L'INTELLIGENCE ARTIFICIELLE (IA) ET LE SECTEUR DE LA DÉFENSE

À PROPOS DE CE DOCUMENT D'INFORMATION SUR LA RSS

Ce document d'information sur la réforme du secteur de la sécurité (RSS) offre une vue d'ensemble de l'intersection entre l'intelligence artificielle (IA) et le secteur de la défense. Utilisée dans de nombreux domaines, dont celui de la sécurité nationale, l'IA peut améliorer l'efficacité des activités du secteur de la défense. Mais elle soulève aussi des défis importants en matière de droits de l'homme et de bonne gouvernance. Le document définit l'IA, présente ses applications dans la défense, en analyse les risques pour la bonne gouvernance, et propose des pistes pour renforcer le contrôle, notamment par un suivi rigoureux, une transparence accrue, des mécanismes de responsabilisation et une collaboration entre parties prenantes.

CE DOCUMENT D'INFORMATION RÉPOND AUX QUESTIONS SUIVANTES :

À PROPOS DE CETTE SÉRIE

Les documents d'information sur la RSS fournissent une introduction concise à certaines questions liées à la bonne gouvernance du secteur de la sécurité (GSS) et à la réforme du secteur de la sécurité (RSS). Cette série résume les débats actuels, définit les termes clés et révèle les tensions centrales dans ces domaines en s'appuyant sur un large éventail d'expériences internationales. Les documents d'information sur la RSS ne cherchent pas à promouvoir des modèles, politiques ou propositions spécifiques en matière de gouvernance ou de réforme, mais proposent une liste de références additionnelles offrant aux personnes intéressées la possibilité d'approfondir leurs connaissances sur chaque sujet. Ils constituent des ressources utiles pour les acteurs de la gouvernance et de la réforme du secteur de la sécurité qui cherchent à comprendre et à appréhender de façon critique les approches actuelles en la matière.

DCAF, le Centre pour la gouvernance du secteur de la sécurité, Genève se consacre à l'amélioration de la sécurité des États et de leurs citoyens dans un cadre de gouvernance démocratique, d'état de droit, de respect des droits de l'homme et d'égalité des genres. Depuis sa création en 2000, le DCAF contribue à rendre la paix et le développement plus durables en aidant les États partenaires et les acteurs internationaux qui soutiennent ces États à améliorer la gouvernance de leur secteur de la sécurité grâce à des réformes inclusives et participatives. Il crée des produits de connaissances innovants, encourage les normes et les bonnes pratiques, fournit des conseils juridiques et politiques et soutient le renforcement des capacités des acteurs étatiques et non étatiques du secteur de la sécurité.

Le DCAF tient à remercier
Alexandru Lazar et Dawn Lui pour la recherche, la conceptualisation et la rédaction de ce document ; Gabriela Manea, Franziska Klopfer, Teodora Fuior, Alexander Walsh et Ken Isaac pour leurs commentaires et révisions ; Alec Crutchley pour la révision linguistique ; Sandra Vojvodic pour la traduction en français ; Ioan Nicolau pour l'édition en français ; Petra Gurtner pour la mise en page et la conception.

Rédactrice de la série
Gabriela Manea

Pour citer cette publication
DCAF – Centre de Genève pour la gouvernance du secteur de la sécurité. Intelligence artificielle (IA) et secteur de la défense.
Série de documents d'information sur la RSS.
Genève : DCAF, 2025.

DCAF – Centre pour la gouvernance du secteur de la sécurité, Genève
Maison de la Paix
Chemin Eugène-Rigot 2E
CH-1202 Geneva
Switzerland

info@dcaf.ch
+41 22 730 94 00

www.dcaf.ch

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE (IA) ?

L'intelligence artificielle (IA) est une technologie à usage général qui a le potentiel d'améliorer le bien-être des populations et de contribuer à répondre à des défis sur une échelle globale. Comme illustré à la Figure 2, l'IA peut être définie comme *un système basé sur une machine* qui déduit, à partir des données d'entrée reçues, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions pouvant influencer des environnements physiques ou virtuels.

Le **terme** *système basé sur une machine* désigne souvent un système utilisant des algorithmes, des logiciels et du matériel informatique pour réaliser des tâches telles que le traitement des données, la prise de décisions ou le contrôle d'autres machines ou processus. Un algorithme est un ensemble d'instructions ou de règles étape par étape, conçu pour effectuer une tâche spécifique ou résoudre un problème particulier. Ces instructions peuvent aller de simples consignes à des séquences complexes d'opérations.

Un 'système d'IA' est une intégration approfondie de modèles d'IA, d'agents, de données et d'infrastructures, conçu afin d'exécuter des tâches spécifiques de manière autonome ou semi-autonome, souvent au sein d'un environnement opérationnel plus large. Dans ce contexte, il est important de distinguer un modèle d'IA d'un agent d'IA, car ces deux termes seront utilisés dans différentes parties de ce document d'information sur la RSS. Un modèle d'IA est une représentation mathématique ou statistique des motifs dans des données, entraîné pour effectuer une tâche spécifique. Il reçoit des données d'entrée, les traite à l'aide de paramètres appris, et produit une sortie (par exemple, ChatGPT). Un 'agent d'IA' interagit avec un environnement pour atteindre des objectifs de manière autonome. Il perçoit, décide et agit, utilisant souvent des modèles d'IA comme composants dans un cadre plus large.

