Royaume du Maroc

Contribution sous le thème :

« Application de l'Intelligence Artificielle dans le Domaine Militaire : Opportunités et Défis (Hors SALA) dans un contexte de paix et de sécurité internationale »

Sommaire

I. Introduction

L'émergence rapide de l'intelligence artificielle (IA) transforme profondément les stratégies et capacités militaires à travers le monde. La communauté internationale observe avec attention cette révolution technologique aux potentialités immenses, tout en s'inquiétant des risques qu'elle comporte.

Cette transformation, poussée par l'IA, laisse entrevoir des opportunités considérables pour prévenir les conflits, protéger les populations et appuyer les missions de paix. Cependant, en parallèle de ces espoirs, l'IA suscite de vives préoccupations éthiques, juridiques et sécuritaires. En l'absence de contrôle humain adéquat, l'IA pourrait échapper à notre gouvernance et menacer nos sociétés. De plus, l'usage militaire de l'IA a principalement retenu l'attention par le prisme des Systèmes d'Armes Létaux Autonomes (SALA), souvent appelés *« robots tueurs»*, qui posent des questions inédites sur la délégation de la décision de vie ou de mort à des machines.

Consciente de ces enjeux, la communauté internationale débat activement pour encadrer ou interdire de tels systèmes. Néanmoins, l'objet du présent document est volontairement circonscrit aux applications de l'IA dans le domaine militaire en excluant les SALA, afin d'examiner plus largement comment l'IA peut soutenir les missions militaires tout en respectant la paix et la sécurité internationales.

Cette démarche s'inscrit dans un contexte d'attention croissante pour l'IA militaire au-delà des seules armes autonomes. Le 24 décembre 2024, l'Assemblée Générale de l'ONU a adopté la première Résolution sur l'IA militaire, intitulée « l'intelligence artificielle dans le domaine militaire et ses conséquences pour la paix et la sécurité internationales ». D'autres processus ont également émergé dans ce cadre, tels que la toute première Conférence Internationale sur le rôle de l'IA dans la mise en œuvre de la Convention sur l'Interdiction des Armes Chimiques (*Octobre 2024, Rabat, Maroc*), les Sommets « REAIM - Responsible AI in the Military Domain » (*Septembre 2024, Séoul, République de Corée et Février 2023, La Haye, Pays-Bas*), la Déclaration Politique sur l'utilisation militaire responsable de l'IA et sur l'autonomie (*lancée au Sommet REAIM 2023, La Haye, Pays-Bas*) et la Déclaration de Paris sur le maintien du contrôle humain dans les systèmes d'armes basés sur l'IA (*Sommet IA pour l'Action, février 2025, Paris, France*).

Le présent document explore les opportunités et défis liés à l'application de l'IA dans le domaine militaire (hors SALA). Il s'appuie prioritairement sur les travaux et rapports onusiens, notamment ceux de l'UNIDIR (Institut des Nations Unies pour la recherche sur le désarmement) [1], de l'ONUDA (Bureau des affaires de désarmement) et d'autres agences pertinentes, pour dégager les meilleures pratiques, illustrations concrètes et préoccupations éthiques afférentes.

Après cette introduction, la section II présentera les principaux domaines d'application positive de l'IA militaire ainsi que des cas d'usage allant de la prévention des conflits aux opérations de maintien de la paix et à la gestion des catastrophes. La section III examinera les défis techniques, éthiques, stratégiques et juridiques qui se posent pour assurer un usage responsable de ces technologies dans le respect du droit des relations internationales, notamment le droit international humanitaire (DIH). Enfin, la conclusion offrira une synthèse des enseignements et recommandations clés, dans un esprit conforme aux principes de la Charte des Nations Unies et de la coopération multilatérale.

II. Opportunités de l'IA dans le domaine militaire (hors SALA)

L'IA appliquée au secteur militaire ne se résume pas aux armes autonomes. Elle englobe un spectre beaucoup plus large de fonctions *« en amont »* (renseignement, logistique, analyse d'information, etc.) susceptibles d'accroître l'efficacité des opérations tout en réduisant les risques pour les militaires et les civils. Dans cette seconde partie, nous distinguons d'une part les grands domaines d'opportunités offertes par l'IA pour améliorer les capacités militaires non létales, et d'autre part des cas d'usage concrets illustrant comment ces technologies peuvent servir des objectifs de paix et de sécurité.

A. Domaines fonctionnels

1. Renseignement, surveillance et reconnaissance (ISR)

Le renseignement, la surveillance et la reconnaissance (ISR) constituent un pilier fondamental des opérations militaires, permettant de collecter et d'analyser des informations sur l'environnement, les forces adverses et les menaces potentielles. L'IA offre des possibilités inédites pour renforcer les capacités ISR en automatisant l'analyse de vastes flux de données et en détectant des motifs qu'un œil humain pourrait manquer. Par exemple, des algorithmes de vision par ordinateur [2] peuvent passer en revue en temps réel des images aériennes ou satellitaires et repérer des éléments d'intérêt (véhicules, infrastructures, mouvements de troupes) avec une rapidité et une précision accrues. Couplée à des drones de surveillance ou à des capteurs terrestres, l'IA permet ainsi de couvrir de larges zones de terrain en continu, fournissant aux analystes militaires une « connaissance de la situation » plus complète et à jour [3].

L'IA embarquée dans les *« UAV »* (Unmanned Aerial Vehicles) peut aider à traiter le flux vidéo et à distinguer, par exemple, un attroupement civil anodin d'un regroupement potentiellement hostile, ce qui offre l'espoir d'alerter préventivement les autorités pour désamorcer des crises probables. Autrement dit, l'analyse prédictive des données ISR par l'IA peut contribuer à la détection précoce des conflits (early warning).

Un des avantages majeurs de l'IA en matière de renseignement est sa capacité à fusionner et corréler des données hétérogènes. Elle peut agréger des informations issues de multiples sources (images satellite, interceptions de communications, renseignements open source sur les réseaux sociaux) et en extraire des renseignements exploitables. Par exemple, des systèmes d'IA peuvent scruter les médias sociaux et détecter une soudaine montée de discours violents ou de fausses rumeurs dans une région donnée, signalant une montée des tensions. Combinées aux données de surveillance physique du terrain, ces analyses peuvent offrir une vision holistique de la situation sécuritaire.

En milieu opérationnel, l'IA améliore également la vitesse de traitement et de diffusion du renseignement. Là où un analyste aurait mis des heures à analyser un lot d'images ou de rapports, un algorithme peut accomplir cette tâche en quelques secondes, permettant aux chefs d'obtenir plus rapidement une image actualisée du théâtre d'opérations. Cette accélération du *temps décisionnel* grâce à l'IA peut s'avérer déterminante pour garder l'initiative sur le terrain. Néanmoins, il convient de souligner que cette même rapidité pose des défis, car des décisions militaires sensibles prises sur la base d'analyses automatisées doivent être soigneusement vérifiées pour éviter les erreurs.

En résumé, dans le domaine ISR, l'IA constitue une opportunité de multiplier les "yeux et oreilles" du militaire. Qu'il s'agisse de surveiller des frontières étendues, de suivre les déplacements de groupes terroristes ou de garder un œil sur des zones de conflit enclavées, les capteurs intelligents peuvent fournir une veille constante là où les moyens humains seraient limités.

2. Logistique et planification opérationnelle

Premièrement, l'IA peut améliorer la planification stratégique et tactique des opérations. En simulant différents scénarios logistiques, des algorithmes peuvent aider les planificateurs à élaborer le déploiement optimal de ressources. Par exemple, des systèmes d'IA sont capables d'optimiser les itinéraires de convoi en fonction de multiples contraintes (sécurité des routes, météo, consommation de carburant) et de recalculer rapidement de nouvelles routes si un imprévu survient.

De même, dans la préparation d'une opération, l'IA peut proposer des arrangements alternatifs des forces et des stocks, détectant par extrapolation des déviations par rapport aux plans initiaux et suggérant des corrections proactives. La planification de scénarios assistée par IA, via des outils de type *wargaming* virtuels, permet aux décideurs d'explorer une multitude d'hypothèses afin d'atténuer les imprévus [4].

Deuxièmement, l'IA contribue à rendre la chaîne d'approvisionnement plus réactive et efficiente. Dans de nombreuses armées, on observe une inflation des données à gérer : inventaires informatisés, capteurs de suivi des véhicules, informations d'état de santé des équipements, etc. Un algorithme peut assimiler en continu ces flux d'information et déclencher des alertes ou décisions automatiques.

Dans des contextes de déploiements internationaux complexes, tels que les missions de l'ONU ou les coalitions multinationales, l'IA peut aussi faciliter la coordination logistique entre partenaires. Elle peut servir de courroie pour partager des informations sur les ressources disponibles, les besoins urgents, et proposer des mutualisations.

On peut imaginer une mission de paix où plusieurs pays contributeurs fournissent des unités : une plateforme IA pourrait indiquer que le contingent A dispose d'un surplus de matériel médical qui pourrait être redéployé vers le contingent B confronté à une crise sanitaire dans sa zone. Ainsi, en améliorant la visibilité globale, l'IA renforce la solidarité logistique entre alliés et optimise l'allocation collective des moyens [5].

Enfin, l'automatisation permise par l'IA touche également la robotique logistique. Des véhicules autonomes de transport de charges (sur terre, ou des drones cargo dans les airs) commencent à être testés pour livrer du ravitaillement en zone dangereuse sans risque humain.

Bien qu'encore émergents, ces systèmes bénéficient de l'IA pour naviguer en terrain incertain, éviter des obstacles ou menaces, et atteindre leur destination de manière fiable. Le fait de pouvoir envoyer un robot convoyeur, réapprovisionner une position avancée sous le feu ennemi, ou évacuer des blessés en véhicule autonome, représenterait un gain humanitaire majeur, en éliminant l'exposition de soldats de logistique à des embuscades.

En somme, l'IA apporte à la logistique et à la planification opérationnelle des gains de précision, d'agilité et de prévoyance. De meilleures prévisions permettent de réduire le *brouillard de la guerre* sur l'état des ressources, tandis que l'optimisation algorithmique accroît la résilience des forces armées face aux imprévus.

3. Aide à la décision

Dans le processus de commandement et de contrôle (C2), l'IA est perçue comme un atout pour assister les décideurs, qu'il s'agisse d'un officier sur le terrain devant réagir en secondes ou d'un Etat-Major planchant sur une stratégie complexe. L'aide à la décision par l'IA se matérialise par des systèmes capables de synthétiser l'information, de proposer des options d'action et d'évaluer les conséquences probables de chaque choix, le tout en soutien (et non en remplacement) du commandement humain [6,7].

L'une des applications phares est le système de support décisionnel en temps réel. Face à une situation tactique évolutive (par exemple la progression inattendue d'une colonne ennemie ou une détérioration soudaine de la météo affectant les hélicoptères), un outil d'IA peut instantanément agréger les données ISR pertinentes, comparer la situation à des cas similaires enregistrés en base de données, et fournir au commandant plusieurs scénarios de réponse avec leurs avantages et risques estimés. L'officier conserve toujours la décision finale, mais il bénéficie d'une *analyse accélérée* et d'une *évaluation objective* fournie par la machine. Cela peut réduire le risque d'une décision prise sous le coup de la panique ou du biais cognitif, en offrant un contrepoids informatif. Par exemple, si l'IA indique qu'un repli stratégique temporaire représente 90% de chances de préserver l'unité intacte en attendant des renforts, alors que l'instinct du moment pousserait à tenir la position, les décideurs pourront reconsidérer sa réaction initiale à la lumière de ces données [9].

Un autre aspect est l'analyse prédictive de haut niveau. Des modèles d'IA peuvent être entraînés sur d'énormes jeux de données historiques (conflits passés, manœuvres militaires, indicateurs politiques et économiques) pour dégager des tendances pouvant influencer les décisions stratégiques. Par exemple, l'IA pourrait être utilisée pour évaluer les probabilités de survenance de différents scénarios de crise régionale dans les cinq prochaines années, en intégrant des variables comme la démographie, les tensions ethniques, les effets du changement climatique, etc. Un tel outil ne prédit pas l'avenir avec certitude, mais il quantifie des risques et attire l'attention des décideurs sur des facteurs parfois négligés.

