

Le droit international humanitaire à l'épreuve de l'utilisation de l'IA militaire

Étude d'un système d'intelligence artificielle destiné au ciblage humain

Romane Malbec

Conception de la couverture:

Ted van Baarda

Le droit international humanitaire à l'épreuve de l'utilisation de l'IA militaire

Étude d'un système d'intelligence artificielle destiné au ciblage humain

Romane Malbec

étudiante à:

Institut Catholique de Paris
France

© EuroISME 2026

Ce prix est soutenu par:



Membres du jury:

1. Aumônier principal Stefan Gugerel (Autriche), président
 2. Général de brigade (à la retraite) Benoît Royal (France)
 3. Dr. Veronika Bock (Allemagne)
 4. Lt.Col. Dr. Patrick Hofstetter (Suisse)
 5. Aumônier principal Lt.Col. Dr. Janne Aalto (Finlande)
 6. Dr. Asta Maskaliūnaitė (Estonie)
 7. Général de division (à la retraite) Dr. Juan Moliner Gonzalez (Espagne)
 8. Prof. Dr. Pieter Vos (Les Pays-Bas)
- Mme. Monique van der Wal - Tempelman (Les Pays-Bas),
secrétaire

Demandes de renseignements:
secretariat.ethicsprize@euroisme.eu

www.euroisme.eu

L'Institut Catholique de Paris n'entend donner aucune approbation ni improbation aux opinions émises dans la thèse. Elles doivent être considérées comme propres à leur auteur.

Remerciements

Je tiens à remercier Madame Marie-des-Neiges Ruffo pour son encadrement et ses précieux conseils qui m'ont permis d'aborder avec enthousiasme un sujet aussi technique que passionnant.

Merci également à toutes les personnes ayant contribué à nourrir ma réflexion au fil de ce travail.

Mes remerciements les plus chaleureux s'adressent ensuite à ma directrice de Master, Madame Diane Galbois-Lehalle qui m'a donné la chance d'intégrer ce Master, m'a accordé sa confiance pour de nombreux projets et m'a permis de trouver ma voie professionnelle.

Enfin, je souhaiterais remercier mes parents, pour leurs conseils avisés et leur soutien indéfectible, qui m'ont accompagnée tout au long de la rédaction de ce mémoire et de mon parcours universitaire.

Table des matières

| | |
|---|--------------|
| Introduction | ...1 |
| Titre 1: Les systèmes d'IA militaire et leur conformité au DIH | ...15 |
| Chapitre 1: Aspects légaux | ...17 |
| Section 1: Les origines du DIH dans le <i>jus in bello</i> | ...17 |
| § 1 Historique du DIH: un droit par essence coutumier | ...17 |
| A) Le <i>jus in bello</i> : héritage des civilisations anciennes | ...17 |
| B) Des sagesses antiques à l'éthique guerrière | ...18 |
| C) De la coutume à la codification du <i>jus in bello</i> | ...19 |
| § 2 Le DIH moderne | ...19 |
| A) La naissance du DIH contemporain | ...19 |
| B) Naissance des fondements du DIH | ...20 |
| C) Des influences étrangères à la définition du DIH moderne | ...21 |
| Section 2: Les principes fondamentaux du DIH | ...23 |
| § 1: L'examen de licéité pour toute nouvelle arme, nouveaux moyens ou nouvelles méthodes | ...23 |
| A) L'IA en tant qu'arme, moyen ou méthode de guerre ? | ...23 |

| | |
|--|-------|
| 1- La règle contre les armes intrinsèquement indiscriminées | ...25 |
| 2- La règle contre les armes qui causent des souffrances inutiles | ...26 |
| 3- Limitation et interdiction de certaines armes de guerre | ...27 |
| B) Article 36 du Protocole Additionnel I | ...28 |
| 1- L'examen de licéité pour les SIA de défense | ...28 |
| 2- Révision de la licéité des armes intégrant un logiciel d'IA dans leur évaluation | ...30 |
| § 2: Principe de Distinction, Proportionnalité et Précaution | ...33 |
| A) Le principe de distinction | ...33 |
| B) Le principe de proportionnalité | ...35 |
| C) Le principe de précaution | ...37 |

**Chapitre 2: Les défis de l'intégration d'un SIA de
ciblage humain dans le DIH** **...39**

Section 1: Les difficultés techniques **...39**

§ 1 Les limites cognitives et décisionnelles des systèmes
d'Intelligence Artificielle **...39**

A) L'opacité et l'imprévisibilité des systèmes
d'Intelligence Artificielle **...39**

1- Les algorithmes d'IA utilisés par les forces
armées **...39**

| | |
|---|-------|
| 2- Explicabilité des algorithmes: l'effet «boîte noire» | ...43 |
| B) L'incapacité de l'IA à interpréter les principes juridiques du DIH | ...46 |
| 1- Fiabilité | ...46 |
| 2- Robustesse | ...46 |
| § 2: Les risques pratiques liés à l'utilisation opérationnelle de l'IA en contexte militaire | ...53 |
| A) Biais des données et erreurs dans l'identification des cibles | ...53 |
| B) Vulnérabilités aux attaques et défis de cybersécurité | ...56 |
| Section 2: Les difficultés géopolitiques | ...59 |
| § 1: Confrontation entre course à l'armement et un néces- saire cadre normatif à l'IA | ...59 |
| A) La course à l'IA dans le domaine militaire selon les pays | ...63 |
| 1- Les États Unis | ...63 |
| 2- La Chine | ...64 |
| 3- L'Europe/France | ...67 |
| B) Les positions des pays sur l'IA militaire et leur respect du DIH | ...70 |
| 1- Les États Unis | ...70 |
| 2- Chine | ...72 |

| | |
|--|-------|
| 3- France | ...74 |
| § 2: Les effets de cette course à l'armement | ...77 |
| A) Implication pour la sécurité mondiale | ...77 |
| 1- L'escalade incontrôlée | ...77 |
| 2- Les risques de prolifération | ...81 |
| B) Défis éthiques et juridiques | ...82 |
| 1- Les problématiques éthiques | ...82 |
| 2- Défis juridiques | ...84 |

**Titre 2: Vers une adaptation du DIH à l'ère de
L'IA militaire** ...89

Chapitre 1: Les enjeux de la responsabilité ...91

| | |
|---|-------|
| Section 1: Les causes de cette responsabilité du fait de l'usage de ce type de SIA | ...91 |
| § 1: Érosion du contrôle humain | ...91 |
| A) Le cas des SIA de ciblage humain sous super- vision humaine (type Lavender) | ...91 |
| 1- Lavender <i>vs</i> SALA | ...92 |
| 2- Les facteurs de l'érosion du contrôle humain | ...94 |
| 3- Tension avec le DIH | ...97 |
| B) Le cas des SIA de ciblage humain intégralement autonome | ...99 |
| 1- La disparition intégrale de l'humain de la | |

| | |
|---|--------|
| boucle |99 |
| 2- À l'avènement de l'hyperguerre par IA | ..101 |
| § 2: Érosion d'une responsabilité humaine | ..103 |
| A) <i>Actus reus</i> et <i>mens rea</i> : fondements de la responsabilité pénale | ..103 |
| B) Disparition de l'agentivité humaine: un changement de paradigme | ..105 |
| 1- De l'agentivité humaine... | ..105 |
| 2- ...À l'agentivité de la machine | ..106 |
| C) Le problème d'attribution de la responsabilité aux SALA | ..106 |
| Section 2: Les moyens pour endiguer cette responsabilité | ..108 |
| § 1: L'attribution de la personnalité juridique aux robots létaux autonomes : une solution illusoire? | ..108 |
| A) Autonomie décisionnelle des IA militaires? | ..109 |
| 1- Positions doctrinales | ..109 |
| 2- Proposition du Parlement | ..110 |
| B) Un patrimoine pour les robots autonomes militaires | ..111 |
| § 2: Vers des mécanismes alternatifs de responsabilisation pour les systèmes d'IA militaires | ..112 |
| A) Un patrimoine d'affectation | ..113 |
| B) Les SIA militaires: ni personne, ni chose | ..114 |
| C) L'humain inévitablement au cœur de la responsa- | |

| | |
|---|--------------|
| bilité? | ..115 |
| 1- Une programmation humaine, une responsa- bilité humaine | ..115 |
| 2- La responsabilité du fait des robots | ..117 |
| | |
| Chapitre 2: Vers un cadre juridique et éthique pour l'IA militaire | ..121 |
| Section 1: Un corpus juridique unique spécifique à l'IA militaire | ..121 |
| § 1: La clause de Martens ou un complément des lois de la guerre | ..121 |
| A) Les différentes interprétations de la Clause de Martens | ..122 |
| 1- Sens restreint | ..122 |
| 2- Interprétation large | ..123 |
| 3- Sens modéré | ..124 |
| B) Un nouveau cadre juridique exclusif | ..125 |
| § 2: Les moyens de rationaliser | ..129 |
| A) Les instruments existants de régulation des SALA | ..129 |
| B) Une efficacité remise en question | ..130 |
| | |
| Section 2: Les limites absolues du DIH | ..132 |
| § 1: Les enjeux stratégiques et de souveraineté | ..132 |
| A) L'IA militaire comme levier de supériorité | |

| | |
|---|--------------|
| technologique et dissuasive | ..132 |
| B) Tensions entre refus d'ingérence juridique et respect des obligations du DIH | ..134 |
| § 2: Volonté des États de respecter le DIH ou non | ..135 |
| A) Les facteurs déterminants du respect du DIH par les États | ..135 |
| B) Le respect sélectif du DIH selon les intérêts politiques et opérationnels | ..137 |
| Conclusion | ..141 |
| Postface | 143 |
| Bibliographie | ..147 |

Liste des abréviations

| | |
|-----------------|--|
| APL | Armée populaire de libération |
| CCAC | Convention sur Certaines Armes Classiques |
| CICR | Comité international de la Croix-Rouge |
| CNIL | Commission nationale de l'informatique et des libertés |
| DARPA | Défense Advanced Research Projects Agency |
| DIH | Droit International Humanitaire |
| DL | Deep Learning |
| DoW | US Department of War |
| GEG | Groupe d'Experts Gouvernementaux |
| IA | Intelligence artificielle |
| <i>Ibid.</i> | Ibidem (même auteur et même ouvrage) |
| JAIC | Joint Artificial Intelligence Center |
| MGCS | Main Ground Combat System |
| ML | Machine Learning |
| NCIA | Agence OTAN d'information et de communication |
| ONU | Organisation des Nations Unies |
| <i>op. cit.</i> | <i>Opus citatum</i> (ouvrage cité) |
| OTAN | Organisation du traité de l'Atlantique Nord |
| p. | Page |
| pp. | Page n° (...) à page n° (...) |
| RIA | Règlement européen sur l'Intelligence Artificielle |
| RLA | Robots létaux autonomes |
| SAA | Systèmes d'armes autonomes |
| SALA | Systèmes d'armes létales autonomes |
| SALIA | Systèmes d'armes létaux intégrant de l'autonomie |
| SCAF | Système de Combat Aérien du Futur |

| | |
|--------|---|
| SIA | Système d'Intelligence Artificielle |
| SLAM-F | Système de Lutte Anti-Mines du Futur |
| UAV | Unmanned Aerial Vehicle (véhicule aérien sans pilote) |
| UE | Union européenne |

Introduction

La technologie de l'intelligence artificielle (IA) a atteint un niveau tel que le déploiement de ces systèmes est — pratiquement, sinon légalement — envisageable d'ici quelques années, et non des décennies, avec des enjeux majeurs: les armes autonomes ont été qualifiées de troisième révolution dans l'art de la guerre, après la poudre à canon et les armes nucléaires.¹

1. Cette formule issue de la lettre ouverte publiée le 28 juillet 2015 à l'occasion de *l'International Joint conference on artificial intelligence*, résume avec force la rupture technologique et éthique à laquelle les sociétés modernes sont confrontées. C'est un fait, l'irruption de l'Intelligence Artificielle dans le champ de la guerre transforme profondément les paradigmes du combat, du commandement et du droit, certains se demandant même si ce n'est pas «*une nouvelle révolution militaire?*»². Désormais, des systèmes autonomes ou semi-autonomes sont capables d'analyser des données, d'identifier des cibles, et, dans certains cas, de prendre des décisions létales sans intervention humaine directe. La perspective du déploiement à grande échelle des systèmes d'armes létaux autonomes (SALA), ou de systèmes de ciblage fondés sur l'IA, n'est plus une simple hypothèse futuriste. Bien que ces derniers n'aient pas encore atteint une pleine autonomie, l'utilisation de l'IA dans les conflits armés

¹ Future of life institute, “Autonomous Weapons Open Letter: AI & Robotics Researchers”, 9 février 2016. Disponible sur [[Le lien](#)].

² Jean-Christophe Noël, «Intelligence artificielle: vers une nouvelle révolution militaire?», Octobre 2018. Disponible en version PDF sur [[Le lien](#)].

est déjà une réalité, comme en témoignent les usages observés en Ukraine et en Israël.³

2. Le droit international humanitaire (DIH), parfois appelé le droit de la guerre ou le droit des conflits armés,⁴ regroupe l'ensemble des normes du droit international destinées à protéger, en temps de conflit armé, les personnes affectées par les hostilités ainsi que les biens qui ne participent pas directement aux combats⁵. Il repose sur des principes fondamentaux tels que les principes de distinction, proportionnalité, précaution, nécessité militaire, et de responsabilité, dont la mise en œuvre suppose généralement une appréciation humaine contextuelle.⁶

3. La technologie qui nous intéresse, l'Intelligence Artificielle, est au cœur des discussions et des débats ces dernières années, en en faisant à la fois un instrument technologique et politique, comme en témoigne le récent sommet sur l'IA organisé à Paris les 10 et 11 février derniers.

4. Toutefois, il n'y a pas de consensus universel sur la définition de l'Intelligence Artificielle. Il existe en réalité plusieurs définitions de ce qu'est l'IA.

³ Amélie Férey et Laure de Roucy-Rochegonde, «De l'Ukraine à Gaza: l'intelligence artificielle en guerre», article publié dans la *Revue Politique Étrangère*, vol.89, n°3, automne 2024. Disponible en version PDF sur [[Le lien](#)].

⁴ Manuel de droit des opérations militaires, édition 2022. Disponible en version PDF.

⁵ *Ibid.*

⁶ Comité éthique de la Défense, principe n°8 de l'avis sur l'usage des technologies d'intelligence artificielle par les forces armées, 14 janvier 2025. Disponible en version PDF.

5. La définition historique de l'IA a été donnée lors d'une proposition pour le projet «Darthmouth summer research on artificial intelligence», en 1956 par 4 chercheurs pionniers en la matière. Ils définissent l'IA ainsi: *“To proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it”*.⁷ Cette conférence constitue un moment fondateur dans l'histoire de l'Intelligence Artificielle, largement reconnue comme l'événement à l'origine de la naissance de la discipline.

6. Par la suite, plusieurs auteurs et chercheurs se sont appliqués à la définir. Ainsi selon Yann Lecun, chercheur français en IA et vision artificielle chez Facebook, l'IA peut être définie comme la *«capacité des machines à prendre des décisions et à adopter des comportements attribués généralement à des humains ou à des animaux, mais avec un moindre niveau de performances»*.⁸

7. D'autres ont préféré définir l'IA en tant que discipline à part entière, c'est le cas de la France qui la définit comme étant un *«Champ interdisciplinaire théorique et pratique qui a pour objet la*

⁷ Dartmouth College, «A proposal for the Dartmouth summer research project on artificial intelligence», August 31, 1955. Traduction «Partir de l'hypothèse selon laquelle chaque aspect de l'apprentissage, ou toute autre caractéristique de l'intelligence, peut en principe être décrit avec une telle précision qu'une machine pourrait être conçue pour le simuler». Disponible sur [\[Le lien\]](#)

⁸ Définition de l'IA par Yann Lecun. Article issu de la revue DéfiS, n°8, 12 novembre 2019, écrit et publié sous la direction de l'Institut national des hautes études de la sécurité et de la justice (INHESJ) qui a rejoint l'Institut des hautes études du ministère de l'Intérieur (IHEMI).

*compréhension de mécanismes de la cognition et de la réflexion, et leur imitation par un dispositif matériel et logiciel, à des fins d'assistance ou de substitution à des activités humaines».*⁹

8. Plus récemment, l'IA en tant que discipline a disparu pour laisser place à une définition sur un Système d'Intelligence Artificielle (SIA). En effet, selon le Règlement européen sur l'Intelligence Artificielle du 13 juin 2024, un SIA est défini comme suit *«un système basé sur une machine qui est conçu pour fonctionner avec différents niveaux d'autonomie et qui peut faire preuve d'adaptabilité après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des données qu'il reçoit, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels».*¹⁰

9. Le secteur de la défense fait partie des principaux champs d'application de l'intelligence artificielle. Il est d'ailleurs identifié comme l'un des quatre domaines prioritaires par le rapport Villani, aux côtés de la santé, des transports et de l'environnement, dans le cadre de la stratégie visant à *«Donner un sens à l'intelligence artificielle»*.¹¹

⁹ Journal Officiel de la République Française, n°0285 du 9 décembre 2018, Commission d'enrichissement de la langue française, «Vocabulaire de l'intelligence artificielle (liste de termes, expressions et définitions adoptés)».

¹⁰ Commission européenne, Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle.

¹¹ Rapport de Cédric Villani, «Donner un sens à l'intelligence artificielle (IA)», 28 mars 2018. Rapport officiel. Mission confiée par le Premier Ministre Édouard Philippe. Mission parlementaire du 8 septembre 2017 au 8 mars 2018. Disponible sur [[Le lien](#)].

10. Ce débat se situe donc au croisement de plusieurs dynamiques historiques, technologiques et juridiques. Il s'inscrit dans une longue trajectoire de transformation des armements et de leur encadrement par le droit international, depuis les conventions de La Haye et de Genève jusqu'aux discussions contemporaines sur les armes nouvelles dans le cadre de la Convention sur Certaines Armes Classiques (CCAC). Cette dynamique est marquée par un mouvement constant de réactivité du droit face aux innovations technologiques où *«il existe parfois un retard significatif entre les ruptures technologiques observées et les normes susceptibles de les réguler»*.¹²

11. Ainsi, le droit s'est progressivement penché sur l'étude des armes chimiques aux armes biologiques, des mines antipersonnel aux armes à laser aveuglantes. La question des systèmes d'IA militaires s'inscrit dans cette lignée, tout en bouleversant la donne par l'ampleur des enjeux juridiques, éthiques et techniques qu'elle implique.

12. L'irruption de l'IA dans le champ militaire révèle un nouveau rapport entre l'homme, la machine et la violence par l'usage de la force létale,¹³ qui interroge les fondements même de

¹² Kirkham, Joanne, *Les systèmes d'armes létaux autonomes: une étude de droit international*, thèse de doctorat en droit public soutenue le 10 décembre 2022, École doctorale de droit international droit européen, relations internationales et droit comparé, Université Paris Panthéon-Assas. Disponible sur le [Le lien].

¹³ United Nations Institute for Disarmament Research, "Table-Top Exercises on the Human Element and Autonomous Weapons Systems, Summary Report", Giacomo Persi Paoli, Alice Spazian, Alisha Anand, This report summarises the findings of UNIDIR's series of regional Table-Top Exercises conducted

la guerre juste, de la chaîne de commandement et du droit à la vie. De plus, ces technologies sont inégalement maîtrisées selon les États, ce qui accentue les tensions géopolitiques. La course à l'IA militaire oppose aujourd'hui des puissances comme les États-Unis, la Chine, Israël, la Russie,¹⁴ ou encore les membres de l'OTAN, et soulève des interrogations stratégiques profondes sur les équilibres de puissance et la stabilité globale.

13. Pour appréhender les enjeux juridiques liés à l'usage de l'Intelligence Artificielle dans le ciblage humain, il est indispensable de délimiter précisément le périmètre de cette étude. Le champ des technologies d'Intelligence Artificielle utilisées à des fins militaires est extrêmement vaste. Il englobe aussi bien des systèmes d'aide au renseignement, ou d'aide à la décision,¹⁵ que des systèmes létaux entièrement autonomes. L'objet de notre étude porte plus spécifiquement sur les systèmes d'intelligence artificielle utilisés pour le ciblage humain, c'est-à-dire l'ensemble des opérations menées envers un individu visant à le neutraliser.¹⁶

between September 2020 and June 2021. Disponible en version PDF sur [\[Le lien\]](#).

¹⁴ Institut National des Affaires Stratégiques et Politiques, «Implications géopolitiques et compétition internationale : la nouvelle course à l'IA militaire», disponible sur [\[Le lien\]](#).

¹⁵ Comité éthique de la Défense, «Avis sur l'usage des technologies d'intelligence artificielle par les forces armées», *op.cit.*

¹⁶ Éric Gomez, «le ciblage dans le domaine militaire», Ouvrage «Les guerres de l'information à l'ère numérique», Sous la direction de Céline Marangé et Maud Quessard, Hors collection, 2021, Presses Universitaires de France, 456 pages.

14. Nous nous intéresserons ici à deux catégories de systèmes qui, bien qu'ils remplissent une fonction de ciblage, diffèrent de manière cruciale en matière de gouvernance humaine.

15. Dès 2013, Christof Heyns, le rapporteur spécial de l'Assemblée générale des Nations Unies avait défini les robots létaux autonomes (RLA) comme étant des *«systèmes d'armes robotiques qui, une fois activés, peuvent sélectionner et attaquer des cibles sans intervention complémentaire d'un opérateur humain. L'élément à retenir est que le robot choisit de façon autonome de viser telle cible et d'utiliser la force meurtrière»*.¹⁷

16. La résolution du Parlement européen du 12 septembre 2018 sur les systèmes d'armes autonomes, les définit quant à elle comme *«des systèmes d'armes dépourvus de véritable contrôle humain sur les fonctions critiques de choix et d'attaque de cibles individuelles»*.¹⁸

17. Le Comité Éthique de la Défense français, évoque que la notion des systèmes d'armes létaux doit être circonscrite à l'intégration de l'autonomie dans ces systèmes d'armes. Pour ce comité, les SALA sont *«des systèmes d'armes létaux programmés pour être capables de faire évoluer, au-delà du cadre d'emploi fixé initialement, leurs règles de fonctionnement, pour être capables d'utiliser cette programmation pour calculer une décision en vue d'une action sans appréciation de situation par le*

¹⁷ «Rapport du Rapporteur spécial sur les exécutions extrajudiciaires, sommaires ou arbitraires, Christof Heyns» du 9 avril 2013, document A/HRC/23/47, §38, p. 8.