Par exemple, un système d'IA peut aller d'un chatbot automatisé à un véhicule autonome avancé. Ces systèmes s'appuient sur des *algorithmes d'apprentissage automatique,* des capteurs et des processeurs pour interpréter les données et prendre des décisions en temps réel. Au lieu d'être uniquement programmés pour exécuter une tâche, les *systèmes d'apprentissage automatique* sont entraînés sur de grandes quantités de données, ce qui leur permet d'identifier des modèles, d'améliorer leurs performances au fil du temps et de s'adapter à de nouvelles informations sans intervention humaine.

L'IA évolue rapidement et contribue à de nombreux progrès économiques et sociétaux. En améliorant la prédiction, en optimisant les opérations et l'allocation des ressources, et en personnalisant les solutions numériques offertes aux individus et aux organisations, l'utilisation de l'IA peut offrir des avantages compétitifs clés. Aujourd'hui, l'IA est déployée dans de nombreux secteurs, allant de l'éducation, la finance et les transports à la santé et la sécurité nationale. En matière de sécurité nationale, le potentiel transformateur de l'IA dans le secteur de la défense suscite une attention importante de la part des décideurs politiques, des stratèges militaires et des chercheurs. Cependant, en parallèle à ses avantages potentiels, l'IA est de plus en plus perçue comme une arme à double tranchant, créant des conditions susceptibles de perpétuer les inégalités sociales, d'éroder les droits de l'homme, de fragiliser la démocratie et la bonne gouvernance, et de causer des dommages. Ces dommages peuvent être tangibles ou intangibles, incluant des préjudices physiques, psychologiques, sociétaux ou environnementaux.

Gouverner l'IA est l'un des plus grands défis auxquels la communauté internationale est confrontée. Alors que le débat sur la gouvernance autour de la réglementation de l'IA est passé de la question de savoir ***si*** elle devait être réglementée à celle de savoir ***comment*** elle devait l'être, le cadre exact de gouvernance reste ouvert au moment de la rédaction de ce document d'information sur la RSS, puisqu'il n'existe pas encore de *cadre réglementaire universellement accepté* pour l'IA dans le secteur de la défense. Réglementer l'IA dans le domaine militaire présente des défis spécifiques par rapport à la gouvernance civile de l'IA. Par exemple, l'application de l'IA dans un contexte militaire implique souvent des informations classifiées et des actifs stratégiques, nécessitant des mesures de sécurité strictes. L'utilisation de l'IA dans les armes autonomes et la prise de décision militaire soulève des préoccupations éthiques uniques, telles que le risque de dommages involontaires résultant d'actions militaires pilotées par l'IA. Dans le même temps, les niveaux croissants de volatilité dans le système mondial poussent les établissements de défense du monde entier à chercher à sécuriser un avantage stratégique dans ce domaine, rendant d'autant plus complexe le défi global de réglementer les systèmes militaires d'IA.

Cependant, des principes tels que « l'IA responsable » ou « l'IA centrée sur l'humain » ont pris de l'ampleur dans le cadre du débat au sens large sur l'IA, et semblent offrir une base solide pour façonner les futurs cadres de gouvernance. L'IA responsable fait référence au développement, au déploiement et à l'utilisation éthiques et responsables des systèmes d'IA. Elle englobe des principes, des lignes directrices et des pratiques visant à garantir que les technologies d'IA sont conçues et mises en œuvre de manière à privilégier l'équité, la transparence, la responsabilité, la confidentialité et le bien-être sociétal. L'approche centrée sur l'humain cherche à assurer que les valeurs humaines soient au cœur de la manière dont les systèmes d'IA sont développés, utilisés et surveillés. Elle implique le respect des droits fondamentaux, de la dignité humaine, ainsi que la prise en compte de l'environnement naturel et des autres êtres vivants en tant que partie intégrante de l'écosystème humain.

Ce document d'information sur la RSS sert d'introduction à l'application de l'IA dans le secteur de la défense. Son objectif est de présenter les bénéfices potentiels et les risques que les systèmes d'IA introduisent pour la bonne gouvernance du secteur de la défense, ainsi que de fournir des suggestions pour des pratiques de contrôle rigoureuses concernant l'utilisation de l'IA par les acteurs de la défense au sens large.

FIGURE 1 CLARIFICATIONS CONCEPTUELLES

Intelligence artificielle (IA) : L'IA est le domaine de l'informatique qui permet aux ordinateurs d'exécuter des tâches qui nécessitent typiquement l'intelligence humaine, telles que le raisonnement, l'apprentissage, la résolution de problèmes, la perception et la compréhension du langage. Elle comprend l'IA au sens étroit (spécifique à une tâche) et l'IA générale (intelligence comparable à celle de l'humain).

