

L'intelligence artificielle (IA) est une discipline fondée notamment sur les mathématiques en vue de simuler des facultés mentales humaines avec des machines. Elle regroupe un ensemble de techniques fondées sur la modélisation de connaissances et sur l'exploitation de données. Elle peut avoir des applications variées dans le domaine militaire : aide à la décision en planification et en conduite ; combat collaboratif ; cyberdéfense et influence ; logistique, soutien et maintien en condition opérationnelle ; renseignement ; robotique et autonomie ; administration et santé. Ainsi l'usage de l'IA dans le domaine militaire est plus vaste que sa seule utilisation dans les systèmes d'armes létaux autonomes (SALA).

I- **L'impact de l'IA dans le domaine militaire sur la paix et la sécurité internationale, entre opportunités et défis**

Alors que l'IA a commencé à transformer notre monde, la révolution technologique qu'elle constitue apporte également de profonds changements dans les affaires militaires. Les innovations qu'elle est susceptible d'entraîner offrent des opportunités sans précédent, tout en soulevant des défis significatifs, auxquels la communauté internationale doit répondre.

A. **Des opportunités à exploiter**

1. Une IA de défense pour renforcer les capacités opérationnelles

i) L'IA pour aider à la planification et à la prise de décision

L'IA permet d'apprécier une situation de manière plus rapide et plus fiable tant au niveau tactique qu'opérationnel. Dans une phase d'élaboration de la manœuvre, des systèmes traitant une quantité importante de données, captées sur une durée longue, permettent de dissiper d'autant le brouillard de la guerre (présence de forces ennemies, configuration du terrain, présence de civils) et de déterminer les points clés de la manœuvre à effectuer. Par exemple, les armées françaises travaillent à l'intégration de l'IA dans les processus de planification. Au niveau tactique, elle travaille aussi sur l'exploitation, au travers de l'IA, des bases de données dont elles disposent en matière d'événements MUNEX (Munitions et explosifs) afin que cette IA (nourrie par de très nombreuses données) puisse leur permettre de développer des outils prédictifs quant aux menaces potentielles dans une zone définie.

ii) L'IA pour soutenir l'humain

Intégrée à la simulation, l'IA permettra d'améliorer la formation différenciée des individus. Par exemple, dans les armées françaises, le système « IA FPN » est mis en avant afin de maximiser les chances de succès de la formation des pilotes. Face à l'exigence, à la durée et au coût de la formation des personnels navigants, l'enjeu de ce système est d'alerter les instructeurs suffisamment tôt et de cibler les efforts à réaliser pour surmonter les difficultés que l'élève pilote rencontre et ainsi maximiser ses chances de succès, grâce à une analyse des données recueillies lors des formations (vol ou simulation).

En libérant l'homme de tâches particulièrement chronophages, l'IA lui permettra aussi de se concentrer sur celles à haute valeur ajoutée et de réduire les erreurs humaines qui surviennent largement lorsqu'il doit traiter une grande quantité de données. Par exemple, les armées françaises développent le système « Oreille d'or », traitant massivement des données acoustiques pour en faire le tri et orienter l'attention des opérateurs oreilles d'or sur les seuls signaux à valeur ajoutée, sur lesquels ils pourront apporter leur compétence métier.

iii) L'IA pour contrer nos vulnérabilités dans le domaine des technologies de l'information et de la communication

Les technologies d'IA peuvent servir à la cybersécurité et prévenir des cyber attaques, en optimisant les processus de lutte informatique défensive, de recherche de failles automatisée ou de traitement de masses de données, dans l'objectif de rendre la défense des systèmes d'information et des systèmes d'armes plus efficace et réactive face à un incident.

Face aux capacités de production de fausses informations et aux graves répercussions que peuvent en avoir les diffusions sur les réseaux sociaux ou Internet, l'IA permet de détecter de manière automatique sur ces derniers des passages falsifiés dans des supports vidéo, images ou audio. Elle offre ainsi la possibilité de détecter plus vite toutes désinformation et d'y répondre de manière plus réactive. Par exemple, un système de détection des *deepfakes* et des informations mensongères contre les forces armées.