¹⁸ Parlement européen, «Résolution du Parlement européen du 12 septembre 2018 sur les systèmes d'armes autonomes». Disponible sur [[Le lien](#)].

commandements». ¹⁹ Ce comité poursuit l'encadrement de la définition de cette notion, en faisant une distinction avec les systèmes d'armes létaux intégrant de l'autonomie, dénommés ci-après «SALIA». Dès lors, les SALIA sont définis par ce comité comme étant «*un système d'armes létal dont les fonctions de décision (au sens de l'autonomie décisionnelle robotique) sont dépendantes d'un cadre d'action précis. Le SALIA ne peut prendre d'initiatives létales qui l'amèneraient à modifier son domaine de fonctionnement*». ²⁰

18. En somme, les systèmes d'armes létaux autonomes (SALA) peuvent se définir comme «*des systèmes qui une fois activés, sont en mesure d'identifier une cible et d'user de la force létale sans être supervisés par un opérateur humain*». ²¹

19. Ce type de système peut prendre la forme de missiles autoguidés, de robots militarisés ou encore de drones, ²² et se caractérise donc par l'autonomie décisionnelle complète dans l'usage de la force létale, y compris contre des cibles humaines. Cette autonomie est rendue possible par les progrès de l'intelligence artificielle, qui permettent désormais de confier à une machine des «*fonctions cognitives telles que la reconnaissance, la*

¹⁹ Comité d'éthique de la défense, «Avis sur l'intégration de l'autonomie dans les systèmes d'armes létaux», 29 avril 2021, Ministère des Armées.

²⁰ *Ibid.*

²¹ De Roucy-Rochegonde, Laure, *La guerre à l'ère de l'intelligence artificielle, quand les machines prennent les armes*, aux Presses Universitaires de France, 1^{ère} édition, octobre 2024, p.12.

²² Pérémarty Léa, «Les armes létales autonomes: entre mythes et confusions», 26 juillet 2023. Disponible sur [[Le lien](#)].

classification, l'apprentissage ou encore la prise de décision».²³ Elle repose sur des algorithmes avancés de traitement de données et d'analyse contextuelle, ainsi que sur l'intégration de technologies comme la reconnaissance faciale via des capteurs,²⁴ ou les réseaux de neurones profonds.

Or, le concept même de SALA reste juridiquement flou. Malgré les tentatives d'encadrement, *«les SALA, parfois assimilés rapidement à des « robots tueurs », ne font pas à ce jour l'objet d'une définition communes*».²⁵ Depuis 2013, le Groupe d'Experts Gouvernementaux sur la question des SALA, réuni au sein de la Convention sur certaines armes classiques, est chargé de *«fournir une définition commune et un cadre juridique international permettant la supervision de son développement*».²⁶ En effet, cette absence de définition unifiée fragilise toute tentative de régulation et accentue la difficulté d'analyser leur compatibilité avec le DIH.

20. À l'inverse, certains systèmes actuels tels que le système israélien *Lavender* intègrent de l'IA pour assister les opérateurs humains dans les opérations de ciblage, sans pour autant se substituer à la prise de décision finale, à savoir la décision létale. Ces systèmes fonctionnent en «boucle humaine», ou «human-in-

²³ Assemblée Nationale, Rapport d'information n°3248 en conclusion des travaux d'une mission d'information sur les systèmes d'armes létaux autonomes, 22 juillet 2020. Disponible sur [Le lien].

²⁴ Future of Life Institute, "A Diplomat's Guide to Autonomous Weapons Systems", 2024. Disponible en version PDF sur [Le lien].

²⁵ Biblioveille «Les systèmes d'armes létales autonomes (SALA)», mars 2021. Disponible en version PDF sur [Le lien].

²⁶ Barrier, Félicité, «Les systèmes armés létaux autonomes (Sala): vers une nouvelle course à l'armement?» pp. 19 à 24 de la *Revue Défense Nationale*, 2018/5, n°810. Disponible sur [Le lien].

the-loop”, où «*le système exécute la tâche assignée de manière autonome, mais nécessite une intervention humaine pour valider et mettre en œuvre certaines actions spécifiques*». ²⁷ L’IA y est utilisée pour «*des capacités d’analyse avancées, de traitement massif de données et de connectivité en temps réel, transformant les dynamiques de prise de décision sur le champ de bataille*», ²⁸ et surtout «*pour générer des milliers de «cibles» potentielles pour des frappes militaires dans le feu de l’action d’une guerre*», ²⁹ mais la frappe finale ne peut être déclenchée qu’après validation par un opérateur humain. Ce type de système, bien qu’intégrant des mécanismes d’automatisation avancés, ne relève donc pas du SALA au sens strict, mais suscite également des débats, notamment sur le degré réel de supervision humaine et sur la confiance excessive que les opérateurs peuvent accorder aux suggestions de l’IA se rapprochant ainsi fortement du SALA. Or, «*Le fait que des machines visent de manière autonome des êtres humains est une ligne morale à ne pas franchir*». ³⁰ Là est tout le débat juridique de notre étude.

21. Ces différentes utilisations des SIA dans les opérations militaires nous amène aussi à faire une distinction entre les différents niveaux d’automatisation des systèmes d’IA militaire.

²⁷ UNIDIR, *op.cit.*

²⁸ Ismailovic, Muzen, «Ciblage algorithmique: le rôle de l’intelligence artificielle dans les frappes israéliennes à Gaza et ses implications éthiques», 20 février 2025. Disponible en version PDF sur [\[Le lien\]](#).

²⁹ Abraham, Yuval, “Lavander: The AI machine directing Israel’s bombing spree in Gaza”, +972 *Magazin*, 3 avril 2024. Disponible sur [\[Le lien\]](#).

³⁰ Nations Unies, “Appel conjoint du Secrétaire Général de l’ONU et de la Présidente du CICR demandant aux États d’énoncer de nouvelles interdictions et limitations pour les systèmes d’armes autonomes», 5 octobre 2023. Disponible sur [\[Le lien\]](#).

Sur ce point, Thierry Berthier en a fait une classification très éclairante.³¹

Ainsi, un système de ciblage humain intégrant de l'IA sous supervision humaine, tel que *Lavender* se situerait dans la catégorie L3 «*Système armé autonome soumis à autorisation de tir*» mais sans la «*composante armée du système*», puisque ce type de système ne va pas jusqu'à utiliser la force létale, c'est bien l'opérateur humain qui le fait.

À l'inverse, un SALA relèverait de la catégorie L5, celle des «*systèmes armés autonomes sans tutelle humaine*»,³² dans laquelle l'opérateur humain ne peut reprendre le contrôle. Le système détermine seul ses déplacements en fonction de sa perception du terrain et de sa mission, détecte et identifie les objets de manière autonome, sélectionne ses cibles via ses capteurs et capacités d'analyse, et peut ouvrir le feu sans intervention humaine, ni possibilité de désactivation, sauf par destruction.

22. Par ailleurs, le rôle croissant des entreprises privées telle que le partenariat conclu entre l'entreprise française Mistral AI avec l'Agence ministérielle pour l'IA de Défense³³ pour la France ou encore l'acquisition par l'OTAN du *Maven Smart System NATO*, développé par Palantir,³⁴ illustre ainsi l'importance de la question

³¹ Berthier, Thierry, «Systèmes armés semi-autonomes: que peut apporter l'autonomie?», *Revue Défense Nationale*, 2019/5, n°820, pp. 74-80. Disponible sur [\[Le lien\]](#).

³² *Ibid.*

³³ Ministère des Armées, Un accord entre l'AMIAD et Mistral IA, «IA de défense: Sébastien Lecornu inaugure le pôle recherche de l'Amiad». Disponible sur [\[Le lien\]](#).

³⁴ Sauveton, Pierre, «IA militaire: l'OTAN choisit Palantir, la France répond avec Artemis, GenIAI... et Comand AI», 23 avril 2025. Disponible sur [\[Le lien\]](#).

pour les décideurs politiques et militaires et soulève des questions sur la régulation de ces nouvelles formes de pouvoir technologique.

Les investissements massifs dans l'IA militaire sont également un facteur de pression. L'IA, en tant que technologie, est perçue par de nombreuses nations comme une clé pour maintenir leur supériorité stratégique, comme en témoigne la déclaration de Vladimir Poutine, «*Celui qui deviendra leader en ce domaine sera le maître du monde*». ³⁵ Les enjeux de sécurité mondiale sont en jeu, et il est impératif de déterminer si le cadre actuel du DIH est suffisamment robuste pour encadrer de manière adéquate l'utilisation de ces nouvelles armes.

23. Problématique. En introduisant une automatisation croissante dans l'analyse, l'identification et potentiellement la neutralisation d'objectifs humains, ces systèmes bouleversent les fondements du droit et de l'éthique militaire. Cette évolution technologique met à l'épreuve les cadres juridiques existants, en particulier le Droit International Humanitaire, dont les principes reposent historiquement sur une appréciation humaine du contexte et des conséquences. L'assemblée générale des Nations Unies l'a affirmé: «*le droit international, en particulier la Charte des Nations Unies, le droit international humanitaire, le droit international des droits humains et le droit pénal international, s'applique aux systèmes d'armes autonome*». ³⁶ Toutefois, la fracture que l'IA entraîne tant d'un point

³⁵ Tanguy, Vincent, «Poutine pense dominer le monde en maîtrisant l'intelligence artificielle», article publié dans la *Revue Sciences et Avenir*, 05/09/2017.

³⁶ Assemblée Générale des Nations Unies, «Résolution adoptée par l'Assemblée générale le 22 décembre 2023», A/RES/78/241. Disponible en version PDF sur [Le lien].

de vue juridique, qu'éthique ou technique, soulève une question centrale. Le Droit International Humanitaire permet-il réellement d'encadrer l'utilisation d'un système d'Intelligence Artificielle destiné au ciblage humain en contexte opérationnel?

24. Méthodologie. L'importance de ce sujet réside dans sa pertinence immédiate dans le contexte géopolitique et militaire actuel. La montée en puissance des technologies d'IA, couplée à une course effrénée à la militarisation de l'Intelligence Artificielle, impose une réflexion urgente sur les cadres juridiques qui gouvernent l'usage de telles technologies. Il s'agit ici d'interroger la capacité du droit international humanitaire à s'appliquer, ou à s'adapter, aux systèmes de ciblage humain fondés sur l'IA en situation opérationnelle. L'analyse portera donc essentiellement sur les SALA et les systèmes de ciblage humain intégrant de l'IA sous supervision humaine. A la fois sous l'angle de leur adaptation au DIH mais aussi sur les difficultés que ces systèmes peuvent poser d'un point de vue juridique et technique.

En concentrant notre attention sur ces deux configurations, nous faisons le choix délibéré de ne pas étudier l'ensemble des usages militaires de l'IA, afin de nous concentrer sur les implications les plus critiques et sujettes à des débats controversés. Ce choix méthodologique s'inscrit dans la logique du DIH, qui accorde une importance centrale à l'humain, tant dans le principe de responsabilité, que dans les principes de distinction, de précaution et de proportionnalité qui nécessitent une mise en œuvre humaine.

25. Annonce de plan. L'introduction croissante de l'Intelligence Artificielle dans les systèmes militaires, notamment dans les fonctions de ciblage, impose un examen rigoureux de leur

compatibilité avec les principes fondamentaux du droit international humanitaire (TITRE 1). Mais face aux limites juridiques actuelles, cette évolution technologique soulève également la nécessité d'une adaptation du DIH, afin d'assurer la protection des personnes dans un contexte marqué par l'automatisation croissante de la violence armée (TITRE 2).

Titre 1: Les systèmes d'IA militaire et leur conformité au DIH

26. En raison de leur autonomie croissante, les systèmes d'Intelligence Artificielle à usage militaire soulèvent d'importantes interrogations quant à leur conformité au droit international humanitaire. Cette partie se propose d'examiner, dans un premier temps, les fondements juridiques encadrant l'usage de la force au regard du DIH (Chapitre 1), avant d'analyser les défis spécifiques posés par l'intégration des systèmes d'IA de ciblage humain dans ce cadre normatif (Chapitre 2).

Chapitre 1: Aspects légaux

27. Afin de mieux cerner les enjeux liés à l'intégration des systèmes d'IA militaires dans le droit international humanitaire, il convient d'abord de revenir sur les origines de ce corpus normatif (Section 1), avant d'en analyser les principes fondamentaux applicables aux forces armées recourant à l'IA (Section 2).

Section 1: Les origines du DIH dans le Jus in Bello

28. Enraciné dans des pratiques anciennes de limitation de la violence en temps de guerre, le DIH s'est d'abord constitué comme un droit essentiellement coutumier, façonné par les usages et les traditions guerrières (§1). Ce n'est que progressivement, notamment à partir du XIXe siècle, qu'un droit humanitaire moderne a émergé (§2).

§ 1 Historique du DIH: un droit par essence coutumier

A) *Le jus in bello: héritage des civilisations anciennes*

29. **Ancrage historique.** Bien avant l'émergence du droit international humanitaire tel que codifié dans les conventions de Genève et les protocoles additionnels du XXe siècle, les sociétés humaines ont, de manière empirique, cherché à imposer des limites à la violence guerrière. Le *jus in bello*, ou droit de la guerre applicable en cours de conflit, puise ses racines dans des pratiques coutumières très anciennes. Dès l'Antiquité, des civilisations comme celles de l'Égypte, de la Grèce ou de Rome adoptaient des règles informelles pour encadrer le traitement des prisonniers, épargner les blessés ou protéger certains lieux sacrés.

B) *Des sagesse antiques à l'éthique guerrière*

30. Pensées fondatrices. Par exemple, dans le Mahabharata, épopée fondatrice de la pensée juridique indienne, il était déjà indiqué qu'il fallait limiter les horreurs de la guerre³⁷ et ne pas attaquer un non-combattant, *«est-ce le grand Indra en personne qui t'a enseigné cet usage des armes? [...] Les sages ne frappent pas quelqu'un qui est distrait ou effrayé, privé de char, ni quelqu'un qui demande grâce, ni quelqu'un qui est dans le malheur?»*³⁸ pouvons-nous lire dans ce récit. De même, dans l'islam ancien, les enseignements de juristes comme Al-Shaybani ont donné lieu à caractériser le comportement idéal du guerrier.³⁹ En Europe, sur une décision du IIème Concile de Latran, en 1139, *«l'usage dans les guerres entre chrétiens, de l'arc et de l'arbalète, considérés comme des engins trop meurtriers»* sont prohibés.⁴⁰ Tous ces textes, souvent issus de la littérature, montrent autant de principes qui annoncent les règles modernes du DIH et montrent que jusqu'au débat encore actuelle des SALA, *«l'impératif de protéger la personne humaine a constamment guidé les tentatives de réguler l'emploi de la force dans les conflits»*.⁴¹

³⁷ François Bugnion, «Droit international humanitaire coutumier», 24 novembre 2007, version PDF disponible sur: [Le lien], consulté le 25/04/2024.

³⁸ Le Mahabharata, «Mort de Bhûrisravas», cité par Gérard Chaliand, «Anthologie mondiale de la stratégie, des origines au nucléaire», p. 455, (1990), Paris, Laffont.

³⁹ Anne de Luca, «Le calife, le soldat et le paysan: l'art de la guerre en islam», *Les Cahiers de l'Orient*, 2017/1 n°125, Article de revue disponible sur: [Le lien], consulté le 25/04/2025.

⁴⁰ Fournier Paul, «Interdiction de l'usage de l'arc et de l'arbalète par le IIe concile de Latran en 1139», *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres*, 1916, 60-3/ pp. 295-296. Disponible sur: [Le lien].

⁴¹ Ministère des Armées, «Manuel de droit des opérations militaires», édition 2022. Disponible sur [Le lien].

C) *De la coutume à la codification du jus in bello*

31. La codification. Toutefois, ces normes restaient souvent morcelées, ancrées dans des contextes culturels spécifiques et fortement imprégnées de références religieuses. C'est seulement à partir du XIXe siècle, avec la 1ère Convention de Genève du 22 août 1864 pour l'amélioration du sort des militaires blessés dans les armées en campagne,⁴² que l'on assiste à une tentative sérieuse de codification universelle du *jus in bello*. Cette évolution témoigne d'une continuité intellectuelle et morale: même dans la guerre, la protection de la dignité humaine ne peut être totalement suspendue. Comme l'affirmait déjà Frédéric de Martens lors de la Conférence de La Haye en 1899, les conflits doivent être régis par «*les lois de l'humanité et les exigences de la conscience publique*»,⁴³ une idée fondatrice qui inspire encore aujourd'hui le droit humanitaire contemporain.

§ 2. *Le DIH moderne*

A) *La naissance du DIH contemporain*

32. Historique du DIH moderne. Le droit international humanitaire moderne, tel qu'il est aujourd'hui reconnu, s'est constitué progressivement à partir de la seconde moitié du XIXe siècle, dans un contexte où les ravages croissants des guerres

⁴² *Convention de Genève du 22 août 1864 pour l'amélioration du sort des militaires blessés dans les armées en campagne*. Genève, 22 août 1864.

⁴³ Rupert Ticehurst, «La clause de Martens et le droit des conflits armés», *Revue Internationale de la Croix-Rouge*. Volume 79, Issue 824, April 1997, pp.133-142. Disponible sur [[Le lien](#)].

industrielles exigeaient une réponse juridique et morale d'envergure.

33. Deux figures déterminantes dans cette entreprise.

C'est sous l'impulsion conjointe d'Henry Dunant, homme d'affaires humaniste, et du général Guillaume-Henri Dufour, officier et stratège militaire respecté, que naît le droit international humanitaire. C'est à la suite de la sanglante bataille de Solferino, à laquelle Dunant assista en 1859, qu'il prit conscience de l'absence totale de secours organisés pour les blessés. De retour à Genève, il publia en 1862 *Un Souvenir de Solferino*, un ouvrage qui relate «*un souvenir de ce que cet homme avait tenté de faire, avec quelques habitants du pays, pour soulager un peu d'innombrables souffrances*». ⁴⁴ Sensible à cette cause, Dufour s'associa à Dunant, et tous deux réunirent, avec Gustave Moynier, Théodore Maunoir et Louis Appia, le «Comité des cinq», qui allait donner naissance au Comité international de la Croix-Rouge (CICR) en 1863. ⁴⁵

B) Naissance des fondements du DIH

34. Adoption de la 1^{ère} Convention de Genève. Sous leur impulsion, le gouvernement suisse convoqua en 1864 une conférence diplomatique réunissant seize États, qui adopta la première Convention de Genève. Celle-ci visait à améliorer le sort des militaires blessés sur les champs de bataille et constitue l'acte de naissance formel du DIH moderne. Cette convention fut la première à poser des normes juridiques contraignantes sur la

⁴⁴ Henry Dunant, *Un souvenir de Solferino*, publié en 1862, préface, p. 5.

⁴⁵ CICR, «Droit international humanitaire, réponses à vos questions», 0703/001 05.2015 15000. Disponible en version PDF.

protection des victimes, tout en reconnaissant la neutralité du personnel médical et des hôpitaux de guerre. Si Dunant en fut l'inspirateur humanitaire, le rôle de Guillaume-Henri Dufour dans la structuration institutionnelle et politique de ce cadre juridique fut tout aussi crucial.

C) Des influences étrangères à la définition du DIH moderne

35. Autres apports internationaux. Parallèlement à cette dynamique européenne, d'autres initiatives ont enrichi le socle normatif du DIH. Pendant la guerre de Sécession (1861-1865), le juriste Francis Lieber élabora dès 1863 un code de conduite militaire, connu sous le nom de «Lieber Code». ⁴⁶ Ce document, considéré comme le premier code militaire codifié et adopté par les Etats-Unis, influencera les futurs développements du droit coutumier en matière de guerre. ⁴⁷ À la croisée des XIXe et XXe siècles, les États ont entamé une véritable codification multilatérale des règles du *jus in bello*. La Convention de Genève de 1906 étend les dispositions initiales de 1864, ⁴⁸ tandis que les Conventions de La Haye de 1899 et 1907 pour «*le règlement pacifique des conflits internationaux*», ⁴⁹ introduisent des règles encadrant l'usage des armes, la conduite des hostilités, et les droits des prisonniers de guerre.

Toutefois, les deux conflits mondiaux du XXe siècle ont mis en lumière les limites criantes de ces instruments, notamment

⁴⁶ Code Lieber, 1863, disponible sur: [**Le lien**].

⁴⁷ Ministère des Armées, «Manuel de droit des opérations militaires», édition 2022. Disponible sur [**Le lien**].

⁴⁸ *Convention pour l'amélioration du sort des blessés et malades dans les armées en campagne*. Genève, 6 juillet 1906. Disponible sur [**Le lien**]

⁴⁹ *Convention de la Haye pour le règlement pacifique des conflits internationaux*, 1899.

en ce qui concerne la protection des populations civiles. À la suite des atrocités commises, et des procès de Nuremberg et de Tokyo où la responsabilité pénale internationale fut pour la première fois consacrée, une nouvelle architecture juridique fut mise en place.

36. L'adoption des Conventions de Genève. Les quatre Conventions de Genève du 12 août 1949, adoptées par la communauté internationale, élargissent de manière décisive le champ de protection du DIH, en intégrant les civils, les blessés, les naufragés et les prisonniers dans des régimes distincts mais complémentaires. Ces textes, régulièrement complétés par des protocoles additionnels, constituent aujourd'hui la pierre angulaire du droit international humanitaire contemporain, fondé sur un principe cardinal, celui du principe de traitement humain qui «est le véritable leitmotiv des quatre Conventions de Genève».⁵⁰

⁵⁰ Jean Pictet (sous la direction de), *Les Conventions de Genève du 12 août 1949: Commentaire, Vol. IV, Convention de Genève relative à la protection des personnes civiles en temps de guerre* (ci-après Commentaire GC IV), CICR, Genève, 1958, p. 219. Cité par Cordula Droegge Original anglais, publié sous le titre "In Truth the *leitmotiv*: the prohibition of torture and other forms of ill-treatment in international humanitarian law", *International Review of the Red Cross*, Vol. 89, No 867, septembre 2007, pp.515-541. Disponible sur [[Le lien](#)]

Section 2: Les principes fondamentaux du DIH

37. Le droit international humanitaire repose sur des principes fondamentaux visant à encadrer l'usage de la force et à en limiter les effets en période de conflit. Il impose notamment un examen préalable de licéité pour toute nouvelle arme ou méthode de guerre (§1), ainsi que le respect des principes de distinction, de proportionnalité et de précaution dans la conduite des hostilités (§2).