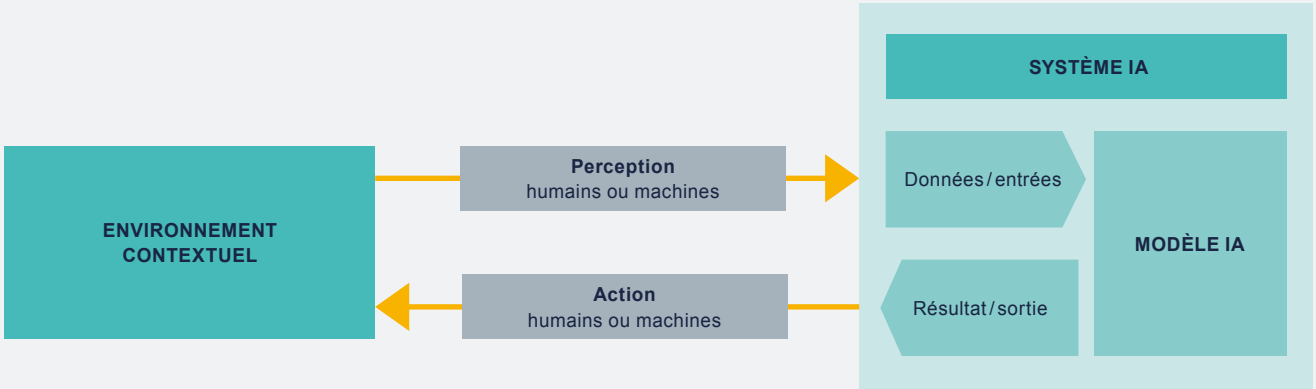
Apprentissage automatique (ML) : Une sous-catégorie de l'IA dans laquelle les ordinateurs apprennent à partir des données pour faire des prédictions sans programmation explicite. Le ML comprend l'apprentissage supervisé (données étiquetées), l'apprentissage non supervisé (découverte de modèles) et l'apprentissage par renforcement (apprentissage par essais et erreurs).

Apprentissage profond et réseaux neuronaux : L'apprentissage profond, une sous-catégorie du ML, utilise des réseaux neuronaux artificiels (RNA) à plusieurs couches pour analyser des motifs complexes dans les données. Les réseaux neuronaux, inspirés par le cerveau humain, traitent les données d'entrée via des nœuds interconnectés afin de faire des prédictions, alimentant des applications telles que la reconnaissance d'images et le traitement du langage naturel.

Traitement du langage naturel (NLP) : Le NLP est une branche de l'IA qui permet aux ordinateurs de comprendre, interpréter et générer le langage humain. Il est utilisé dans des applications telles que les chatbots, la reconnaissance vocale, la traduction et l'analyse des sentiments. Les modèles linguistiques à grande échelle, comme GPT (Generative Pre-trained Transformer), exploitent des techniques d'apprentissage profond pour améliorer les capacités du NLP.

Grands modèles de langage (LLM) et IA générative : Les LLM traitent d'immenses volumes de données textuelles pour comprendre et générer un langage proche de celui des humains. L'IA générative (GenAI) va plus loin en créant du contenu original – texte, images, audio ou vidéo – sur la base de modèles appris à partir d'ensembles vastes de données.

FIGURE 2 FONCTIONNEMENT D'UN SYSTÈME D'IA



Source: OECD.AI Policy Observatory

Pour en savoir plus sur les développements technologiques dans le secteur de la sécurité, veuillez consulter le compte-rendu **Digitalisation et GSS/R : projections vers l'avenir,** ainsi que les documents d'information sur la RSS intitulés **La digitalisation et la gouvernance et la réforme du secteur de la sécurité** et **Le contrôle du renseignement à l'ère de la numérisation.**