2. L'IA pour renforcer le respect du droit et la stabilité internationale

i) L'IA pour favoriser la mise en œuvre du droit international humanitaire et la protection des personnes et biens protégés

L'IA dans le domaine militaire peut présenter des possibilités et des avantages potentiels en matière de respect du droit international humanitaire (DIH), notamment pour la protection des personnes et biens civils en période de conflit armé. En permettant une meilleure appréciation de l'environnement des opérations à un niveau tactique et opératif, l'IA peut fortement contribuer à renforcer la mise en œuvre des principes cardinaux du DIH. Tout d'abord, la distinction entre combattants et non combattants pourrait être facilitée le traitement massif d'un large nombre de données par l'IA. De plus, la proportionnalité pourrait être mieux appréciée grâce à la fiabilité de l'analyse scientifique et au renforcement par l'IA de la méthode d'évaluation des risques de pertes et dommages civils (Collateral Damage Estimation, CDE). Enfin, l'IA permettra d'améliorer le respect du principe de précaution, en offrant une palette différenciée dans le choix des armements et moyens. Ainsi, l'IA pourrait également aider à garantir que les dommages collatéraux civils ne soient pas excessifs par rapport à l'avantage militaire poursuivi, en offrant une large gamme de moyens et de méthodes (allant des moyens non létaux aux moyens plus conventionnels).

L'IA pourra agir directement pour la protection des personnes et biens civils en période de conflit armé et le retour des populations civiles en aidant notamment au déminage des territoires. Par exemple, l'armée de Terre française travaillent aujourd'hui avec des entreprises qui investissent dans le champ de la reconnaissance de zones minées, utilisant des drones, munis de capteurs (photographie, infrarouge, radar de pénétration de sol, photogrammétrie) couplés à une IA, permettant de déterminer la présence ou non de mines, avec l'affichage d'un indice de probabilité.

L'IA peut aussi être utile aux missions d'aide humanitaire et de secours. Par exemple, l'amélioration des capacités de traitement des images satellitaires ou des diverses données montrant les dégâts causés aux infrastructures (contenus publiés sur les médias, sur les réseaux sociaux, etc.) peuvent permettre une évaluation plus précise des dommages en temps réel et ainsi améliorer la qualité de la réponse humanitaire et des efforts de reconstruction. L'IA peut aussi aider à prévoir les mouvements de population pour mieux s'y préparer. Enfin, les drones et les robots pilotés par l'IA ont révolutionné les efforts de recherche et de sauvetage, grâce aux capteurs et algorithmes d'IA qui détectent des signes de vie et transmettent des informations cruciales aux intervenants. Ces systèmes peuvent naviguer sur des terrains difficiles et dans des environnements dangereux, augmentant ainsi la rapidité et l'efficacité de la localisation et du sauvetage des survivants.

ii) L'IA pour renforcer la maîtrise des armements

L'IA pourrait jouer un rôle clé dans la maîtrise des armements et les mécanismes de vérification en rendant ces processus plus rapides, précis et efficaces dans les principaux domaines d'application.

Dans le domaine de la surveillance et la détection des armes, l'IA peut être utilisée pour analyser des images satellites, des données radar et des signaux électromagnétiques pour identifier des installations militaires suspectes, des mouvements de troupes ou des tests d'armes ou en aidant les inspecteurs à vérifier le nombre de matériels militaires déployés. Par exemple, les algorithmes de reconnaissance d'images à base d'IA permettent de détecter des missiles mobiles, des lancements clandestins ou des changements dans des sites de production d'armements d'une manière plus rapide et plus pertinente.