§ 1: L'examen de licéité pour toute nouvelle arme, nouveaux moyens ou nouvelles méthodes

A) *L'IA en tant qu'arme, moyen ou méthode de guerre?*

38. Le droit international encadre les méthodes et moyens de conduite des hostilités. Il impose des restrictions concernant tant le type d'armes employées que les modalités de leur utilisation, ainsi que le comportement général des parties impliquées dans un conflit armé.⁵¹

39. Définition de l'arme. La qualification juridique d'une arme repose principalement sur deux critères: la nature de l'équipement et sa destination,⁵² c'est-à-dire l'usage qui en est fait. Dès lors, une technologie peut être considérée comme un matériel de guerre dès lors qu'elle est intégrée à un système spécifiquement conçu ou modifié à des fins militaires. Cela implique une évaluation fondée à la fois sur ses caractéristiques techniques et sur son application concrète. De la même manière, l'évaluation d'une arme

⁵¹ CICR, «Moyens et méthodes de guerre». Disponible sur [[Le lien](#)].

⁵² Serge Rayne, *Répertoire de droit pénal et de procédure pénal*, Armes-Définition, Chapitre 2-Définition (45-68), février 2023.

ne peut être dissociée du contexte opérationnel et des méthodes de guerre dans lesquelles elle est employée.⁵³ Ce processus relève du contrôle de licéité, qui sera abordé ultérieurement dans ce développement.⁵⁴

40. En France, la définition juridique de l'arme est précisée à l'article 132-75 du Code pénal, selon lequel: «*Est une arme tout objet conçu pour tuer ou blesser*». D'un point de vue plus large, sur le plan technique, une arme est définie comme tout objet, qu'il soit matériel ou immatériel, conçu ou utilisé dans le but de porter atteinte à l'intégrité physique ou psychique d'un individu.⁵⁵ Cette approche élargie ouvre la voie à l'inclusion de nouvelles technologies, telles que les systèmes autonomes reposant sur l'Intelligence Artificielle, dans la catégorie des armes.

41. S'agissant de l'Intelligence Artificielle, il convient de préciser que celle-ci, prise isolément, ne constitue pas une arme au sens traditionnel. En effet, l'IA ne cause pas directement de destruction ou de blessures. Toutefois, elle peut être intégrée dans des systèmes d'armement autonomes, comme les drones de combat autonomes, les systèmes de ciblage automatique ou les systèmes d'armes létaux autonomes, dans lesquels elle joue un rôle déterminant. Ainsi, pour la recherche qui nous intéresse, l'usage de l'IA dans le domaine militaire, il est nécessaire de faire une

⁵³ Caroline Brandao, «L'analyse des nouveaux systèmes d'armes robotisés sous le prisme de l'article 36 du Protocole additionnel I», publié dans la revue *Défense Nationale*, 2018/HS1 n°Hors-Série, pp. 133 à 143.

⁵⁴ Voir *infra*, paragraphe 49.

⁵⁵ Cumin, David, *Manuel de droit de la guerre*, Bruxelles: Bruylant, Collection Paradigme, 2^{ème} ed., 2020, p.188.

distinction entre une technologie d'IA qui n'est pas une arme en soi et les usages que nous pouvons faire de cette technologie, notamment à des fins militaires, qui peuvent être vus comme des méthodes et moyens de combats. Dans ce cadre, l'IA devient une composante essentielle, permettant au système d'identifier des cibles, de prendre des décisions et, potentiellement, d'exécuter des frappes sans intervention humaine directe,⁵⁶ ce qui fait de cet ensemble, une arme.

42. Ainsi, si l'IA, en tant que telle, ne répond pas seule à la définition d'une arme, son intégration au sein de dispositifs capables d'exercer une force létale la positionne comme une composante qui rend possible l'autonomie de certaines fonctions d'un système d'armement. Combinée à des éléments létaux, l'IA contribue directement à la conduite des hostilités et peut, à ce titre, être considérée comme une arme au regard des principes du DIH.

43. Le cadre juridique. Les deux règles substantielles du droit des armes qui doivent être prises en compte sont la règle contre les armes intrinsèquement indiscriminées et la règle contre les armes qui causent des souffrances inutiles.

1- La règle contre les armes intrinsèquement indiscriminées

44. L'article 51, 5, b du Protocole additionnel n°1 aux Conventions de Genève dispose que «5. *Seront, entres autres, considérés*

⁵⁶ Sans intervention humaine directe entendu comme le fait pour le système d'exécuter sa mission une fois activé.

comme effectués sans discrimination les types d'attaques suivants: b) les attaques dont on peut attendre qu'elles causent incidemment des pertes en vies humaines dans la population civile, des blessures aux personnes civiles, des dommages aux biens de caractère civil, ou une combinaison de ces pertes et dommages, qui seraient excessifs par rapport à l'avantage militaire concret et direct attendu».

En faisant l'analogie de cet article aux systèmes d'armes autonomes intégrant de l'IA, ces derniers pourraient être qualifiés d'«*armes de destruction indiscriminées*» selon l'informaticien N. Sharkey.⁵⁷ C'est-à-dire que l'arme est indiscriminée par nature si elle ne peut être dirigée vers une cible spécifique et est aussi susceptible de toucher des civils que des combattants.

Toutefois, selon Kenneth Anderson et Matthew C. Waxman l'autonomie en soi des armes létales autonomes ne viole pas cette règle, tant que le système peut être alimenté avec des informations de ciblage fiables.⁵⁸

2- La règle contre les armes qui causent des souffrances inutiles

45. L'article 35, 2) du Protocole Additionnel n°1 aux Conventions de Genève dispose qu' «2. *Il est interdit d'employer des armes, des projectiles et des matières ainsi que des méthodes de guerre de nature*

⁵⁷ Noël Sharkey, "Weapons of Indiscriminate Lethality", in *Forum des informaticiens pour la paix et la responsabilité sociale* (FIFF Kommunikation), n°1/09, pp. 26-29. Cité par J. Ancelin: «Les systèmes d'armes létaux autonomes (SALA): Enjeux juridiques de l'émergence d'un moyen de combat déshumanisé», 2016, paragraphe 3. Disponible en version PDF.

⁵⁸ Kenneth Anderson et Matthew C. Waxman, "Law and ethics for autonomous weapons systems: why a ban won't work and how laws of war can", *American University Washington College of Law*, Articles in law reviews & other academic journals, Research Paper No. 2013-11, 2013. Disponible sur [[Le lien](#)].

à causer des maux superflus». Toujours selon l'étude réalisée par Kenneth Anderson et Matthew C. Waxman, le fait qu'un système autonome sélectionne la cible ou attaque ne viole pas cette règle.⁵⁹

3- *Limitation et interdiction de certaines armes de guerre*

46. Le Droit International Humanitaire régit strictement les méthodes et moyens de guerre afin de limiter les souffrances humaines lors des conflits armés. En vertu de ses principes fondamentaux, il interdit les armes et tactiques qui causent des maux superflus ou des souffrances inutiles, ainsi que celles qui ne font pas de distinction entre civils et combattants. Cette interdiction s'applique non seulement à l'emploi d'armes aux caractéristiques intrinsèquement indiscriminées, mais également à certaines méthodes spécifiques de conduite des hostilités.

47. Dans cette dynamique de limitation des moyens de guerre, l'attention internationale s'est tournée vers les systèmes d'armes létaux autonomes souvent qualifiés de «robots tueurs». En effet, dès 2013, plusieurs ONG se sont saisies de la question, notamment Human Rights Watch, et ont lancé la campagne to “Stop Killer Robots” qui appelait déjà de ses vœux *«à une interdiction préventive sur le développement, la production et l'utilisation d'armes totalement autonomes»*.⁶⁰

48. Ces technologies, encore à l'heure actuelle, suscitent de vives préoccupations juridiques, éthiques et sécuritaires. En effet,

⁵⁹ *Ibid.*

⁶⁰ Human Rights Watch, «Robots tueurs», disponible sur [Le lien].

confier à des machines la capacité de décider, sans intervention humaine directe, du recours à une force létale,⁶¹ c'est-à-dire de tuer ou non un individu, soulève des interrogations majeures quant au respect du principe de distinction, de proportionnalité et de responsabilité en droit international humanitaire que nous développerons par la suite.⁶²

B) *Article 36 du Protocole Additionnel I*

1- *L'examen de licéité pour les SLA de défense*

49. L'article 36 du Protocole Additionnel I aux Conventions de Genève joue un rôle crucial en imposant un examen de licéité pour toute nouvelle arme, ce qui inclut les systèmes basés sur l'IA. Cet article dispose 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*». ⁶³ En ce sens, il en va de la responsabilité des États de déterminer si l'emploi d'une nouvelle méthode de guerre, telle que l'usage de l'Intelligence Artificielle dans les conflits armés, pourrait être interdit ou non. Plusieurs États ont à ce sujet, déjà pris parti tels que les Etats-Unis et le Royaume-Uni en «*refusant d'adopter des systèmes létaux entièrement*

⁶¹ Amnesty international, «Quelles armes sont interdites par le droit international humanitaire?» disponible sur [[Le lien](#)].

⁶² Voir *infra*, paragraphes n°54 et suivants.

⁶³ Article 36 du *Protocole Addiitionnel aux Conventions de Genève du 12 août 1949 relatif à la protection des victimes des conflits armés internationaux* (protocole 1), du 8 juin 1977.

autonome».⁶⁴ Néanmoins, ce n'est qu'une partie des systèmes intégrant de l'IA. En effet, l'IA pourrait aussi s'envisager comme une technologie de ciblage humain seulement sans avoir la dimension létale. C'est notamment le cas du système d'IA utilisé par l'armée israélienne dans le conflit Israélo-Palestinien avec leur programme nommé *Lavender*. Toutefois, au lieu d'interdire ces systèmes, il pourrait être envisageable de programmer des intelligences artificielles qui seront utilisées dans le cadre opérationnel pour qu'elles soient conformes aux prescriptions du DIH, thèse que nous reverrons plus tard dans le développement (n°57 à 60). Ainsi, ces «*nouveaux moyens de guerre*»⁶⁵ réperandraient à l'examen de licéité exigé par l'article 36 du Protocole n°1 aux Conventions de Genève.

50. Comme évoqué précédemment, il est nécessaire d'effectuer une distinction entre une technologie d'IA, qui ne constitue pas en soi une arme et qui, par conséquent n'est pas concernée par l'examen de licéité prévu par cet article 36 et une IA qui serait intégrée dans une arme. Dans ce cas, si nous considérons que certaines technologies d'IA sont utilisées à des fins militaires et ayant des conditions d'emploi qui peuvent être regardées comme étant des méthodes et moyens de combat, l'examen de licéité de l'article 36 s'impose. Dès lors, dans son avis sur l'usage des technologies d'Intelligence Artificielle par les forces armées, le

⁶⁴ Nevejans, Nathalie, «La légalité des robots de guerre dans les conflits internationaux», *Recueil Dalloz* 2016, p. 1273.

⁶⁵ Smarandescu, Ioana-Alexandra et VARZARU, Nina-Maria, "The new means of waging war. Artificial intelligence in the field of international humanitarian law", *Romanian Journal of International Law*, RJIL No.30/2024, p. 35-54.

Comité Éthique de la Défense préconise dans sa recommandation n°1 que *«cet examen est conduit, en tant que de besoin, lors des différentes phases du cycle de vie d'un système d'armes: en phase de préparation, de réalisation mais aussi d'utilisation opérationnelles»*.⁶⁶

2- *Révision de la licéité des armes intégrant un logiciel d'IA dans leur évaluation*

51. L'exigence de l'article 36 devrait s'étendre également aux armes existantes dès lors qu'elles subissent une modification substantielle, notamment par l'intégration d'un logiciel d'IA, ce qui en ferait potentiellement une nouvelle arme, pour une tout autre utilisation. En effet, comme le souligne Joanne Kirkham, *«toute arme faisant l'objet d'une modification substantielle devrait nécessairement faire l'objet d'une nouvelle procédure d'examen»*.⁶⁷ L'adjonction de l'IA peut modifier de manière significative les caractéristiques opérationnelles d'une arme, notamment en termes de prise de décision autonome, de vitesse de traitement des informations et d'engagement des cibles, des éléments qui justifient un réexamen rigoureux au regard du droit international humanitaire. À cet égard, l'intégration du missile *Hellfire* aux drones *Predator* en constitue un exemple. Ces drones, initialement conçus pour des missions de surveillance, ont acquis une dimension létale lorsqu'ils ont été équipés de ces mis-

⁶⁶ Comité éthique de la Défense, «Avis sur l'usage des technologies d'intelligence artificielle par les forces armées», 14 janvier 2025.

⁶⁷ Joanne Kirkham, *Les systèmes d'armes létaux autonomes: une étude de droit international*, thèse de doctorat en droit public soutenue le 10 décembre 2022, École doctorale de droit international droit européen, relations internationales et droit comparé, Université Paris Panthéon-Assas. p. 99, paragraphe 123. *Op.cit.* Disponible sur le [Le lien].

siles.⁶⁸ Or, la condition de licéité ne s'applique pas aux systèmes de surveillance, il est donc *«facile de les armer, les transformant de fait en armements, sans avoir jamais passé le test»*.⁶⁹ Dès lors, cette transformation aurait dû, en vertu de l'article 36, mener à une réévaluation du respect de leurs conditions d'usage en vertu du droit international.⁷⁰

52. L'intégration d'une IA dans un système d'armement soulève la question de l'autonomie de ces dispositifs, qui, selon Nathalie Nevejans, est définie par leur capacité à fonctionner sans intervention humaine directe.⁷¹ L'autonomie des systèmes d'armement basés sur l'IA peut être envisagée comme un continuum dans lequel les systèmes varient en fonction du degré de contrôle humain, ce dernier n'étant pas absolu mais modulé en fonction des phases d'engagement des armes.⁷² Dès lors, il devient crucial de déterminer si un humain conserve la capacité d'intervenir pour stopper ou modifier l'action de ces systèmes, ce qui permet

⁶⁸ Must Paris, «MQ-1 Predator – Rôle et évolution du drone de combat US», 1 décembre 2024, disponible sur: [[Le lien](#)], consulté le 27/04/2024.

⁶⁹ Jean-Baptiste Jeangène Vilmer, «Robots létaux autonomes: amorce d'un débat», décembre 2012. Disponible en version PDF.

⁷⁰ Il est précisé dans l'article susmentionné que «Les Etats-Unis ont accepté l'obligation d'évaluer la légalité de leurs armes nouvelles (dans une directive du *Defense Acquisition Guidebook* depuis 2003), mais rien n'indique pour l'instant qu'ils interdiraient des armes avantageuses à la légalité discutable».

⁷¹ Nathalie Nevejans, «La légalité des robots de guerre dans les conflits internationaux» (EA n 2471), *Recueil Dalloz* 2016, p.1273.

⁷² *Convention sur l'interdiction ou la limitation de l'emploi de certaines armes classiques qui peuvent être considérées comme produisant des effets traumatiques excessifs ou comme frappant sans discrimination*, «Rapport du Groupe d'experts gouvernementaux sur les technologies émergentes dans le domaine des systèmes d'armes létaux autonomes sur sa session de 2021», 22 février 2022.

de garantir que l'usage de telles armes ne contrevient pas aux principes du droit international humanitaire et serait donc licite en vertu de l'article 36. En outre, le concept de «*contrôle humain significatif*»,⁷³ exige que l'interaction humaine avec ces systèmes ne se limite pas à une supervision marginale mais implique une possibilité d'intervention effective à chaque étape clé du processus de ciblage et d'engagement.⁷⁴

53. Enfin, conformément à l'esprit de l'article 36, l'examen de licéité ne concerne pas uniquement la conformité d'une arme dans un cadre statique, mais doit également prendre en compte les modalités d'utilisation concrètes. Ainsi, une arme qui aurait été initialement jugée conforme aux règles du droit international humanitaire pourrait, en raison de sa modification, devenir illégale si son emploi évolue de manière substantielle. Cela est particulièrement pertinent lorsque l'intégration d'une IA confère à l'arme une autonomie décisionnelle accrue, comme cela pourrait être le cas avec des armes susceptibles de prendre des décisions de tir sans intervention humaine directe, à savoir les SALA. Ce type de changement exige une révision continue de l'examen de licéité afin de garantir que le système reste en conformité avec les principes fondamentaux du droit de la guerre, en particulier la distinction et la proportionnalité.⁷⁵

⁷³ Webinar of United Nations, Office for Disarmament Affairs, "Experts Reflect on Novel Approaches to "Meaningful Human Control", 7 mars 2023. Disponible sur [[Le lien](#)].

⁷⁴ Campaign Stop Killer Robot, Note d'information, «Maintenir le contrôle sur les systèmes d'armes». Disponible en version PDF sur [[Le lien](#)]

⁷⁵ CCAC, Rapport du 22 février 2022, *op.cit.*

§ 2. Principe de Distinction, Proportionnalité et Précaution

54. Pour comprendre comment le Droit international humanitaire peut encadrer l'utilisation de l'IA dans les conflits armés, il est essentiel de commencer par examiner ses principes fondamentaux, notamment ceux de distinction, proportionnalité et précaution, qui guident les actions militaires, principes issus directement de la tradition du droit de la guerre, le *jus in bello* (cf Section 1).

A) Le principe de distinction

55. Le principe de distinction implique que *«les Parties au conflit doivent en tout temps faire la distinction entre la population civile et les combattants ainsi qu'entre les biens de caractère civil et les objectifs militaires et, par conséquent, ne diriger leurs opérations que contre des objectifs militaires»*.⁷⁶ Dans un conflit armé international, les combattants sont autorisés à prendre part directement aux hostilités, ce qui signifie qu'ils peuvent mener des actions de guerre légitimes pour atteindre un objectif militaire de manière efficace. Toutefois, le principe de distinction impose que ces actions ne visent que les combattants ennemis et les objectifs militaires, tout en limitant les dommages inutiles et excessifs aux civils.⁷⁷

56. **Distinction inefficace.** Dès lors, au regard du principe de distinction tel que défini en droit international humanitaire, il apparaît que l'incapacité des Robots Létaux Autonomes à distinguer de manière fiable entre combattants et civils pourrait

⁷⁶ *Protocole additionnel aux Conventions de Genève relatif à la protection des victimes des conflits armés internationaux* (Protocole I), 1977, art. 48.

⁷⁷ Comité International de la Croix-Rouge, «Moyens et méthodes de guerre» disponible sur: [[Le lien](#)]

rendre leur utilisation juridiquement illicite. En effet, le risque que des civils soient pris pour cibles en raison d'erreurs de ciblage imputables à des systèmes automatisés⁷⁸ constitue une atteinte grave à ce principe fondamental.

57. **Programmer pour distinguer.** En revanche si nous pouvons programmer les RLA en les obligeant à respecter les prescriptions du DIH, alors leur utilisation en serait totalement conforme et permettrait même que *«dans certains contextes, la technologie offre une meilleure précision»*.⁷⁹ En effet, Christof Heyns, ancien rapporteur de l'Assemblée Générale des Nations Unies, disait déjà en 2013 que *«les robots ont ainsi la capacité d'agir avec prudence et de ne tirer qu'en riposte. En outre, dans certains cas, les capteurs puissants des RLA et leur capacité de traitement de l'information peuvent aider les soldats à y voir plus clair dans le « brouillard de la guerre » et ainsi prévenir le genre d'erreur qui, dans les conflits armés, mène souvent à des atrocités et épargner des vies»*.⁸⁰ Nous pourrions donc ainsi envisager que les systèmes intégrant de l'IA puissent respecter ce principe de distinction, principe essentiel du DIH. Toutefois, l'argumentaire inverse pourrait aussi être envisagé. Pour les besoins de cette étude, cet examen sera effectué dans la section technique.⁸¹

⁷⁸ Anthony Cardia, «Les conflits armés à l'époque de l'intelligence artificielle», 26 novembre 2024. Disponible sur [\[Le lien\]](#).

⁷⁹ Assemblée générale des Nations Unies, Rapport du Rapporteur spécial sur les exécutions extrajudiciaires, sommaires ou arbitraires, Christof Heyns, 9 avril 2013

⁸⁰ *Ibid.*

⁸¹ Voir Section 1, par. 65 et suivants.

B) *Le principe de proportionnalité*

58. Concernant le principe de proportionnalité, ce dernier est défini par le Protocole additionnel aux Conventions de Genève relatif à la protection des victimes des conflits armés internationaux comme étant un principe qui *«interdit de lancer des attaques dont on peut attendre qu'elles causent incidemment des pertes en vies humaines dans la population civile, des blessures aux personnes civiles, des dommages aux biens de caractère civil, ou une combinaison de ces pertes et dommages, qui seraient excessifs par rapport à l'avantage militaire concret et direct attendu»*.⁸²

59. Sur ce principe, le rapporteur général de la Commission des Nations Unies a soulevé que *«l'incapacité des RL A d'appréhender la situation et de la situer dans son contexte pourrait avoir pour conséquence qu'un robot décide de lancer une attaque en se fondant sur une lecture non seulement incomplète, mais aussi erronée des circonstances. Il convient cependant de garder à l'esprit que cela peut également être le cas chez l'homme»*.⁸³

60. En ce sens, si l'on admet qu'un individu peut également commettre des fautes lors des opérations militaires, et que les machines sont susceptibles d'effectuer des erreurs similaires, il semble difficile de justifier leur interdiction. Dès lors, il paraît pertinent de considérer que certains systèmes d'armement, sous réserve qu'ils soient programmés conformément aux principes du DIH et utilisés exclusivement dans des situations strictement

⁸² Protocole I additionnel aux Conventions de Genève relatif à la protection des victimes des conflits armés internationaux, 1977, art. 51 5) b) et 57 2) iii.

⁸³ Assemblée Générale des Nations Unies, Rapport du Rapporteur spécial sur les exécutions extrajudiciaires, sommaires ou arbitraires, Christof Heyns, 9 avril 2013.

encadrées, pourraient être compatibles avec ce cadre juridique. Cette position est notamment défendue par Ronald Arkin, qui soutient qu’ *«il y a des raisons de croire qu’à l’avenir, les systèmes robotiques seront capables de percer le brouillard de la guerre plus efficacement que les humains ne pourraient jamais le faire»*.⁸⁴ Dès lors, c’est précisément parce que l’humain fait des erreurs, commet des crimes, est soumis à des émotions fortes, et que le robot étant neutre, il serait plus moral et donc il serait plus favorable de n’envoyer qu’eux sur le terrain. R. Arkin va encore plus loin en disant qu’ *«il pourrait être possible de créer à terme des systèmes militaires robotiques autonomes intelligents capables de réduire les pertes civiles et les dommages matériels par rapport aux performances des combattants humains»*.⁸⁵ Cela serait donc dans l’intérêt du respect des principes de DIH que de réduire les pertes civiles.