FIGURE 3 ACTEURS DU SECTEUR DE LA DÉFENSE

ACTEUR	RÔLE
 **Forces armées**	• Mener des opérations militaires • Mettre en œuvre des stratégies et tactiques pour défendre l'État contre les menaces extérieures et assurer la sécurité nationale • Exécuter les ordres des autorités exécutives de contrôle
 **Autorités exécutives de contrôle**	• Par exemple, le président ou le premier ministre, le ministre de la défense, le conseil national de sécurité, les chefs d'état-major, les ministères de la défense et autres agences gouvernementales concernées • Définir la politique et les objectifs globaux de la défense • Allouer des ressources pour les activités de défense • Prendre des décisions concernant les stratégies militaires et les déploiements
 **Prestataires nationaux de services de sécurité**	• Par exemple, les autorités chargées de l'application de la loi (comme les forces de police), les services de renseignement, les garde-frontières, les services d'urgence et les sociétés privées de sécurité • Coopérer avec les forces de défense pour collecter et partager des informations relatives aux menaces à la sécurité nationale • Fournir des informations et des analyses aux décideurs pour appuyer la politique et les opérations de défense
 **Comités de contrôle internes et externes**	• Par exemple, les commissions parlementaires de la défense, les organismes gouvernementaux de contrôle, les unités d'audit interne, les institutions de médiation, les commissions indépendantes et les organisations internationales • Assurer un contrôle et un examen indépendants des activités et des dépenses de défense • Vérifier le respect des lois, règlements et normes éthiques au sein du secteur de la défense • Identifier les inefficacités, les risques et les domaines d'amélioration dans les opérations de défense
 **Prestataires commerciaux de services de défense**	• Par exemple, les contractants militaires privés, les entreprises de technologie de défense, les fabricants d'armement, les sociétés de cybersécurité, ainsi que les entreprises de logistique ou d'infrastructures • Développer et fournir le matériel et la technologie nécessaires • Offrir des services spécialisés, tels que la logistique, le renseignement, la cybersécurité et la maintenance • Fournir des solutions et services flexibles pour répondre aux menaces émergentes, aux situations de crise ou aux engagements militaires
 **Organisations de la société civile**	• Par exemple, les organisations non gouvernementales (ONG), les groupes de réflexion, les groupes de plaidoyer, les institutions académiques, les organisations de défense des droits de l'homme et les médias • Les médias militent pour la transparence, la responsabilité et des pratiques éthiques dans les politiques et opérations de défense nationale • Les groupes de réflexion mènent des recherches et encouragent la participation citoyenne éclairée aux questions de défense nationale • Les ONG surveillent et rapportent les éventuelles violations des droits de l'homme

FIGURE 4 AVANTAGES DE L'IA POUR LE SECTEUR DE LA DÉFENSE

AUTOMATISATION DES TÂCHES ROUTINIÈRES	CONNAISSANCE DE LA SITUATION	COMMANDE ET CONTRÔLE
• L'IA peut alléger les charges administratives en automatisant les processus routiniers tels que les tâches financières. • L'IA peut améliorer la précision des prévisions budgétaires et détecter efficacement les anomalies dans les données financières. • En automatisant ces tâches, les agences de défense peuvent libérer des ressources humaines pour des initiatives plus complexes et stratégiques.	• L'IA peut analyser d'immenses sources de données, fournissant aux agences de défense des informations en temps réel sur l'évolution des menaces. • Les systèmes d'IA peuvent rapidement identifier des menaces potentielles qui pourraient échapper aux opérateurs humains. • Une meilleure connaissance de la situation peut minimiser les risques pour le personnel de défense.	• L'IA peut analyser rapidement d'immenses flux de données, offrant aux commandants des informations en temps réel. • Cela permet des réponses agiles et éclairées aux menaces et améliore l'efficacité opérationnelle. • L'IA dans le commandement et le contrôle peut renforcer la supériorité et la préparation des forces de défense dans des environnements complexes.

QU'EST-CE QUE LE SECTEUR DE LA DÉFENSE ?

Le secteur de la défense est le domaine de l'administration publique responsable de la puissance militaire. Il fait partie du secteur plus large de la sécurité nationale, et comprend les forces armées, leurs autorités exécutives de contrôle, d'autres agences étatiques impliquées de manière permanente ou occasionnelle dans les affaires de défense, les prestataires commerciaux de services de défense, et les organisations de la société civile impliquées dans la supervision des activités de défense. Pour plus d'informations sur le secteur de la défense, veuillez consulter les documents d'information sur la réforme de la défense et sur les forces armées.

Le rôle du secteur de la défense est double :

1. **Conduite opérationnelle** (c'est-à-dire le commandement et l'exécution des opérations militaires).

2. **Logistique, gestion des ressources et supervision** (c'est-à-dire la gestion financière, l'acquisition de matériel militaire, le recrutement, la formation et l'administration du personnel, ainsi que le contrôle des opérations militaires).

POURQUOI L'IA EST-ELLE UTILISÉE PAR LE SECTEUR DE LA DÉFENSE ?

Le secteur de la défense est à la pointe de l'innovation et de la technologie. Les organisations de défense comptent parmi les plus grands consommateurs et producteurs de données, car elles détiennent déjà une quantité disproportionnée d'informations par rapport à d'autres, en raison de la nature de leur travail. Compte tenu du nombre d'actifs impliqués — y compris les personnes, le matériel, l'information et la technologie — ainsi que du fait que les ministères de la défense sont de plus en plus sollicités pour faire davantage avec des budgets réduits, les systèmes d'IA sont perçus comme des solutions attrayantes. L'utilisation de l'IA par le secteur de la défense est principalement motivée par son potentiel à offrir un avantage stratégique en transformant les capacités militaires existantes. Plus précisément, les technologies d'IA promettent une meilleure connaissance de la situation et une prise de décision améliorée pour les autorités exécutives tout en minimisant les risques pour le personnel de défense.