Des systèmes basés sur l'IA peuvent aussi surveiller les essais d'armes chimiques et biologiques, en repérant des anomalies dans les données environnementales. L'IA pourrait également accroître la pertinence des algorithmes de sélection des sites à inspecter, et, dans le cas des armes chimiques, l'IA pourrait anticiper des voies de synthèses et précurseurs non connus, afin de mieux les identifier et d'en prévenir la mise au point. Ainsi, l'IA pourra être utilisée pour repérer d'éventuelles violations des traités de désarmement et de maîtrise des armements.

Pour contrôler les exportations d'armes, l'IA pourrait améliorer la traçabilité des armements en analysant la masse de données relatives aux flux commerciaux, les bases de données douanières, et même les images satellites de convois pour détecter d'éventuelles livraisons illégales d'armes. Cela permettrait d'identifier et de prévenir le trafic d'armes vers des groupes terroristes ou des États sous embargo.

iii) L'IA pour renforcer la prévention, le maintien et la consolidation de la paix

L'IA pourrait être utilisée pour réduire le risque d'escalade involontaire, en analysant les crises militaires, en fournissant des scénarios prédictifs ou en détectant et corrigeant les erreurs de classification des menaces, évitant ainsi les fausses alertes et les escalades involontaires.

L'IA permettrait également de mener des opérations de maintien de la paix mieux adaptées et donc plus efficace. Le traitement du langage naturel (NLP) et les outils de traduction sont susceptibles d'améliorer l'interopérabilité des forces multinationales et de faciliter la communication avec la population locale et ainsi de lutter contre la désinformation. Par exemple, le système « Resistance » de traduction instantanée des langues étrangères mis en avant par les armées françaises a pour objectif de permettre, en opération, hors-connexion et sans réseau, de communiquer avec les populations civiles et les autorités locales.

B. Des risques à atténuer

1. Des risques propres à la technologie

Les caractéristiques de l'IA supposent des risques propres aux technologies qui l'intègrent contre lesquels des mesures spécifiques d'atténuation doivent être mises en place.

Les techniques d'apprentissage présentent différents risques de biais: de biais involontaires en particulier lorsque les données d'apprentissage ne sont pas représentatives (par exemple, biais ethnique dans des données de population) ; de biais volontaires si un tiers malveillant a pu modifier les données d'apprentissage ou le modèle afin de produire un résultat anormal éventuellement sur demande ; de reconstitution de données particulièrement sensibles d'apprentissage (rétro-conception), en particulier si un tiers dispose de connaissances sur les techniques ou les outils d'apprentissage ayant été utilisés. Les techniques d'apprentissage présentent également des risques de résultats opaques ou peu explicables auxquels il sera difficile pour l'homme de faire confiance dans des systèmes critiques.

Enfin se pose la question d'une consommation exponentielle de ressources en énergie (composants électroniques, eau pour le refroidissement, fonctionnement des modèles, etc.) engendrée par une compétition pour maximiser la puissance des modèles génératifs. Or ces ressources ne sont pas infinies et leur exploitation soulève des questions cruciales sur la soutenabilité de ce modèle.

2. L'aggravation des risques pour la sécurité et la stabilité internationale

Dans de mauvaises mains, l'IA peut aggraver certains risques à la sécurité et la stabilité internationale, pour lesquels il faudra adapter les mesures d'atténuation des risques (mesures de confiance, maîtrise des armements, lutte contre la prolifération). Des ruptures technologiques dans les domaines de la détection, de l'agression ou de la prise de décision, pourraient favoriser de nouveaux déséquilibres propices à des scénarios d'escalade (attaques plus sophistiquées, mieux coordonnées, plus offensives, rapides et massives).

La course technologique engagée dans l'IA pourrait entraîner une compétition pour la supériorité militaire et une course aux armements, phénomène amplifié par la dualité et l'extension des applications technologiques de l'IA. Ainsi, elle pourrait permettre à des nations dotées de capacités militaires traditionnelles peu développées de mener plus facilement des actions de combat ou de déstabilisation (aspect niveleur de la technologie). L'IA pourrait aussi amener à une prolifération vers des acteurs non-étatiques, qui pourraient exploiter de futures technologies civiles sur étagère pour élaborer des surprises tactiques à partir de moyens inédits (sans nécessairement devoir être en pointe sur l'ensemble du spectre technologique de l'IA, la maîtrise de quelques technologies-clés peut suffire).