En contexte militaire national, on voit émerger des centres de commandement augmentés par l'IA, où des assistants virtuels font partie de l'équipe de planification. Ces assistants peuvent répondre à des requêtes en langage naturel, par exemple : *« Quelle est la probabilité que l'ennemi contre-attaque si nous prenons la ville X ? »*, en fournissant une analyse structurée basée sur les données disponibles. Ils peuvent également surveiller de manière autonome l'évolution d'une opération et signaler immédiatement aux chefs des anomalies ou opportunités. En somme, ils agissent comme des *anges gardiens informationnels*, prêts à alerter ou conseiller, tout en déchargeant les analystes humains de tâches fastidieuses, telles que parcourir des centaines de rapports d'incident pour en extraire une tendance.

Il est important de noter que dans cette vision d'intelligence augmentée, l'IA n'a pas vocation à se substituer au jugement humain, mais à le renforcer. Les valeurs éthiques et le bon sens du commandement restent primordiaux. L'IA ne fait que fournir des *éléments factuels et des inférences probables*.

En tirant profit de l'IA pour l'aide à la décision, les armées espèrent également réduire les erreurs et minimiser les dommages collatéraux. Mieux informé et conseillé, un chef militaire sera moins susceptible de prendre une décision hasardeuse pouvant mettre des civils en danger ou compromettre la mission. Par exemple, avant de lancer une opération nocturne dans un village, un système d'IA pourrait alerter : *« forte probabilité de présence de civils à cette heure d'après les données de téléphonie mobile, ajustez l'horaire de l'assaut »*. Ce type d'apport concourt à l'intégration du droit international humanitaire dans le processus décisionnel lui-même.

En définitive, l'aide à la décision par l'IA constitue une opportunité de faire face à la complexité croissante des conflits modernes. Sur des théâtres d'opérations hybrides, où s'entremêlent actions militaires, cyberattaques, bataille de l'information, pression de l'opinion publique, le décideur est saturé de données et de contraintes.

4. Cybersécurité et protection des systèmes numériques critiques

À l'ère du tout-numérique, la cybersécurité est devenue un enjeu militaire aussi vital que la défense physique du territoire. Les forces armées, les infrastructures de défense et même les

équipements sur le terrain reposent largement sur des réseaux informatiques et des logiciels susceptibles d'être la cible de cyberattaques. L'IA intervient ici comme une double opportunité : renforcer la sécurité des systèmes critiques face à des menaces sophistiquées, et aider à contrer les offensives menées par des acteurs malveillants dans le cyberespace.

Premièrement, l'IA améliore les capacités de détection d'intrusion et de menaces (AI-powered IDS) dans les réseaux [8,9]. Des systèmes d'IA dotés d'apprentissage automatique peuvent établir en temps réel une base de référence du fonctionnement normal d'un réseau informatique militaire (flux de données habituels, processus légitimes en cours, connexions autorisées), puis détecter toute déviation suspecte pouvant indiquer une attaque. Par exemple, si un malware inconnu commence à se propager sur le réseau logistique, l'IA repèrera des signatures ou comportements anormaux (comme une série de connexions inhabituelles ou une exécution de code non autorisé) et pourra immédiatement alerter les opérateurs humains, voire enclencher des mesures d'isolement de la menace. Cette réactivité accrue est essentielle face à des attaques qui se déroulent parfois en quelques secondes. De même, l'IA excelle à analyser de grandes quantités de journaux informatiques (*logs*) pour y trouver des indices faibles d'une campagne d'intrusion lente et furtive qu'un analyste humain, submergé de données, pourrait rater [10,11].

Les infrastructures critiques nationales, telles que centrales électriques, réseaux de transport, communications stratégiques, peuvent également être protégées par l'IA. Par exemple, un réseau électrique *smart grid* peut utiliser l'IA pour distinguer une fluctuation due à une panne technique d'une fluctuation possiblement induite par une cyberattaque qui cherche à provoquer une panne. En cas de soupçon d'attaque, le système peut enclencher des procédures de secours, informer le commandement militaire et isoler la partie du réseau compromise. De plus, des systèmes IA peuvent simuler l'impact d'une cyberattaque sur l'ensemble d'une infrastructure afin d'aider les autorités à planifier des parades et des redondances. Cette résilience cyber est particulièrement cruciale dans des pays en développement où les infrastructures sont parfois moins robustes.

Sur le plan offensif, même si ce document se concentre sur les usages défensifs et de soutien, notons que l'IA peut être utilisée pour contrer les campagnes de désinformation menées par des adversaires, ce qui impacte la sécurité globale. Des algorithmes NLP (traitement du langage naturel) peuvent détecter la diffusion coordonnée de fausses nouvelles, d'incitation à la haine ou de propagande en ligne, souvent orchestrée par des bots. L'IA fournit des moyens de surveiller ces manœuvres et d'y répondre par des contre-discours vérifiés, contribuant à protéger les populations locales de manipulations qui pourraient attiser la violence et les foyers de tension.

En somme, sécuriser le cyberespace militaire par l'IA est devenu incontournable. Les cyberattaques dopées à l'IA visent déjà des infrastructures critiques et même des opérations de paix. Face à cela, l'IA du côté des défenseurs constitue un *bouclier adaptatif* : plus les attaques se perfectionnent, plus elle apprend à les détecter.

5. Entraînement et simulation

Des soldats s'entraînent dans un environnement virtuel immersif, illustrant comment l'IA et la simulation améliorent la préparation opérationnelle tout en réduisant les coûts et les risques. La préparation des troupes à affronter des situations de combat ou de crise repose traditionnellement sur des entraînements intensifs, des exercices tactiques et des simulations de manœuvres. L'IA révolutionne ce domaine en permettant des entraîneurs virtuels intelligents et des scénarios de simulation réalistes à un niveau jamais atteint auparavant. L'objectif est double : améliorer l'efficacité de la formation des militaires, tout en économisant des ressources (munitions, carburant) et en évitant les dangers inhérents aux exercices réels.

L'une des avancées notables est le développement de simulateurs d'entraînement immersifs utilisant la réalité virtuelle (VR) ou augmentée, combinés à des moteurs d'IA pour peupler et animer l'environnement virtuel. Par exemple, un groupe de soldats peut évoluer dans un champ de bataille simulé en VR, confronté à des ennemis gérés par l'IA qui réagissent de manière crédible aux actions des stagiaires. Ces ennemis virtuels ne se contentent pas de comportements prédéfinis : grâce à l'Apprentissage par Renforcement, ils peuvent adapter leurs tactiques, apprendre des erreurs du joueur et présenter un défi renouvelé à chaque session. Cela signifie qu'un soldat ou une unité en formation ne répétera pas simplement un scénario figé, mais pourra être surprise par des embuscades imprévues ou des ruses élaborées de l'IA, tout comme dans la réalité. Ainsi, la courbe d'apprentissage s'en trouve accrue, car les entraînés doivent affiner leurs réflexes et leur coordination pour triompher d'un adversaire virtuel évolutif.

Les instructeurs humains bénéficient aussi de l'IA dans le suivi et l'évaluation des performances. Des systèmes peuvent analyser automatiquement les actions entreprises pendant l'exercice (déplacements, temps de réaction, précision de tir, respect des protocoles de communication) et générer un retour d'information détaillé. Par exemple, après une simulation de mission de maintien de la paix en environnement urbain, l'IA pourrait signaler que l'équipe a négligé de surveiller un angle mort pendant 30 secondes, ou que la communication radio était saturée d'informations redondantes, des points à améliorer que l'instructeur pourra ensuite aborder. Cette objectivité et granularité des évaluations aident à identifier des faiblesses subtiles difficilement détectables à l'œil nu.

Un autre apport est la personnalisation de l'entraînement. Grâce à l'IA, il est possible d'adapter les scénarios en temps réel au niveau du stagiaire, un peu comme un jeu vidéo adapte sa difficulté. Si un peloton excelle dans un domaine, l'IA peut augmenter la complexité (par exemple en introduisant des aléas supplémentaires, comme une panne soudaine d'équipement dans le scénario) pour le pousser à se dépasser. Au contraire, si un élève-pilote d'hélicoptère a des difficultés dans les manœuvres d'atterrissage, le simulateur intelligent peut lui proposer davantage d'exercices ciblés sur cet aspect jusqu'à ce qu'il le maîtrise. On parle parfois de "tutorat intelligent", l'IA endossant le rôle d'un coach virtuel qui ajuste le programme en fonction des progrès individuels.

Enfin, outre les soldats, l'IA sert également à entraîner les décideurs et planificateurs via des simulations stratégiques. Les jeux de guerre (war games) assistés par IA permettent aux chefs militaires de tester des plans de campagne dans un environnement sans risque. L'IA peut jouer le rôle de tous les acteurs (ennemi, alliés, acteurs civils) et fournir un retour accéléré sur les conséquences d'une stratégie ; identifiant, par exemple, que tel plan de déploiement aboutirait à une impasse logistique au bout de X jours. Ces enseignements virtuels, acquis *avant* l'action réelle, sont inestimables pour éviter des erreurs coûteuses en vies humaines.

En conclusion, l'IA redéfinit l'entraînement militaire comme un laboratoire virtuel où l'on peut apprendre de ses erreurs sans en payer le prix sanglant, et affiner ses compétences dans des contextes variés à volonté. C'est une opportunité de préparer des forces plus polyvalentes, mieux aguerries aux surprises, et donc plus à même de remplir leurs missions avec succès et humanité.

Naturellement, cela nécessite des investissements en équipements de simulation, en élaboration de contenus scénaristiques fidèles, et en formation des instructeurs aux nouvelles méthodes. Mais ces coûts initiaux sont largement compensés par les économies réalisées (moins de munitions réelles consommées en entraînement, moins d'accidents en manœuvres réelles) et par l'élévation du niveau opérationnel des forces.

B. Cas d'usages

Après avoir passé en revue les grands domaines fonctionnels où l'IA offre des gains pour les forces armées, intéressons-nous maintenant à des cas d'emploi spécifiques de ces technologies dans des contextes liés à la paix et la sécurité internationales. Il s'agit d'illustrer concrètement comment l'IA peut être mise au service d'objectifs tels que la prévention des conflits, le soutien aux opérations de paix, la réponse aux catastrophes ou la gestion des phénomènes migratoires. Chaque sous-section aborde un scénario ou un domaine particulier où l'IA apporte des solutions novatrices. Ces exemples permettront également de faire apparaître la transversalité des opportunités décrites précédemment : on verra que les capacités en ISR, en aide à la décision, en analyse de données, etc., se combinent dans la pratique pour relever des défis complexes sur le terrain.

1. Prévention des conflits et diplomatie

Prévenir un conflit avant qu'il n'éclate est sans doute la mission la plus noble et la plus difficile de la diplomatie internationale. La *diplomatie préventive* cherche à identifier tôt les signes avant-coureurs de crises et à favoriser le dialogue pour désamorcer les tensions. L'intelligence artificielle devient un allié précieux dans cet effort, en offrant des outils pour détecter les dynamiques conflictuelles latentes et même pour faciliter la médiation entre parties prenantes.

Un premier apport concret de l'IA est la détection automatisée des signaux faibles de conflit. Des modèles de big data analytics ingèrent d'énormes quantités de données hétérogènes, rapports d'ONG, données économiques (chômage, inflation), événements climatiques, échanges sur les réseaux sociaux, incidents de violence locale, afin de repérer des combinaisons de facteurs qui, historiquement, ont souvent précédé des crises.

Par exemple, une hausse soudaine du prix des denrées de base combinée à une rhétorique de plus en plus polarisée sur les réseaux sociaux, dans une région où coexistent différentes communautés ethniques, pourrait être signalée par l'IA comme un motif d'alerte.

De tels systèmes d'alerte précoce assistés par IA pourraient faire l'objet d'études et certains programmes pilotes ont montré qu'ils pouvaient anticiper de quelques mois des flambées de violence, donnant une fenêtre d'action aux diplomates pour intervenir.