Voici quelques exemples de la manière dont l'IA est actuellement utilisée, ainsi que pourrait être utilisée, dans le secteur de la défense :

- **Systèmes de guerre :** Intégrés dans des armes autonomes, des drones et des munitions intelligentes, pour n'en citer que quelques-uns. Par exemple, un agent d'IA peut coordonner plusieurs drones pour agir de manière collaborative lors d'opérations de reconnaissance, de combat ou de soutien logistique.
- **Cybersécurité :** Renforcer les défenses cybernétiques en détectant, prédisant et contrant des menaces informatiques dans des réseaux et infrastructures militaires.
- **Simulation et formation :** Fournir des environnements d'entraînement réalistes pour les soldats, améliorant la préparation et les compétences tactiques dans divers scénarios.

La figure 4 illustre comment les solutions basées sur l'IA peuvent améliorer les connaissances de situation et la planification, optimiser l'allocation des ressources et rationaliser les processus opérationnels pour divers acteurs de la défense, dans le but d'accroître l'efficacité et la performance des organisations de défense dans la protection des intérêts de la sécurité nationale.

Par conséquent, l'utilisation de l'IA dans le secteur de la défense offre la possibilité d'éliminer des activités superflues et peu productives qui finissent par détourner l'attention de la mission principale, tout en aidant les différents acteurs à rationaliser et centraliser les processus bureaucratiques. Néanmoins, le manque de transparence de nombreux systèmes d'IA et de la technologie qui les sous-tend soulève d'importantes préoccupations éthiques, en particulier si ces systèmes contrôlent des équipements militaires létaux valant plusieurs millions de dollars.

FIGURE 5 UTILISATIONS DE L'IA DANS LE SECTEUR DE LA DÉFENSE

ACTEUR	DÉTECTION	PLANNING	OPÉRATIONS	LOGISTIQUE
 **Forces armées**	Collection et analyse de données pertinentes provenant de diverses sources afin d'identifier les mouvements ennemis, les menaces et les anomalies.	Soutien à la planification stratégique et tactique par l'analyse de grandes quantités de données, en anticipant les scénarios futurs et en optimisant l'allocation des ressources.	Soutien aux opérations militaires en temps réel ; apport de renseignements exploitables aux commandants.	Automatisation de tâches administratives telles que la logistique, la gestion de la chaîne d'approvisionnement et la planification du personnel.
 **Autorités exécutives de contrôle**	Soutien aux dirigeants dans la rédaction de la législation et des politiques de sécurité.	Analyse de risques et évaluation de différentes options pour les capacités et la préparation de la défense.	Apport d'analyses en temps réel sur l'efficacité opérationnelle, permettant d'ajuster les ressources en fonction de l'évolution des besoins.	Automatisation de tâches administratives telles que la planification budgétaire, la gestion des contrats et le suivi des actifs ; amélioration de l'efficacité et de la responsabilité dans les opérations d'approvisionnement et de logistique de la défense.
 **Prestataires nationaux de services de sécurité**	Traitement et analyse de grands volumes de données de sécurité intérieure afin d'identifier des schémas, des tendances et des menaces potentielles.	Identification de cibles prioritaires, hiérarchisation des menaces et élaboration de rapports stratégiques.	Les systèmes de surveillance et de reconnaissance propulsés par l'IA permettent une précision accrue.	Automatisation du traitement des données, de la génération de rapports et la diffusion de l'information ; rationalisation des flux de travail.
 **Comités de contrôle internes et externes**	Surveillance et analyse des dépenses de défense, des processus d'approvisionnement et de la conformité aux réglementations afin de détecter la fraude, le gaspillage ou les abus au sein des organisations de défense.	Soutien aux activités de contrôle en analysant les politiques de défense, en évaluant les plans stratégiques et en appréciant l'efficacité des programmes et initiatives de défense.	Apport d'analyses en temps réel sur les opérations de défense, l'exécution budgétaire et les indicateurs de performance ; favorisation de la responsabilité et la transparence.	Automatisation des procédures d'audit, des évaluations des risques et des contrôles de conformité ; amélioration de l'efficacité et de la performance des fonctions de surveillance ; et garantie du respect des exigences légales et réglementaires.
 **Prestataires commerciaux de services de défense**	Développement d'outils d'intelligence artificielle pour détecter et prévenir les cybermenaces.	Soutien à la planification stratégique et l'évaluation des risques ; automatisation des outils de prévision pour optimiser l'allocation des ressources.	Développement et utilisation de drones équipés d'intelligence artificielle pour la surveillance.	Optimisation de la maintenance et de la disponibilité des équipements grâce à l'IA prédictive ; utilisation des algorithmes d'IA pour optimiser la gestion de la chaîne d'approvisionnement.
 **Organisations de la société civile**	Identification et signalement de potentielles violations de droits de l'homme.	Plaidoyer en faveur d'une utilisation transparente et éthique de l'IA dans le secteur de la défense.	Soutien aux initiatives de sensibilisation et de mobilisation communautaire concernant les systèmes de défense intégrant l'intelligence artificielle.	Renforcement de la responsabilité dans les dépenses de défense et l'utilisation des ressources afin de développer ou acquérir des outils d'intelligence artificielle.

QUELS SONT LES RISQUES QUE L'IA PRÉSENTE POUR LE SECTEUR DE LA DÉFENSE ?