Cette technologie pourrait également jouer sur la stabilité internationale via des risques d'extension des opérations d'influence et des actions dans le champ informationnel (détournement de flux d'informations, décrédibilisation médiatique, *fake news*, etc.) ; de multiplication des actions hostiles « à haute-fréquence » dans le domaine cyber (attaques coordonnées, actions de déception), ou de paralysie des capacités de commandement (par la neutralisation, le leurrage ou le détournement des technologies).

Enfin, les applications militaires de l'IA pourraient soulever des problèmes s'agissant du respect du droit international, en particulier le DIH qui repose sur des principes cardinaux. Il s'agit des principes de distinction, de proportionnalité, de nécessité militaire et de précaution dans les attaques. Ensemble, ils forment « *le tissu* » du DIH qui régit la conduite des hostilités. Les applications militaires dotées d'IA comme tout autre armes, moyens et méthodes de combat, devront être développées et employées conformément au DIH. Face aux risques de déresponsabilisation induits par la dépendance aux technologies, le maintien de la responsabilité humaine doit donc être garantie (cf. développements *infra* C(1)(ii)). En effet, le risque de dépendance à une technologie ne doit pas être écarté, nécessitant d'éviter toute déresponsabilisation humaine.

Les technologies basées sur l'IA sont un outil majeur pour les forces armées, en soutien à l'accomplissement de leurs missions. Elles soulèvent d'ores et déjà également des risques et des préoccupations légitimes. Dans ce contexte, notre approche doit être pratique et double : exploiter pleinement le potentiel de l'IA tout en veillant à ce que son développement et son utilisation soient conçus et poursuivis de manière responsable.

II- Les principes et mesures clés pour une « IA responsable » durant tout le cycle de vie

L'IA présentant des avantages mais faisant également peser des risques, il convient d'assurer une application responsable de l'IA dans le domaine militaire, centrée sur l'humain, responsable, sûre, sécurisée et digne de confiance, utilisée dans le respect du droit international, et cela sur l'ensemble du cycle de vie des capacités d'IA appliquées dans le domaine militaire.

A. Développer une IA respectueuse du droit international humanitaire

1. Adapter les examens de licéité

Les technologies d'IA conçues et utilisées dans le cadre d'opérations militaires pour directement ou indirectement blesser, tuer, endommager ou neutraliser des personnes ou des biens sont soumises au même régime juridique, que tous les autres moyens de guerre.

Afin de veiller à ce que les technologies d'IA répondant à la définition des armes, moyens et méthodes de guerre soient conformes au DIH, l'article 36 du Protocole Additionnel I aux Conventions de Genève de 1949 prévoit que « *Dans l'étude, la mise au point, l'acquisition ou l'adoption d'une nouvelle arme, de nouveaux moyens ou d'une nouvelle méthode de guerre, une Haute Partie contractante a l'obligation de déterminer si l'emploi en serait interdit, dans certaines circonstances ou en toutes circonstances, par les dispositions du présent Protocole ou par toute autre règle du droit international applicable à cette Haute Partie contractante.* » Ainsi, cet examen permet de prévenir l'acquisition de technologies d'IA militaire qui seraient contraires au droit international en toutes circonstances ou, le cas échéant, d'imposer les restrictions d'emploi nécessaires pour assurer la conformité de leur utilisation.

Si cet examen s'applique pleinement aux nouvelles technologies, notamment l'IA militaire, les modalités précises de cet examen devront s'adapter aux spécificités de ces technologies (*machine learning* et modification de l'algorithme ; compétences techniques, ré-évaluation après déploiement sur le terrain).