Bien sûr, l'IA n'est pas infaillible dans ses prédictions, mais elle permet de traiter l'information bien plus vite et d'attirer l'attention sur des zones que l'analyse humaine classique n'aurait peut-être pas considérées prioritaires.

2. Missions de maintien de la paix et de stabilisation

Les opérations de maintien de la paix de l'ONU, déployées dans des environnements post-conflit fragiles, font face à des défis considérables : protection des civils, surveillance des cessez-le-feu, soutien à la stabilisation et à la reconstruction, souvent dans des zones vastes et peu équipées en infrastructures. L'introduction de l'IA dans ces missions ouvre de nouvelles perspectives pour renforcer l'efficacité des Casques bleus et améliorer la sécurité tant des populations que du personnel onusien sur le terrain.

Les missions de paix évoluent de plus en plus vers une approche intégrée civilo-militaire, où comprendre le contexte local est crucial. L'IA peut aider les équipes de la composante civile de la mission à effectuer des analyses socio-économiques rapides pour orienter les projets de stabilisation. Par exemple, en croisant des données démographiques, l'historique des violences et les plaintes de la population recueillies par les bureaux de la mission, un algorithme pourrait indiquer quelles localités sont les plus vulnérables à une résurgence du conflit faute de services publics suffisants. Les responsables de la mission peuvent alors plaider auprès du gouvernement ou des donateurs pour cibler prioritairement ces zones en aide au développement. Ainsi, l'IA

agit indirectement en consolidant la paix par une allocation plus fine des efforts de reconstruction.

Un autre cas d'usage concret concerne la sécurité du personnel onusien lui-même. Malheureusement, les Casques bleus et humanitaires sont parfois pris pour cible. L'IA peut être utilisée pour évaluer les risques d'attaque contre la mission en analysant l'activité en ligne de groupes hostiles, en détectant des patterns autour des bases onusiennes (reconnaissance armée, drones d'observation adverses) ou en anticipant des troubles (manifestations orchestrées contre la mission via les réseaux sociaux). Mieux informée, la mission peut adapter sa posture de protection, renforcer certaines positions ou limiter les déplacements non essentiels lors de périodes à risque. On peut y voir une forme de "casque bleu virtuel" qui veille en permanence sur la mission.

Enfin, l'IA peut améliorer la gestion interne de la mission. Par exemple, optimiser les patrouilles : un algorithme peut proposer des schémas de patrouille qui couvrent plus efficacement une zone donnée en fonction des incidents passés et des retours d'information communautaires. Ou bien gérer l'information : une mission de paix produit chaque jour de nombreux rapports, comptes-rendus, notes d'analyse. Un assistant IA peut trier et synthétiser ces informations pour le leadership de la mission, mettant en avant les points critiques, évitant ainsi que des alertes importantes se perdent dans la masse de documents. Cela rejoint l'objectif de modernisation numérique des opérations de paix [12].

En résumé, dans les missions de maintien de la paix et de stabilisation, l'IA agit comme un multiplicateur de force "bleue" : elle étend la portée des yeux et des oreilles des Casques bleus, accélère la circulation du renseignement, anticipe les menaces et aide à gagner les cœurs et les esprits en adaptant les efforts au plus près des besoins locaux. Tout ceci, bien sûr, doit s'inscrire dans un usage responsable : l'ONU se doit veiller à ce que ces technologies soient employées dans le respect de son mandat et du droit international. Par exemple, la surveillance par drones se fait sans empiéter sur la souveraineté nationale puisque les données recueillies sont partagées avec les autorités concernées pour renforcer la confiance. De même, les données personnelles éventuellement traitées par l'IA (opinions exprimées, etc.) sont protégées pour éviter tout stigmate sur les individus.

3. Gestion de catastrophes

Les forces armées sont souvent mobilisées en appui lors de catastrophes naturelles ou de crises humanitaires majeures, apportant logistique, secours et coordination. L'intelligence artificielle, avec sa capacité d'analyse rapide et prédictive, s'avère précieuse à toutes les phases de la gestion de catastrophes : avant (prévention et préparation), pendant (réponse d'urgence) et après (relèvement). Pour les armées et la protection civile, elle offre des outils pour sauver des vies plus efficacement et cibler au mieux l'aide apportée.

En phase de prévention/préparation, l'IA peut aider à cartographier les zones à risque. Par exemple, en combinant des données climatologiques, topographiques et d'usage des sols, des modèles prédictifs anticipent les endroits les plus susceptibles d'être inondés lors de la prochaine saison des pluies ou les localités côtières vulnérables aux ondes de tempête. Ces informations, fournies aux autorités militaires et civiles, permettent de renforcer les digues, de planifier des évacuations potentielles ou de pré-positionner des stocks de secours (tentes, vivres) à proximité des populations menacées.

Pendant la phase aiguë d'une catastrophe, l'IA amplifie la capacité de réaction et de coordination. Prenons l'exemple d'un séisme majeur frappant une région. Immédiatement, des drones survolent la zone pour évaluer les dégâts. En temps réel, des algorithmes de vision artificielle interprètent les images : ils identifient les bâtiments effondrés, les routes coupées, les attroupements de personnes en détresse, avec une précision géolocalisée. En quelques heures, une carte des destructions et des besoins prioritaires peut être dressée, là où il fallait autrefois des jours d'évaluation humaine. Les équipes de secours militaires et civiles peuvent ainsi cibler leurs efforts : envoyer d'urgence les éléments de la protection civile là où un immeuble effondré présente encore des poches de survie probables, dépêcher des hélicoptères sur des quartiers isolés par des décombres, etc. L'IA peut également analyser les communications d'urgence : par exemple trier des milliers de messages sur les réseaux sociaux ou appels d'aide pour repérer ceux mentionnant des situations critiques (incendies, personnes piégées) et les signaler aux opérateurs.

Dans une crise humanitaire complexe (suite à un conflit, une famine), l'IA peut aider les militaires à gérer les flux logistiques de l'aide. En optimisant les itinéraires de convoi (comme vu en section logistique), mais aussi en surveillant l'efficacité de la distribution. Par exemple, en analysant les données de mobilité des populations déplacées via les signaux mobiles anonymisés, l'IA pourrait estimer si un camp de réfugiés reçoit un afflux soudain de personnes, indiquant qu'une autre zone est sous-aidée, ce qui permet de rééquilibrer l'effort. De plus, l'IA peut prédire les effets domino d'une catastrophe : si un pont logistique est détruit, quelles localités seront isolées et dans combien de temps auront-elles besoin d'un ravitaillement alternatif ? Ce type d'anticipation est essentiel pour éviter une crise secondaire, comme des épidémies faute d'accès à l'eau potable.

Dans la phase de relèvement, après l'urgence immédiate, l'IA continue d'être utile. Par exemple, pour évaluer les dommages de manière exhaustive. Des images satellite post-catastrophe traitées par IA peuvent donner un bilan précis : X ponts détruits, Y kilomètres de routes inutilisables, Z habitations rasées. Ceci aide à planifier la reconstruction et à estimer les besoins financiers, éléments cruciaux pour mobiliser la communauté internationale. Par ailleurs, l'IA peut aider à bâtir mieux en reconstruisant : en simulant, via des modèles, l'efficacité de différentes architectures pour résister aux futures secousses, orientant ainsi les choix de reconstruction (matériaux, normes).

En conclusion, l'IA dans la gestion de catastrophes permet de gagner un temps précieux, d'allouer au mieux les moyens limités et d'éviter des sur-crises. Elle donne une vue d'ensemble quasi instantanée aux décideurs en pleine crise, là où jadis régnait l'incertitude. Bien sûr, l'IA ne remplace pas la mobilisation humaine et la solidarité, elle les potentialise. Les secouristes doivent toujours aller sur le terrain, mais ils le font plus efficacement en étant guidés par les analyses IA.

4. Cartographie des risques selon le contexte

Chaque contexte géopolitique présente un ensemble unique de risques et de menaces. La cartographie des risques consiste à identifier, spatialiser et évaluer ces menaces potentielles pour mieux s'en prémunir. L'IA se révèle particulièrement efficace pour établir et mettre à jour ces cartographies dynamiques, en tenant compte d'une multitude de facteurs complexes. Que ce soit dans un contexte de conflit, de tensions transfrontalières ou de criminalité organisée, l'IA aide les planificateurs militaires et les autorités à *voir* sur la carte ce qui est invisible à l'œil nu : les zones de risques émergents.

Dans le domaine de la criminalité transnationale (trafics d'armes, drogue, piraterie maritime), l'IA sert à cartographier les itinéraires probables et les bases arrière. Par exemple, en mer,

l'analyse des trajectoires de navires couplée à de la détection d'anomalies (comme un cargo qui éteint son transpondeur AIS) permet de repérer les zones où des transferts illicites en haute mer pourraient se produire. Une carte des points de transbordement suspects peut être produite, orientant les marines nationales ou les missions navales vers une surveillance accrue de ces lieux. De même, sur terre, des algorithmes identifient des schémas dans les itinéraires de contrebande (par exemple, tel col de montagne voit anormalement souvent des déplacements la nuit indiquant un trafic transfrontalier).

Un cas concret intéressant est la cartographie des risques de conflits intercommunautaires dans un pays en tension. Imaginons un pays fictif avec des provinces multiethniques où la stabilité est précaire. En scrutant les indicateurs socio-économiques (chômage, accès à la terre), les griefs exprimés (via médias locaux), les incidents violents récents, l'IA pourrait produire une carte montrant quelles provinces ou districts sont au bord de l'explosion.

La gestion des flux migratoires est un autre domaine où cartographier le risque est crucial. Par exemple, des systèmes IA pourraient être utilisés pour évaluer quelles routes migratoires pourraient voir un afflux soudain (dû à un conflit ou à un changement de politique). Avoir cela sur une carte stratégique permet de se préparer en termes d'accueil ou de contrôle, et pour les pays de départ ou de transit, d'anticiper l'impact interne (campements informels, tensions possibles avec les communautés locales).

Ce qui rend l'IA puissante pour la cartographie des risques, c'est son aptitude à prendre en compte des corrélations complexes dans l'espace et le temps. Elle peut révéler par exemple que *« lorsque la saison sèche atteint son pic et que l'accès à l'eau baisse dans telle région, le risque de conflit entre éleveurs et agriculteurs monte de 80% dans les 2 mois qui suivent »*. Cette corrélation spatio-temporelle, une fois identifiée, permet sur la carte de voir apparaître des *fenêtres de risque* avec un chronomètre. Les autorités peuvent alors mener des actions préventives (distribution d'eau, dialogue intercommunautaire) précisément dans ces fenêtres. Sans l'IA, de telles corrélations pourraient rester inaperçues ou être découvertes trop tard.

Bien sûr, une carte n'est qu'une représentation à un instant T ; l'avantage de l'IA est de la mettre à jour en continu, comme une cartographie vivante. Cette capacité est un changement de paradigme : on passe de rapports statiques annuels sur les risques à une sorte de tableau de bord quasi-temps réel. Cela exige une bonne intégration des systèmes d'information et le partage inter-agences (renseignement, forces armées, protection civile…), ce qui en soi est un défi organisationnel.

En résumé, la cartographie des risques contextuels grâce à l'IA offre aux décideurs une vision d'ensemble éclairée par les données, leur permettant d'être proactifs plutôt que réactifs. C'est en quelque sorte un prolongement moderne de la cartographie stratégique traditionnelle, mais enrichie par l'intelligence des machines. Tant que ces cartes sont interprétées avec discernement (l'IA peut se tromper ou omettre des facteurs intangibles), elles constituent un support précieux pour la planification sécuritaire, diplomatique et humanitaire.

5. Gestion des flux migratoires

Les migrations massives, qu'elles soient provoquées par les conflits, l'instabilité ou les catastrophes, constituent un défi majeur pour la stabilité régionale et la sécurité humaine. La gestion des flux migratoires implique de sauver des vies, d'assurer un accueil digne aux réfugiés, tout en prévenant les infiltrations de combattants et en maintenant la cohésion sociale dans les zones d'arrivée. L'IA peut aider à mieux comprendre, prévoir et gérer ces mouvements de population complexes, en complément des efforts diplomatiques et humanitaires.