Préjugés et discrimination : Étant donné que les systèmes d'IA sont conçus par des humains et entraînés sur des données humaines, ils comportent le risque de perpétuer les biais sociaux existants, aggravant ainsi les problèmes de discrimination et de ciblage injuste lors des opérations militaires.

Risques liés à la cybersécurité : L'IA transforme le paysage des menaces cybernétiques en réduisant les barrières d'accès pour les cybercriminels. Le potentiel de l'IA à améliorer les techniques d'ingénierie sociale, telles que le phishing et la technologie des deepfakes, devient de plus en plus évident et constitue un défi considérable pour le secteur de la défense. De plus, comme beaucoup d'autres outils numériques, les systèmes d'IA sont eux-mêmes vulnérables aux cyberattaques. De telles attaques courantes incluent :

- **Attaques par injection de consignes,** où un acteur malveillant manipule la consigne (entrée) d'un système d'IA pour modifier son comportement prévu, contourner les mesures de sécurité ou extraire des informations sensibles. Cela peut manipuler les systèmes décisionnels militaires d'IA, entraînant des actions non autorisées, de la désinformation ou la compromission de renseignements classifiés.
- **Attaques adversariales,** où des données manipulées sont injectées dans les systèmes d'IA pour provoquer des erreurs ou des erreurs de classification. Ces erreurs peuvent entraîner des défaillances opérationnelles ou des incidents de tirs amis pour les acteurs de la défense.
- **Empoisonnement du modèle,** où des acteurs malveillants compromettent les données d'entraînement afin de corrompre le modèle d'IA. Une telle attaque pourrait permettre aux adversaires de manipuler les résultats, perturber les opérations ou obtenir des avantages stratégiques.
- **Attaques d'extraction de données,** où des informations sensibles sont divulguées en exploitant les sorties du système d'IA. Cela peut entraîner la fuite de renseignements classifiés, de stratégies opérationnelles ou de détails sur les infrastructures de défense, compromettant la sécurité nationale et offrant un avantage tactique aux adversaires.

Désinformation : Les outils d'IA peuvent être utilisés pour générer ou diffuser de la désinformation, comme la création de vidéos deepfake dans lesquelles des responsables militaires font des déclarations trompeuses, ou la génération automatisée de fausses informations ou de publications sur les réseaux sociaux qui déforment le récit autour des opérations militaires.

Vie privée : Les outils de surveillance basés sur l'IA dans le secteur de la défense peuvent porter atteinte à la sphère privée des individus en collectant et en analysant de grandes quantités de données personnelles sans garanties adéquates.

QUEL EST L'IMPACT DE L'INTELLIGENCE ARTIFICIELLE SUR LA BONNE GOUVERNANCE DU SECTEUR DE LA DÉFENSE ?

En plus de ces risques, l'utilisation de l'intelligence artificielle par le secteur de la défense présente d'autres défis pour la mise en œuvre d'une bonne gouvernance du secteur de la sécurité, en particulier :

- **Responsabilité :** L'opacité de la technologie et des algorithmes de décision fondés sur l'IA peut entraver la capacité à retracer et à comprendre la manière dont les décisions sont prises, notamment en ce qui concerne la définition des cibles ou les décisions d'action létale. Par conséquent, il pourrait devenir difficile de tenir des individus ou des institutions responsables des erreurs, des biais ou d'un usage abusif des systèmes d'IA.

- **Transparence :** Le manque de transparence des systèmes d'IA peut engendrer incertitude et méfiance parmi les citoyens. Cela pourrait, à son tour, éroder la confiance dans les institutions de défense, avec un double impact, à la fois sur les opérations militaires en cours et sur le soutien intérieur.

- **État de droit :** Les systèmes d'IA peuvent, de manière involontaire, enfreindre des principes juridiques et éthiques, tels que le droit international humanitaire, en raison de biais dans les algorithmes et la conception, d'erreurs dans la prise de décision ou d'un manque de supervision humaine. Le non-respect de l'État de droit et des cadres juridiques établis peut avoir de graves conséquences, tant pour les individus que pour les États.

CARACTÉRISTIQUES D'UNE BONNE GOUVERNANCE DU SECTEUR DE LA SÉCURITÉ

Lorsque les principes de bonne gouvernance sont appliqués au secteur de la défense, l'État assure la sécurité de la population de manière efficace et responsable, dans un cadre de contrôle démocratique et civil, de respect de l'état de droit et des droits humains. La bonne gouvernance du secteur de la sécurité (GSS) est un ensemble de principes, et non un modèle institutionnel, et par conséquent les mêmes principes fondamentaux s'appliquent différemment selon le contexte de chaque secteur national de la défense. L'établissement d'une bonne GSS nécessite un ajustement constant, car les menaces et les besoins en matière de sécurité évoluent. Aucun secteur de la défense n'est à l'abri du besoin d'amélioration. Bien que chaque secteur de la défense soit confronté à des menaces et besoins spécifiques, certaines caractéristiques institutionnelles typiques définissent une bonne GSS.

- **Participation :** Le manque de transparence et de compréhension de la technologie de l'IA peut entraver la participation et l'engagement véritables des parties prenantes, y compris le personnel de la défense, les organisations de la société civile et les communautés marginalisées.

