



Institut
EGA

Gouverner l'intelligence artificielle dans le domaine militaire : défis juridiques et stratégiques

Romane Wanner

Directrice du département Sécurité internationale et forces armées de l'Institut d'études de géopolitique appliquée

Maéva Obiang Ndong

Juriste en droit international

17 juillet 2025

Les opinions exprimées dans ce texte n'engagent que la responsabilité de l'auteur.

ISSN : 2739-3283

© Tous droits réservés, Paris, Institut d'études de géopolitique appliquée, 2025.

Comment citer cette publication :

Romane Wanner, Maéva Obiang Ndong, *Gouverner l'intelligence artificielle dans le domaine militaire : défis juridiques et stratégiques*, Institut d'études de géopolitique appliquée, Paris, 17 juillet 2025.

121 rue du Vieux Pont de Sèvres 92100 Boulogne-Billancourt

Courriel : secretariat@institut-ega.org

Site internet : www.institut-ega.org

SOMMAIRE

Introduction – **P. 1**

L'intégration de l'intelligence artificielle au service de la défense – **P. 3**

L'encadrement de l'intelligence artificielle par le cadre juridique existant – **P. 5**

La responsabilité en cas de dommage causé par un système d'intelligence artificielle militaire – **P. 8**

Introduction

L'intelligence artificielle (IA) dans le domaine militaire constitue désormais un enjeu mondial de premier plan. En raison de ses implications directes sur la sécurité des États et sur la stabilité internationale, elle exige une coopération renforcée entre les acteurs étatiques. L'intégration de l'IA dans les systèmes de défense est devenue un facteur stratégique majeur, influençant les rapports de force, les choix politiques et les orientations technologiques. À ce titre, elle représente à la fois un défi et une opportunité pour les États.

Selon l'Union européenne, l'IA est définie comme un « *système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels* »¹. Dans le domaine militaire, l'IA peut être mobilisée pour diverses fonctions : reconnaissance et analyse du terrain, évaluation et hiérarchisation des options tactiques, optimisation des trajectoires de missiles, gestion en temps réel des ressources ou encore appui à la prise de décision stratégique grâce à l'analyse de données provenant de multiples sources. Certaines applications vont au-delà de l'aide à la décision : certains systèmes sont conçus pour détecter, identifier et, dans certains cas, engager des cibles de manière autonome, sans intervention humaine directe. Ces systèmes, communément appelés « systèmes d'armes létaux autonomes », suscitent aujourd'hui un débat international majeur, tant sur le plan éthique que juridique.

L'IA, particulièrement dans le domaine militaire, demeure une technologie complexe et en constante évolution, ce qui rend sa compréhension et son encadrement juridique particulièrement délicats. Élaborer un cadre législatif clair pour une technologie aux contours encore flous représente un défi considérable. Ce défi s'accroît lorsqu'il s'agit d'établir des normes à l'échelle internationale, en raison des divergences d'approche, des intérêts stratégiques et des niveaux de développement technologique entre les États. Même dans

¹ Article 3 du RÈGLEMENT (UE) 2024/1689 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle), [L_202401689FR.000101.fmx.xml](https://eur-lex.europa.eu/eli/reg/2024/1689/oj)

l'hypothèse d'un consensus international, la mise en œuvre et le contrôle effectif de ces règles pourraient se heurter à des obstacles importants, notamment en matière de contournement ou d'application inégale.

L'IA est appelée à transformer en profondeur le champ de bataille, en apportant des capacités accrues en matière de reconnaissance autonome, de prise de décision rapide et d'automatisation des systèmes d'armement.

Il devient alors essentiel d'encadrer rigoureusement l'utilisation de ces technologies. Un tel encadrement doit garantir leur conformité aux normes éthiques, prévenir toute dérive stratégique et réduire les risques tant pour les civils que pour les combattants. Il est également indispensable de veiller au strict respect des principes fondamentaux du droit international humanitaire.

L'intégration de l'intelligence artificielle au service de la défense

Au-delà de son omniprésence dans les négociations civiles, l'intégration et la régulation de l'IA sont également au cœur des discussions à des fins militaires. Or ce domaine présente des spécificités qui lui sont propres.

La défense militaire désigne l'ensemble des politiques et décisions ayant pour but « d'assurer par la force des armes la défense de la patrie et des intérêts supérieurs de la Nation »². Cela inclut, entre autres, la défense de l'intégrité du territoire, la protection de la population contre les agressions armées, la lutte contre les autres menaces susceptibles de compromettre la sécurité nationale et la garantie du respect des alliances européennes et internationales³. Pour atteindre ces objectifs de défense et de protection, ce domaine est soumis à de nombreux impératifs exceptionnels. À titre d'exemple, la confidentialité absolue y est indispensable. La défense ne se limite pas à la protection des intérêts et du bon fonctionnement d'une entité, comme c'est le cas dans le domaine civil. La divulgation d'informations classées secret-défense porterait intrinsèquement atteinte, parfois gravement, à la défense et à la sécurité nationales⁴.