61. Toujours dans cette optique de programmer les SIA pour qu’ils soient conformes au DIH, le Comité Éthique de la Défense a publié un avis recensant divers *«cas d’usage possibles pour les armées»*,⁸⁶ couvrant un large éventail d’activités, qu’il s’agisse de systèmes d’Intelligence Artificielle embarqués ou non dans les plateformes et armements. Parmi ces cas, on trouve des applications telles que *«l’aide à la préparation opérationnelle»*, *«l’aide au renseignement»*, *«l’aide à la*

⁸⁴ Ronald Arkin, “Lethal Autonomous Systems and the Plight of the Non-combatant”, No 137, juillet 2013. Publié dans le livre *The Political Economy of Robots: Prospects for Prosperity and Peace in the Automated 21st Century* (International Political Economy Series), 1st ed. 2018 Edition, par Ryan Kiggins (Editeur). Disponible sur [[Le lien](#)].

⁸⁵ *Ibid.*

⁸⁶ Comité éthique de la Défense, «Avis sur l’usage des technologies d’intelligence artificielle par les forces armées», 14 janvier 2025.

surveillance et la sûreté», ainsi que «*l'aide à la décision, voire la décision pour le combat*». Le ciblage, qui figure également dans cette liste, pourrait ainsi, dans le cadre d'une programmation conforme au DIH, être encadré et, par conséquent, être considéré comme une utilisation licite de ces technologies.

C) *Le principe de précaution*

62. Prévu par l'article 57(1) du Protocole Additionnel I, le principe de précaution prévoit que les «*opérations militaires doivent être conduites en faisant tout ce qui est pratiquement possible pour vérifier que les objectifs à attaquer ne sont ni la population civile, ni des biens de caractère civils*»,⁸⁷ ou des catégories de personnes protégées au titre des Conventions de Genève.

63. Ce principe repose avant tout sur un jugement éthique et humain, impliquant la nécessité d'un contrôle humain. Toutefois, l'utilisation des SALA et de l'Intelligence Artificielle n'est pas en soi incompatible avec ce principe. En effet, pour être conforme à cette exigence, «*l'action des SALA doit être prévisible. Les utilisateurs doivent être en mesure d'ajuster ou d'annuler les effets des systèmes d'armement si nécessaire, ce qui n'est possible que si l'action et la réaction de ces systèmes peuvent être raisonnablement anticipées*».⁸⁸ Dès lors, si un SIA est totalement transparent, et qu'il est possible à tous les niveaux du cycle de vie de ce SIA de l'expliquer, il pourrait tout à fait être

⁸⁷ Article 57 (1) du *Protocole Additionnel n°1 aux Conventions de Genève du 12 août 1949 relatif à la protection des victimes des conflits armés internationaux*, 8 juin 1977.

⁸⁸ Conseil de l'Europe, Commission des questions juridiques et des droits de l'homme, «*L'Émergence des systèmes d'armes létales autonomes (SALA) et leur nécessaire appréhension par le droit européen des droits humains*», Projet de résolution adopté à l'unanimité par la Commission le 14 novembre 2022.

conforme au principe de précaution. En outre, dans le cas d'un simple système d'IA de ciblage n'allant pas jusqu'à tuer une cible potentielle, le contrôle humain est normalement indiscutable puisque c'est bien l'humain qui prend la décision ou non de tuer. Néanmoins, il apparaît que sur ces simples systèmes, le contrôle humain ne soit pas si évident, c'est ce que nous verrons plus tard.⁸⁹

⁸⁹ Le cas des SIA de ciblage humain sous supervision humaine, par.160 et suivants.

Chapitre 2: Les défis de l'intégration d'un SIA de ciblage humain dans le DIH

64. L'introduction de systèmes d'Intelligence Artificielle dans les opérations de ciblage soulève de nombreux défis quant à leur conformité au droit international humanitaire. Ces enjeux sont d'abord d'ordre technique (Section 1) mais aussi géopolitique (Section 2).

Section 1: Les difficultés techniques

65. L'intégration des systèmes d'IA dans les fonctions de ciblage militaire se heurte à d'importantes limites techniques. D'une part, ces systèmes présentent des lacunes cognitives et décisionnelles qui compromettent leur capacité à comprendre des environnements complexes et à réagir de manière conforme aux exigences du DIH (§1). D'autre part, leur déploiement opérationnel soulève des risques concrets, notamment en matière de biais algorithmique et de vulnérabilité (§2).

§ 1. Les limites cognitives et décisionnelles des systèmes d'Intelligence Artificielle

A) L'opacité et l'imprévisibilité des systèmes d'Intelligence Artificielle

1- Les algorithmes d'IA par les forces armées

66. **Les IA faibles.** Premièrement, il n'existe actuellement qu'un type d'IA: l'IA «faible». Ce sont celles qui sont programmables par les humains. Ce type d'IA là permet de faire des tâches spécifiques, il est donc considéré comme intelligent, car l'algo-

rithme peut résoudre de lui-même une tâche qui lui a été préalablement assignée.

L'IA faible a notamment pu contribuer à plusieurs tâches précises telles que la reconnaissance faciale, la reconnaissance vocale, etc.⁹⁰

67. **Les IA fortes.** Par ailleurs, avec l'essor de plus en plus fulgurant des données, les investissements colossaux injectés dans l'IA, un autre type pourrait potentiellement voir le jour. Il s'agit des IA dites «fortes». Ce type d'IA n'est pas encore une réalité concrète mais d'après les scientifiques, l'IA forte serait capable d'imiter l'intelligence et le raisonnement humain. Dès lors, *«l'ordinateur convenablement programmé est véritablement un esprit»*⁹¹ qui pense et qui peut prendre des décisions seul, résoudre des problèmes seul comme le ferait un humain.

68. Ainsi, dans un avenir plus ou moins proche, les IA fortes seront vraiment des «intelligences» artificielles en ce qu'elles seront capables de résoudre tout type de problèmes. Jean Mosconi dira même que *«Tout comportement humain qui peut être décrit avec précision peut être simulé par un ordinateur convenablement programmé»*.⁹² Dès lors

⁹⁰ Article de blog, Datasulting, «Algorithme, Intelligence Artificielle, Machine Learning, Deep Learning: quelles différences?», 22 juin 2023. Disponible sur [Le lien].

⁹¹ Patrick Goutefangea, «Alan Turing et l'intelligence artificielle: le «jeu de l'imitation» et l'IA forte», 24 juin 2017, Hal open science. Disponible sur: [Le lien].

⁹² Jean Mosconi, «Sur quelques capacités et incapacités des machines», *Bulletin de la Société Française de Philosophie*, 3, juillet-septembre 1991, p. 86. Cité par Patrick Goutefangea, *op.cit.*

ces IA, à l'image de l'intelligence humaine, seront soumises aux principes du DIH.

69. La *superintelligence artificielle*. Certains estiment même qu'il existerait un troisième type d'IA: les supers intelligences artificielles.⁹³ Ces formes d'IA totalement futuristes dépasseraient l'intelligence humaine. Dans cette hypothèse, la *superintelligence artificielle* n'imiterait plus l'humain, elle deviendrait une nouvelle espèce, une «espèce numérique».

70. Globalement, la communauté scientifique s'accorde à dire qu'il existe deux types d'IA: les IA faibles, celles que nous connaissons actuellement, et les IA fortes, un idéal pour certains, une dystopie pour d'autres.

71. Les composantes de l'IA. Dans le domaine de l'intelligence artificielle, le Machine Learning (ML), ou apprentissage automatique, constitue un sous-ensemble essentiel. Il repose sur la capacité des machines à apprendre à partir de données, en passant par une phase d'entraînement qui consiste pour le système à «*apprendre des règles implicites en se basant sur ces données*»,⁹⁴ sans être explicitement programmées pour chaque tâche. Le ML englobe plusieurs types d'apprentissage: supervisé, basé sur des données étiquetées par des humains, auto-supervisé, où le système génère

⁹³ Tim Mucci et Cole Stryker, «Qu'est-ce que la superintelligence artificielle», IBM, 18 décembre 2023. Disponible sur [[Le lien](#)].

⁹⁴ Ludovic De Matteis, Steeven Janny, Solal Nathan et Wenqi Shu-Quartier, «Introduction à l'apprentissage automatique», École normale supérieure, Paris-Saclay, édité le 24/05/2022. Disponible sur [[Le lien](#)].

lui-même les étiquettes et par renforcement, dans lequel l'agent apprend par essai-erreur en fonction de récompenses ou punitions.

Le *Deep Learning* (DL), ou apprentissage profond, est une branche du Machine Learning qui s'appuie sur des réseaux de neurones artificiels composés de nombreuses couches successives. Ce type d'apprentissage est particulièrement performant pour traiter des données complexes et non structurées, comme les images ou le langage. Aujourd'hui, le Deep Learning est couramment utilisé dans le développement des véhicules autonomes,⁹⁵ et par extension, il pourrait s'appliquer aux systèmes d'armes létales autonomes.

72. La transition du ML classique vers DL marque une rupture significative dans la manière dont les modèles traitent les données. Alors que les approches traditionnelles reposaient encore fortement sur une intervention humaine pour la sélection des caractéristiques pertinentes et la structuration préalable des données, l'apprentissage profond permet désormais aux algorithmes d'extraire eux-mêmes des représentations complexes et pertinentes, sans nécessiter d'a priori humain sur la nature ou la structure des informations traitées.⁹⁶

73. Ces deux sous-ensembles que sont le Machine Learning et le Deep Learning forment l'IA et sont donc utilisés pour la création de SIA militaire. Le Comité International de la Croix-Rouge

⁹⁵ Muhammad Tuhin, "7 examples of How Deep Learning Is Used Today", *Science News Today*, 30 avril 2025. Disponible sur [[Le lien](#)].

⁹⁶ Auguste Lehuger, «Applications récentes de l'apprentissage machine», Cours de M2 Droit de l'Intelligence artificielle, Année 2024-2025.

s'intéresse particulièrement à l'utilisation du Machine Learning dans son utilisation dans la conduite de la guerre. Pour le CICR, «*Les logiciels d'IA et d'apprentissage automatique – notamment ceux développés pour la «reconnaissance automatique des cibles» – pourraient constituer la base des futurs systèmes d'armes autonomes, apportant une nouvelle dimension d'imprévisibilité à ces armes*». Par ailleurs, le CICR reconnaît que ce type de logiciel pourrait aussi être utilisé pour fabriquer des systèmes qui n'iraient pas jusqu'à la létalité, mais seulement d'aide à la décision pour le ciblage.⁹⁷

2- *Explicabilité des algorithmes: l'effet «boîte noire»*

74. Enjeu. L'explicabilité des algorithmes d'Intelligence Artificielle, en particulier ceux fondés sur l'apprentissage automatique et plus spécifiquement l'apprentissage profond, constitue aujourd'hui un enjeu scientifique, technique et éthique de premier ordre. En effet, si les systèmes d'IA atteignent des performances remarquables dans divers domaines, ils présentent néanmoins un défaut structurel majeur : leur opacité intrinsèque.⁹⁸

75. Définition. Ce phénomène, souvent qualifié de «problème de la boîte noire», désigne l'incapacité, même pour les concepteurs, à comprendre ou à expliquer de manière intelligible le raisonnement sous-jacent à une décision algorithmique. Comme le souligne le rapport Villani, «*l'explicabilité des systèmes à base*

⁹⁷ ICRC Position Paper: “Artificial intelligence and machine learning in armed conflict: A human-centred approach”, mars 2021. Disponible sur [**Le lien**].

⁹⁸ Maël Pégy, Mohamed Issam Ibnouhsein, «Quelle transparence pour les algorithmes d'apprentissage machine?», mai 2018, pp. 7-8. HAL Id: hal-01791021. Disponible sur [**Le lien**].

d'apprentissage constitue donc un véritable défi scientifiques.⁹⁹ À la différence des programmes déterministes classiques, où les règles de décision sont explicites, traçables et compréhensibles, les modèles issus du Machine Learning reposent sur la manipulation de millions de paramètres, notamment dans les réseaux de neurones profonds, évoluant dans des espaces de données à très haute dimension. Cette complexité rend quasi impossible l'identification des variables déterminantes ayant conduit à une décision donnée, en particulier dans des cas individuels.

76. L'absence de transparence ne soulève pas seulement des obstacles cognitifs pour les utilisateurs ou les régulateurs, elle entraîne également des risques éthiques et juridiques considérables. L'opacité des systèmes peut en effet dissimuler l'origine de biais discriminatoires, sans qu'il soit possible de déterminer si ceux-ci proviennent de l'algorithme, des données d'entraînement, ou d'une interaction entre les deux. Comme le souligne la Professeure Marie-des-Neiges Ruffo, *«l'emploi de l'apprentissage machine peut potentiellement laisser dans la "black box", autrement dit dans l'opacité, la manière dont la machine a procédé à sa décision»*,¹⁰⁰ ouvrant ainsi la voie à des dérives où l'absence d'explication technique pourrait servir de prétexte à l'absence de justification juridique.¹⁰¹

⁹⁹ Cédric Villani, «Donner un sens à l'intelligence artificielle», rapport du 28 mars 2018. Rapport officiel. Mission confiée par le Premier Ministre Édouard Philippe. Mission parlementaire du 8 septembre 2017 au 8 mars 2018. Disponible sur [[Le lien](#)].

¹⁰⁰ MdN Ruffo, «Avons-nous le choix d'utiliser l'IA en temps de guerre?», 47^e numéro des Cahiers de la sécurité et de la justice, 12 mars 2020. Disponible sur [[Le lien](#)]

¹⁰¹ *Ibid.*

77. L'opacité des systèmes d'armes létaux autonomes.

Ces enjeux deviennent particulièrement critiques lorsqu'il s'agit de systèmes d'armes autonomes. Le recours à des dispositifs décisionnels opaques dans des contextes militaires soulève une série de questions fondamentales relatives à la responsabilité, à la conformité au droit international humanitaire et à l'éthique de la délégation létale. Dans un contexte de conflit armé, l'impossibilité de retracer la logique d'un tir ou d'une cible sélectionnée de manière autonome compromet non seulement la possibilité d'établir des responsabilités juridiques, mais aussi celle de corriger, prévenir ou même reconnaître des erreurs techniques ou morales. Dès lors, l'absence d'explicabilité ne peut être considérée comme une simple limite technique, mais bien comme une contrainte systémique susceptible d'affaiblir les fondements mêmes de la gouvernance humaine sur la force armée.

78. Contrôle humain. Dans ce contexte, la mise en place d'un contrôle humain significatif sur les systèmes d'IA déployés dans des situations sensibles, notamment militaires, apparaît comme un principe de précaution essentiel. Ce contrôle viserait à garantir la traçabilité, la compréhension et, le cas échéant, la justification des décisions prises par les systèmes autonomes, dans une logique de responsabilisation et de conformité au cadre normatif existant. C'est exactement ce que préconise la recommandation n°7 du comité éthique de la défense, qui dispose que *«La formation des militaires doit favoriser leur compréhension des documents de transparence et surtout favoriser leur capacité à détecter que la fonction assurée par une technologie d'Intelligence Artificielle est dégradée, polluée ou insuffisante et qu'il*

est nécessaire de reprendre la main». ¹⁰² La formation des militaires conseillée par le Comité Éthique de la Défense rejoint le principe de maîtrise de l'IA formulé à l'article 4 du Règlement européen sur l'Intelligence Artificielle qui impose aux responsables utilisant des SIA de former leur personnel sur le fonctionnement et l'utilisation des SIA dans le contexte dans lequel ils sont amenés à les utiliser. ¹⁰³

*B) L'incapacité de l'IA à interpréter les principes juridiques du
DIH*

1- Fiabilité

79. Définition de la fiabilité. Le recours à l'Intelligence Artificielle, fondée principalement sur des algorithmes et des traitements mathématiques, soulève des enjeux spécifiques en matière de fiabilité des systèmes qui l'intègrent. La norme ISO/IEC TS 5723 définit la fiabilité comme étant la « *capacité d'un élément à fonctionner comme requis, sans défaillance, pendant un intervalle de temps donné et dans des conditions données* ». ¹⁰⁴ Cette définition met l'accent sur la constance et la stabilité des performances attendues d'un système. Dans la même perspective,

¹⁰² Comité d'éthique de la défense, «Avis sur l'usage des technologies d'intelligence artificielle par les forces armées», 14 janvier 2025. Disponible sur [\[Le lien\]](#).

¹⁰³ 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).

¹⁰⁴ Technical specification, ISO/IEC TS 5723, "Trustworthiness – Vocabulary", First edition, 2022-07. Disponible en version PDF sur [\[Le lien\]](#).

Dominique Lambert, docteur en sciences et philosophie, précise que la fiabilité d'un système repose sur la prévisibilité de son comportement dans l'ensemble des environnements où il est destiné à opérer.¹⁰⁵

80. La fiabilité d'un SIA de ciblage humain. Si nous devons reprendre cette définition et l'appliquer à un SIA militaire de ciblage humain, cela reviendrait à le définir comme la capacité de ce système à cibler un ou plusieurs individus préalablement enregistrés par les commanditaires de l'action.

81. Cause de ce manque de fiabilité. Ce manque de fiabilité des algorithmes d'Intelligence Artificielle peut être le résultat de plusieurs causes. En effet, surtout dans un contexte opérationnel, les algorithmes doivent pouvoir résister aux perturbations climatiques, électriques, ennemies, etc. Ainsi des drones autonomes de ciblage doivent pouvoir fonctionner même en cas de brouillage (GPS ou autres) ou de capteurs dégradés. Pour les systèmes de ciblage humain intégrant de l'IA, cela signifie qu'ils ne doivent pas arrêter la liaison avec le commandant humain qui les dirige. Par ailleurs, le manque de fiabilité peut aussi venir de la mauvaise qualité des données qui ont permis l'entraînement du système d'IA. Ce problème peut être d'autant plus flagrant en contexte militaire où les données sont rares ou classifiées.

82. Manque de fiabilité appliqué aux SIA militaires. Une interrogation essentielle émerge dans les débats sur l'usage de

¹⁰⁵ Dominique Lambert, *La robotique et l'intelligence artificielle, Collection Que penser de...?* p. 55. Namur/Paris: Fidélité: Éditions Jésuites, DL 2019. - 1 vol. (138 p.).

l'intelligence artificielle en contexte militaire: «étant donné un algorithme général et une série de données qu'on lui propose, peut-on assurer que cet algorithme produira des réponses qui sont acceptables par rapport aux critères que se sont fixés les concepteurs?»,¹⁰⁶ souligne le Dr Lambert. Appliquée à l'IA dans un contexte militaire opérationnel, nous avons vu qu'il serait possible de programmer les SIA pour qu'ils soient conforme au DIH. Mais cette question remet cette thèse en cause. Cela reviendrait à poser une autre question, est-ce que les critères de proportionnalité, de précaution et de distinction, critères fondamentaux du DIH pourraient être respecté par la machine? En effet, ces principes impliquent non seulement des analyses contextuelles, mais également des jugements de valeur et des interprétations normatives, que les systèmes d'IA sont aujourd'hui incapables de réaliser de manière autonome. Fondés sur des algorithmes statistiques, ces systèmes n'ont ni compréhension du contexte éthique, ni accès aux nuances morales ou juridiques que ces principes exigent.

83. Le cas S. Petrov. Cette limite est parfaitement illustrée par le célèbre cas de Stanislav Petrov, officier soviétique qui, le 26 septembre 1983, a refusé de déclencher une riposte nucléaire malgré l'alerte d'un système «Oko» signalant une attaque américaine. Grâce à son jugement humain, et notamment permis par un examen approfondi d'environ 17 minutes,¹⁰⁷ Petrov a estimé que les signaux détectés étaient probablement erronés, ce qui s'est avéré exact. Un système d'IA, même techniquement fiable,

¹⁰⁶ *Ibid.*

¹⁰⁷ «Le Russe qui a sauvé le monde, article de site interne Watson», 14/04/2024. Disponible sur [[Le lien](#)].

aurait pu obéir à la logique programmée sans remettre en question la validité des données ni en évaluer la vraisemblance stratégique ou l'intention réelle de l'adversaire. Ainsi, Petrov en défiant le système montre que l'intervention humaine efficace ne consiste pas simplement à «rester dans la boucle» mais à être capable de remettre en cause les résultats de la machine, de les défier ou de désobéir au protocole.¹⁰⁸ Et c'est précisément cette défiance en la machine et ce jugement éthique humain qui a permis d'éviter une catastrophe mondiale par la machine. C'est d'ailleurs ce que rappelle le Groupe d'Experts Gouvernementaux sur les SALA qui explique que *«le jugement humain est essentiel pour garantir que les systèmes d'armes basés sur les technologies émergentes dans le domaine des systèmes d'armes létaux autonomes seront potentiellement utilisés dans le respect du droit international et en particulier du droit international humanitaire»*.¹⁰⁹

84. Ainsi, la fiabilité d'un système militaire ne peut se réduire à son bon fonctionnement algorithmique. Elle doit inclure la capacité d'interprétation juridique et morale dans des situations ambiguës et critiques, c'est-à-dire un contrôle humain puisque selon le CICR *«Ce sont des humains – et non pas des machines, des programmes informatiques ou des systèmes d'armes – qui appliquent le droit et sont tenus de le respecter»*.¹¹⁰

¹⁰⁸ Antón Barba-Kay, "There Is No Ethical Automation: Sanislav Petrov's Ordeal by Protocol", *Journal of military ethics*, 2024, Vol 23, n°3-4, pp. 277-288. Disponible sur [Le lien].

¹⁰⁹ CCW/GGE.1/2019/3. Disponible sur [Le lien]

¹¹⁰ CICR, «Un nouveau pas vers l'imposition de limites à l'autonomie des systèmes d'armes», 12/04/2018. Disponible sur [Le lien].

Ce cas montre ainsi que la fiabilité d'un système d'IA ne garantit en rien sa légitimité opérationnelle ou juridique, surtout lorsqu'il s'agit de respecter les contraintes du droit international humanitaire.

2- Robustesse

85. Définition. Une des limites fondamentales des algorithmes d'Intelligence Artificielle réside dans leur robustesse. La Commission Nationale de l'Informatique et des Liberté (CNIL) a défini cette notion comme étant *«la capacité du système à maintenir sa conformité à des exigences de performance et/ou de sécurité en présence de données d'entrée extérieures à son domaine d'emploi, par exemple en raison d'un défaut sur un capteur»*.¹¹¹

86. Exemple. Dans ce contexte, Dominique Lambert, nous éclaire par un exemple. Un algorithme de reconnaissance de formes, est considéré comme robuste s'il fournit des résultats identiques lorsqu'il est confronté à des données légèrement différentes mais indiscernables pour l'œil humain.¹¹² Cette notion est cruciale dans le cadre de l'utilisation des systèmes d'Intelligence Artificielle pour des applications de ciblage militaire, notamment les systèmes de ciblage automatique d'individus.