- **Réactivité :** Une dépendance excessive envers les systèmes d'IA peut diminuer la place du jugement et de l'intuition de l'humain, pourtant essentiels pour une prise de décision rapide et adaptée dans des situations en constante évolution. Cela peut potentiellement retarder ou compromettre les réponses efficaces aux menaces ou crises émergentes.

- **Efficacité :** Les préjugés présents dans les systèmes d'IA peuvent entraîner des pratiques discriminatoires, réduisant l'efficacité des opérations de défense et de l'allocation des ressources, ainsi que compromettant l'atteinte des objectifs stratégiques.

- **Efficience :** Le manque de transparence, de responsabilité et de supervision dans les systèmes d'IA peut entraîner des déficiences, telles que des erreurs dans la prise de décision, des duplications d'efforts ou une mauvaise utilisation des ressources, compromettant ainsi l'efficience des opérations de défense et de la gestion des ressources.

Deux autres enjeux transversaux à prendre en considération sont :

- **Droits de l'homme :** L'absence de réglementation, de supervision et de responsabilité des systèmes d'IA peut entraîner des violations des droits de l'homme, telles que des atteintes à la vie privée, des abus liés à la surveillance, des discriminations et le ciblage de populations vulnérables, sapant ainsi les droits et libertés fondamentaux protégés par le droit international.

- **Égalité de genre :** Les systèmes d'IA peuvent perpétuer les biais sexistes présents dans les données d'apprentissage, entraînant des résultats discriminatoires dans les processus de recrutement, de promotion et de prise de décision au sein du secteur de la défense. Cela peut renforcer les inégalités entre les genres et freiner l'avancement des femmes dans les professions militaires et liées à la défense.

Pour en savoir plus sur l'égalité de genre et les droits de l'homme dans le contexte de la gouvernance et de la réforme du secteur de la sécurité, veuillez consulter les fiches d'information du DCAF sur l'égalité des genres et la bonne gouvernance du secteur de la sécurité, l'égalité des genres et la réforme du secteur de la sécurité, ainsi que sur les droits humains et la G/RSS.

COMMENT LE SECTEUR DE LA DÉFENSE PEUT-IL RENFORCER LA SURVEILLANCE DE L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE ?

Les organisations de défense peuvent atténuer les risques et maximiser les avantages des technologies d'intelligence artificielle tout en respectant les normes éthiques ainsi que les cadres juridiques et réglementaires. Elles pourraient mettre en place des mécanismes de contrôle rigoureux, renforcer la transparence et la responsabilité, et favoriser la collaboration entre les organes internes de surveillance, les parties prenantes externes ainsi que leurs homologues internationaux.

La figure 6 ci-dessous présente comment les différents acteurs du secteur de la défense peuvent renforcer la surveillance de leur utilisation de l'intelligence artificielle.

FIGURE 6 SURVEILLANCE DES UTILISATIONS DE L'INTELLIGENCE ARTIFICIELLE DANS LE SECTEUR DE LA DÉFENSE

ACTEUR	CADRE RÉGLEMENTAIRE NATIONAL	TRANSPARENCE ET RESPONSABILITÉ	PARTENARIATS ET COLLABORATIONS
 **Forces armées**	Mise en œuvre de processus d'audit et de révision spécifiques à l'IA afin de surveiller le développement, l'utilisation et la performance des systèmes d'IA.	Renforcement de la transparence des systèmes d'IA en divulguant aux parties prenantes, y compris le personnel de la défense et les organes externes de contrôle, des informations non classifiées sur les sources de données, les algorithmes et les processus décisionnels.	Collaboration avec des organisations de la société civile, des institutions académiques, des organismes de recherche et des partenaires industriels pour échanger les meilleures pratiques, partager les enseignements tirés et encourager l'innovation dans la gouvernance et la surveillance de l'IA.
 **Autorités exécutives de contrôle**	Mise en place des organes ou comités de surveillance dédiés chargés de mettre en œuvre des cadres de gestion des risques visant à identifier, évaluer et atténuer les risques potentiels liés à l'adoption de l'IA, y compris les considérations éthiques, juridiques et sécuritaires.	Publication d'évaluations d'impact de l'IA et des rapports détaillant le déploiement, la performance et les résultats des systèmes d'IA dans les opérations de défense.	Renforcement de la collaboration avec les commissions parlementaires, les organismes de contrôle gouvernementaux et les auditeurs indépendants afin de fournir des mises à jour et des rapports réguliers sur les initiatives et les investissements en matière d'IA dans le secteur de la défense.
 **Prestataires nationaux de services de sécurité**	Mise en place de comités d'examen indépendants pour surveiller et évaluer les implications éthiques et juridiques de l'utilisation de l'IA dans les activités de sécurité nationale.	Mise en œuvre de contrôles internes et des audits afin de garantir le respect des cadres juridiques et éthiques régissant l'utilisation de l'IA dans la sécurité nationale.	Développement de partenariats avec des organisations de défense des droits humains, des défenseurs de la vie privée et des experts universitaires pour réaliser des évaluations et des examens indépendants des systèmes d'IA utilisés dans les activités de sécurité intérieure ; promouvoir la transparence et la responsabilité dans la gouvernance de l'IA.
 **Comités de contrôle internes et externes**	Mise en place de mécanismes de protection des lanceurs d'alerte afin d'encourager la dénonciation des préoccupations ou des comportements répréhensibles liés à l'IA au sein des organisations de défense.	Develop open-source platforms for citizens to submit any issues or abuses committed by defence actors when using AI technology.	Continuous engagement between internal and external stakeholders, including civil society groups and industry experts, to solicit feedback and recommendations for enhancing AI oversight mechanisms.
 **Prestataires commerciaux de services de défense**	Test, évaluation et mise à jour réguliers des modèles d'IA pour garantir leur conformité aux dernières réglementations et normes éthiques.	Garantie de la transparence des algorithmes d'IA et des processus décisionnels ; établir des structures internes claires de responsabilité pour le développement et la commercialisation de l'IA.	Collaboration avec le gouvernement, la société civile et les entreprises technologiques pour renforcer l'éthique de l'IA.
 **Organisations de la société civile**	Surveillance et audit indépendants des activités et résultats de défense reposant sur l'IA.	Plaidoyer en faveur de la divulgation publique de l'utilisation et de l'impact de l'IA dans la défense ; promouvoir la responsabilité à travers des rapports publics et le rôle des organismes de contrôle.	Engagement auprès des entités de défense pour garantir des pratiques éthiques en matière d'IA ; travail avec des organisations internationales sur la réglementation de l'IA dans le domaine de la défense.