² Code de la défense, art. L4111-1.

³ Id., art. L1111-1.

⁴ Id., art. R2311-3.

De même, des atteintes exceptionnelles aux droits humains sont admises⁵. En témoigne le fait qu'il puisse être attendu d'un militaire qu'il accomplisse une mission qui lui est confiée, même au péril de sa vie⁶ ; une exigence sans équivalent dans le domaine civil.

C'est dans ce contexte particulier que se pose la question des différentes formes d'intégration de l'IA. Le recours à celle-ci se fait essentiellement dans deux espaces distincts : l'espace physique (ou réel) et l'espace virtuel (ou cyberspace).

Dans le cadre de l'espace physique, l'IA peut, par exemple, prendre la forme de systèmes d'apprentissage automatique capables d'analyser une grande quantité de données et d'en extraire des caractéristiques utilisables pour faciliter la prise de décision⁷. En matière de défense, ces décisions peuvent aller de la détermination de la meilleure stratégie militaire à la protection optimale des personnes et des bâtiments civils⁸, cette dernière étant particulièrement importante dans un contexte où les conflits armés ne se déroulent plus sur un champ de bataille éloigné, mais souvent dans des zones urbaines et densément peuplées⁹. Toujours dans cet espace, l'IA peut également se matérialiser sous la forme de systèmes d'armes létales intégrant un certain degré d'autonomie tout en conservant un contrôle humain significatif. Conformément aux impératifs de rapidité et de réactivité propres au domaine militaire, ces systèmes peuvent s'avérer précieux en raison de leur précision lors de frappes ciblées ou de leur capacité à intervenir dans des conditions ou des environnements difficilement accessibles pour l'être humain¹⁰.

Dans le cadre du cyberspace, l'apprentissage automatique antagoniste, ou *adversarial machine learning*, peut être utilisé à des fins militaires. Cet apprentissage « consiste à entraîner un modèle à partir d'un jeu de données comprenant des données trompeuses ou ambiguës [souvent produites par un autre modèle], de manière qu'il apprenne à discerner de telles données »¹¹. À des fins stratégiques, il serait par exemple possible de perturber la qualité de l'entraînement d'un modèle « ennemi », pratique connue sous le nom de data poisoning, en

⁵ Voir par exemple Convention Européenne de sauvegarde des Droits de l'Homme et des Libertés fondamentales, ouverte à la signature le 4 novembre 1950, entrée en vigueur le 3 septembre 1953, STCE 005 (ci-après « CESDH »), art. 8(2)

⁶ Id., art. D4122-4. Voir aussi art. 4(3)b) CESDH.

⁷ C. VILLANI, « Les enjeux de l'IA pour la Défense de demain », *Revue Défense Nationale*, 2019/5, n° 820, p.24.

⁸ R. STEWART et G. HINDS, « Les algorithmes de la guerre : le recours à l'intelligence artificielle pour la prise de décision dans les conflits armés », *CICR - Droit & Politique Humanitaires*, 2024, p. 3.

⁹ CICR, « Le droit international humanitaire et les défis posés par les conflits contemporains », *Rapport à l'intention de la Conférence internationale de la Croix-Rouge et du Croissant-Rouge (Conférence internationale)*, 2019, p. 16.

¹⁰ COMITÉ D'ÉTHIQUE DE LA DÉFENSE, Avis sur l'intégration de l'autonomie dans les systèmes d'armes létaux, 14 janvier 2025, p. 5.

¹¹ MINISTÈRE DE LA CULTURE, « 50 termes clés de l'intelligence artificielle », *Délégation générale à la langue française et aux langues de France*, 2025, vol. 1, p. 5.

introduisant à son insu des données trompeuses dans sa base d'apprentissage ou en associant des caractéristiques volontairement erronées à des données authentiques¹².

Si les avantages liés à l'intégration progressive de l'IA dans le domaine de la défense sont indéniables, son utilisation soulève également des interrogations quant à ses effets potentiellement négatifs sur la conduite et la nature des conflits armés. De manière générale, il pourrait être soutenu qu'au regard des avancées rapides en matière d'IA, les risques de détournement pourraient être perçus comme une menace ou que des dysfonctionnements indépendants de la volonté humaine contribuent à l'escalade de nouveaux conflits armés, qu'ils soient internationaux ou non¹³. En outre, l'extension des conflits au cyberspace augmente inévitablement la surface d'attaque et, *de facto*, les enjeux de défense et de sécurité qui en découlent¹⁴. À la protection externe des infrastructures contre la destruction physique s'ajoute ainsi la nécessité d'une protection interne contre les cyberattaques.