87. Application aux SIA de ciblage humain. Les SIA de ciblage humain sont généralement conçus pour identifier des individus spécifiques dans des environnements complexes, souvent à des fins militaires. Leur objectif est de distinguer les

¹¹¹ Définition de robustesse d'une intelligence artificielle par la CNIL.

¹¹² *Op. cit.* p.56

cibles légitimes, telles que les combattants ennemis, des civils. Cependant, ces systèmes peuvent être vulnérables aux attaques des adversaires, comme des leurres, où l'ennemi introduit délibérément des perturbations dans les données, par exemple, un soldat ennemi pourrait porter un uniforme non conventionnel, tel qu'un uniforme de couleur rose, afin de tromper le système de ciblage.¹¹³ Cette forme de manipulation pourrait entraîner des erreurs de classification, où un civil, confondu avec un combattant ennemi, ne serait pas reconnu comme une cible inoffensive. Une telle situation pourrait compromettre l'efficacité et la précision des missions, voire entraîner des conséquences tragiques, notamment des pertes civiles injustifiées.

88. À l'inverse, si un SIA de ciblage est réentraîné sur la base de cette anomalie, c'est-à-dire qu'il apprend que l'uniforme rose peut être associé à un combattant ennemi, pour reprendre notre exemple, le système pourrait alors classer un civil portant des vêtements de cette couleur comme une cible légitime, ce qui soulève de graves préoccupations éthiques et opérationnelles. Ce processus de réapprentissage des données, bien qu'il puisse améliorer la robustesse du système dans certains contextes, peut également aggraver la situation si l'algorithme apprend des associations erronées.

89. Scénario avec supervision humaine. Dans les scénarios où un SIA de ciblage humain est sous contrôle direct de l'opérateur, la supervision humaine constitue une forme de

¹¹³ Exemple tiré de l'ouvrage de Laure de Roucy-Rochegonde, «La guerre à l'ère de l'intelligence artificielle», p.132, reformulé et adapté à notre argumentation.

sauvegarde. En effet, l'opérateur humain reste responsable de la décision finale, en l'occurrence celle de tirer ou non sur la cible identifiée par le système. Cette capacité de supervision humaine permet de maintenir une certaine rigueur éthique et opérationnelle, garantissant que les erreurs éventuelles de classification ne mènent pas à des actions irréversibles.

90. Scénario sans supervision humaine. En revanche, lorsque l'on passe à des systèmes d'armement léthal autonome, la supervision humaine devient moins évidente, voire inexistante. Dans ce cas, un SIA entièrement autonome pourrait, par exemple, identifier un civil qu'il aurait de façon erronée classé comme un combattant, et prendre la décision d'ouvrir le feu sans intervention humaine. Cette absence de contrôle humain soulève des questions éthiques et juridiques majeures, en particulier en ce qui concerne le respect des principes du Droit International Humanitaire. En effet, l'un des principes majeurs du DIH est le principe de la distinction entre civils et combattants, et aussi la nécessité de limiter les effets des conflits armés à des objectifs militaires légitimes. Dans un tel scénario, l'utilisation de systèmes autonomes qui échappent au contrôle humain pourrait entrer en conflit direct avec ces principes fondamentaux.

91. La «non-robustesse» des algorithmes d'Intelligence Artificielle pose de sérieux défis lorsqu'elle est confrontée à des scénarios de guerre avec l'usage de l'IA, où la prise de décision rapide et précise est essentielle. L'absence de supervision humaine, notamment dans le cadre des SALA, pourrait mener à des violations graves des droits humains et des normes internationales,

mettant en péril la légitimité et la moralité de l'utilisation de ces technologies en contexte militaire.

§ 2. Les risques pratiques liés à l'utilisation opérationnelle de l'IA en contexte militaire

A) Biais des données et erreurs dans l'identification des cibles

92. Dans le domaine des SIA de ciblage humain, les biais dans les systèmes d'Intelligence Artificielle prennent une dimension particulièrement critique.

93. Définitions. D'un point de vue technique, selon Chira Lindloff and Ingo Siegert, le biais décrit une modulation d'entrée dans les réseaux neuronaux.¹¹⁴ Dans leur étude sur la définition des biais dans les systèmes d'IA, les auteurs proposent d'identifier deux catégories distinctes de biais. D'une part, le biais technique, qui est *«un concept neutre architectural, faisant référence à la modulation d'entrée dans les réseaux neuronaux»*¹¹⁵ et d'autre part, le biais social qui est une *«différenciation équitable qui reconnaît les différences pertinentes pour le résultat»*.¹¹⁶ Dans le cadre de notre étude, nous nous pencherons sur cette définition du biais où il est défini *«the systematic errors that occur in decision-making processes, leading to unfair outcomes»*.¹¹⁷ Les risques de biais sont dus à une multitude de facteur que nous allons détailler.

¹¹⁴ Chiara Lindloff and Ingo Siegert, "Defining bias in AI-systems: Biased models are fair models", 25 février 2025. Disponible sur [Le lien].

¹¹⁵ *Ibid.*

¹¹⁶ *Ibid.*

¹¹⁷ Emilio Ferrara, "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies", 26 décembre 2023, Article from

94. Échantillon d'apprentissage biaisé. Le premier des biais tient du fait que l'échantillon d'apprentissage est biaisé. C'est-à-dire que lors de l'entraînement de l'IA, le système s'entraîne sur un échantillon de données biaisé. Par exemple si nous faisons fonctionner une IA du style ChatGPT en lui donnant comme échantillon d'apprentissage la littérature mondiale et que nous lui demandions la différence homme/femme, il se pourrait que l'IA produise des réponses statistiquement misogynes. Finalement, un entraînement d'un SIA sur des échantillons biaisés ne permettrait pas d'arriver à une bonne solution, ou tout du moins, une solution satisfaisante.

95. Application à un SIA de ciblage humain. Un échantillon d'apprentissage biaisé pourrait entraîner une mauvaise identification des cibles. Par exemple, si un système d'arme autonome a été entraîné avec des données principalement issues de certains environnements culturels ou géographiques, il pourrait être moins précis, voire discriminatoire, envers des populations sous-représentées, augmentant ainsi le risque de tirs injustifiés.

96. Biais dès la conception. Ensuite il peut y avoir des biais qui peuvent être introduit dans la conception même de l'algorithme. Les biais de conception peuvent être dramatiques. Comme pour l'algorithme COMPAS,¹¹⁸ utilisé dans le domaine

MDPI. Traduction: une erreur systématique dans les processus de prise de décision, qui conduit à des résultats injustes. Disponible sur [[Le lien](#)].

¹¹⁸ L'algorithme COMPAS pour Correctional Offender Management Profiling for Alternative Sanctions, est un algorithme qui mobilise des méthodes

judiciaire, qui reproduisait des préjugés raciaux, un système d'armes autonome pourrait intégrer les biais inconscients de ses concepteurs, affectant la sélection des cibles et les décisions d'engagement, sans que cela ne soit intentionnel au départ.

97. Environnement biaisé. Ensuite il y a des IA qui sont exposées à un environnement biaisé. Ce dernier peut en effet fausser les décisions de l'IA embarquée: tout comme l'affaire TAY¹¹⁹ a montré comment une IA pouvait être rapidement corrompue par un environnement toxique,¹²⁰ un SALA placé dans un contexte chaotique ou conflictuel pourrait interpréter de manière erronée les comportements humains, menant à des frappes injustifiées.

98. Biais de fonctionnement. Enfin, il peut exister un biais qui est dû au fonctionnement de l'IA, à la structure de l'IA et notamment à sa nature statistique. Les biais de fonctionnement et la nature statistique des IA soulignent une limite fondamentale: une IA ne comprend pas, elle corrèle. Ainsi, tout comme une IA a confondu husky et loup¹²¹ à cause de la présence de neige, une IA

d'apprentissage automatique pour prédire la probabilité de récidive d'une personne. Une enquête menée par ProPublica en 2016 révèle que cet outil contient des biais raciaux.

¹¹⁹ TAY est un chatbot conversationnel créé par Microsoft.

¹²⁰ TUAL, Morgane, «A peine lancée, une intelligence artificielle de Microsoft Dérape sur Twitter» article publié dans *Le Monde*, 24 mars 2016. Disponible sur [[Le lien](#)].

¹²¹ Le modèle confondait les loups et les huskies car il associait la neige en arrière-plan à la présence d'un loup. Voir l'article de Sammy Soetaert «Une IA explicable est-elle la clé d'un développement ultérieur», publié le 11/03/2022. Disponible sur [[Le lien](#)].

militaire pourrait mal interpréter des signaux contextuels sur le terrain (gestes, habits, regroupements de personnes), et agir de manière incorrecte. Ces erreurs ne peuvent jamais être complètement éliminées, car elles sont inhérentes à la nature statistique de l'apprentissage machine.¹²²

B) *Vulnérabilités aux attaques et défis de cybersécurité*

99. Les systèmes d'Intelligence Artificielle, en particulier ceux utilisés dans des contextes militaires pour le ciblage humain, qu'ils soient entièrement autonomes ou sous supervision humaine, reposent sur des infrastructures complexes et interconnectées. Ces systèmes sont vulnérables à divers types d'attaques numériques, que l'on appelle des cyberattaques. Les vulnérabilités, peuvent être définies comme des failles techniques, logicielles ou structurelles qu'un adversaire peut exploiter¹²³ pour altérer le comportement, la performance ou les résultats d'un système autonome.

Comme le souligne Paul Scharre, *«une fois qu'un adversaire découvre une vulnérabilité dans un système autonome, il est libre de l'exploiter jusqu'à ce qu'un humain découvre la vulnérabilité et apporte un correctif au système ou modifie son utilisation»*.¹²⁴ Cela montre à quel point une vulnérabilité peut être dangereuse surtout en contexte opérationnel étant donné que la machine ne peut s'autoréparer ou s'auto-adapter. Cela impliquerait donc des conséquences terribles pour les systèmes de ciblage humain qui pourraient, par exemple, se retourner «contre leur camp» du fait d'une cyberattaque. D'autre

¹²² Cours de Mr Raphael Dechaux, M1 Droit de l'Intelligence artificielle.

¹²³ IBM, «Qu'est-ce que la gestion des vulnérabilités?». Disponible sur [Le lien]

¹²⁴ Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, New York, W. W. Norton and Company, 2018, p.178.

part, le fait que l'opérateur humain doive se rendre compte de la faille, de réparer l'algorithme, voire de le réentraîner, cette latence dans la réaction peut engendrer des conséquences désastreuses notamment à l'égard des civils qui pourraient être ciblés. De plus, une réévaluation de la licéité pourrait s'avérer nécessaire une fois les vulnérabilités corrigées. (Voir par. 51 à 53).

100. Multitude de cyberattaques. Les cyberattaques peuvent être très variées allant de la manipulation des données d'entrée, de l'empoisonnement des données d'entraînement ou du modèle, des attaques de confidentialité, à l'intrusion dans les réseaux de communications ou encore en passant par des attaques par déni de service.

101. Les implications des vulnérabilités pour les systèmes de ciblage humain autonomes ou supervisés. Les vulnérabilités techniques et les cyberattaques posent des enjeux majeurs dans le contexte de systèmes d'armement autonomes. Premièrement, elles compromettent la fiabilité de ces systèmes dans les opérations militaires. Une attaque réussie pourrait désactiver un système, manipuler ses décisions de ciblage, ou encore entraîner une escalade non désirée du conflit.

Deuxièmement, ces vulnérabilités fragilisent la responsabilité juridique. En effet, dans le cadre du droit international humanitaire, le respect des principes de distinction, de proportionnalité et de précaution est une exigence fondamentale. Or, une IA dont le comportement a été altéré par une cyberattaque pourrait violer ces principes sans que la chaîne de commandement en ait conscience ni contrôle effectif. Cette absence de traçabilité ou d'intelligibilité des erreurs soulève également des problèmes pour rendre

des comptes qui est un des 11 principes directeurs affirmés par le GGE sur les technologies émergentes dans le domaine des SALA.¹²⁵

En outre, le risque de prolifération technologique vers des acteurs non étatiques, notamment des groupes terroristes, aggrave la menace.¹²⁶ Un système piraté ou détourné pourrait être réutilisé contre ses créateurs initiaux ou contre des populations civiles, dans une transgression totale des normes du DIH.

Enfin, le risque de perte de contrôle opérationnel est accentué lorsque la supervision humaine repose sur des communications vulnérables. Une attaque perturbant la liaison entre l'IA et son opérateur peut rendre inopérante toute tentative humaine de reprendre le contrôle en cas de défaillance.¹²⁷

102. Propositions pour renforcer la cybersécurité dans les systèmes d'IA de ciblage humain. Face à ces menaces, plusieurs pistes doivent être envisagées pour assurer une cybersécurité de niveau militaire adaptée à la criticité des systèmes de ciblage humain. Tout d'abord, il est indispensable que les systèmes d'IA soient conçus dès l'origine avec une approche dite *secure by design*.¹²⁸ Cela implique l'intégration de mécanismes de défense en

¹²⁵ Assemblée nationale, Rapport d'information n°3248 en conclusion des travaux d'une mission d'information sur les systèmes d'armes létaux autonomes, 22 juillet 2020. Disponible sur [Le lien].

¹²⁶ *Ibid.*

¹²⁷ Mathias Comandré et Mattéo Mevellec, «Les défis éthiques et les risques de l'intelligence artificielle militaire», Institut National des Affaires Stratégiques et Politiques. Disponible sur [Le lien].

¹²⁸ ANSSI, «Recommandations de sécurité pour un système d'IA générative», ANSSI-PA-102, 29/04/2024. Disponible sur [Le lien]

profondeur, combinant pare-feux, détection d'anomalies, et vérifications cryptographiques.

Par ailleurs, une solution serait de doter ces systèmes eux-mêmes de mécanismes d'auto-défense basées sur l'IA.¹²⁹ Ces mécanismes permettraient de détecter et de contrer en temps réel les tentatives d'intrusion, de manipulation de données ou de falsification de capteurs.

Section 2: Les difficultés géopolitiques

103. Au-delà des enjeux techniques, l'intégration des systèmes d'Intelligence Artificielle militaires s'inscrit dans un contexte géopolitique tendu. D'un côté, une dynamique de course à l'armement alimente l'innovation accélérée, souvent au détriment d'un encadrement juridique clair et partagé (§1); de l'autre, cette rivalité stratégique entraîne des effets délétères sur la stabilité internationale, en augmentant les risques d'escalade incontrôlée, et de prolifération (§2).

§ 1: Confrontation entre course à l'armement et un nécessaire cadre normatif à l'IA

104. Les transformations technologiques dans le domaine militaire, telles que l'avènement des systèmes d'armement létaux autonomes, ne sont pas simplement une question de progrès technique, mais découlent aussi de transformations politiques.

¹²⁹ MdN Ruffo, «Avons-nous le choix d'utiliser l'IA en temps de guerre?», 47^e numéro des Cahiers de la sécurité et de la justice, 12 mars 2020. Disponible sur [\[Le lien\]](#)

Comme l'indique Clausewitz, «*les transformations de l'art de la guerre proviennent donc des transformations de la politiques. Loin de montrer qu'elles sont dissociables, elles démontrent au contraire leur intime union*». ¹³⁰ Dans ce contexte, la course à l'armement, en particulier dans le domaine de l'IA, doit être mise en balance avec la nécessité de créer un cadre normatif rigoureux. Il apparaît ainsi indispensable que les SALA soient programmés conformément aux principes du DIH, non seulement pour répondre à des considérations stratégiques, mais aussi pour garantir qu'elles servent des objectifs politiques légitimes, respectant l'éthique de la guerre et les protections humaines fondamentales. Cela souligne l'intime union entre la politique et la technologie, où les choix techniques doivent être alignés sur des objectifs politiques plus larges.

105. Ainsi, l'Intelligence Artificielle, qu'elle soit utilisée seule ou en combinaison avec d'autres technologies émergentes ou de rupture, est en mesure de transformer profondément le caractère, voire l'essence même de la guerre. ¹³¹ Son intégration dans les systèmes et processus militaires représente un levier de supériorité stratégique et politique, susceptible de bouleverser en profondeur l'art de la guerre. Comme le soulignent certains experts, *“the integration of AI into military systems has the potential to revolutionise*

¹³⁰ Carl Von Clausewitz, *De la Guerre*, 1832, édition abrégée et présentée par Gérard Chaliand, collection tempus, éditions Perrin, p. 402.

¹³¹ Payne Kenneth, “Artificial Intelligence and the Nature of War”, in: Sweijts Tim et Michaels Jeffrey H. (dir.), *Beyond Ukraine: Debating the Future of War*, Oxford University Press, 2024, p. 223-240. Cité par la *Revue Nationale de Défense*, n°879, avril 2025.

warfare,¹³² mettant en évidence l'ampleur des transformations à venir. Il est donc clair que les pays se bousculent pour participer à cette course à l'IA militaire pour justement prendre le pas de ce «nouveau style de guerre».

106. Lors du sommet européen sur l'Intelligence Artificielle qui s'est tenu à Paris les 10 et 11 février dernier, la «*Déclaration de Paris sur le Maintien du Contrôle Humain dans les Systèmes d'Armes Dotés d'Intelligence Artificielle*»¹³³ a été signée par pas moins de 27 pays dont la France mais sont absents les Etats-Unis, la Chine et la Russie. Cette déclaration montre à quel point la question des systèmes d'armes incluant de l'Intelligence Artificielle est majeur dans le domaine de la défense.

107. Par ailleurs, l'actualité récente montre aussi que la course à l'IA dans le domaine militaire est en route. Très récemment, Singapour et l'entreprise française Thales ont créé un laboratoire commun dédié à l'Intelligence Artificielle et dont l'objectif serait de renforcer les systèmes de défense des Forces armées de Singapour (SAF), en particulier face à la menace croissante des drones.¹³⁴

¹³² Clement Sven, «NATO and AI: Navigating the Challenges and Opportunities», *Special Report for Science and Technology Committee of NATO Parliamentary Assembly*, 24 novembre 2024, disponible sur [Le lien]. Traduction: «L'intégration de l'intelligence artificielle dans les systèmes militaires a le potentiel de révolutionner l'art de la guerre».

¹³³ Paris Declaration on Maintaining Human Control in AI enabled Weapon Systems. Publié le 11 février 2025. Disponible sur [Le lien].

¹³⁴ Thales Group, «La DSTA et Thales annoncent la création d'un laboratoire commun axé sur l'Intelligence Artificielle pour renforcer les systèmes de défense de Singapour», publié le 04 avril 2025. Disponible sur [Le lien].

108. Le 25 mars 2025, l'Agence OTAN d'information et de communication (NCIA) et l'entreprise *Palantir* ont conclu l'acquisition du système *Maven Smart System NATO*, une plateforme d'IA permettant l'aide à l'analyse de données opérationnelles visant à «accélérer la prise de décision sur le champ de bataille». ¹³⁵ Ce système, conçu pour être déployé au sein du Commandement allié des opérations, représente une étape déterminante dans la modernisation des capacités militaires de l'Alliance et qui permettra de lui apporter «des capacités d'intelligence artificielle de pointe, sur mesure, et dote nos forces des outils nécessaires pour opérer efficacement et avec détermination sur le champ de bataille moderne». ¹³⁶ L'acquisition très récente de ce système par l'OTAN démontre à quel point, les organisations militaires, les pays, se saisissent de la question pour se doter de l'IA dans leur armée.

109. Outre le fait que cette course soit indéniable, cette dernière se joue sur quatre champs de bataille selon Paul Scharre: «des données, la puissance de calcul, les talents et les institutions». ¹³⁷ La compétition mondiale en matière d'armement fondé sur l'Intelligence Artificielle mobilise plusieurs puissances majeures, chacune investissant de manière substantielle dans la recherche et le développement de cette technologie afin d'acquérir un avantage stratégique sur le plan militaire. Les États-Unis et la Chine

¹³⁵ Sauveton, Pierre, «IA militaire : l'OTAN choisit Palantir, la France répond avec Artemis, GenIAI... et Comand AI». Disponible sur [[Le lien](#)].

¹³⁶ NATO Communications and Information Agency, 14 avril 2025. Disponible sur : [[Le lien](#)].

¹³⁷ Scharre Paul, *Four Battlegrounds—Power in the Age of Artificial Intelligence*, W.W. Norton & Company, 2023, 496 pages. Cité par Laurent Saunois dans la *Revue Nationale de Défense*, n°879, avril 2025.

occupent une position dominante dans cette dynamique, tandis que d'autres États poursuivent également des avancées notables dans ce domaine.

A) La course à l'IA dans le domaine militaire selon les pays

1- Les Etats-Unis

110. Les États-Unis sont considérés comme l'un des pays les plus avancés sur le sujet de l'Intelligence Artificielle dans le secteur de la défense, notamment en termes d'investissement injecté en la matière qui se chiffre en milliards de dollars. Au cœur de cette démarche se trouve la DARPA (Defense Advanced Research Projects Agency), un programme spécialisé dans l'innovation des nouvelles technologies et le Joint Artificial Intelligence Center (JAIC), désigné comme le «centre d'excellence en Intelligence Artificielle du ministère de la Défense (DoW)». ¹³⁸ Le JAIC a pour mission de piloter l'intégration stratégique de l'intelligence artificielle au sein du Département de la Défense américain. Il vise à exploiter pleinement le potentiel de l'IA pour renforcer l'efficacité opérationnelle, tout en garantissant une utilisation éthique, légale et sécurisée de ces technologies. Pour ce faire, il soutient la recherche, développe des outils communs, forme les effectifs et renforce les partenariats pour moderniser les capacités militaires tout en protégeant les infrastructures critiques. ¹³⁹ Le JAIC travaille activement à l'intégration de l'IA pour concevoir des outils comme une «machine de guerre» combinant plusieurs

¹³⁸ Actu IA, «Etats-Unis : les dernières nouvelles autour de l'IA du Pentagone et du Joint Artificial Intelligence Center». Disponible sur: [**Le lien**].

¹³⁹ STIP COMPASS, Joint artificial intelligence center, last up dated: June 3, 2024. Disponible sur [**Le lien**].

services, «capable de combiner l'analyse rapide de données avec l'action militaire en temps réels»¹⁴⁰ pour améliorer la prise de décision sur le champ de bataille. Cette structure dédiée et ces initiatives ambitieuses soulignent l'investissement considérable et l'engagement des États-Unis dans la course à l'IA de défense.