POUR APPROFONDIR

- Centre for Humanitarian Dialogue
  Code of Conduct on Artificial Intelligence in Military Systems
  Genève : HD 2021

- Cheong, Sandra
  Intelligence Oversight in the Age of Digitalization
  Genève : DCAF 2024

- Davison, Neil
  A legal perspective: Autonomous weapon systems under international humanitarian law
  Genève : ICRC 2016

- Deloitte
  The Age of With™ – The AI advantage in defence and security
  Ontario : Deloitte: 2024

- Parlement européen et Conseil de l'Union européenne
  Règlement de l'Union européenne sur l'intelligence artificielle
  Bruxelles : UE, 2024

- Evans, Thamy
  La réforme de la défense
  Genève : DCAF 2019

- Herd, Graeme P., Puhl, Detfel, et Costigan, Sean
  Emerging Security Challenges: Framing the Policy Context
  Genève : GCSP, 2013

- Huhtanen, Heather et Triquet, Veerle
  L'égalité des genres et la bonne gouvernance du secteur de la sécurité
  Genève : DCAF 2015

- ISSAT Advisory Note
  Artificial Intelligence and SSG/R
  Genève : DCAF 2023

- Lui, Dawn et Lazar, Alexandru
  Digitalization and SSG/R: Projections into the Future
  Genève : DCAF 2023

- Lui, Dawn
  La digitalisation et la gouvernance et réforme du secteur de la sécurité
  Genève : DCAF 2022

- OECD.AI Policy Observatory
  Ressources sur l'IA
  Paris: OECD 2024

- Stanford University
  The AI Index Report: Measuring trends in AI
  Stanford: Stanford University, 2024

- Triquet, Veerle et Watson, Callum
  L'égalité des genres et la réforme du secteur de la sécurité
  Genève : DCAF 2015

- Ulnicane, Inga
  Intersectionality in Artificial Intelligence: Framing Concerns and Recommendations for Action
  Warwick: Social Inclusion, vol. 12, 7543

- US Department of State
  Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy
  2024

AUTRES RESSOURCES DU DCAF SUR LA RSS

Les publications du DCAF comprennent une large gamme de manuels et outils spécifiques permettant de guider les praticiens oeuvrant dans le domaine de la RSS et de la bonne GSS, téléchargeables gratuitement à l'adresse suivante : **www.dcaf.ch**

NOTRE MISSION

Aider les États à améliorer le mode de gouvernance de leur secteur de la sécurité.

Donner des conseils sur l'élaboration de mesures de gouvernance du secteur de la sécurité à la fois efficaces et viables.

Favoriser la mise en œuvre par les États de réformes participatives, valorisant la contribution de tous et intégrant la dimension de genre.

NOS ACTIONS

Fournir une expertise technique aux processus de RSS/G menés au niveau national.

Renforcer les capacités des acteurs étatiques et non étatiques.

Diffuser en libre accès des ressources et des résultats de travaux de recherche.

Promouvoir les bonnes pratiques de gouvernance recommandées au niveau international.

Conseiller sur les questions juridiques et politiques liées au secteur de la sécurité.

DCAF – Geneva Centre for Security Sector Governance

Maison de la Paix
Chemin Eugène-Rigot 2E
1202 Geneva
Switzerland

+41 22 730 94 00
info@dcaf.ch
@DCAF_Geneva

www.dcaf.ch