2. Effectuer des réexamens appropriés

La mise en œuvre d'un processus d'examen de licéité de manière itérative, pourrait s'avérer particulièrement pertinente pour encadrer le développement ou l'intégration d'une technologie d'IA dans un système utilisé par les armées, en particulier dans des fonctions comme l'identification, la sélection ou l'engagement. Cet examen doit être conduit, en tant que de besoin, lors des différentes phases du cycle de vie d'un système d'arme : en phase de préparation, de réalisation mais aussi d'utilisation opérationnelle. Un réexamen doit se tenir lorsqu'un même dispositif subit des innovations ou intègre de nouveaux composants susceptibles de modifier significativement les effets provoqués.

B. Développer une IA fiable et sécurisée

1. Evaluer, qualifier et certifier les systèmes

Une IA de confiance repose notamment sur la rigueur de conception des systèmes. Ces systèmes doivent être évalués et qualifiés, au juste niveau, c'est-à-dire selon des exigences proportionnées aux bénéfices que l'on veut tirer et aux risques à éviter. Ce « juste niveau » est établi en fonction de la criticité des fonctions remplies et découle d'une analyse de risque systématiquement conduite dès la phase de conception. Cette analyse de risque doit permettre d'identifier parmi les différentes fonctions, celles qui seront les plus critiques pour en déduire des exigences en termes de développement, de qualification et de suivi en service.

De plus, les systèmes intégrant des technologies d'IA doivent être associés à des cas d'utilisation définis, pour lesquels ils ont été testés, vérifiés et validés, voire certifiés. La révision de ces vérifications

avec les évolutions des systèmes et de leurs usages doit être envisagée avec une fréquence adaptée aux enjeux. Il importe de mener des études d'ergonomie prenant en compte les conditions réelles d'utilisation de systèmes intégrant des technologies d'IA dans le cadre des opérations, y compris en conditions dégradées.

2. S'appuyer sur des données maîtrisées et souveraines

En amont ou en cours d'utilisation des systèmes, il existe des risques d'atteinte : à l'intégrité des données (par une pollution intentionnelle ou non, qui fausserait les résultats issus de l'apprentissage) ; à la disponibilité des données (qui obérerait le fonctionnement d'une technologie d'IA, nécessitant la comparaison avec des données de référence) ; à la confidentialité des données et/ou des résultats (par le vol d'informations sensibles).

Il convient de mettre en place des parades (tatouages, cryptographie, entraînement spécifique des modèles) et des défenses adaptées. De façon générale, les formations accompagnant la mise en place de systèmes intégrant des technologies d'IA doivent permettre la sensibilisation aux risques de sécurité des systèmes d'information.

3. Corriger et réentraîner les systèmes

Il importe de relever et de caractériser les erreurs rencontrées lors des tests ou de l'usage opérationnel d'un système intégrant des technologies d'IA (relever le contexte d'emploi, la nature de l'erreur, etc.). Il est également nécessaire de sensibiliser les opérateurs à la nécessité du retour d'expérience sur l'utilisation des systèmes d'IA. L'aptitude au réentraînement et à la correction du système, avec des données réelles annotées est cruciale. Elle le sera d'autant plus dans une dynamique où les progrès de la recherche et de la technologie permettront que la fiabilité des technologies d'IA aille croissant. De plus, comme expliqué précédemment, il est nécessaire de penser un processus d'examen de licéité « itératif », afin de vérifier continuellement si le système est conforme à nos obligations internationales.

C. Soumettre l'IA à un contrôle humain approprié et une chaîne de commandement responsable

Le principe de maintien du contrôle humain revêt une importance cardinale eu égard à la nécessité d'assurer le respect du droit international humanitaire, raison pour laquelle la France a promu un engagement à ne pas autoriser qu'une décision de vie ou de mort soit prise par un système d'arme autonome fonctionnant totalement en dehors du contrôle humain et d'une chaîne de commandement responsable, dans une déclaration sur le maintien du contrôle humain dans les systèmes d'armes basés sur l'IA, publiée à l'occasion du Sommet de Paris pour l'action sur l'IA et soutenue par 28 pays.