111. Par ailleurs, les États-Unis disposent d'une avance significative en matière de capacités de calcul, grâce à un avantage technologique marqué et à une maîtrise approfondie du cycle de production des semi-conducteurs. Aussi, leur capacité à attirer les talents, y compris parmi les chercheurs étrangers, renforce leur position dominante dans le domaine de l'intelligence artificielle.¹⁴¹

2- La Chine

112. La Chine, à travers son armée populaire de libération (APL), joue un rôle stratégique central dans l'intégration de l'intelligence artificielle au sein de ses stratégies de défense et de militarisation, participant activement à la dynamique mondiale de la course à l'armement technologique dans le domaine militaire. Depuis la déclaration du lieutenant-général Liu Guozhi en 2017, qui prédisait que l'IA provoquerait une révolution majeure dans la guerre moderne,¹⁴² la Chine a intensifié ses efforts pour se positionner comme un leader mondial dans ce domaine. Cette

¹⁴⁰ *Ibid.*

¹⁴¹ Revue de Défense Nationale, «Enjeux et Defis de l'Intelligence artificielle (IA) pour l'OTAN», n°879, Avril 2025.

¹⁴² [人大代表刘国治：人工智能将加速军事变革进程], cité par Dr Michel Valantin, «Militarisation de l'Intelligence artificielle en Chine», The Red Team Analysis Society, 23 avril 2018.

ambition se manifeste à travers des investissements considérables, notamment pour développer des armes autonomes,¹⁴³ ainsi que la création d'instituts de recherche spécialisés qui favorisent le développement de technologies d'armement basées sur l'IA.

113. Dès 2011, la Chine avait déjà inclus dans son dictionnaire officiel la définition des armes intégrant de l'IA comme étant “*a weapon that utilizes AI to pursue, distinguish, and destroy enemy targets automatically ; often composed of information collection and management systems, knowledge base systems, decision assistance systems, mission implementation systems, etc.*”¹⁴⁴ Cela montre à quel point la Chine se positionnait sur ces questions.

114. Par ailleurs, le gouvernement chinois a mis en place une stratégie de “fusion civil-militaire,” visant à maximiser les synergies entre la recherche civile et les besoins militaires.¹⁴⁵ Ce modèle permet une intégration fluide des découvertes technologiques

¹⁴³ L'IA, le nouvel outil économique et géopolitique de la Chine, Article publié le 06/02/2025 dans le journal numérique rfi. Disponible sur [\[Le lien\]](#).

¹⁴⁴ 全军军事术语管理委员会 [All-Military Military Terminology Management Committee], 中国人民解放军军语 [People's Liberation Army Military Terminology], (Beijing: 军事科学出版社 [Military Science Press], 2011). Cité par Elsa B. Kania, ““AI weapons” in China's Military Innovation”, Avril 2020. Traduction française: «Une arme qui utilise l'intelligence artificielle pour poursuivre, identifier et détruire automatiquement des cibles ennemies; elle est souvent composée de systèmes de collecte et de gestion de l'information, de bases de connaissances, de systèmes d'aide à la décision, de systèmes d'exécution de mission, etc.»

¹⁴⁵ Rapport de la Task Force, «L'intelligence artificielle au service de la défense», 2019.

civiles dans des applications militaires, accélérant ainsi la militarisation des technologies émergentes telles que les véhicules aériens sans pilote (UAV), les missiles de croisière autonomes dotés de capacités de cognition avancée, ainsi que des sous-marins autonomes.¹⁴⁶ Ces systèmes, qui reposent sur des algorithmes d'apprentissage automatique et des processus décisionnels en temps réel, illustrent la convergence des efforts civils et militaires sous l'égide d'un contrôle étatique centralisé, en particulier sous la direction de Xi Jinping.

115. Le concept d'"intelligentisation", qui repose sur l'automatisation des opérations militaires par des systèmes autonomes, constitue un élément clé dans la modernisation de l'APL. L'objectif à long terme est de transcender les concepts de guerre "informatisée" pour adopter une guerre "intelligentiée",¹⁴⁷ dans laquelle les décisions stratégiques et tactiques sont en grande partie prises par des systèmes autonomes capables de s'adapter en temps réel aux évolutions du champ de bataille.¹⁴⁸ Ce changement vise non seulement à augmenter l'efficacité opérationnelle, mais aussi à renforcer la capacité de projection de puissance de la Chine, en vue de soutenir ses ambitions géopolitiques.

116. Dans cette dynamique, la Chine ne se contente pas de développer des capacités d'IA pour des fins purement militaires.

¹⁴⁶ Dr Michel Valantin, «Militarisation de l'Intelligence artificielle en Chine», The Red Team Analysis Society, 23 avril 2018. Disponible sur [[Le lien](#)].

¹⁴⁷ Elsa B. Kania "The PLA's trajectory from informatized to "intelligitized" warfare", 8 juin 2017. Disponible sur [[Le lien](#)].

¹⁴⁸ Dr Michel Valantin, «Militarisation de l'Intelligence artificielle en Chine», The Red Team Analysis Society, 23 avril 2018. Disponible sur [[Le lien](#)].

En cherchant à instaurer une domination dans la recherche et l'innovation en IA, elle entend également utiliser cette technologie comme un outil stratégique de coercition militaire, amplifiant ainsi sa position sur la scène mondiale. Par le biais de cette course à l'armement technologique, la Chine s'efforce de se doter d'une "puissance IA", afin de devenir une superpuissance technologique et militaire et ambitionne de dépasser les Etats-Unis en la matière.¹⁴⁹

117. Les mécanismes de coopération entre le secteur privé et l'État chinois permettent une centralisation efficace des efforts de recherche, en particulier dans le domaine militaire. En ce sens, le principe de "fusion civil-militaire" est devenu un pilier stratégique de la politique chinoise, favorisant l'intégration des industries privées dans des projets de défense de plus en plus sophistiqués. Cette approche reflète une vision résolument orientée vers l'utilisation des technologies de l'IA à des fins sécuritaires et militaires, avec une implication directe du gouvernement dans la gestion et la régulation des acteurs privés.¹⁵⁰

3- L'Europe/ France

118. L'Europe: une puissance intermédiaire. A côté des deux «superpuissances» dans cette course à l'IA que sont les Etats-Unis et la Chine, l'Europe est considérée comme une puissance intermédiaire.¹⁵¹

¹⁴⁹ Elsa B. Kania, *op.cit.*

¹⁵⁰ Rapport de la Task Force, 2019, *op.cit.*

¹⁵¹ *Ibid.*

Contrairement aux deux superpuissances dont le modèle économique repose sur le marché de la donnée, disposant d'un écosystème fondé autour des géants du numérique, les GAFAM et les BATX, l'Europe se veut plus timide. Toutefois, sa force, mais aussi sa faiblesse, peut résider dans sa position stricte sur l'éthique et le respect du droit¹⁵² et notamment sa ferme volonté à respecter le droit international humanitaire.

119. Par ailleurs, les investissements dans l'IA varient nettement entre les États européens pris individuellement et les États-Unis ou la Chine, ce qui pourrait constituer un facteur déterminant dans cette course à l'armement militaire. Pour pouvoir rivaliser face à ces investissements colossaux, l'Europe doit se montrer unie. Pour les rapporteurs de la mission d'information sur les systèmes d'armes létaux autonomes, *«il ne fait guère de doute que c'est à l'échelle européenne que se joue l'autonomie stratégique nationale et européenne dans le domaine du numérique et, par conséquent, dans celui de l'intelligence artificielle de défense. Les enjeux dépassent en effet largement la sphère nationale»*.¹⁵³

120. L'enjeu de cette course à l'IA militaire se joue aussi sur le terrain des entreprises technologiques et militaires. Il est essentiel que l'Europe redevienne une terre d'accueil pour les grandes entreprises technologiques, notamment dans le domaine de la défense, afin de pouvoir rivaliser efficacement avec des puissances telles que les États-Unis et la Chine. À cet égard, plusieurs projets conjoints en matière d'intelligence artificielle de défense, tels que

¹⁵² *Ibid.*

¹⁵³ Assemblée nationale, Rapport d'information n°3248. *Op.cit.* Disponible sur [Le lien].

le MGCS (Main Ground Combat System), le SLAM-F (Système de Lutte Anti-Mines du Futur), le SCAF (Système de Combat Aérien du Futur) ou encore le programme Space Situational Awareness, constituent autant d'opportunités de renforcer une vision stratégique européenne commune sur les enjeux liés à l'IA.¹⁵⁴

121. Position des pays européens sur les SALA. Par ailleurs, sur la question des SALA, le rapport de l'Assemblée nationale sur les systèmes d'armes létaux autonomes permet de voir certaines des positions communes des pays membres de l'UE.¹⁵⁵ L'Autriche milite en faveur d'un instrument juridiquement contraignant imposant un contrôle humain substantiel sur les fonctions critiques des SALA. La Finlande pour sa part soutient une interdiction du développement et de la production de systèmes d'armes basés sur l'IA, tout en reconnaissant l'évolution inévitable de l'autonomie dans les systèmes militaires. L'Italie défend l'importance du contrôle humain comme garantie du respect du DIH et considère la CCAC comme le forum légitime pour faire progresser la régulation des SALA.

De son côté l'Espagne rejette l'utilisation de systèmes où la décision de tirer ne serait pas imputable à un humain et insiste sur un usage strictement défensif de ces technologies. Les États baltes tels que l'Estonie, la Lettonie et la Lituanie privilégient la CCAC comme cadre légitime pour les discussions sur les SALA, en lien avec leurs intérêts industriels dans la robotique.

Concernant le Parlement européen, ce dernier appelle à un instrument juridiquement contraignant interdisant les SALA

¹⁵⁴ *Ibid.*

¹⁵⁵ *Ibid.*

dénués de contrôle humain, et à une position commune affirmée au sein des enceintes multilatérales.

122. Enfin, de son côté, la France reste consciente de l'apport de l'intelligence artificielle dans le domaine de la défense. C'est pourquoi plusieurs programmes utilisant de l'IA sont en cours de création. C'est le cas par exemple des programmes cités susmentionnés, à savoir SCORPION, SLAM-F ET SCAF utilisant largement l'IA pour améliorer des aspects tels que la connaissance de la situation tactique, la protection, la fusion de données, le guidage, l'évitement d'obstacles, la gestion de capteurs/effecteurs, la connectivité et la coopération entre aéronefs habités et drones.¹⁵⁶

B) *Les positions des pays sur l'IA militaire et leur respect du DIH*

1- USA

123. Les États-Unis ont déclaré que les systèmes d'armes létaux autonomes soulèvent d'*«importantes questions juridiques, politiques et éthiques»*, et ont recommandé de poursuivre les discussions sur ce sujet dans un forum de droit international humanitaire.¹⁵⁷ Bien qu'une directive politique du ministère de la Défense autorise le développement de tels systèmes, les États-Unis insistent sur le fait qu'elle *«n'encourage ni n'interdit le développement de tels systèmes à l'avenir»*.¹⁵⁸ Concernant le respect du Droit international humanitaire,

¹⁵⁶ *Ibid.*

¹⁵⁷ Campaign to stop killer robot, "report on outreach on the UN report on "lethal autonomous robotics"", 31 juillet 2013. Disponible en version PDF sur [Le lien].

¹⁵⁸ Human Rights Watch, «Stopper les robots tueurs, positions des pays sur l'interdiction des armes entièrement autonomes et sur le maintien du contrôle humain». Disponible sur [Le lien].

les États-Unis estiment que le DIH fournit un système de réglementation robuste et cohérent pour l'utilisation des armes.¹⁵⁹ Ils considèrent que les propositions visant à négocier un nouveau traité international sur de tels systèmes d'armes sont «prématurées».

124. Par ailleurs, l'US Department of State a publié une déclaration politique sur l'utilisation responsable de l'IA et de l'autonomie artificielle, où il est souligné que l'utilisation militaire de l'IA, y compris pour des fonctions et systèmes autonomes, doit respecter le droit international en vigueur et, en situation de conflit armé, être conforme aux obligations des États au titre du droit international humanitaire, y compris ses principes fondamentaux. Les États devraient donc prendre des mesures pour s'assurer que leurs capacités d'IA militaires sont utilisées conformément au droit international, en particulier le DIH, et même chercher à les utiliser pour améliorer la mise en pratique du DIH et la protection des civils. L'utilisation militaire de l'IA doit également s'effectuer sous une chaîne de commandement et un contrôle humain responsable.¹⁶⁰ Cette déclaration a notamment été signée par la France.

¹⁵⁹ Gouvernement des États-Unis, Déclaration au Groupe d'experts gouvernementaux sur les systèmes d'armes létaux autonomes, "Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the CCW", Genève, 10 avril 2018. Disponible en version PDF sur [\[Le lien\]](#)

¹⁶⁰ U.S. Department of State, "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy", 9 novembre 2023. Disponible sur [\[Le lien\]](#).

125. Toutefois, il est à noter que les Etats-Unis n'ont pas signé la «*Declaration on Maintaining Human Control in AI enabled Weapon Systems*» lors du Sommet sur l'IA qui s'est tenu à Paris en 2025. Cette déclaration stipule notamment que les Etats signataires s'engagent à utiliser les «*capacités de l'IA dans le domaine militaire de manière responsable, conformément au droit international, en particulier au droit international humanitaire*». ¹⁶¹

2- Chine

126. La République populaire de Chine adopte une approche à la fois prudente mais aussi stratégique à l'égard des applications militaires de l'intelligence artificielle, notamment des SALA, en particulier dans le contexte du droit international humanitaire. Dès mai 2013, lors du Conseil des droits de l'homme des Nations Unies, la Chine a exprimé son soutien à l'ouverture de négociations multilatérales sur les SALA qu'elle a qualifiés d'«*excessivement complexes*» dans leur conception et leur encadrement juridique. ¹⁶² Elle a alors souligné les risques que ces technologies représentent pour la stabilité stratégique mondiale ainsi que pour les régimes de maîtrise des armements.

127. En décembre 2016, à l'occasion des discussions tenues dans le cadre de la CCAC, la Chine a renforcé ses réserves en déclarant que les SALA «*présentent des incertitudes considérables*» quant à leur compatibilité avec le DIH, et a plaidé pour l'adoption de mesures de précaution inspirées du précédent établi par

¹⁶¹ Paris Declaration on Maintaining Human Control in AI enabled Weapon Systems. *Op. cit.*

¹⁶² Human Rights Watch, *Op. cit.*

l'interdiction des armes à laser aveuglantes.¹⁶³ Ces déclarations témoignent d'une volonté de prévenir, plutôt que de réagir, face aux dangers potentiels posés par l'autonomie complète des systèmes d'armes, notamment en ce qui concerne le respect des principes de distinction, de proportionnalité et de nécessité.

128. En avril 2018, la Chine a franchi une étape supplémentaire en appelant à l'interdiction de l'utilisation des armes entièrement autonomes. Toutefois, elle a précisé que cette interdiction ne devait pas s'étendre au développement ou à la production de telles armes,¹⁶⁴ ce qui laisse transparaître une position ambivalente, à la fois soucieuse de prévenir leur emploi immédiat sur le champ de bataille, mais déterminée à ne pas entraver ses propres capacités technologiques. Depuis cette déclaration, la Chine s'est prononcée en faveur de la négociation d'un instrument juridique contraignant sur les SALA mais uniquement au sein de la CCAC,¹⁶⁵ ce qui témoigne d'un engagement soutenu dans les processus de discussion multilatérale.

129. Enfin, le 14 décembre 2021, la Chine a publié un document officiel exposant sa position sur la régulation des applications militaires de l'intelligence artificielle.¹⁶⁶ Ce texte énonce que les

¹⁶³ *Ibid.*

¹⁶⁴ *Ibid.*

¹⁶⁵ Nations Unies, communiqué de presse, «Première Commission: les pays soulignent l'importance d'une prise en compte des dimensions éthiques et humanitaires de certaines armes classiques». AG/DSI/3748, 24 octobre 2024. Disponible sur [[Le lien](#)].

¹⁶⁶ Ministère des Affaires étrangères de la République Populaire de Chine, «Position Paper of the People's Republic of China on Regulating Military

États ont la responsabilité de promouvoir les valeurs universelles de l'humanité et de veiller à ce que les systèmes d'armement basés sur l'IA soient conçus et utilisés conformément aux exigences du DIH. Il insiste sur la nécessité de respecter les normes éthiques dans les phases de développement, de déploiement et d'utilisation des technologies militaires concernées. La Chine appelle ainsi à garantir la conformité des armes émergentes aux principes humanitaires fondamentaux, à minimiser les pertes collatérales, et à prévenir les usages malveillants ou incontrôlés de ces systèmes.¹⁶⁷ Ce positionnement reflète une volonté de légitimation éthique sur la scène internationale, tout en préservant sa marge de manœuvre stratégique dans le domaine de l'armement intelligent.

3- France

130. La France a pris l'initiative, dès 2013, d'introduire des discussions sur les SALA aux Nations unies, dans le cadre de la Convention sur certaines armes classiques, où elle a joué un rôle actif. Elle se mobilise avec ses partenaires pour bâtir un consensus international et promeut «*l'adoption d'un ensemble de principes visant à régir le développement et l'usage des futurs systèmes d'armes autonomes*»¹⁶⁸ notamment dans le respect du DIH. C'est d'ailleurs dans ce contexte-là de coopération que la France et l'Allemagne ont obtenu l'approbation en novembre 2019 de 11 principes

Applications of Artificial Intelligence (AI)», 14 décembre 2021. Disponible sur: **[Le lien]**.

¹⁶⁷ *Ibid.*

¹⁶⁸ Ministère de l'Europe et des Affaires étrangères, «11 principes sur les systèmes d'armes létaux autonomes». Disponible sur **[Le lien]**.

directeurs¹⁶⁹ par les États parties à la CCAC, qui affirment notamment que le DIH «s'applique à tous les systèmes d'armes, y compris ceux dotés d'autonomie, et que l'emploi de ces systèmes doit toujours dépendre d'une responsabilité humaine».¹⁷⁰ Ces principes évoquent également la nécessité d'une interaction entre l'humain et la machine pour assurer le respect du DIH. La France continue de soutenir activement les discussions au sein de la CCAC pour enrichir ces principes afin de les intégrer dans un «cadre normatif et opérationnel» applicable aux SALA.¹⁷¹

131. La position de la France sur les SALA est clairement définie et s'appuie sur plusieurs principes fondamentaux liés au respect du droit international, en particulier le droit international humanitaire, et au maintien d'un contrôle humain suffisant.¹⁷² La France considère les SALA comme des systèmes d'armes «capables de recourir à la force létale sans aucune intervention ou supervision humaine»,¹⁷³ et susceptibles de réorienter ou de s'attribuer de nouvelles missions.¹⁷⁴ Consciente des enjeux éthiques et juridiques importants soulevés par le développement potentiel de telles armes, la France a «clairement affirmé qu'elle ne développera pas de systèmes d'armes capables de recourir à la force de façon totalement autonome, considérant qu'ils seraient fondamen-

¹⁶⁹ Voir la liste des 11 principes directeurs: ASSEMBLÉE NATIONALE, Rapport d'information n°3248 en conclusion des travaux d'une mission d'information sur les systèmes d'armes létaux autonomes, 22 juillet 2020, p.53. Disponible sur [[Le lien](#)].

¹⁷⁰ *Ibid.*

¹⁷¹ *Ibid.*

¹⁷² Ministère de l'Europe et des Affaires étrangères, «Systèmes d'armes létales autonomes, quelle est la position de la France». Disponible sur [[Le lien](#)].

¹⁷³ *Ibid.*

¹⁷⁴ *Ibid.*

talement contraires à ses principes». ¹⁷⁵ Elle ne jugerait pas acceptable que la décision de vie ou de mort soit intégralement confiée à des machines échappant à toute forme de supervision humaine. C'est ce qu'avait rappelé dans son discours du 5 avril 2019, Madame Florence Parly, alors Ministre de la Défense, que «*la position française est sans ambiguïté, le président de la République a été parfaitement clair: la France refuse de confier la décision de vie ou de mort à une machine qui agirait de façon pleinement autonome et échapperait à tout contrôle humain*». ¹⁷⁶

132. D'un point de vue juridique, la France souligne que pour que le droit international humanitaire s'applique, l'usage d'une arme doit pouvoir être rattaché à une responsabilité humaine, ¹⁷⁷ et l'homme doit conserver la capacité de prendre les décisions ultimes s'agissant du recours à la force létale. Sa stratégie de développement de l'intelligence artificielle de défense repose sur trois grands principes. Le respect du droit international, la responsabilité du commandement dans l'emploi des armes, et le maintien d'un contrôle humain suffisant, avec une interaction homme-machine permettant de superviser et de contrôler tout système d'armes. ¹⁷⁸ La France n'autorisera pas une machine à s'assigner ou modifier une mission sans validation humaine, et le commandement humain reste le seul responsable légitime pour définir les règles d'emploi et d'engagement des systèmes d'armes.

¹⁷⁵ «Systèmes d'armes létales autonomes, quelle est l'action de la France?», *op.cit.*

¹⁷⁶ *Revue Défense Nationale*, 2019/5 n°820, Avant-propos – Intelligence artificielle et défense, par Florence Parly, pp. 9 à 17. Disponible sur [\[Le lien\]](#).

¹⁷⁷ «Systèmes d'armes létales autonomes, quelle est l'action de la France?», *op.cit.*

¹⁷⁸ Assemblée nationale, Rapport d'information n°3248 en conclusion des travaux d'une mission d'information sur les systèmes d'armes létaux autonomes, 22 juillet 2020, p. 87. Disponible sur [\[Le lien\]](#).

§ 2. Les effets de cette course à l'armement

133. La course à l'IA militaire et notamment aux systèmes autonomes, létaux ou non, peuvent avoir de sérieuses implications pour la sécurité mondiale avec des risques d'escalade incontrôlée et de prolifération des systèmes d'armes létaux autonomes (A) mais aussi des risques éthiques et juridiques (B).

A) Implication pour la sécurité mondiale

134. Concernant les risques sécuritaires, la course à l'IA militaire entraîne vraisemblablement des risques de conflit et d'instabilité stratégique. L'ancien secrétaire général de l'OTAN, Anders Fogh Rasmussen, disait que *«le développement et le déploiement rapides de l'IA dans les systèmes militaires pourraient fondamentalement modifier la nature de la guerre et la stabilité stratégiques»*.¹⁷⁹

135. **Les deux principaux risques.** En effet, dans un climat géopolitique instable, cette course peut amener certains États avec des positions strictes sur les armes autonomes, à revoir leur position pour s'adapter aux stratégies adverses qui auraient des positions plus souples. Dès lors, nous pourrions identifier deux risques majeurs: les escalades incontrôlées et le risque de prolifération de ces armes.