Il est dès lors indispensable de maintenir un équilibre entre la préservation de la paix et de la sécurité, tant à l'échelle nationale qu'internationale, et le respect des normes du droit international, en particulier celles établies dans le cadre du droit international humanitaire et des droits de l'Homme.

L'encadrement de l'intelligence artificielle par le cadre juridique existant

Avant toute chose, il est pertinent de rappeler que l'encadrement de l'utilisation de l'IA ne fait pas l'objet d'un vide juridique en droit international. Les États ont, par exemple, l'obligation de veiller à ce que l'emploi de celle-ci demeure « cohérent avec les obligations de protection des droits de l'homme, telles qu'elles sont consacrées par le droit international applicable et par [leur] droit interne »¹⁵. Le recours à l'IA à des fins militaires n'est pas interdit par le droit international : le Protocole additionnel (I) aux Conventions de Genève du 12 août 1949, relatif

¹² D. ARAYA et M. KING, « The Impact of Artificial Intelligence on Military, Defence and Security », *CIGI Papers*, 2022, n° 263, p. 11.

¹³ IHEDN, « Les ruptures stratégiques contemporaines : Les enjeux de l'intelligence artificielle pour la défense », *Fiche d'actualité - Département des études et de la recherche*, 2024, n° 11, p. 5.

¹⁴ A. LÉVÊQUE., « Données et intelligence artificielle : quels enjeux pour la cybersécurité des États ? », *Cahiers français*, 2021/1, n° 419, p. 78.

¹⁵ Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, ouverte à la signature le 5 septembre 2024, STCE 225, art. 4.

à la protection des victimes des conflits armés internationaux, garantit aux parties à un conflit armé le droit de choisir leurs méthodes et moyens de guerre, tout en rappelant que ce droit n'est pas illimité¹⁶, dès lors que ces moyens ne sont pas de « nature à causer des maux superflus »¹⁷. Cela inclut, entre autres, le développement et l'emploi « de nouveaux moyens ou d'une nouvelle méthode de guerre »¹⁸.

La Cour internationale de Justice a par ailleurs rappelé que les principes et règles établis en droit des conflits armés (DCA) s'appliquent « à toutes les formes de guerre et à toutes les armes, celles du passé, comme celles du présent et de l'avenir »¹⁹. Le recours à l'intelligence artificielle est dès lors soumis aux principes du DCA relatifs à la conduite des hostilités : la nécessité militaire ; la distinction entre civils et combattants ainsi qu'entre biens civils et objectifs militaires²⁰ ; la proportionnalité entre les dommages civils collatéraux attendus et « l'avantage militaire concret et direct attendu »²¹ ; et le principe de précaution (aussi appelé principe d'humanité), afin d'éviter de causer des dommages et blessures superflus et inutiles²². À ce corpus juridique contraignant (ou *Hard Law*) relativement général se superpose un corpus juridique non contraignant (ou *Soft Law*) spécifique au recours à l'IA à des fins militaires.

Dans un objectif d'encadrement, ce dernier peut constituer une alternative sérieuse à l'adoption d'un instrument juridique contraignant, dont le formalisme est parfois lourd et chronophage. À titre d'exemple, en 2017, le Manuel de Tallinn 2.0 sur le droit international applicable aux cyberopérations²³ a été publié à l'initiative du Centre d'excellence en coopération pour la cyberdéfense de l'OTAN (CCDCOE). Ce document ne se limite pas au droit des conflits armés ; il s'intéresse à la transposition globale du droit international public et des principes de maintien de la paix et de la sécurité internationales à ces nouvelles opérations. Cela inclut tant les normes propres au régime général de droit international, telles que la souveraineté²⁴, les questions de juridiction²⁵ et la responsabilité internationale des États²⁶, que

¹⁶ Protocole additionnel aux Conventions de Genève du 12 août 1949 relatif à la protection des victimes des conflits armés internationaux, ouvert à la signature le 8 juin 1977, entré en vigueur le 7 décembre 1978 (ci-après « Protocole I »), art. 35(1).

¹⁷ Art. 35(2) Protocole I.

¹⁸ Art. 36 Protocole I.

¹⁹ CIJ, *Licéité de la menace ou de l'emploi d'armes nucléaire*, Avis consultatif, 8 juillet 1996, Rec. p. 226, par. 86.

²⁰ Art. 48 Protocole I.

²¹ Art. 57(2)iii) Protocole I.