1. Garantir un contrôle humain indispensable au respect de nos obligations juridiques

i) Garantir la conformité de la prise de décision et de l'action avec le droit

L'opérateur ou le chef militaire doit pouvoir exercer son propre discernement, pour vérifier si les résultats proposés sont justes et conformes aux ordres donnés. De plus, l'appréciation de la situation est nécessaire afin de prendre en compte des considérations juridiques (obligation de distinguer, de prendre des décisions éclairées au regard du droit, de récupérer toutes les informations actualisées, de demander un conseil juridique, de faire ce qui est pratiquement possibles pour minimiser les effets incidents contre les personnes et biens protégés). Le contrôle humain est donc indispensable pour avoir un esprit critique des remontées d'informations produites par les systèmes d'IA et être en mesure de les écarter si nécessaire.

ii) Garantir la responsabilité humaine

La responsabilité humaine dans la conception, le déploiement et l'emploi de technologies d'IA constitue un principe indérogeable. En effet, en cas de violations graves du droit des conflits armés,

l'opérateur ou le chef militaire peut voir sa responsabilité engagée. Il importe donc de formaliser les chaînes de responsabilité de ceux en charges des fonctions de commandement, du contrôle et de l'exécution quelles que soient les fonctions réalisées par le système intégrant des technologies d'IA. Par ailleurs, les autres responsabilités (que ce soit dans le développement, la programmation, l'acquisition ou le déploiement de ces moyens) doivent être clairement définies afin qu'elles puissent être assumées.

2. Appliquer un contrôle humain approprié et évolutif

Le principe de maintien du contrôle humain revêt une importance cardinale eu égard à la nécessité d'assurer le respect du droit international humanitaire, raison pour laquelle la France a promu un engagement à ne pas autoriser qu'une décision de vie ou de mort soit prise par un système d'arme autonome fonctionnant totalement en dehors du contrôle humain et d'une chaîne de commandement responsable, dans une déclaration sur le maintien du contrôle humain dans les systèmes d'armes basés sur l'IA, publiée à l'occasion du Sommet de Paris pour l'action sur l'IA et soutenue par 28 pays.

i) Adapter le contrôle humain aux différentes variables contextuelles et techniques

Analyser et caractériser le contrôle humain approprié, sans pour autant brider les capacités du système intégrant des technologies d'IA (au risque d'en perdre le bénéfice), est une problématique complexe qui doit prendre en compte différents facteurs, humains (tels que la fatigue, le stress, la compétence, etc.), techniques (transparence, aide à la validation, etc.) et contextuels (dynamique et complexité de l'environnement). En effet, dans un contexte de conflit armé, l'environnement opérationnel ainsi que l'intensité du conflit influent, de manière évolutive, sur le niveau de délégation accordé par le commandement. Les technologies d'IA devront permettre, le cas échéant, de faire varier le niveau d'autonomie de certaines fonctions, selon l'appréciation du commandement et de ses délégataires.

ii) Former les chefs militaires et le personnel pour maîtriser ces systèmes

Pour s'assurer que ceux qui auront à utiliser des systèmes intégrant des fonctions d'IA en mesurent bien les tenants et aboutissants, une phase de formation et d'entraînement doit être instaurée avant leur emploi opérationnel. Il s'agit de sensibiliser individuellement l'ensemble du personnel aux apports mais aussi aux risques induits par cette technologie. La formation doit développer le discernement face aux résultats proposés par les systèmes d'IA. Afin que ces systèmes s'insèrent pleinement dans une doctrine d'emploi des forces, il sera nécessaire de former les opérateurs, les chefs militaires et leurs conseillers, aux apports et aux limites de ces systèmes, pour favoriser leur capacité à détecter que la fonction assurée par une technologie d'IA est dégradée, polluée ou insuffisante et qu'il est nécessaire de reprendre la main.