1- L'escalade incontrôlée

136. L'un des enjeux les plus préoccupants soulevés par le déploiement des systèmes d'armes létaux autonomes réside dans

¹⁷⁹ «L'avenir de l'IA dans la guerre: concepts clés et implications». Disponible sur [[Le lien](#)].

leur potentiel à engendrer une escalade incontrôlée des conflits armés, à la fois sur le plan opérationnel, stratégique et géopolitique. Ces technologies, en modifiant en profondeur la manière dont la force est projetée, introduisent de nouveaux déséquilibres dans les relations internationales.

137. En premier lieu, les SALA *«augmentent le risque d'escalade des conflits étant donné que l'usage de la force serait facilité»*,¹⁸⁰ notamment en rendant les actions militaires plus rapides, moins coûteuses en vies humaines pour l'agresseur, et donc politiquement plus acceptables. Cette dynamique pourrait abaisser le seuil de déclenchement des hostilités,¹⁸¹ en particulier dans des périodes de tensions internationales accrues, où la prise de décision peut être précipitée par la rapidité des systèmes autonomes, sans intervention humaine suffisante. En l'absence de jugement contextuel, un mauvais ciblage ou une erreur d'interprétation de la part de l'algorithme peut conduire à une escalade involontaire, voire au déclenchement d'un conflit non souhaité.

138. Par ailleurs, le risque d'escalade dû à cette course aux armements technologiques, alimenterait la peur de subir une «surprise technologique», et inciterait donc certains États à adopter

¹⁸⁰ Frédéric Casier, «L'appel d'urgence à de nouvelles règles internationales sur les systèmes d'armes autonomes face aux risques humanitaires croissants», 11/01/2024.

¹⁸¹ Assemblée Générale des Nations Unies, Soixante-dix-neuvième session Première Commission Point 98 ss) de l'ordre du jour Désarmement général et complet: systèmes d'armes létaux autonomes, A/C.1/79/L.77, 18 octobre 2024. Disponible sur [[Le lien](#)].

des postures préventives ou préemptives.¹⁸² Cette logique peut favoriser des prises de décision hâtives, fondées sur des systèmes non encore pleinement maîtrisés, dans un contexte où la diplomatie et la négociation n'ont pas encore eu le temps de s'exercer.¹⁸³ Le rythme rapide du développement de l'intelligence artificielle militaire dépasse parfois les mécanismes politiques et juridiques traditionnels de construction de la confiance et de contrôle des armements qui passe notamment par l'examen de licéité de l'arme.

139. En outre, les SALA exposent les personnes affectées par les conflits, civils comme combattants hors de combat, à des risques accrus d'usage disproportionné de la force. L'automatisation du ciblage pourrait, dans certaines situations, conduire à la sélection de cibles inadéquates, faute d'éléments contextuels que seule une intelligence humaine peut apprécier. De telles erreurs pourraient provoquer des «représailles, des actes de rétorsion»¹⁸⁴ ou des campagnes de terreur, surtout si des civils sont perçus comme les seules cibles disponibles.

140. Ces dynamiques d'escalade ne sont pas une fatalité. Certains États, à l'image de la France, affirment dans leurs doctrines nationales la nécessité de privilégier la désescalade, en

¹⁸² Rapport de la Task Force, 2019, p. 7 *op.cit.*

¹⁸³ Institut National des Affaires Stratégiques et Politiques, «Implications géopolitiques et compétition internationale: la nouvelle course à l'IA militaire», disponible sur [\[Le lien\]](#).

¹⁸⁴ Assemblée Générale des Nations Unies, Rapport du Rapporteur spécial sur les exécutions extrajudiciaires, sommaires ou arbitraires, Christof Heyns, A/HRC/23/47, 9 avril 2013, p. 18.

promouvant des mécanismes de coopération internationale, des réponses proportionnées et un encadrement rigoureux de l'usage de la force.¹⁸⁵ Par cette position, la France rappelle que la stabilité stratégique passe par une approche éthique et responsable du recours aux nouvelles technologies militaires, en cohérence avec le droit international humanitaire.

141. Dans la même veine, certains auteurs pensent que les armes létales autonomes, entendu comme des robots, pourraient amener au contraire un processus de désescalade.¹⁸⁶ En ce sens, il faudrait que les États créent un climat de confiance autour de leur utilisation afin d'apaiser le climat qui pourrait mener à des escalades incontrôlées. En ce sens c'est déjà ce qu'avait mentionné le secrétaire général des Nations Unies en 1984 à propos de l'arme nucléaire où «le processus nécessaire pour amorcer la désescalade et inverser la course aux armements pourrait être facilité si les États prenaient des initiatives unilatérales visant à réduire les tensions internationales en créant progressivement un climat de confiance mutuelle et améliorant de façon générale l'atmosphère en vue de négociations sur la limitation des armements et le désarmement».¹⁸⁷

¹⁸⁵ Ministère des Armées, «Manuel de droit des opérations militaires», Édition 2022. Disponible en version PDF.

¹⁸⁶ Ysens de la Panouse – de France, «Le droit à l'épreuve de l'automatisation des systèmes militaires robotisés terrestres», 2019. Disponible sur [[Le lien](#)].

¹⁸⁷ A/39/516, Review of the implementation of the recommendations and decisions adopted by General Assembly at its 10th special session – Unilateral nuclear disarmament measures, Report of the Secretary General, 5 octobre 1984, §2. Cité par Grégory Bouterin, «Maîtrise des armements non conventionnels: le salut viendra-t-il du soft disarmament?», *AFDI*, LIII, 2007, p. 243. Disponible en version PDF sur [[Le lien](#)].

2- *Les risques de prolifération*

142. Un autre danger relatif à la course à l'armement de l'IA est corrélé à la prolifération des systèmes d'intelligence artificielle. En effet l'IA est utilisée par tous parce que son coût est relativement faible comparativement à d'autres armes de guerre comme l'arme nucléaire. Dès lors, il peut être facile de constituer ses propres logiciels d'IA pour les intégrer dans des systèmes d'armes autonomes.

Du fait de cette facilité, l'IA intégrée dans des systèmes d'armes autonomes qu'elles soient létales ou non «pourraient tomber entre les mains d'acteurs non étatiques, comme des groupes terroristes, qui cherchent à exploiter les failles des puissances établies».¹⁸⁸

Cette prolifération des armes autonomes intégrant de l'IA pose donc un grand défi sécuritaire qui rend complexe sa régulation. En effet, quand bien même les pays avancés sont en discussion pour leur encadrement, certains groupes terroristes, non étatiques ou de simples hackers pourraient créer encore plus de dégâts humanitaires.

Sur la complexité de la régulation des systèmes d'IA militaire du fait de leur prolifération, Laure de Roucy, directrice du Centre géopolitique des technologies de l'Institut français des relations internationales, fait une comparaison utile sur la question. Ainsi, «dans les années précédant la Convention d'Ottawa, il était estimé que 110 millions de mines antipersonnel étaient déployées dans 64 pays, coûtant chacune 3 dollars – ce qui les rendait

¹⁸⁸ Institut National des Affaires Stratégiques et Politiques, «Implications géopolitiques et compétition internationale: la nouvelle course à l'IA militaire». Disponible sur [[Le lien](#)].

accessibles pour n'importe quel pays», mais finalement «même avec des armements largement répandus et peu coûteux, la régulation s'est avérée possible».¹⁸⁹ Par cet exemple, la prolifération des IA militaires montre que leur régulation est un défi complexe mais pas insurmontable.

B) Défis éthiques et juridiques

143. La course à l'IA militaire pose aussi de nombreux défis. En effet, bien que des discussions pour l'encadrement des SALA soient enclenchées depuis 2013 auprès de la CCAC, il n'en demeure pas moins que la décision de tuer provenant d'une machine pose des questions d'ordre éthique et juridique.

1- Les problématiques éthiques

144. Origine de l'éthique. Le terme *éthique* dérive du grec *êthos*, et le mot *moral* du latin *mos-moris*, tous deux désignant les mœurs, la conduite ou les règles de vie. Étymologiquement, ils partagent un sens proche. Toutefois, dans la pensée contemporaine, une distinction s'est progressivement imposée. La morale est souvent perçue comme un ensemble de règles universelles, rigides et prescriptives, tandis que l'éthique, plus contextuelle, renvoie à une réflexion sur le *bien agir*, ancrée dans les situations concrètes.¹⁹⁰ Elle représente, selon la formule de Paul Ricœur, «le mouvement même

¹⁸⁹ Laure de Roucy-Rochegonde, «La guerre à l'ère de l'intelligence artificielle: quand les machines prennent les armes», édition PUF, 1^{ère} édition, édition 2024, p. 306.

¹⁹⁰ Laure Tabouy, Cours de Transhumanisme, M2 Droit de l'Intelligence Artificielle. Année 2024-2025.

de la liberté qui cherche une vie bonne, dans la sollicitude envers autrui et dans un juste usage des institutions sociales». ¹⁹¹

145. Éthique militaire. Dans le champ militaire, cette approche éthique est essentielle lorsqu'il s'agit de l'emploi de systèmes d'armes létaux autonomes. Leur autonomie décisionnelle soulève des interrogations profondes sur la délégation de la force létale à une machine. La position de la France est claire sur ce point: «d'un point de vue éthique, il n'est pas acceptable que la décision de vie ou de mort puisse être intégralement confiée à des machines échappant à toute forme de supervision humaine». ¹⁹²

146. L'exemple du colonel soviétique Stanislav Petrov, que nous avons vu précédemment (n°83), illustre avec force la nécessité du discernement humain dans des contextes à haute intensité. Si l'erreur humaine reste possible, elle s'accompagne d'un jugement moral, d'une capacité à contextualiser, à douter, à désobéir parfois, au nom d'un bien supérieur. Une machine, aussi perfectionnée soit-elle, ne peut encore reproduire cette faculté éthique.

147. Néanmoins, certains experts soutiennent qu'il serait envisageable, dans certains cas, de concevoir des robots capables de respecter les normes du droit international humanitaire. Le spécialiste en robotique Ronald Arkin a notamment proposé l'intégration d'un dispositif de contrôle éthique, nommé le «*Test*

¹⁹¹ Paul Ricoeur, *Soi-même comme un autre*, éditions du Seuil, 1990 p. 202. Disponible en version PDF sur [\[le lien\]](#)

¹⁹² Ministère de l'Europe et des affaires étrangères, «systèmes d'armes létales autonomes, quelles est l'action de la France?», *op.cit.*

d'Arkin» dans les systèmes autonomes, visant à garantir leur conformité aux principes du DIH au cours des opérations militaires.¹⁹³

148. Ainsi une réflexion structurée sur l'éthique au sein des forces armées est nécessaire pour équilibrer technologie et valeurs sociétales et apparaît comme la «*seule solution pragmatique pour encadrer la montée en puissance des SALA, au contraire d'une interdiction générale inefficace, voire contre-productive*». ¹⁹⁴

2- Défis juridiques

149. L'introduction de l'Intelligence Artificielle dans les Systèmes d'Armes Létaux Autonomes soulève de profondes interrogations quant à leur conformité au droit international humanitaire, en particulier en ce qui concerne le respect des principes fondamentaux que sont la distinction, la proportionnalité et la précaution. Une première question essentielle est de savoir si des machines, même avancées, peuvent réellement garantir le respect de ces normes, qui impliquent des jugements contextuels et éthiques complexes.

¹⁹³ Ronald Arkin, *Governing Lethal Behaviour in Autonomous Robots* (CRC Press, 2009); Kenneth Anderson et Matthew Waxman, «Law and ethics for robot soldiers», *Policy Review*, no 176 (2012). Cité par Christof Heyns, rapporteur spécial sur les exécutions extrajudiciaires, sommaires ou arbitraire. Assemblée générale des Nations Unies, A/HRC/23/47, 9 avril 2013.

¹⁹⁴ Nathalie Durhin, «Sytèmes d'armes létaux autonomes: ne pas mélanger juridique et éthique», *Revue Défense Nationale*, 2018/HS1 n°Hors-série, pp. 167 à 176. Disponible sur [Le lien].

150. Certaines analyses soutiennent que, du fait de leur capacité à opérer sans émotions humaines, les SALA pourraient théoriquement mieux appliquer les règles du DIH, en particulier en discriminant plus précisément entre combattants et non-combattants, ou en calibrant la force utilisée en fonction de la menace. Ainsi, le rapport de la *Task Force IA et défense* du ministère des Armées souligne que l'intelligence artificielle pourrait « *favoriser le respect du droit international humanitaire en permettant une meilleure appréciation de l'environnement des opérations à un niveau tactique, opératif et stratégique* », ¹⁹⁵ tout en renforçant la proportionnalité et la stricte nécessité des actions militaires.

151. Toutefois, ces potentialités se heurtent à des limites techniques majeures. Un enjeu central réside dans la conception même de ces systèmes, largement tributaire des partis-pris humains. En effet, comme l'ont souligné plusieurs rapports, les SALA pourraient s'avérer discriminatoires pour de multiples raisons. Les biais cognitifs des développeurs, leurs choix de programmation ou encore les données utilisées pour entraîner les algorithmes peuvent influencer profondément le comportement décisionnel des systèmes. Ce biais algorithmique, parfois invisible, remet directement en cause la capacité des SALA à appliquer de manière neutre les principes de distinction ou de proportionnalité, notamment dans des contextes opérationnels complexes et dynamiques. ¹⁹⁶ Le rapport d'information de l'Assemblée nationale

¹⁹⁵ Rapport de la Task Force, 2019. *Op.cit.*

¹⁹⁶ Human Rights Watch, «Un danger pour les droits humains, systèmes d'armes autonomes et processus décisionnels numériques», résumé et recommandations, 28 avril 2025. Disponible en version PDF sur [Le lien].

rappelle ainsi que les capacités de perception des SALA, reposant uniquement sur des algorithmes préétablis, sont susceptibles d'être inadaptées à la complexité du champ de bataille. Il note notamment que *«la capacité d'un SALA à distinguer un civil d'un combattant interroge directement la capacité à respecter le principe de distinction»*,¹⁹⁷ et que la modélisation algorithmique des principes de proportionnalité ou d'interdiction des souffrances inutiles demeure extrêmement délicate, car elle suppose un jugement contextuel subtil.

152. Le Conseil de l'Europe, dans son projet de résolution de 2022, insiste quant à lui sur l'importance de la prévisibilité des SALA pour garantir leur conformité au principe de précaution. Selon lui, *«pour être conforme au principe de précaution, l'action des SALA doit être prévisible. Les utilisatrices et les utilisateurs doivent pouvoir ajuster ou annuler les effets des systèmes d'armes si nécessaire, ce qui n'est possible que s'ils peuvent raisonnablement prévoir l'action et la réaction d'un système d'armes»*.¹⁹⁸ Sur la prévisibilité des systèmes le CICR s'est positionné en disant que *«les systèmes d'armes autonomes imprévisibles devraient être formellement exclus, notamment en raison de leurs effets indiscriminés»*¹⁹⁹.

153. Dans cette optique, le contrôle humain reste crucial, à la fois dans la phase de programmation et au moment du

¹⁹⁷ Assemblée nationale, rapport d'information n°3248. *Op.cit.*

¹⁹⁸ Conseil de l'Europe, Commission des questions juridiques et des droits de l'homme, «L'Emergence des systèmes d'armes létales autonomes (SALA) et leur nécessaire appréhension par le droit européen des droits humains», Projet de résolution adopté à l'unanimité par la Commission le 14 novembre 2022. Disponible en version PDF sur [\[Le lien\]](#).

¹⁹⁹ *Revue internationale de la Croix-Rouge*, «Position du Comité International de la Croix-Rouge sur les systèmes d'armes autonomes et documentations de référence», IRCC No.915, janvier 2022. Disponible sur [\[Le lien\]](#).

déclenchement de l'action létale. A ce titre, la Professeure Marie-des-Neiges Ruffo propose que *«l'humain demeure l'expert qui vérifiera les critères de "tri", et fera la distinction finale pour épargner les innocents»*, en soulignant la nécessité d'interventions humaines a priori, via la définition des cibles et a posteriori, dans l'exécution de l'attaque. Et propose aussi que pour respecter le principe de distinction entre les combattants et non-combattants une zone soit clairement identifiée et délimitée *«pour prévenir les civils de ne pas s'approcher»*.²⁰⁰

154. Sur la question de la responsabilité, le CICR maintient sa position sur le fait que ce sont les personnes qui *«planifient, approuvent et réalisent des opérations militaires»*²⁰¹ qui sont responsables des obligations juridiques et des principes prévus par le DIH. Toutefois, *«l'augmentation de la vitesse de ciblage, de l'ampleur de la force utilisable et de la dépendance à l'égard de l'intelligence artificielle et de l'apprentissage automatique pour le contrôle du ciblage et de l'exercice de la force – entraînent une réduction de la capacité des utilisateurs humains à interpréter les informations reçues, à analyser de manière critique leurs choix et à agir rapidement en accord avec les principes humanitaires, juridiques et éthiques»*²⁰². Ce constat met en lumière l'un des défis juridiques les plus sensibles liés à l'autonomisation croissante des décisions létales: celui de la dilution de la responsabilité que nous reverrons

²⁰⁰ MdN Ruffo, «Avons-nous le choix d'utiliser l'IA en temps de guerre?», 47^e numéro des Cahiers de la sécurité et de la justice, 12 mars 2020. *Op.cit.*

²⁰¹ CICR, Rapport, «Le droit internationale humanitaire et les défis posés par les conflits armés contemporains: engagement renouvelé en faveur de la protection dans les conflits armés à l'occasion du 70^e anniversaire des Conventions de Genève», 2020. Disponible en version PDF sur [\[Le lien\]](#).

²⁰² *Revue Internationale de la Croix-Rouge*, «Position du Comité International de la Croix-Rouge sur les systèmes d'armes autonomes et documentations de référence». *Op.cit.*

plus tard dans le développement. Bien que «*les États soient d'accord sur l'importance du contrôle humain, ou de la "responsabilité humaine" pour garantir le respect du droit, les avis divergent sur sa signification dans la pratique*». ²⁰³

155. Conclusion du Titre 1. En fin de compte, la question de la conformité au DIH des systèmes d'armes de ciblage humain intégrant de l'IA qu'ils soient autonomes ou non restent en suspens. En effet, les SIA militaires posent de nombreux défis, juridiques, techniques et éthiques qui font obstacle à leur conformité avec les principes du DIH. Toutefois, cette mise en conformité n'est pas impossible si les États travaillent de concert pour adapter le DIH à l'ère de l'IA militaire.

²⁰³ CICR, Rapport, «Le droit internationale humanitaire et les défis posés par les conflits armés contemporains: engagement renouvelé en faveur de la protection dans les conflits armés à l'occasion du 70^e anniversaire des Conventions de Genève», 2020. *Op.cit.*

TITRE 2: Vers une adaptation du DIH à l'ère de l'IA militaire

156. Face aux mutations technologiques apportées par l'IA militaire, une question centrale se pose: peut-on adapter efficacement ces systèmes aux normes existantes du droit international humanitaire, ou les limites techniques de l'IA exigent-elles une évolution du cadre juridique? Cette partie se propose d'explorer, dans un premier temps, les enjeux liés à la responsabilité juridique engendrée par l'usage de ces technologies (Chapitre 1), avant d'examiner la nécessité d'un cadre juridique et éthique renouvelé pour encadrer leur déploiement (Chapitre 2).

Chapitre 1: Les enjeux de la responsabilité

157. L'utilisation de systèmes d'armes autonomes soulève d'importants enjeux de responsabilité. Ces enjeux tiennent à leurs causes juridiques et opérationnelles (Section 1), mais aussi aux moyens de les encadrer (Section 2).

Section 1: Les causes de cette responsabilité du fait de l'usage de ce type de SIA

158. Les causes de la responsabilité liée aux SALA tiennent notamment à une double érosion. D'une part celle du contrôle humain effectif (§1), que ces systèmes soient sous supervision humaine ou en dehors de tout contrôle, et d'autre part, celle de la responsabilité humaine directe (§2) notamment par la disparition de l'agentivité humaine au profit d'une agentivité de la machine.

§ 1: *Érosion du contrôle humain*

A) Le cas des SIA de ciblage humain sous supervision humaine (type Lavender)

159. Dans les débats dans l'utilisation de l'IA par les forces militaires, il est souvent mentionné l'aspect du contrôle humain significatif ou «*meaningful human control*». Cette notion désigne en fait la capacité d'un opérateur humain à comprendre les décisions prises par le système, intervenir pour modifier, arrêter ou valider

une action et assumer la responsabilité de cette dernière.²⁰⁴ C'est donc ce contrôle qui est censé être opéré dans les SIA de ciblage humain sous supervision humaine, comme le système *Lavender*, utilisé par l'armée israélienne.

1- *Lavender vs SALA*

160. Présentation de *Lavender* et des autres utilisations de l'IA par l'armée israélienne. Le rôle de l'intelligence artificielle dans les conflits contemporains est devenu un sujet crucial de débat. L'armée israélienne, pionnière dans l'adoption de ces technologies, utilise diverses applications de l'IA pour soutenir ses stratégies militaires. Parmi ces technologies, le système *Lavender* s'est particulièrement distingué dans l'optimisation du ciblage humain. Ce système est conçu pour identifier des individus en fonction de critères spécifiques, comme leur affiliation présumée à des groupes armés palestiniens, en attribuant des scores aux habitants de Gaza.²⁰⁵ Ces scores déterminent ensuite la liste des personnes potentiellement visées pour des frappes aériennes. Selon une enquête du magazine *+972*, l'influence de *Lavender* «était telle que les résultats de l'IA étaient traités comme s'il s'agissait d'une décision humaine»,²⁰⁶ avec peu de remise en question de la machine par les opérateurs. En complément de ce système, des outils comme «*Where's Daddy?*» localisent précisément ces cibles avant leur élimination par les forces israéliennes. Toutefois, bien que la

²⁰⁴ CICR, «Réunion d'experts sur les systèmes d'armes létales autonomes», 15/11/2017. Disponible sur [Le lien].

²⁰⁵ Human Rights Watch, «Gaza: Les outils numériques de l'armée israélienne mettent en danger les civils», 10 septembre 2024. Disponible sur [Le lien].

²⁰⁶ Yuval Abraham, «Lavander: The AI machine directing Israel's bombing spree in Gaza», *+972 Magazine*, 3 avril 2024. Disponible sur [Le lien].

technologie de *Lavender* ne prenne pas directement la décision de la frappe létale, elle joue un rôle décisif dans la préparation de ces attaques, en soumettant une liste de cibles que l'humain ne fait que valider.

Cette dynamique soulève une question importante: dans quelle mesure l'humain conserve-t-il un contrôle effectif sur ces systèmes d'IA, notamment quand la machine devient le principal acteur dans la prise de décision?