Un apport fondamental de l'IA est la prévision des flux migratoires. En identifiant les déterminants qui poussent les populations à partir (violence, effondrement économique,

sécheresse), des modèles prédictifs peuvent estimer l'ampleur et la direction des migrations futures. Par exemple, en surveillant les indicateurs dans un pays fragile (combats s'intensifiant dans telle province, récoltes catastrophiques signalées par imagerie satellite), l'IA pourrait alerter qu'un exode d'un groupe de personnes vers des frontières avoisinantes est probable dans les prochaines semaines. Du point de vue sécuritaire, les pays de destination peuvent aussi anticiper la pression sur leurs frontières et coordonner avec les agences onusiennes une réponse équilibrant contrôle et assistance.

L'IA contribue aussi à la détection en temps réel des mouvements. Des systèmes combinant images satellite, capteurs frontaliers et données mobiles peuvent repérer les colonnes de migrants en marche ou l'afflux soudain à un poste-frontière. Cela permet de déclencher des opérations de secours rapidement (par exemple, dépêcher des unités pour porter assistance à un groupe bloqué dans le désert avant qu'il ne meure de soif). Évidemment, cela doit se faire avec précaution pour ne pas donner l'impression d'une surveillance intrusive des populations vulnérables, le but étant humanitaire avant tout.

Dans un cadre plus structurel, l'IA peut aider à optimiser la gestion des camps et des itinéraires migratoires. Par exemple, en étudiant les données sur la durée de séjour dans les centres de transit, le taux de renvoi ou de réinstallation, l'IA peut recommander des améliorations : tels campements temporaires deviennent surchargés, mieux vaut en ouvrir un autre en prévision ; tel processus administratif crée un goulot d'étranglement provoquant des attroupements.

Un angle intéressant est la lutte contre les passeurs et les trafiquants qui exploitent les flux migratoires. Les réseaux criminels profitent souvent du désespoir des migrants, organisent des traversées dangereuses, voire du trafic d'êtres humains. L'IA peut analyser les communications interceptées ou les schémas financiers pour aider à démanteler ces réseaux. Par exemple, en identifiant sur les réseaux sociaux les publications ciblant les candidats à l'exil avec de fausses promesses de passage, et en soutenant les forces de l'ordre pour remonter jusqu'aux recruteurs et passeurs. Cela concourt à la sécurité des migrants eux-mêmes et à la stabilité, car ces activités criminelles alimentent la corruption et la violence.

En synthèse, la gestion des flux migratoires par l'IA est un exercice d'équilibre entre humanité et sécurité. L'IA offre une sorte de *radar humanitaire* permettant de voir venir et de mieux organiser l'accueil, ce qui profite tant aux migrants (conditions d'accueil dignes) qu'aux pays hôtes ou de transit (moins de chaos, meilleure allocation des ressources). Elle offre aussi un *filtre de sûreté* pour isoler les éléments dangereux se cachant dans ces flux, sans criminaliser les masses. Utilisée sagement, l'IA peut aider à transformer un phénomène souvent géré dans l'urgence en un processus plus gouverné, prévisible et coopératif, réduisant ainsi les tensions internationales autour de la question migratoire.

6. Surveillance de frontières

Le contrôle et la surveillance des frontières sont des attributs essentiels de la souveraineté étatique, mais ils deviennent de plus en plus ardus face aux menaces transfrontalières modernes: infiltrations de groupes armés, trafics illicites, migrations irrégulières, etc. Les frontières s'étendant parfois sur des milliers de kilomètres de terrains divers (déserts, montagnes, forêts) ne peuvent être surveillées efficacement par des moyens humains classiques seuls [13-14]. C'est là qu'intervient l'IA, au cœur de ce qu'on appelle souvent les *« frontières intelligentes »* [13].

L'IA, couplée à des réseaux de capteurs, permet d'établir une vigilance permanente le long des frontières. Des caméras à détection de mouvement dotées d'algorithmes de vision peuvent distinguer entre un animal, un véhicule ou un être humain traversant une ligne frontière, de jour comme de nuit, et alerter immédiatement les postes de garde en cas d'intrusion suspecte.

Au-delà de la détection, l'IA est utile pour anticiper les tentatives de passage. En analysant les tendances (périodes de l'année, points de passage favoris) et en combinant avec des informations de renseignement, un logiciel peut pointer les segments frontaliers les plus probables pour une tentative à venir. Par exemple, noter qu'à chaque veille de nouvel an, une augmentation de passages clandestins est observée dans tel secteur boisé. Avec cette connaissance, le commandement peut renforcer la veille IA/humaine sur cette période critique.

Sur les frontières maritimes, l'IA s'avère tout aussi cruciale. Les garde-côtes utilisent des systèmes de radar et d'imagerie pour surveiller les côtes ; l'IA peut filtrer le bruit de la mer et détecter de petits canots souvent indétectables par radar classique. L'agence Frontex en Europe teste des algorithmes pour repérer les *boat people* sur la Méditerranée via des drones ou satellites, afin de guider les navires de secours [15]. De même, pour la lutte contre la pêche illégale ou l'infiltration de commandos par mer, l'IA aide à discriminer une embarcation anormale parmi des milliers de signaux marins.

Un autre aspect est la gestion automatisée des postes-frontières officiels. Les systèmes d'IA de reconnaissance faciale et d'analyse du comportement peuvent accélérer les contrôles tout en repérant de potentiels individus suspects (par ex, quelqu'un faisant l'objet d'un avis de recherche ou d'une notice de sécurité).

Ces technologies, déjà présentes dans certains aéroports, posent des questions de vie privée mais peuvent augmenter l'efficacité. À un poste terrestre très fréquenté (comme entre deux pays voisins avec beaucoup d'échanges), l'IA peut fluidifier les passages réguliers (couloir rapide pour les camions connus, etc.) tout en allouant l'attention humaine aux cas atypiques ou à risque.

Là où l'IA excelle aussi, c'est dans la fusion des données frontalières. Par exemple, agréger les informations d'un capteur sismique enterré (détectant des pas ou véhicules), d'un drone aérien et d'une patrouille au sol, pour donner une image unifiée d'une tentative d'intrusion. Plutôt que chaque source séparée, l'IA combine et réduit le *brouillard informationnel*, ne donnant l'alerte que lorsque l'ensemble converge vers un événement significatif. Ceci évite les fausses alertes à répétition qui pourraient sinon saturer les gardes et les conduire à l'erreur.

Un cas d'usage sophistiqué est celui de la frontière virtuelle. Par exemple, dans des zones montagneuses sans clôture physique, on peut imaginer une frontière "délimitée" par un réseau invisible de capteurs et de drones. L'IA gère ce réseau : un drone suit automatiquement une cible repérée par un capteur au sol, pendant qu'un autre va se recharger, etc. Cela crée une sorte de barrière active modulable. Le but n'est pas de militariser outrancièrement les frontières, mais de multiplier les capacités de détection sans exiger une présence humaine partout.

En définitive, l'IA redéfinit la *frontière* non plus comme une simple ligne défensive statique, mais comme un espace intelligent où l'information circule et où la détection-prévention prime sur la réaction tardive. Cela permet d'améliorer la sécurité tout en potentiellement réduisant l'usage de la force (car on peut intercepter tôt, sans confrontation, ou dissuader via la vigilance affichée).

Après cette exploration détaillée des opportunités et cas d'usage de l'IA dans le domaine militaire hors systèmes d'armes létaux autonomes, il apparaît clairement que ces technologies peuvent jouer un rôle transformateur positif. Du renseignement à la logistique, de la prévention des conflits à la gestion des crises humanitaires, l'IA offre aux décideurs et aux intervenants de nouveaux moyens d'agir plus efficacement et avec davantage d'anticipation. Pour autant, la réalisation de ces promesses n'est pas automatique : elle dépend de choix judicieux, d'un encadrement éthique et d'une gestion attentive des risques. C'est pourquoi la section suivante sera consacrée aux défis, techniques, éthiques, stratégiques et juridiques, qu'il convient de relever afin de garantir une intégration responsable de l'IA dans le domaine militaire.

III. Défis de l'IA dans le domaine militaire (hors SALA)

Malgré les atouts manifestes de l'intelligence artificielle dans le domaine militaire, son intégration soulève de nombreux défis qu'il est impératif d'identifier et d'affronter. Ceux-ci sont de diverses natures et souvent interconnectés : techniques (fiabilité des systèmes, vulnérabilités informatiques, etc.), éthiques (respect de l'humain, biais, vie privée), stratégiques (déséquilibres de pouvoir, risque d'escalade) et juridiques (vide normatif, responsabilisation).

Cette section examine chacun de ces quatre grands types de défis, en détaillant leurs manifestations et en s'appuyant sur des analyses et recommandations issues de rapports onusiens. L'objectif est double : prévenir les dérives potentielles liées à l'IA militaire et proposer des pistes pour s'assurer qu'elle demeure un outil au service de la paix et non un facteur de risque supplémentaire [16].

1. Défis techniques

a. Opacité algorithmique et biais cognitifs

Un des premiers obstacles techniques à l'adoption sereine de l'IA en milieu militaire est ce que l'on appelle souvent la « boîte noire » algorithmique. De nombreux systèmes d'IA, en particulier ceux fondés sur des techniques d'apprentissage profond (deep learning), fonctionnent de manière interne trop complexe pour être comprise aisément par un humain. Ils peuvent fournir une recommandation ou prendre une décision (par exemple identifier une cible sur une image) sans qu'il soit clair *comment* ils sont parvenus à ce résultat. Cette opacité algorithmique pose problème dans un contexte de défense où la confiance et l'explicabilité sont essentielles. Les analystes et responsables militaires ont naturellement du mal à se fier à une analyse qu'ils ne peuvent pas expliquer ou vérifier par des raisonnements clairs. Imaginons un logiciel annonçant : *« ce véhicule est hostile avec une probabilité de 92% »*. Si on ignore que c'est parce qu'il a confondu une ombre sur l'image avec une arme (cas d'erreur classique), on pourrait agir sur une base erronée [17].

Cette opacité se double souvent de biais dans les algorithmes. Les biais peuvent provenir des données d'entraînement ou de la conception même du modèle. Ainsi, un système de reconnaissance faciale entraîné majoritairement sur des visages d'une origine ethnique précise identifiera moins bien les visages d'autres origines ethniques, un biais documenté dans de nombreuses études. Transposé au militaire, cela pourrait signifier qu'un système d'identification de personnes fonctionne mal sur certaines populations, entraînant des erreurs potentiellement graves (confondre un civil avec un suspect parce que son visage est mal reconnu, par exemple).

Les biais cognitifs des opérateurs humains interagissent également avec l'opacité de l'IA. Un phénomène connu est celui de la confiance excessive en l'automatisation. Si un système IA a souvent donné de bons résultats, les opérateurs pourraient lui faire aveuglément confiance même lorsqu'il se trompe. Cette tendance est renforcée si l'algorithme ne donne pas ses raisons, ce qui empêche l'humain de détecter une éventuelle absurdité. Par exemple, lors d'une identification d'objectif par IA, des militaires pourraient suivre la recommandation de tir sans oser la remettre en question, surtout sous stress, alors qu'une analyse humaine plus poussée aurait pu révéler une erreur (cible civile mal classée). Des experts ont souligné que l'interaction humain-machine en contexte militaire est délicate : trop de défiance vis-à-vis de l'IA annule ses bénéfices, trop de confiance aveugle peut conduire à des drames.

Sur le plan technique, la recherche en IA explicable (XAI) s'efforce de développer des modèles plus transparents ou capables de fournir des justifications en langage clair de leurs résultats. Par exemple, un système de vision pourrait indiquer : *« j'ai détecté cette personne comme hostile*

car elle porte un uniforme semblable à tel modèle », ce qui permettrait de vérifier la validité de la raison. Ce champ doit être encouragé dans les applications militaires.

Sur la question des biais, il s'agit d'assurer une diversité des données d'entraînement et de tester systématiquement les modèles sur des scénarios variés. Les forces armées qui développent des IA devraient intégrer dans leurs équipes des spécialistes des sciences sociales ou des droits humains pour détecter les biais potentiels et les corriger.