D. Développer une IA durable

1. Protéger la recherche

L'objet et l'horizon des programmes de recherche doivent être ouverts sans que soient édictés a priori des interdictions trop larges. La France considère que pour protéger ses forces, il importe de connaître les dispositifs que ses adversaires pourraient utiliser (même si elle-même s'empêcherait de les utiliser). Il conviendrait ainsi de prêter attention à la surenchère normative qui pourrait constituer un frein à la recherche.

2. Soumettre la recherche à une réflexion éthique

Les recherches relatives aux systèmes d'IA doivent pouvoir être conduites en toute responsabilité. Ces recherches doivent s'accompagner d'une réflexion holistique (juridique, éthique, sécuritaire), en particulier sur la portée et les applications possibles de leurs résultats. La France est devenue la première grande puissance militaire à se doter d'une structure de réflexion permanente sur les enjeux éthiques des nouvelles technologies dans le domaine de la défense. En créant le Comité d'Éthique de

la Défense (COMEDF), elle a démontré son intention de nourrir la réflexion et d'investir dans la recherche, pour rendre l'intelligence artificielle robuste et sûre, et développer des systèmes de façon maîtrisée.

3. Développer une IA frugale

La frugalité peut être définie comme la nécessité d'optimiser un système basé sur l'IA dans un environnement où les ressources sont limitées. Ceci est particulièrement pertinent pour les applications militaires de l'IA, étant donné le contexte dans lequel les forces armées opèrent. Cela implique de privilégier un comportement frugal, c'est-à-dire de ne recourir à l'IA qu'en cas de nécessité, et aussi d'améliorer la résilience et la durabilité des systèmes, ainsi que de contrôler les coûts.

Pour bénéficier des avantages de l'IA tout en atténuant ces risques, nous devons respecter des principes d'une IA responsable, de confiance et contrôlée par l'humain, développée et utilisée dans le respect du droit international, sans obérer la recherche et l'innovation.

III. Un processus dédié pour mettre en place une gouvernance globale visant à opérationnaliser les principes d'une IA responsable

Compte tenu des défis que le développement de l'IA dans le domaine militaire peut représenter pour la paix et la sécurité internationales, la France estime nécessaire que la communauté internationale travaille à la mise en place d'une gouvernance appropriée. Le lancement d'un processus dédié dans les enceintes multilatérales devra répondre aux paramètres suivants.

A. Un processus universel et inclusif

Afin de renforcer la stabilité stratégique mondiale, ce processus doit permettre l'élaboration d'un cadre applicable à tous.

Il conviendra dès lors que les discussions rassemblent toutes les parties prenantes, en particulier les Etats qui développent et emploient les systèmes militaires dotés d'IA, dont la participation est une nécessité absolue. Eu égard à cet impératif, le processus décisionnel devra tenir compte des diverses positions et adopter des règles en ce sens en vue d'assurer le consensus.

Le processus devra également permettre un engagement de tous les acteurs impliqués : le monde industriel qui conçoit et fabrique les systèmes, le domaine scientifique et académique qui recherche et informe dans ce domaine et la société civile qui analyse les risques et propose des solutions. L'engagement avec le secteur privé sera fondamental pour comprendre le fonctionnement des systèmes mobilisant de l'IA, permettant d'éviter des discussions déconnectées du réel et de préserver l'innovation. Donner la possibilité aux acteurs industriels de s'engager participe également à leur responsabilisation. Enfin, cette inclusivité doit permettre aux Etats de fonder leurs discussions sur une connaissance précise des développements technologiques et des enjeux qui en découlent, en mobilisant notamment l'expertise que l'UNIDIR contribue à enrichir par ses activités et études.

L'universalisation et l'inclusivité sont donc essentielles et permettront de renforcer la légitimité des discussions. La Première Commission de l'Assemblée générale des Nations unies peut constituer une

enceinte appropriée, dans le cadre de laquelle un processus dédié pourrait discuter des enjeux liés aux applications militaires de l'IA.