²² Art. 58 Protocole I.

²³ M. N. SCHMITT (dir.), *Tallin Manual 2.0 on the International Law Applicable to Cyber Operations*, Cambridge, éd., Cambridge University Press, 2017.

²⁴ *Id.*, pp. 11-27.

²⁵ *Id.*, pp. 51-75.

²⁶ *Id.*, pp. 79-157.

les normes relevant de régimes spécialisés du droit international, par exemple le droit de la mer²⁷, le droit spatial²⁸ ou encore le droit des relations diplomatiques et consulaires²⁹. Deux ans plus tard, en 2019, le CCDCOE a également publié les Principes directeurs affirmés par le Groupe d'experts gouvernementaux sur les technologies émergentes dans le domaine des systèmes d'armes létaux autonomes. Parmi ceux-ci figurent l'importance de garantir que l'utilisation de ces nouvelles armes s'inscrive dans une chaîne de commandement et de contrôle humains responsable, ainsi que la nécessité de mettre en place des mesures d'évaluation et d'atténuation des risques associés au recours à de telles technologies, notamment les risques de prolifération et d'acquisition par des groupes terroristes.

S'il est inexact de parler de vide juridique quant à l'encadrement de l'utilisation de l'IA à des fins militaires, de nombreux flous demeurent. Dans le cas de systèmes d'armes létaux intégrant de l'autonomie, voire entièrement autonomes, la question centrale est celle de l'imputabilité en cas de violation des principes du DCA dans le contexte de la relation humain/machine. C'est pour cette raison que les négociations internationales et plusieurs initiatives civiles convergent vers l'interdiction préventive des systèmes létaux entièrement autonomes. En novembre 2024, un projet de résolution sur les systèmes d'armes létaux autonomes a été adopté au sein de la Première Commission (désarmement et sécurité internationale) de l'Assemblée générale des Nations unies. Néanmoins, comme l'a rappelé le Secrétaire général des Nations unies, António Guterres, dans son Nouvel Agenda pour la paix, l'adoption d'un instrument juridique contraignant demeure indispensable pour assurer une interdiction effective des systèmes d'armes létaux autonomes « qui fonctionnent sans supervision ou contrôle humain ». Il a donc appelé les États à poursuivre les négociations internationales en vue d'adopter un tel instrument juridiquement contraignant d'ici 2026.

²⁷ *Id.*, pp. 232-239.

²⁸ *Id.*, pp. 270-279

²⁹ *Id.* pp. 209-230.

La responsabilité en cas de dommage causé par un système d'intelligence artificielle militaire

L'intégration croissante de l'IA dans les systèmes d'armement transforme en profondeur la conduite des opérations militaires. Cette évolution soulève de nombreuses interrogations, notamment en matière de responsabilité en cas de dommages causés par ces technologies. À mesure que certains systèmes deviennent partiellement ou totalement autonomes, l'attribution claire des responsabilités se complexifie considérablement.

Parmi les préoccupations majeures figure l'utilisation des systèmes d'armes létaux autonomes. Ces systèmes, capables de détecter, sélectionner et engager des cibles sans intervention humaine directe, posent des questions juridiques fondamentales. Le droit international humanitaire repose notamment sur deux principes essentiels : la distinction entre combattants et civils, et la proportionnalité dans l'usage de la force. Or, une IA, aussi avancée soit-elle, ne possède ni conscience morale ni capacité de discernement contextuel, lui permettant d'appréhender des situations complexes sur le terrain, telles que la reddition d'un ennemi ou la présence de civils dans une zone de combat.

Dès lors, une incertitude juridique majeure subsiste : qui est responsable en cas de faute ou de dommage ? Si un système d'armes létaux autonome cause des pertes civiles injustifiables ou agit en violation du droit humanitaire, la responsabilité incombe-t-elle au programmeur, à l'opérateur, à l'État ? Peut-on aller jusqu'à considérer la machine elle-même comme « responsable » ? Cette absence de clarté crée un vide juridique préoccupant, qui fragilise à la fois la légitimité du recours à ces technologies et la confiance du public à leur égard.

Il devient donc urgent que la communauté internationale engage une réflexion approfondie pour définir un cadre de responsabilité clair, adapté à la spécificité des systèmes d'IA militaires, garantissant en toutes circonstances le respect du droit humanitaire.



Institut EGA

Institut d'études de géopolitique appliquée

121 rue du Vieux Pont de Sèvres

92100 Boulogne-Billancourt

Tél. +33 3 85 51 27 78

www.institut-ega.org

ISSN 2739-3283

En libre accès

Dépôt légal : Juillet 2025

Tous droits réservés

Imprimé en France