En somme, le défi de l'opacité et des biais est celui de la fiabilité intelligible. Pour que l'IA gagne la confiance des militaires tout en préservant les principes d'éthique et de non-discrimination, il faut qu'elle soit aussi *comprise* et *maîtrisée* que possible.

b. Dépendance technologique

L'introduction massive de l'IA dans les systèmes militaires s'accompagne d'un risque de dépendance technologique accru.

Du point de vue opérationnel, s'appuyer sur l'IA pour des tâches critiques peut conduire à une érosion des compétences humaines. Si, par exemple, la planification logistique est entièrement optimisée par un logiciel IA, que se passera-t-il le jour où le logiciel est indisponible ou défaillant ? Les militaires sauront-ils encore élaborer un plan logistique complexe "à l'ancienne" ? Dans l'histoire militaire, on a vu des exemples de troupes surspécialisées perdre des aptitudes de base (par exemple, la navigation astrale presque oubliée à l'ère du GPS, ce qui devient problématique si le GPS est brouillé). Une inquiétude est qu'en s'accoutumant à l'assistance de l'IA, les forces perdent en flexibilité cognitive et en capacité d'improvisation sans ces béquilles numériques. Le défi est donc de conserver un bon équilibre homme-machine : l'IA doit être une aide, non un substitut complet. L'entraînement devrait intégrer des scénarios de "retour au manuel" au cas où la technologie est indisponible [19-20].

Un autre aspect est la dépendance à l'infrastructure technologique. Les IA militaires reposent sur des réseaux de télécommunications, des bases de données, des infrastructures de calcul (HPC). En situation de conflit intense, ces éléments peuvent être détruits, brouillés ou sabotés. Si toute la chaîne de commandement et de renseignement est construite autour d'une colonne vertébrale IA, la neutralisation de cette dernière par l'ennemi pourrait paralyser les opérations. Les armées doivent donc prévoir des modes dégradés de fonctionnement. Ce constat rejoint la problématique de la résilience : s'assurer que, même sans IA, on peut continuer la mission (certes moins efficacement). De plus, cela pose la question du secours en cas de bug : un bug critique dans un système d'IA embarqué (par exemple un véhicule autonome bloqué à cause d'une mauvaise mise à jour logicielle) ne doit pas mettre des vies en danger faute de solution de contournement.

Il est recommandé de s'assurer qu'il existe toujours un plan B sans IA ou avec une autre IA de secours. Par analogie, dans l'aviation civile, bien que le pilote automatique soit omniprésent, les pilotes sont entraînés à piloter manuellement si besoin et l'avion a des systèmes de secours. Les forces armées devraient adopter une approche similaire (par exemple, conserver des compétences de navigation traditionnelle, ou avoir des équipements analogiques de réserve).

c. Vulnérabilité aux cyberattaques adverses

En déployant des systèmes d'IA, les forces armées s'exposent à de nouvelles formes d'attaques de la part d'adversaires cherchant à déjouer ou corrompre ces systèmes. L'IA elle-même peut devenir la cible de tactiques hostiles spécifiques, qu'il s'agisse de manipuler ses capteurs, d'empoisonner ses données ou de pirater ses algorithmes.

On parle d'attaques adverses (*adversarial attacks*) [21-22] lorsqu'un acteur malveillant introduit des perturbations conçues pour tromper l'algorithme : par exemple, peindre un véhicule avec un motif particulier pour le faire passer inaperçu d'un système de vision, ou diffuser de fausses données dans les canaux de renseignement pour induire l'IA en erreur.

Des chercheurs ont montré qu'il est possible de duper une IA de reconnaissance visuelle en lui présentant des images altérées de façon imperceptible pour l'œil humain, mais que le modèle classifie à tort. Dans un contexte militaire, un ennemi pourrait exploiter ce procédé pour faire confondre à un système autonome une cible légitime avec un objet anodin, ou inversement.

Les systèmes d'IA connectés (drones, robots, centres de fusion de données) présentent en outre des surfaces d'attaque cyber. Un adversaire sophistiqué pourrait tenter de pirater le réseau de communication reliant les capteurs et l'IA, afin de soit le neutraliser (déni de service) soit en prendre le contrôle. Si une IA militaire est infiltrée, l'ennemi pourrait s'en servir pour diffuser de fausses informations aux commandants, créant chaos et méfiance interne.

La robustesse des systèmes d'IA face à ces attaques est donc un défi technique de premier plan. Les développeurs doivent anticiper les ruses adverses et entraîner les modèles avec des *données bruitées* ou des *exemples adverses* connus pour les immuniser au mieux.

De plus, sur le terrain, les IA doivent être entourées de contre-mesures de cybersécurité traditionnelles (chiffrement des communications, pare-feu, détection d'intrusion) pour empêcher un accès non autorisé. Un rapport de l'UNIDIR sur la sécurité de l'IA [18] souligne que les problèmes de sûreté (safety) et de sécurité (security) des systèmes automatisés sont étroitement liés, et que la capacité d'adaptation d'une IA à des situations imprévues doit être améliorée pour éviter les effondrements face à des environnements légèrement modifiés.

En parallèle, la communauté internationale commence à échanger sur les bonnes pratiques pour contrer les attaques adverses. Des exercices conjoints entre alliés simulent par exemple des cyberattaques sur des réseaux militaires intelligents afin d'en tirer des enseignements. Mais tant que les normes globales de maîtrise de l'IA ne seront pas en place, ce défi technique persistera: chaque nouvelle capacité IA devra être testée non seulement dans des conditions nominales, mais aussi sous l'épreuve d'un ennemi astucieux. Cette réalité impose de concevoir des IA résilientes, capables de continuer à fonctionner de manière dégradée en environnement hostile.

d. Manque de capacités humaines et infrastructurelles

L'adoption de l'IA à grande échelle dans les armées se heurte dans de nombreux pays à un manque de ressources humaines qualifiées et d'infrastructures adaptées. Concevoir, déployer et maintenir des systèmes d'IA performants requiert des spécialistes en informatique, en science des données, en cybersécurité, que toutes les forces armées ne possèdent pas en nombre suffisant.

À cela s'ajoute le déficit d'infrastructures technologiques dans certains États : centres de données sécurisés, connectivité haut débit sur tout le territoire, capteurs et plateformes numériques de collecte de données. Sans ces prérequis, les meilleurs algorithmes du monde ne pourront être exploités. Par exemple, si une armée d'un pays en développement ne dispose pas de satellites ou de drones pour recueillir des images, elle ne pourra pas alimenter un système d'IA d'analyse géospatiale ; ou si les réseaux dans une zone d'opération sont instables, une IA nécessitant des mises à jour continues risque de ne pas fonctionner correctement.

Les pays les moins avancés pourraient se retrouver incapables de suivre le rythme. L'investissement requis pour implémenter l'IA militaire dans des pays en voie de développement, déjà confrontés à des défis d'infrastructures de base, est significatif, d'où

l'importance des programmes de coopération internationale : partage de connaissances, transfert de technologie encadré, aide à la formation de data scientistes locaux, etc.

Par ailleurs, même dans des armées technologiquement avancées, il faut éviter la perte de compétences traditionnelles. Les militaires doivent conserver des savoir-faire fondamentaux non technologiques pour pallier une éventuelle défaillance (capacité de navigation sans GPS, discernement terrain sans aide informatique, etc.). Cela nécessite un effort de formation double, qui peut être lourd : d'un côté apprendre à utiliser de nouveaux outils IA, de l'autre continuer à exercer l'esprit critique "manuel".

Au niveau infrastructurel, mettre en place l'IA implique aussi de gérer les données de manière adéquate. Beaucoup d'armées ne disposent pas encore de *data centers* dédiés ou de politiques de gouvernance des données (stockage, classification, partage). Sans infrastructure solide pour collecter, nettoyer et stocker les données, l'IA ne peut être déployée efficacement. Or, bâtir ces infrastructures est coûteux et peut prendre du temps (installation de serveurs, adoption de clouds sécurisés, etc.). Des partenariats public-privé peuvent être une solution, à condition de bien négocier la propriété et la confidentialité des données de défense.

En résumé, le défi des capacités humaines et infrastructurelles est un appel à l'investissement et à la formation. Il souligne que l'introduction réussie de l'IA n'est pas qu'une affaire d'algorithmes, mais requiert une transformation organisationnelle profonde. L'élaboration des *plans nationaux sur l'IA* incluant la dimension défense est à encourager, de sorte à planifier sur le long terme le développement des compétences et des infrastructures nécessaires. Cela s'accompagne d'une coordination internationale pour que ces avancées profitent à tous. Dans ce sens, l'ONU pourrait investir pour favoriser l'échanger de bonnes pratiques entre les Etats.

e. Disponibilité et qualité des données

La performance d'un système d'IA dépend avant tout de la qualité, de la quantité et de la pertinence des données sur lesquelles il est entraîné et qu'il exploite en opération. Or, dans le domaine de la défense, obtenir des données fiables et suffisamment abondantes est un défi en soi.

Premièrement, certaines situations sont peu fréquentes ou inédites, ce qui limite la possibilité d'entraîner l'IA. Par exemple, une IA chargée de prévoir les réactions d'un adversaire lors d'une crise ne dispose souvent que de quelques cas historiques pour apprendre (chaque guerre ou confrontation diplomatique est unique). Cela peut conduire l'algorithme à sur-généraliser à partir de trop peu d'exemples.

De même, pour entraîner une IA à reconnaître des engins explosifs improvisés (EEI), il faut idéalement des milliers d'images d'EEI, or chaque théâtre produit un nombre limité de modèles différents. Les militaires doivent donc parfois recourir à de la synthèse de données (générer artificiellement des données similaires) pour augmenter les jeux d'entraînement, mais cela peut introduire des biais.

Deuxièmement, les données disponibles peuvent être partiales ou biaisées. Par exemple, si l'on construit un modèle prédictif de troubles civils en se basant sur les rapports de renseignement d'un pays, ceux-ci peuvent refléter les préjugés ou centres d'intérêt de ce service de renseignement, et non la réalité complète. L'IA apprendra alors ces biais. La représentativité des données d'entraînement est donc cruciale, elles doivent couvrir la diversité des environnements et comportements possibles, sans quoi l'IA sera en défaut dès qu'elle sera hors de sa zone de confort.

Un autre problème est la pénurie de données étiquetées. Beaucoup d'algorithmes supervisés nécessitent des données annotées par des humains (par exemple, des heures de vidéo où chaque

image a été labellisée : “civil”, “combattant”, etc.). Or ce travail d’annotation est titanesque et souvent classifié dans un contexte militaire. Il faut mobiliser des analystes pour le faire, ce qui prend du temps et mobilise des ressources. Des méthodes d’auto-apprentissage (non supervisé) ou d’apprentissage faiblement supervisé sont en développement pour atténuer ce besoin, mais elles restent moins précises. La disponibilité de données fiables devient ainsi un goulot d’étranglement : il ne suffit pas d’avoir des senseurs et des archives, il faut que ces données soient exploitables.

Beaucoup de pays en développement manquent de bases de données numérisées exploitables dans le domaine de la sécurité. Par exemple, pas de registre complet des incidents passés, pas de cartographie numérique fine du terrain, etc. Ils partent donc quasiment de zéro pour nourrir une IA, contrairement à des pays qui depuis des décennies accumulent des renseignements numérisés. Ce retard ne se rattrape pas aisément. D’où l’importance de programmes internationaux de partage de données lorsqu’ils peuvent être mutualisés sans risque pour la souveraineté.

Enfin, même lorsque les données existent, un défi technique est d’assurer leur intégrité et leur actualité. Des données obsolètes peuvent induire l’IA en erreur (par exemple., un plan de ville qui ne tient pas compte des nouvelles constructions). D’où la nécessité de chaînons de vérification humaine et d’hygiène des données : traçabilité de la source, recoupement entre sources multiples pour repérer d’éventuelles incohérences (exemple : un capteur signale un événement qu’aucun autre ne confirme, suspicion de leurre).