B. Une architecture de gouvernance rationalisée et cohérente

Alors que la communauté internationale s'est récemment saisie de la question de l'IA militaire, et que plusieurs formats et initiatives ont vu le jour, il conviendrait de les rationaliser pour éviter une duplication des efforts et une multiplication des discussions.

La fragmentation des discussions (sur la terminologie, les principes directeurs, les mesures à adopter, etc.) doit être évitée pour ne pas aboutir à des résultats concurrents, voire contradictoires. Un cadre unique permet également de rationaliser les efforts pour gagner en efficacité et de renforcer l'impact des résultats.

La gouvernance doit répondre à un impératif de cohérence, en veillant à ce que des processus existants abordant des thématiques connexes ne soient pas fragilisés. En particulier, il est prioritaire que le GGE sur les systèmes d'armes létaux autonomes (sujet qui recoupe largement, sans le recouvrir, celui de l'IA militaire) rattaché à la Convention sur certaines armes classiques puisse remplir son mandat, dont l'échéance est fixée à 2026. La France considère que le lancement d'un nouveau processus ne doit pas intervenir avant que le mandat actuel du GGE SALA arrive à échéance, et devra prendre en compte les résultats de ce dernier. Le GGE SALA doit pouvoir poursuivre ses travaux après 2026 via un nouveau mandat. Il conviendra également que le processus dédié à l'IA militaire assure la complémentarité et la non-duplication des discussions du GGE, qui auront de fait pour objet la régulation d'une partie centrale des enjeux liés au développement de l'IA dans le domaine militaire.

C. Un processus à visée opérationnelle, centré sur les enjeux propres au domaine militaire

Alors que la résolution touche à l'utilisation de l'IA à des fins militaires, la gouvernance doit s'attacher au corpus juridique applicable aux conflits armés, et donc en premier lieu au droit international humanitaire. Le DIH régissant l'utilisation de tous les moyens et méthodes de guerre, y compris ceux du futur, il est ainsi applicable indépendamment de la technologie militaire utilisée. Dès lors, les technologies d'intelligence artificielle utilisées dans ou pour les opérations militaires sont soumises au même régime juridique que tous les autres moyens de guerre. En d'autres termes, le droit international humanitaire s'applique pleinement à l'IA dans le domaine militaire. De par son champ d'application, sa capacité d'adaptation aux nouvelles technologies et sa reconnaissance par la communauté internationale, la déclinaison des principes cardinaux du DIH et leur respect doivent constituer l'horizon essentiel des discussions sur l'IA dans le domaine militaire.

De manière générale, la France considère que la priorité de tout processus international doit être de discuter de la manière d'opérationnaliser et d'assurer le respect des normes juridiques existantes, qui s'appliquent à l'IA dans le domaine militaire. L'opérationnalisation de ces normes implique deux dimensions :

- Une discussion sur l'établissement de principes directeurs et sur les moyens de les mettre en œuvre, à l'instar des principes détaillés en 2^{ème} partie. La France considère qu'il sera opportun de prendre appui sur les discussions ayant eu lieu dans le cadre du processus REAIM, qui contribue de façon significative, dans un format *ad hoc* à faire progresser la compréhension de la communauté internationale sur les principes qui permettront d'assurer un développement de l'IA militaire respectueux du droit international.
- Une facilitation de la mise en œuvre des principes volontaires, laquelle est à la charge des Etats, à qui il incombe de prendre les mesures concrètes à cet effet. Pour les aider dans cette entreprise, le processus international dédié à l'IA militaire doit permettre d'enrichir cette « boîte à outils » et favoriser sa diffusion ainsi que son appropriation, en particulier en facilitant l'échange de bonnes pratiques (sur l'évaluation et la certification des systèmes, les

examens de licéité, etc.), et en favorisant la coopération et l'assistance internationale, selon des modalités adaptées aux affaires militaires.

Enfin, le processus dédié à l'IA devra promouvoir l'élaboration et la mise en œuvre de mesures de confiance et de réduction des risques adaptées au développement de l'IA dans le domaine militaire.