets nationaux

La budgétisation électronique fait référence à la digitalisation des procédures ou des services budgétaires afin d'améliorer la gouvernance financière des acteurs du secteur de la sécurité.

La budgétisation peut être optimisée grâce à l'utilisation des TIC en digitalisant les procédures budgétaires et en diffusant les informations budgétaires au large public.

La budgétisation électronique est un outil puissant, en tant que mesure efficace de réduction des coûts, mais également un pas vers une meilleure transparence financière. L'automatisation de certaines parties des processus d'approvisionnement et de budgétisation devrait, en principe, permettre de gagner du temps sur les tâches auxiliaires, d'optimiser les processus et d'offrir des garanties technologiques contre la corruption.

IA et redevabilité

L'utilisation de l'IA et de ses solutions technologiques offre de précieux avantages pour prévoir et prévenir les crimes, la violence et les conflits. Toutefois, le mécanisme de surveillance joue un rôle-clé pour assurer une utilisation redevable et équitable de ces technologies.

Redevabilité technologique

Pour qu'elle soit techniquement redevable, la conception de l'IA doit garantir l'intégrité du système, l'efficacité des tâches, la transparence et l'interopérabilité avec les autres systèmes. La technologie ne doit fonctionner que dans le cadre de son domaine de surveillance défini par l'homme et ne pas dépasser les capacités des autorités compétentes à surveiller et évaluer son fonctionnement.

Plus concrètement, les données utilisées aux fins de construire des outils d'IA doivent provenir de sources certifiées, tout comme le stockage et l'exécution des algorithmes doivent se dérouler dans des environnements sécurisés afin de garantir l'intégrité et l'intangibilité du système.

⁶ <https://www.ses.com/press-release/belgian-development-agency-partners-ses-connect-foreign-aid-projects-africa-satellite>

La transparence technique peut être mise en œuvre, par exemple, au moyen d'un code et d'une documentation à source ouverte, avec des explications détaillées et accessibles au public sur les systèmes d'IA, y compris le fonctionnement des services et les marges d'erreur.

Redevabilité juridique

La surveillance juridique des solutions d'IA peut être un processus très complexe associant des acteurs publics et privés impliqués dans les technologies numériques, la cybernétique et l'IA et dans la prestation de services publics. Il est essentiel donc de veiller à ce que les systèmes d'IA respectent l'Etat du droit pour préserver la confiance du public.

Étant donné que les systèmes d'IA traitent des données sensibles et personnelles, ils doivent être pleinement conformes aux exigences légales en matière de protection des données et de traitement des informations personnelles. L'IA pose des défis aux cadres juridiques dans la mesure où elle peut conduire à une « confusion d'attribution » puisque l'IA n'est ni une personne morale ni une personne physique.

Par conséquent, la redevabilité devrait être fondée sur un processus de prise de décision pour s'assurer qu'une personne identifiée demeure responsable d'une certaine ligne de conduite. Lorsqu'un système d'IA est fourni par un tiers, les contrats devraient définir clairement quelle organisation serait tenue responsable et dans quelles circonstances.

Redevabilité politique

Tandis que la surveillance technologique et juridique devrait contrôler les performances des applications d'IA, la surveillance politique s'avère nécessaire pour s'assurer que la « bonne » technologie a été utilisée dans le « bon » usage, en utilisant les données « nécessaires » à travers un processus décisionnel adéquat, et de manière transparente, impartiale et équitable.

Les processus décisionnels de l'IA devraient rester explicables et traçables pour la prise de décision humaine, laquelle doit être autonomisée plutôt que limitée par l'utilisation d'outils et de services d'IA.

Ainsi, les autorités publiques devraient délivrer une certification sur l'utilisation des outils d'IA tandis que les autorités indépendantes devraient pouvoir auditer les méthodes de traitement.

Une commission spécialisée pourrait également surveiller et évaluer régulièrement l'utilisation de l'IA dans le secteur de la sécurité en organisant des audits planifiés et aléatoires. Une pareille commission pourrait être composée de membres issus de divers domaines pertinents pour les G/RSS et l'IA, tels que les télécommunications, l'éducation et les droits humains.

Redevabilité éthique

Les garanties éthiques peuvent reposer sur plusieurs fondements, dans la mesure où elles ont des applications différentes dans les secteurs de la police, de la défense et de la justice.

Les systèmes d'IA dans le secteur de la sécurité pourraient mettre en exergue ce que l'on appelle « Approche éthique dès la conception », qui signifie que les données traitées sont examinées par un comité consultatif d'éthique, et que les données et les algorithmes sont rendus facilement compréhensibles pour les personnes.

Un comité d'éthique ministériel pourrait superviser le développement de la technologie de l'IA dans le secteur de la sécurité et œuvrer en étroite collaboration avec un comité consultatif national d'éthique.

Une autre façon d'intégrer des considérations éthiques serait à travers ce qu'on appelle « Approche des droits humains dès la conception », laquelle est une façon d'orienter la Déclaration des droits de l'homme de l'ONU vers l'univers numérique. Cette approche est centrée sur l'utilisateur et se focalise sur son consentement et lui accorde à ce titre de nombreuses options et alternatives.

Dans le secteur de la défense, il est d'une importance majeure que le développement de l'IA respecte les principes du droit international humanitaire régissant les conflits (nécessité, humanité, proportionnalité, distinction). L'adoption de systèmes d'IA entièrement autonomes dans le secteur de la défense peut violer les principes du droit international humanitaire.

S'agissant des secteurs de la police et de la justice, les principes d'équité et de non-discrimination sont essentiels dans le développement des systèmes d'IA.

Où irons-nous ?

Le croisement entre la G/RSS et le secteur de la sécurité est multiforme, d'autant que des praticiens et des chercheurs ont commencé à faire le point sur ce qui s'est déjà produit et où les choses pourront aller. L'IA n'a de cesse de métamorphoser le paysage de la sécurité, les capacités des institutions de sécurité, contribuant ainsi à enrichir la boîte à outils des acteurs de la réforme.

De la reconnaissance faciale dans les zones de conflit à la détection des risques de corruption, l'IA s'impose en tant qu'outil puissant dont se servent les acteurs étatiques et non étatiques, suscitant des interrogations délicates sur les bénéfices, l'autonomie, la confidentialité et le respect des droits humains.

La nature de l'IA, qui permet un apprentissage autonome à partir d'énormes ensembles de données, laisse à dire sans nul doute que les secteurs de la sécurité et de la justice dans le monde entier vont massivement adopter cette technologie.

La cadence de cette adoption et la vitesse des progrès de l'IA rendent la tâche de sa bonne gouvernance une œuvre ardue à entreprendre. En effet, les cadres de gouvernance ont déjà du mal à rattraper leur retard.

Même si la vitesse du changement peut sembler vertigineuse, ce qui est à retenir, du point de vue de la gouvernance de la sécurité, ce sont les principes essentiels de la sécurité centrée sur les individus, la redevabilité et les droits de l'homme.

Les questions à poser au sujet de l'application des technologies de l'IA sont donc les suivantes : *Comment la technologie peut-elle rendre la sécurité plus centrée sur les personnes ? Comment cette technologie peut-elle améliorer la redevabilité ainsi que l'efficacité ? Comment cette technologie peut-elle être utilisée pour renforcer les normes relatives aux droits humains ?* C'est l'essence même de l'optique G/RSS appliquée à l'IA.