De plus, le manque de données de bonne qualité est un frein direct à l’efficacité et à la précision des algorithmes. Par conséquent, il est recommandé d’investir dès maintenant dans la gestion des données de défense : numérisation sécurisée des archives, standardisation des formats de rapport pour faciliter l’analyse automatisée, et création de *data lakes* (réservoirs de données) interopérables entre les différentes branches armées.

En conclusion, le vieux problème du renseignement, disposer d’une information fiable, à jour, exhaustive, reste présent à l’ère de l’IA, simplement transposé sur le terrain numérique. L’IA peut certes aider à combler des trous (en extrapolant des tendances malgré des données incomplètes) mais cela a ses limites. Ce défi technique appelle à la patience et à la rigueur : prendre le temps de construire des bases de données solides conditionnera le succès de l’IA en situation opérationnelle. Les armées et la communauté du renseignement doivent collaborer étroitement avec les spécialistes de la donnée pour créer cette fondation, et les Nations Unies peuvent jouer un rôle de facilitateur en élaborant des normes sur le partage sécurisé de données à des fins de paix et de sécurité.

2. Défis éthiques

a. Primauté du facteur humain et de l’éthique

Au cœur des préoccupations éthiques se trouve la question fondamentale du rôle de l’humain dans l’usage de la force et la conduite de la guerre. Les principes du droit international humanitaire (DIH) et de la Charte des Nations Unies reposent sur l’idée que la décision d’employer la force létale, en particulier, doit être prise avec discernement par des êtres humains responsables. L’émergence de l’IA soulève la crainte d’une érosion de cette primauté humaine. Même en dehors du cas extrême des SALA, l’IA peut influencer des choix d’engagement, de ciblage, de neutralisation de menaces. Il est impératif de rappeler que l’éthique doit guider l’IA, et non l’inverse.

L’IA devrait rester un outil d’aide et non de substitution totale. Par exemple, une IA pourra prioriser des cibles potentielles sur une liste, mais la décision finale de frapper devrait être validée par un opérateur humain entraîné, capable d’appliquer les règles d’engagement et de

faire la part des choses (annuler si doute, vérifier la proportionnalité de l'attaque, etc.). De même, dans la chaîne de commandement, on peut accepter qu'une IA propose un plan tactique, mais c'est au commandant d'en évaluer la conformité avec le droit et les objectifs politiques avant de l'approuver.

Insister sur la primauté humaine, c'est aussi veiller à ce que les valeurs éthiques fondamentales, respect de la vie, dignité de la personne, nécessité et proportionnalité en conflit, soient intégrées comme contraintes dans la conception des systèmes. Par exemple, un véhicule autonome patrouillant doit être programmé pour éviter à tout prix de heurter des civils, quitte à laisser s'échapper un assaillant. Un algorithme de tri de cibles doit pouvoir *reconnaître l'abstention* : détecter qu'il n'y a pas de cible légitime claire et recommander de ne pas tirer. Cela implique un encodage de *règles éthiques* dans les IA (on parle parfois de "gouverneurs éthiques").

Par ailleurs, la formation éthique du personnel devient d'autant plus cruciale à l'ère de l'IA. Si un militaire comprend mal comment fonctionne l'IA, il pourrait être tenté de lui déléguer la responsabilité morale : "c'est la machine qui a décidé". Cela est inacceptable du point de vue du droit de la guerre, la responsabilité ne peut être diluée ou évacuée sur une entité non-humaine. Les militaires doivent être éduqués à conserver leur sens critique et leur conscience morale en présence d'IA. Il s'agit d'éviter le phénomène de *« dilution de responsabilité »* qui pourrait autrement survenir (chacun suivant les recommandations automatiques sans se sentir comptable du résultat final).

Enfin, la primauté de l'éthique implique de fixer des limites rouges à l'utilisation de l'IA. Par exemple, la communauté internationale pourrait convenir que certaines décisions ne doivent jamais être prises par une IA seule, peu importe les avancées techniques. C'est en partie l'objet du débat sur les SALA, où tracer la limite de l'autonomie. En dehors des SALA, on peut penser à des interdits comme : pas d'IA pour décider de traitements dégradants sur des prisonniers (toujours un humain doit valider tout acte sur un prisonnier, en vertu des Conventions de Genève), ou pas d'IA pour évaluer la "valeur" d'une vie humaine contre une autre (un calcul froid de dommage collatéral acceptable ne saurait se substituer à une appréciation humaine). Ces discussions éthiques doivent réunir militaires, juristes, philosophes et sociétés civiles.

L'UNESCO a produit en 2021 une Recommandation sur l'éthique de l'IA qui, bien que générale, pose des principes (transparence, justice, supervision humaine, responsabilité) applicables au secteur militaire [23]. Les États pourraient s'en inspirer pour leur doctrine de défense.

En somme, la guerre n'est pas un simple problème d'optimisation que l'IA peut résoudre, mais un phénomène profondément humain qui devrait rester sous contrôle humain. Les tragédies d'erreurs passées (bombardements de civils par méprise, etc.) appellent non pas à éliminer l'humain du processus, mais au contraire à l'y recentrer avec de meilleurs outils d'aide à la décision. Préserver l'humain, c'est en définitive préserver la possibilité de la compassion, de la désescalade et du pardon, qualités qu'aucune machine ne possédera jamais.

b. Discrimination et biais sociétaux

Malgré l'apparente objectivité des machines, l'IA peut involontairement reproduire ou amplifier des discriminations existantes dans la société. C'est un enjeu éthique majeur : s'assurer que l'adoption de l'IA ne vienne pas renforcer des inégalités ou des injustices, que ce soit en temps de paix ou dans la conduite des opérations militaires.

Dans le contexte des conflits armés, la discrimination algorithmique peut se manifester de manière tragique. Imaginons une IA qui aide à choisir des cibles à frapper : si par construction elle prend moins en compte la présence de certaines catégories de personnes (par exemple, elle détecte mal les femmes et enfants parce que les données d'entraînement se concentraient sur

des silhouettes d'hommes armés), elle pourrait conduire à des frappes moins précautionneuses dans des zones où se trouvent majoritairement des femmes et enfants, donc à des dommages collatéraux plus élevés pour ces populations, c'est une forme de discrimination par l'outil. Inversement, une IA pourrait surévaluer la dangerosité de jeunes hommes d'une certaine origine sur un checkpoint, menant à des traitements plus musclés à leur égard, ce qui viole le principe d'égalité de protection.

La non-discrimination est un principe cardinal du droit international des droits de l'Homme. Son respect doit être assuré même dans les algorithmes. Pour ce faire, il faut mettre en place des procédures d'audit algorithmique régulières : tester le système avec des cas variés, vérifier les taux de faux positifs/négatifs pour différents groupes démographiques, etc. Si l'on détecte un biais, il faut corriger le modèle ou ajuster son usage.

Il convient aussi d'être prudent sur l'usage de certaines caractéristiques dans les modèles. Par exemple, inclure l'ethnie ou la religion comme variable d'entrée d'un algorithme prédictif de comportement serait extrêmement délicat, car cela officialiserait une discrimination.

Toutefois, même sans les inclure explicitement, l'IA peut les déduire indirectement via d'autres données (lieu de résidence, style vestimentaire…). D'où l'importance de concevoir des IA centrées sur des critères pertinents et légitimes (antécédents judiciaires avérés par exemple) et d'exclure au maximum les attributs sensibles.

En définitive, garantir la justice algorithmique est un impératif moral qui recoupe l'intérêt pratique : une IA discriminante sape la confiance du public et peut même provoquer des troubles. Les forces armées doivent au contraire apparaître exemplaires dans l'usage équitable de la technologie. Cela nécessite transparence, audits indépendants, et correction proactive des biais.

c. Inexplicabilité des décisions algorithmiques

Lorsque des décisions prises par ou avec l'aide de l'IA ne peuvent être expliquées de manière compréhensible, cela soulève un problème éthique de transparence et d'imputabilité.

Du point de vue éthique, chaque acte de guerre ou de coercition devrait pouvoir être justifié a posteriori. Si un tir de drone tue des civils, il faut expliquer ce qui a mené à cette bavure pour en répondre et pour éviter sa répétition. Or, si l'analyse de situation était largement automatisée, on risque d'entendre *« c'est la faute à l'algorithme, on ne sait pas pourquoi il a fait ça »*. Cette situation crée un vide de responsabilité inacceptable. Les victimes ou leurs familles sont en droit d'exiger des comptes (droit à la justice, à un recours effectif). Ne pas pouvoir leur fournir une explication compréhensible serait une double peine : non seulement il y a préjudice, mais en plus on leur dénie la clarté sur les causes de ce préjudice.

L'inexplicabilité complique aussi les procédures légales. Imaginons un tribunal qui enquête sur un incident militaire. Si l'officier dit "le système d'IA m'a dit que c'était une cible légitime", le juge va demander "sur quel fondement le système a conclu cela ?" Si personne ne sait (pas de logiciel intelligible, modèle complexe), comment déterminer s'il y a eu négligence, ou violation du DIH ? Cela peut entraver l'application du droit. Dans un contexte national, cela heurterait aussi le principe de redevabilité démocratique : les parlements qui contrôlent l'action de l'exécutif en matière de défense doivent pouvoir examiner et comprendre les informations et processus de décision qui ont conduit à telle opération. S'ils se heurtent à un mur technique, le contrôle démocratique s'affaiblit.

Il y a également une dimension de confiance morale. Les militaires sur le terrain, tout comme le public, auront du mal à faire confiance à un système s'ils ont le sentiment qu'il agit de façon incompréhensible. La confiance est essentielle pour l'acceptation éthique : on accepte plus

volontiers une décision (même négative) si on en comprend les raisons. C'est un principe de base en justice : la décision d'un juge est motivée. De même, on pourrait estimer que chaque décision importante où l'IA a joué un rôle devrait être *motivée* de façon intelligible.

Face à ce défi, l'éthique rejoint la technique : développer des IA explicables (*Explainable AI* ou XAI) [24]. Par exemple, des chercheurs travaillent sur des réseaux de neurones capables de fournir en sortie non seulement un résultat mais aussi une "attention map" mettant en évidence quelles parties des données ont le plus contribué à la décision. Dans l'analyse d'images, cela pourrait se traduire par un surlignage : "j'ai identifié cette personne comme combattant parce que j'ai détecté la forme d'une arme dans sa main". Ce type d'explication visuelle ou textuelle permet à un opérateur de valider ou non l'évaluation de l'IA. Si l'algorithme surligne en fait un objet anodin (par exemple une pelle confondue avec un fusil), l'humain peut corriger.

De même, on promeut des modèles hybrides mêlant règles logiques et apprentissage machine, où la partie règles (exprimées en "si…alors") encadre l'algorithme et peut être auditée facilement. Par exemple, on pourrait avoir une règle intangible : "ne jamais classer comme cible un individu levant les mains ou portant un symbole médical", couplée à une IA identifiant les formes. L'explicabilité vient ici de la règle.

Les organisations internationales encouragent aussi la création de registres d'algorithmes et de journaux d'événements : chaque recommandation ou décision majeure prise par un système d'IA devrait être enregistrée avec son contexte, afin de pouvoir être analysée ensuite par des experts indépendants. Par analogie aux "boîtes noires" des avions, on aurait les "boîtes noires" des IA militaires. En cas d'incident, on examine ces logs pour reconstituer le cheminement décisionnel. Si on découvre que l'algorithme a fait X après l'entrée Y, on pourra chercher à comprendre pourquoi en laboratoire, même si sur le moment ce n'était pas clair.

En définitive, lever le voile sur la boîte noire est non seulement un défi technique mais un impératif éthique pour maintenir la responsabilité et la confiance. C'est en alignant l'innovation sur la transparence que l'on s'assurera que l'IA demeure un outil au service du droit et non un facteur d'opacité ou d'arbitraire.

d. Exacerbation des campagnes de désinformation

L'IA offre des moyens redoutablement efficaces pour créer et propager de la désinformation, ce qui pose un défi éthique et sécuritaire majeur. Dans les conflits modernes, et même en deçà du seuil de la guerre, la bataille de l'information fait rage. La manipulation de l'opinion publique, les rumeurs virales, les faux contenus visant à semer la panique ou la haine peuvent amplifier dramatiquement les impacts d'un conflit ou entraver les missions de paix de l'ONU. L'apparition des *deepfakes* (vidéos ultraréalistes truquées par IA) a marqué un tournant : on ne peut plus se fier à l'évidence de nos sens pour juger de l'authenticité d'une information.

On imagine en effet le scénario cauchemardesque d'une fausse vidéo montrant un chef d'État prononcer une déclaration de guerre, ou un commandant de Casques bleus commettre des exactions, la réaction immédiate pourrait être violente avant que le faux ne soit démasqué. De tels artifices exploitent les émotions et la rapidité de diffusion sur les réseaux sociaux pour créer des "déflagrations informationnelles" difficiles à rattraper.

Éthiquement, l'usage délibéré de la désinformation est condamnable, en particulier quand il vise à faire du tort aux civils (propagande haineuse, etc.). Une guerre cognitive entièrement pilotée par IA, visant à briser la volonté de l'adversaire en inondant son peuple de fausses nouvelles, serait une perspective très sombre qu'il faudrait éviter.

Au contraire, sur le plan positif, l'IA peut servir à révéler la désinformation. Des algorithmes parcourent déjà les réseaux pour détecter les signes de campagnes coordonnées (comptes automatisés, diffusion massive d'un même message, fausses images identifiées par analyse de pixel ou par comparaison à des bases de données d'images originales). L'équilibre éthique est délicat : il faut contrer la désinformation sans basculer dans la censure injustifiée.

Les forces armées et missions de paix peuvent utiliser l'IA pour circuler des informations vérifiées rapidement après un incident, pour couper court aux rumeurs malveillantes. L'IA peut même aider à prévoir la propagation d'une infox : quels groupes vont probablement la reprendre, dans quelles régions elle va se répandre, permettant de cibler la communication corrective.

Cependant, un risque éthique est que l'opinion perde confiance en toute information, vraie ou fausse, face à cette bataille algorithmique. On parle de *"société post-vérité"* où plus rien n'est cru. C'est en soi un danger pour la démocratie et la paix, car la défiance généralisée profite aux fauteurs de troubles. Il est donc crucial que les acteurs légitimes (États, ONU) conservent une réputation de fiabilité. S'ils devaient eux-mêmes être pris à manipuler l'info via IA, leur crédibilité serait ruinée. D'où la nécessité d'une charte éthique : par exemple, certaines armées se sont engagées publiquement à ne pas fabriquer de deepfakes dans leur communication.

Sur le plan juridique, il y a des discussions émergentes pour assimiler certaines attaques informationnelles graves à des violations du droit international (voire à des actes d'agression lorsqu'elles causent des dommages comparables à des attaques armées) [25].

En conclusion, l'éthique et la sécurité convergent sur ce point : il faut prévenir une "infodémie" militaire. L'IA, instrument à double tranchant, doit être canalisée pour défendre la vérité plutôt que semer le mensonge. Il s'agit d'une lutte asymétrique car la fabrication de mensonges attire plus que leur démystification, mais c'est une cause vitale pour la stabilité internationale. Les États peuvent coopérer, dans ce sens, en partageant des algorithmes de détection d'infox, alertant mutuellement en cas de campagne toxique transfrontalière, et en éduquant le public (la "résilience cognitive" des populations est aussi un bouclier) [26]. La technologie doit aller de pair avec l'élévation du sens critique citoyen, afin que l'IA ne trouve pas un terreau facile dans des sociétés vulnérables à la manipulation.

e. Responsabilité et redevabilité

L'introduction de l'IA dans la prise de décision militaire complexifie la question de la responsabilité en cas d'erreur ou de violation. Néanmoins, dans la pratique, déterminer *qui*, au sein de l'appareil d'État ou de la chaîne de commandement, doit être tenu pour responsable peut devenir épineux. Par exemple, si un algorithme biaisé conduit à une bavure, pointe-t-on la responsabilité du soldat opérateur (qui a peut-être suivi son entraînement et ne pouvait pas détecter le biais) ? Du chef qui a validé l'usage de ce système ? Du fabricant privé de l'IA qui a fourni un outil défectueux ? De l'officier qui n'a pas assez supervisé la décision automatique ? Une situation de flou peut créer ce qu'on appelle un "accountability gap", un vide de responsabilité où chacun peut se renvoyer la faute.

Pour faire face à ce défi, plusieurs approches sont envisagées :

- **Traçabilité des décisions** : garder des logs qui permettent d'identifier qui a validé quoi, à quel moment. Cela aide à remonter la chaîne des responsabilités. Par exemple, si une IA a proposé une cible et qu'un analyste l'a explicitement confirmée, la responsabilité de l'analyste est engagée en plus de celle du commandant.

- **Clarté doctrinale** : établir dans la doctrine d'emploi de l'IA que l'humain est toujours responsable. Certains pays ont déjà intégré dans leurs principes d'IA de défense que l'IA n'enlève en rien la responsabilité des opérateurs et des commandants.

- **Mécanismes de revue et de contrôle** : mettre en place au sein des armées des comités d'éthique ou de redevabilité qui examinent les incidents liés à l'IA. Un peu comme les commissions d'enquête, mais spécialisées. Cela permet d'identifier s'il y a eu faute humaine (par exemple le non-respect d'une procédure de double vérification) ou si le problème est technique (appeler alors le fournisseur à corriger, voire engager sa responsabilité contractuelle).

- **Mise à jour des règles d'engagement** : celles-ci peuvent préciser explicitement les marges de manœuvre de l'IA et les moments où l'aval humain est requis. Ainsi, si ces règles sont violées (par ex un opérateur laisse l'IA cibler sans autorisation supérieure), la responsabilité disciplinaire est clairement attribuable.

En somme, l'éthique de la responsabilité exige de ne pas diluer l'imputabilité dans les méandres technologiques. Au contraire, plus la technologie est présente, plus il faut renforcer les garde-fous de responsabilité.

Par exemple, on pourrait définir que toute unité utilisant de l'IA doit avoir un officier spécifiquement chargé de superviser son emploi (un "officier de contrôle IA") et qui sera redevable en cas de souci, comme un officier de sécurité l'est pour des manipulations d'armes. Ce genre d'innovations institutionnelles peut aider à garder un visage humain (et responsable) derrière chaque décision, évitant que la machine ne devienne un bouc émissaire commode ou, pire, un acteur incontrôlé.

3. Défis stratégiques

a. Risques de surconfiance et erreurs stratégiques

L'introduction de l'IA dans la planification et la conduite des opérations peut engendrer un excès de confiance dans les prédictions ou recommandations automatiques, menant potentiellement à des erreurs aux conséquences stratégiques graves. L'histoire militaire est jalonnée de cas où une foi excessive dans une nouvelle technologie ou doctrine a conduit à des déconvenues majeures. Avec l'IA, cette dynamique pourrait se répéter si les décideurs surévaluent la fiabilité de leurs systèmes.

Un premier risque est la mauvaise interprétation des capacités réelles de l'IA. Par exemple, si une armée pense que son système prédictif peut anticiper à coup sûr les mouvements adverses, elle pourrait prendre des décisions audacieuses (comme diviser ses forces, ou lancer une offensive préventive) en se basant sur ces prédictions. Si celles-ci se révèlent fausses, la posture prise devient dangereuse. En 2020, un rapport notait que *« surestimer ce que l'IA peut accomplir »* peut pousser à l'utiliser au-delà de ses capacités et conduire à des échecs retentissants [27]. Ainsi, l'IA pourrait inciter à la témérité stratégique si on la considère comme une sorte d'oracle infaillible.

Le deuxième risque est lié à l'automatisation du temps décisionnel. L'IA permet d'accélérer la boucle « OODA » (Observation, Orientation, Decision, Action) au point que des décisions stratégiques pourraient être prises très vite. Cela peut être un avantage tactique, mais au niveau stratégique, la précipitation est souvent mauvaise conseillère. Un système qui indiquerait en quelques secondes "il faut riposter maintenant" peut créer une pression politique pour agir sans avoir laissé le temps à la diplomatie ou à la réflexion. On parle du problème de la *compression du temps de réaction* : par peur d'être dominées par la vitesse de l'IA de l'autre camp, les parties pourraient se mettre à réagir quasi automatiquement les unes aux autres, avec un risque

d'escalade fulgurante. Au plus haut niveau, cela pourrait signifier perdre des opportunités de désamorcer une crise par le dialogue, parce que l'IA "a déjà agi".

La complexité accrue des systèmes peut aussi générer des erreurs stratégiques. Plus il y a d'automatisation, plus la chaîne logistique et opérationnelle dépend de conditions techniques. Un adversaire intelligent cherchera à *saboter l'IA* de manière stratégique, par exemple en leurrant les satellites ou en hackant une base de données clé, de sorte à provoquer une décision erronée de grande ampleur (comme engager des forces au mauvais endroit). Ce type de stratagème peut causer des *erreurs de calcul* dramatiques si le haut commandement ne détecte pas la supercherie à temps.

Que faire pour mitiger ces risques ? D'abord, cultiver une culture du questionnement vis-à-vis de l'IA. En temps de paix, faire des « *red team »* qui simulent ce qui se passe si l'IA se trompe complètement. En temps de crise, ne pas désactiver l'esprit critique sous prétexte que "la machine l'a dit".

Ensuite, au niveau international, pourrait émerger une sorte de mesure de confiance : un accord tacite pour garder l'humain dans la boucle des décisions stratégiques, un peu comme des "lignes directes" diplomatiques pour clarifier les intentions et éviter les accidents. Par exemple, une transparence partielle sur le fait que tel système d'alerte est supervisé humainement pourrait rassurer l'adversaire.

On peut aussi imaginer l'IA mise au service de la *stabilité* : par exemple, des algorithmes pour détecter automatiquement les signaux d'une erreur et bloquer les réactions automatiques. Des sortes de gardiens d'escalade programmés pour inciter à la retenue si une décision rapide va clairement à l'encontre d'indicateurs diplomatiques (ex : si les canaux diplomatiques sont actifs, c'est qu'un accord est en cours, donc ne pas lancer d'attaque même si la modélisation tactique dit le contraire).

Enfin, un sursaut stratégique pourrait venir de la reconnaissance de la faillibilité de l'IA. Tout comme on a admis que la foudre pouvait brouiller un radar, il faut toujours garder en tête que l'IA peut se tromper. Ce principe de précaution stratégique doit être intégré dans les war games des Etats-Majors : tester "que fait-on si notre IA nous induit en erreur ?", un peu comme on fait des exercices de perte de GPS. Avoir un *Plan B manuel* atténuera la surconfiance.

En conclusion, l'IA ne supprime pas le "brouillard de la guerre", elle en change juste la nature. Le risque de surprise ou d'erreur ne disparaît pas, il peut même s'accroître par excès de certitudes numériques.

b. Accessibilité aux acteurs non étatiques et groupes terroristes

Contrairement aux systèmes d'armes sophistiqués traditionnels (missiles balistiques, avions de combat…), l'IA est largement disponible sous forme de logiciels commerciaux bon marché, ce qui la rend accessible à des acteurs non étatiques, y compris criminels et terroristes.

C'est un défi stratégique de taille : des groupes armés peuvent utiliser des drones équipés d'une IA rudimentaire pour des attaques ciblées, ou des réseaux extrémistes peuvent déployer des bots et deepfakes afin de radicaliser et coordonner des partisans à distance.

On a déjà vu des cas de l'emploi de quadrirotors armés par des organisations terroristes, l'ajout de capacités d'autonomie et de vision automatique pourrait accroître la létalité de ces engins improvisés. De même, un groupuscule pourrait tenter de pirater un système militaire national à l'aide d'outils d'IA disponibles sur le dark web, menaçant la sécurité collective.

Cette démocratisation de l'IA bouscule le monopole étatique de la violence légitime. Elle complexifie les opérations de maintien de la paix (des Casques bleus pourraient faire face à des adversaires dotés d'IA d'attaque ou de camouflage) et la lutte antiterroriste (les services de renseignement doivent intégrer ce paramètre). Stratégie et éthique convergent ici : il faut priver les acteurs malveillants des moyens d'abuser de l'IA.

Si la communauté internationale s'accorde sur des normes d'usage responsable de l'IA, les comportements déviants isolés de terroristes apparaîtront d'autant plus clairement comme criminels et illégitimes, facilitant un consensus pour les combattre.

En outre, en développant en amont des contre-mesures IA (IA défensive pour filtrer la propagande, brouiller les IA adverses, etc.), les États pourront garder l'ascendant technologique sur ces acteurs non étatiques.

4. Défis juridiques

Sur le plan du droit international, l'IA militaire évolue aujourd'hui dans un vide normatif relatif. Aucune convention universelle ne traite spécifiquement de son développement ou de son usage (discussions en cours sur les SALA dans le cadre de la Convention sur certaines armes classiques, qui restent limitées au domaine létal autonome, ou encore la résolution de l'Assemblée Générale de l'ONU sur l'intelligence artificielle dans le domaine militaire et ses conséquences pour la paix et la sécurité internationales adoptée le 24 décembre 2024).

Cette absence de cadre juridique clair soulève des inquiétudes quant à la possibilité de voir émerger des pratiques divergentes, voire contraires aux valeurs internationales, sans moyen de les proscrire efficacement.

Certes, le droit existant, en particulier le droit international humanitaire, s'applique à *toutes* les armes et méthodes de guerre, qu'elles impliquent de l'IA ou non. Les principes de distinction, proportionnalité, précaution dans l'attaque, ainsi que l'interdiction de causer des maux superflus ou de cibler des civils, demeurent en vigueur.

Toutefois, ces règles ont été conçues dans le contexte d'actions humaines ; leur interprétation dans des scénarios hautement automatisés n'est pas consensuelle. Par exemple, comment juge-t-on la proportionnalité d'une frappe si l'évaluation des dommages reposait sur un calcul d'IA? Jusqu'où le devoir de précaution impose-t-il de vérifier les résultats fournis par une IA avant d'agir ? Ce sont des questions débattues que la communauté internationale continue d'aborder.

IV. Conclusion

L'intelligence artificielle est en passe de transformer profondément le domaine militaire, apportant des opportunités considérables pour mieux prévoir les crises, protéger les populations et assister les soldats sur le terrain. Des applications prometteuses se profilent dans le renseignement, la logistique, la prise de décision, la cybersécurité ou l'entraînement, y compris dans des contextes tels que le maintien de la paix, la gestion des catastrophes ou la prévention des conflits ethniques.

Cependant, l'IA n'est pas une panacée magique dénuée de risques. Mal utilisée ou déployée sans garde-fous, elle pourrait exacerber les tensions, éroder les principes humanitaires et engendrer de nouvelles menaces pour la stabilité mondiale.

Les défis techniques (fiabilité des systèmes, protection contre le piratage, qualité des données), les défis éthiques (transparence, absence de biais, responsabilité), les défis stratégiques (éviter

l’escalade incontrôlée, réduire la fracture technologique) et les défis juridiques (combler le vide normatif et inventer des mécanismes de contrôle) sont immenses.

Chacun de ces défis appelle une réponse concertée de la communauté internationale. Face à des technologies transversales et globales, c’est par la coopération, le dialogue et la mise en place de règles communes que nous pourrons minimiser les risques tout en maximisant les bénéfices de l’IA.

Dans un monde déjà ébranlé par les conflits, les pandémies et les changements climatiques, l’IA ne doit pas devenir un nouveau facteur de division ou de destruction. Au contraire, maîtrisée collectivement, elle peut contribuer à désamorcer les conflits (par une meilleure prévention), à réduire les bavures et dommages collatéraux (par une précision accrue et une aide à la décision plus rationnelle), et à renforcer les opérations de paix (par une efficacité logistique et une sécurité des personnels améliorées).

En conclusion, l’application de l’intelligence artificielle dans le domaine militaire est un tournant historique. Elle s’accompagne d’opportunités inédites pour construire un monde plus sûr, où les décisions seraient mieux informées et où les soldats comme les civils seraient mieux protégés.

Références

[1] M.A.-É. internationales, undefined 1989, Institut des Nations Unies pour la Recherche sur le Désarmement (UNIDIR). Désarmement: Problèmes relatif à l'espace extra-atmosphérique. New York, UNIDIR, Erudit.Org M Albert Études Internationales, 1989 erudit.Org (n.d.). https://doi.org/10.7202/702531ar.

[2] S. Zeriouh, M. Boutahri, S. El Yamani, A. Roukhe, Targets Classification on Multispectral Images using Connectionists Methods, Academia.Edu (n.d.). https://www.academia.edu/download/82589859/10.11648.j.ajpa.20150303.14.pdf (accessed April 4, 2025).

[3] O. bin Laden, ISR offensif dans la guerre contre le terrorisme, Airuniversity.Af.Edu (n.d.). https://www.airuniversity.af.edu/Portals/10/ASPJ_French/journals_O/Volume-02_Issue-1/danskine.pdf (accessed April 4, 2025).

[4] D. Micle, F. Deiac, A. Olar, R. Drența, C.F.- Agriculture, undefined 2021, Research on innovative business plan. Smart cattle farming using artificial intelligent robotic process automation, Mdpi.Com (n.d.). https://www.mdpi.com/2077-0472/11/5/430 (accessed April 4, 2025).

[5] W. Aryatwijuka, H. Mutebi, P. Nagawa, B. Tukamuhabwa, Artificial Intelligence and Humanitarian Supply Chain Resilience: Mediating Effect of Localized Logistics Capacity, (2024). https://www.researchgate.net/profile/Samuel-Mayanja/publication/384941643_Artificial_Intelligence_and_Humanitarian_Supply_Chain_Resilience_Mediating_Effect_of_Localized_Logistics_Capacity/links/67122cc709ba2d0c7606f856/Artificial-Intelligence-and-Humanitarian-Supply-Chain-Resilience-Mediating-Effect-of-Localized-Logistics-Capacity.pdf (accessed April 4, 2025).

[6] B. Claverie, G.D.-A.J. of Management, undefined 2022, C2-Command and Control: A System of Systems to Control Complexity, Researchgate.Net (n.d.). https://www.researchgate.net/profile/Gilles-Desclaux/publication/363583535_C2_-_Command_and_Control_A_System_of_Systems_to_Control_Complexity/links/639cb203e42faa7e75cb0d61/C2-Command-and-Control-A-System-of-Systems-to-Control-Complexity.pdf (accessed April 4, 2025).

[7] M.A.-I.T. on automatic control, undefined 2003, Command and control (C2) theory: A challenge to control science, Ieeexplore.Ieee.Org (n.d.). https://ieeexplore.ieee.org/abstract/document/1104607/ (accessed April 4, 2025).

[8] A. Hamdouchi, A. Idri, Enhancing IoT security through boosting and feature reduction techniques for multiclass intrusion detection, Neural Computing and Applications 2025 (2025) 1–24. https://doi.org/10.1007/S00521-025-11001-2.

[9] A. Hamdouchi, A. Idri, Comprehensive Evaluation of Federated Learning Configurations for Intrusion Detection in IoT Contexts, 2024 World Conference on Complex Systems (WCCS) (2024) 1–6. https://doi.org/10.1109/WCCS62745.2024.10765581.

[10] N. Kandybko, O. Zemlina, … V.P.-A. conference, undefined 2023, Application of artificial intelligence technologies in the context of military security and civil infrastructure protection, Pubs.Aip.Org (n.d.). https://pubs.aip.org/aip/acp/article-abstract/2476/1/030019/2891077 (accessed April 4, 2025).

[11] G. Sakr, O. Hakme, The Impact of Artificial Intelligence on Military and Security in the MENA, AI in the Middle East for Growth and Business (2025) 363–383.

https://doi.org/10.1007/978-3-031-75589-7_21.

[12] Strategy for the Digital Transformation of UN Peacekeeping | United Nations Peacekeeping, (n.d.). https://peacekeeping.un.org/en/strategy-digital-transformation-of-un-peacekeeping (accessed April 5, 2025).

[13] T. Ige, A. Kolade, O. Kolade, Enhancing Border Security and Countering Terrorism Through Computer Vision: A Field of Artificial Intelligence, Lecture Notes in Networks and Systems 597 LNNS (2023) 656–666. https://doi.org/10.1007/978-3-031-21438-7_54.

[14] M. Sharma, C.R.S. Kumar, Machine Learning-Based Smart Surveillance and Intrusion Detection System for National Geographic Borders, Lecture Notes in Electrical Engineering 806 (2022) 165–176. https://doi.org/10.1007/978-981-16-6448-9_19.

[15] Search and Rescue, (n.d.). https://www.frontex.europa.eu/what-we-do/operations/search-and-rescue/ (accessed April 5, 2025).

[16] A. Bin Rashid, A.K. Kausik, A. Al Hassan Sunny, M.H. Bappy, Artificial Intelligence in the Military: An Overview of the Capabilities, Applications, and Challenges, International Journal of Intelligent Systems 2023 (2023). https://doi.org/10.1155/2023/8676366.

[17] M. Hutson, The opacity of artificial intelligence makes it hard to tell when decision-making is biased, IEEE Spectr 58 (2021) 40–45. https://doi.org/10.1109/MSPEC.2021.9340114.

[18] A. Mensah-Sackey, H. Giezendanner, P. Holtom, Aperçu de la gestion des armes et des munitions en Afrique: rapport sur l'état d'avancement 2023, (2023). https://unidir.org/publication/apercu-de-la-gestion-des-armes-et-des-munitions-en-afrique-rapport-sur-letat-davancement-2023/ (accessed April 5, 2025).

[19] J. Souza, R. Avelino, … S. da S.-A. and B.D. in E., undefined 2023, 11. Artificial intelligence: dependency, Books.Google.Com (n.d.). https://books.google.com/books?hl=en&lr=&id=HuDoEAAAQBAJ&oi=fnd&pg=PA228&dq=Artificial+intelligence+Technological+dependence&ots=KmmIbW0VSc&sig=D-COzHMgstsFux8ZiN0F6ZC6Nnk (accessed April 5, 2025).

[20] M.W. Khan, M.A. Destek, Z. Khan, Income Inequality and Artificial Intelligence: Globalization and age dependency for developed countries, Soc Indic Res (2025). https://doi.org/10.1007/S11205-024-03493-7.

[21] H. Baniecki, P.B.-I. Fusion, undefined 2024, Adversarial attacks and defenses in explainable artificial intelligence: A survey, Elsevier (n.d.). https://www.sciencedirect.com/science/article/pii/S1566253524000812 (accessed April 5, 2025).

[22] S. Qiu, Q. Liu, S. Zhou, C.W.-A. Sciences, undefined 2019, Review of artificial intelligence adversarial attack and defense technologies, Mdpi.Com (n.d.). https://www.mdpi.com/2076-3417/9/5/909 (accessed April 5, 2025).

[23] Intelligence artificielle au Maroc: l'UNESCO énonce des recommandations | SNRT News, (n.d.). https://snrtnews.com/fr/article/intelligence-artificielle-au-maroc-lunesco-enonce-des-recommandations-97072 (accessed April 5, 2025).

[24] A. Das, G. Student Member, P. Rad, S. Member, Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey, (2020). https://arxiv.org/abs/2006.11371v2 (accessed April 5, 2025).

[25] UNESCO - EU Partnership Grows: Tackling Disinformation and Hate Speech Globally | UNESCO, (n.d.). https://www.unesco.org/en/articles/unesco-eu-partnership-grows-tackling-disinformation-and-hate-speech-globally (accessed April 5, 2025).

[26] U. Nations, Pacte numérique mondial | Nations Unies, (n.d.). https://www.un.org/fr/summit-of-the-future/global-digital-compact (accessed April 5, 2025).

[27] Zouinar, Moustafa, Évolutions de l'Intelligence Artificielle : quels enjeux pour l'activité humaine et la relation Humain-Machine au travail ?, Http://Journals.Openedition.Org/Activites (2020). https://doi.org/10.4000/ACTIVITES.4941.