161. Distinction entre ce type de système et les SALA. Il est essentiel de faire une distinction entre les systèmes comme *Lavender* et les systèmes d'armes létales autonomes, qui, eux, prennent des décisions en totale autonomie, sans intervention humaine directe. Tandis que *Lavender* soutient une décision humaine, son rôle reste limité à la phase de ciblage, le militaire restant responsable de l'activation ou non de la frappe. En revanche, les SALA, comme leur nom l'indique, permettent à l'IA de prendre la décision de tuer sans supervision humaine, en analysant les situations de manière autonome et en engageant l'attaque sans l'aval final d'un opérateur. Dans ses rapports, le CICR met en garde contre le développement de telles technologies, soulignant que l'absence de prise de décision humaine compromet la responsabilité et pourrait porter atteinte aux principes fondamentaux du droit international humanitaire.²⁰⁷ Ainsi, la frontière entre l'assistance de l'IA à l'opération humaine et l'automatisation complète de l'action est mince mais fondamentale.

²⁰⁷ Mise en garde que nous pouvons trouver dans la majorité des rapports du CICR sur les armes autonomes et l'utilisation de l'intelligence artificielle.

La distinction entre *Lavender* et les SALA repose donc sur la question du contrôle humain.

2- Les facteurs de l'érosion du contrôle humain

162. Érosion du contrôle humain dû à l'accélération du rythme de la guerre. En 2020, le CIRC alertait déjà sur l'utilisation de l'IA dans des systèmes d'aide à la décision qui «pourrait avoir de profondes implications sur le rôle des êtres humains dans un conflit armés». ²⁰⁸ A ce titre, le Comité préconisait une approche centrée sur l'humain avec un jugement humain prépondérant pour toutes les décisions susceptibles d'avoir «de graves conséquences sur la vie des personnes» ²⁰⁹ et susceptibles de violer le DIH. Toutefois, dans un contexte opérationnel, la réalité est toute autre et mettent à mal cette recommandation.

163. L'un des défis majeurs réside dans l'accélération du rythme des opérations, où les décisions doivent être prises dans des délais de plus en plus courts. Dans ce contexte, le rôle de l'opérateur humain devient de plus en plus réduit. À cet égard, l'utilisation de *Lavender* par l'armée israélienne en est un exemple frappant. Selon l'enquête menée par Yuval Abraham, un agent de l'armée israélienne a révélé qu'il passait à peine 20 secondes à examiner chaque cible, avant de valider la décision de la machine. ²¹⁰ Cette

²⁰⁸ CIRC, Rapport, «Le droit internationale humanitaire et les défis posés par les conflits armés contemporains: engagement renouvelé en faveur de la protection dans les conflits armés à l'occasion du 70^e anniversaire des Conventions de Genève», 2020. *Op.cit.*

²⁰⁹ *Ibid.*

²¹⁰ Yuval Abraham, "Lavender: The AI machine directing Israel's bombing spree in Gaza", +972 Magazine. *Op.cit.*

courte période de validation remet en question la notion de contrôle humain effectif prôné par le CICR. La rapidité des décisions, imposée par l'urgence du conflit, empêche une véritable réflexion sur la proportionnalité et la distinction des cibles. L'homme, ainsi réduit à une formalité administrative, se trouve dans une situation où sa capacité à intervenir de manière significative est diluée. En effet, le CICR souligne que *«l'augmentation de la vitesse de ciblage, de l'ampleur de la force utilisable et de la dépendance à l'égard de l'intelligence artificielle et de l'apprentissage automatique pour le contrôle du ciblage et de l'exercice de la force – entraînent une réduction de la capacité des utilisateurs humains à interpréter les informations reçues, à analyser de manière critique leurs choix et à agir rapidement en accord avec les principes humanitaires, juridiques et éthiques»*.²¹¹ Cette accélération du processus décisionnel par la machine crée une pression qui fragilise le respect des principes de précaution et de proportionnalité, deux valeurs fondamentales du droit international humanitaire.

Si l'IA offre un gain de vitesse dans le cycle décisionnel militaire, cet avantage peut accroître les risques pour les civils. À l'inverse, ralentir le processus par une « patience stratégique » permet une meilleure évaluation des options, mais l'idée que l'IA faciliterait cette patience en accélérant les phases préalables reste une vision simplificatrice du ciblage.²¹² Cela nous amène à nous

²¹¹ *Revue Internationale de la Croix-Rouge*, «Position du Comité International de la Croix-Rouge sur les systèmes d'armes autonomes et documentations de référence». *Op.cit.*

²¹² Stewart, Ruben et Hinds Georgia, «Les algorithmes de la guerre: le recours à l'intelligence artificielle pour la prise de décision dans les conflits armés», 22 février 2024. Disponible sur [\[Le lien\]](#).

interroger sur les dangers du biais de sur-confiance, qui est amplifié par la dépendance à ces systèmes d'IA.

164. Érosion du contrôle humain dû à un biais de sur-confiance. Le biais de sur-confiance est une autre conséquence inquiétante de l'utilisation de l'IA dans les décisions militaires. L'opérateur humain, dans un environnement où les machines fournissent des résultats censés être fiables et objectifs basés sur des algorithmes complexes, est susceptible de devenir excessivement dépendant de ces outils et de leur faire une confiance aveugle. L'usage de *Lavender*, malgré ses marges d'erreur, illustre bien ce phénomène. Selon l'enquête *Lavender* aurait recensé jusqu'à 37 000 Palestiniens comme cibles potentielles, avec une marge d'erreur de 10%,²¹³ ce qui signifie que jusqu'à 3 700 civils pourraient être inclus à tort dans cette liste. Pourtant, dans ce type de situation, l'humain valide ces décisions sans remettre en question les résultats de l'IA, convaincu que la machine, en raison de son apparente objectivité, ne peut se tromper. Le biais de sur-confiance conduit à une dilution du contrôle humain, d'autant plus que l'humain n'a plus qu'un rôle de validation plutôt que de décision critique.

165. L'interaction homme-machine pose des défis liés aux limites cognitives humaines, notamment le «biais d'automatisation». Cela se réfère à la tendance humaine à ne pas examiner de manière critique les diagnostics générés par le système ou à ne pas chercher d'informations contraires, surtout en situation

²¹³ Yuval Abraham, "Lavender: The AI machine directing Israel's bombing spree in Gaza", +972 *Magazin*. *Op.cit.*

d'urgence.²¹⁴ Dans un conflit armé, une confiance excessive peut avoir des conséquences fatales. A ce titre, l'incident du système Patriot américain en 2003, tirant sur des avions alliés car considérés à tort comme ennemis, a montré que la confiance des opérateurs dans le logiciel était une insuffisance majeure.²¹⁵ Dans ce contexte, ces modes de fonctionnement et l'interaction homme-machine peuvent accroître la probabilité que les évaluations produites par une IA s'écartent des intentions réelles du décideur humain, exposant ainsi à un risque d'escalade involontaire et à un danger accru pour les civils.²¹⁶

166. Ce phénomène soulève d'importantes préoccupations en matière de respect du droit international humanitaire, notamment en ce qui concerne la distinction entre civils et combattants, ainsi que la proportionnalité des attaques.

3- Tension avec le DIH

167. L'utilisation de systèmes d'IA dans les conflits militaires, bien que visant à rendre les opérations plus efficaces, pose de graves questions concernant le respect des principes du Droit International Humanitaire.

En effet, le principe de distinction, qui impose de distinguer les civils des combattants et de cibler exclusivement ces derniers, se voit compromis lorsque des systèmes comme *Lavender* identifient des individus en fonction d'algorithmes qui peuvent être biaisés ou erronés. De plus, le principe de proportionnalité,

²¹⁴ *Ibid.*

²¹⁵ *Ibid.*

²¹⁶ *Ibid.*

qui exige que les attaques ne causent pas de dommages excessifs aux civils par rapport à l'avantage militaire attendu, est également mis en péril. Les frappes basées sur des données incomplètes ou faussées peuvent conduire à des pertes civiles disproportionnées. Aussi, le lourd bilan civil observé s'explique par des décisions humaines liées à la doctrine de ciblage israélienne.²¹⁷ Après le 7 octobre 2023, la «kill list» a été considérablement élargie, affaiblissant le principe de proportionnalité. Dans certains cas, l'armée israélienne a accepté la mort de dizaines, voire de centaines de civils pour atteindre un seul membre du Hamas.²¹⁸

Enfin, le principe de précaution, qui oblige à prendre des mesures pour minimiser les risques pour les civils, est aussi fragilisé par la rapidité des décisions et l'incapacité de réévaluer chaque situation en temps réel. Dans l'esprit du CIRC, l'absence de jugement humain dans la prise de décision pourrait entraîner une violation des principes de précaution, de proportionnalité et de distinction. Ces préoccupations sont d'autant plus importantes face à des technologies comme *Lavender*, où l'humain est réduit à une fonction de validation rapide, sans possibilité réelle d'intervenir.

168. Ce cas montre vraiment que le contrôle humain est dilué au profit d'un contrôle de la machine. Finalement, ce sont les humains qui deviennent des machines, en agréant les décisions de cette dernière et en tuant automatiquement, sans prise de recul, les

²¹⁷ Laure de Roucy-Rochegonde et Amélie Férey, «Bombardement israéliens à Gaza: comment l'intelligence artificielle annihile les préventions éthiques», 07/05/2024. Disponible sur [Le lien].

²¹⁸ *Ibid.*

cibles détectées par le système. Alors que les systèmes d'aide à la prise de décision devraient être «utiles pour éclairer la décision humaine, mais ne sauraient la remplacer dans des situations qui mettent en péril la vie et la dignité des personnes dans les conflits armés».²¹⁹

Cela pose donc une réelle problématique en termes de respect du DIH notamment dans les principes de distinction, de proportionnalité et de précaution.

Alors que le contrôle humain est censé faire en sorte de respecter ces principes, nous voyons bien que dans l'utilisation de *Lavender*, ce n'est pas tout à fait le cas.

B) *Le cas des SLA de ciblage humain intégralement*

169. Nous avons vu que dans le cas de *Lavender*, le contrôle humain est de plus en plus faible, alors que la machine prend de plus en plus d'autonomie. C'est ainsi que le glissement s'effectue pour arriver à des armes entièrement automatisées pour tuer, les SALA.

1- *La disparition intégrale de l'humain de la boucle...*

170. La tendance actuelle dans le développement des systèmes d'armement autonomes laisse entrevoir un effacement progressif, puis total, de la présence humaine dans la boucle décisionnelle. Si des systèmes comme *Lavender* témoignent déjà d'un affaiblissement notable du contrôle humain au profit de l'automatisation, d'autres modèles dits «Terminator»²²⁰ vont bien plus loin. Ces derniers excluent totalement l'homme de la boucle décisionnelle et

²¹⁹ Stewart, Ruben et Hinds, Georgia, *op.cit.*

²²⁰ Noël, Jean-Christophe, «L'IA a-t-elle un avenir pour les armées?», article paru dans la *Revue Défense Nationale*, 2019/N°820. Disponible sur [Le lien].

selon la grille des niveaux d'automatisation des systèmes armés de Thierry Berthier, ces systèmes correspondraient au niveau L5, c'est-à-dire un «*système armé autonome sans tutelle humaine, où l'opérateur humain n'a pas la possibilité de reprendre le contrôle du système pleinement autonome*». ²²¹ L'ensemble du processus, depuis la détection des objets, des formes et des cibles jusqu'à l'ouverture du feu, est exécuté par le système lui-même. ²²² Ainsi le système entier, qu'il prenne la forme d'un robot ou d'un essaim de drones, serait chargé d'une tâche en totale autonomie. L'intelligence artificielle y assure la planification comme l'exécution de la mission, privilégiant certains modes d'action «*en fonction du contexte ou de la situation tactique connus*», ²²³ grâce à l'analyse de «*combats virtuels qu'elle a testés contre elle-même, à la manière d'AlphaGo Zero*». ²²⁴ Dans cette configuration, l'homme ne participe plus à la sélection des cibles ni à la coordination des frappes. Il peut tout au plus, en amont, autoriser le déclenchement du système, dans un cadre parfois défini, parfois non.

171. Ce modèle, s'il semble encore hypothétique d'un point de vue technique mais aussi éthique et moral, est néanmoins envisagé dans des systèmes déjà opérationnels. Ainsi, le système *Aegis* américain intègre «*des modes de tir complètement automatisés*», ²²⁵ soulevant la question cruciale du remplacement de la décision humaine dans les conflits armés. Ce glissement technologique

²²¹ Thierry Berthier, «Systèmes armés semi-autonomes: que peut apporter l'autonomie?», *Revue Défense Nationale*, 2019/5, n°820, pp. 74- 80. Disponible sur [Le lien].

²²² *Ibid.*

²²³ Noël, Jean-Christophe, «L'IA a-t-elle un avenir pour les armées?», *op.cit.*

²²⁴ *Ibid.*

²²⁵ *Ibid.*

nourrit un débat juridique et éthique de plus en plus pressant autour de la notion de «contrôle humain significatif». Les réflexions internationales sur le «contrôle humain suffisant» ou «significatif» sur les SALA témoignent bien de cette inquiétude croissante. Si ce débat existe, c'est qu'un basculement technologique et doctrinal est déjà en cours. Ce constat pose les fondements d'une transformation doctrinale profonde, dans laquelle l'homme devient périphérique dans un système de combat dominé par des logiques algorithmiques.

2- À l'avènement de l'hyperguerre par IA

172. Ce processus d'autonomisation totale des systèmes d'armes débouche logiquement sur ce que certains auteurs nomment l'hyperguerre, c'est-à-dire une forme de conflit échappant presque entièrement à l'intervention humaine. Selon le général John Allen et A. Husain, *«en termes militaires, l'hyperguerre peut être redéfinie comme un type de conflit où la prise de décision humaine est presque entièrement absente de la boucle observation-orientation-décision-action»*.²²⁶ Dans ce modèle, des *«systèmes d'essaims complexes et largement autonomes conduisent des frappes cinétiques coordonnées, imposant des réactions quasi-instantanées automatisées»*.²²⁷ La vitesse d'exécution des actions sur le champ de bataille devient telle que l'homme est dépassé cognitivement, au point qu'*«une IA pourrait seule commander la riposte»*,²²⁸ pilotée par *«des logiciels fonctionnant toujours plus rapidement»*.²²⁹

²²⁶ General John R. Allen and Amir Husain, "On hyperwar", U.S Naval Institute, juillet 2017. Disponible sur [Le lien].

²²⁷ Thomas Noizet, «Les essaims de drones, graal ou chimère?», Revue Défense Nationale 2023/HS13 N° Hors-série, pp. 239-256. Disponible sur [Le lien].

²²⁸ Noël, Jean-Christophe. *Op.cit.*

²²⁹ *Ibid.*

173. L'hyperguerre s'appuie sur une doctrine d'essaimage,²³⁰ où l'IA permet de surmonter les limites cognitives humaines en matière de coordination. À terme, «*l'IA pourrait permettre à l'essaimage de devenir le mode d'opération dominant, avec des milliers d'unités dispersées sur un champ de bataille puis convergeant pour attaquer*». Une telle approche rendrait les structures hiérarchiques traditionnelles obsolètes, et l'intervention humaine serait réduite à la définition d'objectifs généraux. L'action tactique serait confiée à la machine, dans un modèle non plus de collaboration mais de supervision distanciée où «*les humains définiraient les objectifs et superviseraient, mais confieraient l'exécution du comportement de l'essaim aux systèmes d'IA*».²³¹

174. Ce changement de paradigme n'est pas sans rappeler des précédents historiques. L'intégration de l'IA dans la guerre pourrait ainsi reproduire les erreurs de la Première Guerre mondiale, où les militaires ont eu du mal à adapter leurs tactiques à la létalité accrue introduite par la mitrailleuse. Cette comparaison souligne le risque d'une inadéquation doctrinale face à une transformation technique radicale. À long terme, une *singularité*²³² militaire pourrait survenir, où les conflits, accélérés au-delà des capacités humaines de supervision, seraient décidés et menés intégralement par des intelligences artificielles. Une telle évolution pourrait marquer la fin

²³⁰ Paul Scharre, *Four battlegrounds: power in the age of artificial intelligence*, New York, NY : W.W. Norton & Company, 2023, xvi, 473p. Extrait du passage disponible sur [\[Le lien\]](#).

²³¹ Paul Scharre. *Op. Cit.*

²³² *Ibid.*

du contrôle humain sur la guerre, non par accident, mais par incapacité à suivre son rythme.

175. Une telle évolution, bien que située à un horizon de long terme, impliquerait un abandon progressif du contrôle stratégique, tactique et opérationnel de l'homme sur les conflits armés, avec des conséquences potentiellement irréversibles, c'est-à-dire que même en choisissant de déclencher un conflit, les humains pourraient perdre «*la capacité de contrôler l'escalade ou d'y mettre fin à tout moments*».²³³ Ainsi, comme le mentionnait Jacques Ellul, «*c'est maintenant la technique qui opère le choix [...]. L'homme n'est absolument plus l'agent du choix*».²³⁴

176. Ainsi, que ce soit dans le cas des SIA sous supervision humaine où des SIA totalement autonome, l'érosion du contrôle humain est de plus en plus flagrante et cette «*absence de contrôle par l'humain entraînerait donc une dilution de la responsabilité, qui rendrait pratiquement impossible le jugement et la réparation de violations du DIH*».²³⁵

§ 2: Érosion d'une responsabilité humaine

177. Le mot *responsabilité* vient du latin *respondere*, qui signifie «répondre de».²³⁶ Il évoque une obligation de rendre compte de ses

²³³ *Ibid.*

²³⁴ Jacques Ellul, *La technique ou l'enjeu du siècle*, 1954, Ed Economica, 1990, coll. *classiques des sciences sociales* dirigée par Hervé Coutau-Bégarie. p.74, II. Caractères de la technique moderne, Automatisation du choix technique.

²³⁵ Laure de Roucy-Rochegonde, *La guerre à l'ère de l'intelligence artificielle, quand les machines prennent les armées*, Edition Puf, 2024, p.142.

²³⁶ La Toupie, *Dictionnaire de la politique*, «responsabilité», disponible sur [Le lien].

actes, fondement essentiel de tout ordre juridique. En droit pénal, cela suppose que l'individu ayant commis un acte répréhensible soit identifié, jugé et, le cas échéant, sanctionné. Or, les progrès de l'intelligence artificielle ont conduit à son intégration dans les systèmes d'armes létales autonomes, capables de décider seuls d'ouvrir le feu, bouleversant ainsi les catégories traditionnelles du droit. Ces technologies soulèvent une interrogation majeure: à qui imputer la responsabilité lorsqu'une machine tue ?

A) Actus reus et mens rea: fondements de la responsabilité pénale

178. La responsabilité pénale repose traditionnellement sur deux piliers: l'*actus reus*, l'acte matériel illégal, et la *mens rea*, l'intention coupable.²³⁷ Ces deux éléments sont indispensables pour qu'un individu puisse être jugé responsable. Le simple fait de causer un dommage ne suffit pas, encore faut-il avoir agi avec conscience, volonté ou négligence.

Les juridictions pénales internationales ont clarifié ces conditions. Ainsi, le tribunal penal international pour l'ex-Yougoslavie, a affirmé dans l'affaire Kavishema et Ruzindana du 21 mai 1999 que "*there is a further two stage test which must be satisfied in order to establish individual criminal responsibility [...]. This test required the demonstration of (i) participation, that is that the accused's conduct contributed to the commission of an illegal act, and (ii) knowledge or intent, that is awareness by the actor of his participation in a crime*".²³⁸ Ainsi, la

²³⁷ Laure de Roucy-Rochegonde, *op.cit.*, p.138.

²³⁸ TPIY, Trial Chamber II, *The prosecutor vs. Clément Kayishema and Obed Ruzindana*, 21 mai 1999, §198. Disponible sur [Le lien]. Traduction: «il existe un test supplémentaire en deux étapes qui doit être satisfait afin d'établir la responsabilité pénale individuelle [...]. Ce test exige la démonstration de :

responsabilité pénale individuelle suppose la combinaison de la participation effective de l'accusé à l'acte incriminé et la conscience qu'il avait de participer à un crime. Sans ces deux critères réunis, aucune condamnation n'est possible.

Ce cadre s'applique difficilement aux SALA. Une machine peut-elle avoir une intention? Peut-elle faire preuve de discernement? Les SALA, dénués de conscience, d'émotions ou de moralité, ne satisfont pas à ces critères. Ils ne peuvent donc pas être jugés selon les principes fondamentaux du droit pénal.

B) Disparition de l'agentivité humaine: un changement de paradigme

1- De l'agentivité humaine...

179. Le concept d'agentivité désigne la capacité d'un individu à agir intentionnellement, en fonction de choix rationnels orientés vers des objectifs. En anglais, ce concept est résumé par le terme *agency*, qui fait référence à un acteur autonome capable d'influer sur le monde. Selon la Stanford Encyclopedia of Philosophy, «*un agent est un être doté de la capacité d'agir, et l'agentivité désigne l'exercice ou la manifestation de cette capacité.*»²³⁹.

180. Dans le cadre du droit international humanitaire, c'est bien cette agentivité humaine qui structure l'ensemble des règles avec notamment la distinction entre civils et combattants, le principe de

-
- (i) la participation, c'est-à-dire que la conduite de l'accusé a contribué à la commission d'un acte illégal ; et
 - (ii) la connaissance ou l'intention, c'est-à-dire la conscience par l'auteur de sa participation à un crime».

²³⁹ *Stanford Encyclopedia of Philosophy*, "Agency", First published Mon Aug 10, 2015; substantive revision Mon Oct 28, 2019. Disponible sur [Le lien].

proportionnalité, ou encore celui de précaution. Tout repose sur la capacité d'un soldat ou d'un commandement à apprécier la situation et à prendre des décisions morales et juridiques.

2- ... *À l'agentivité de la machine*

181. L'émergence des SALA fait glisser cette agentivité vers la machine. Ce n'est plus l'humain qui décide, mais un système algorithmique qui exécute une mission de manière totalement autonome. Dès lors, l'humain est complètement absent de la décision d'engager la force au moment crucial.

182. Le DIH, pensé pour un monde de guerre «humaine», ne prend pas en compte ce glissement. En effet, comme le souligne Nathalie Nevejans, «*Comment une machine pourrait-elle appliquer des règles bâties à la mesure de l'homme?*».²⁴⁰ Cela rend l'application du droit problématique, les SALA ne perçoivent pas l'intention, n'évaluent pas les émotions, et n'ont aucun sens moral. Cette déshumanisation des conflits rend difficile le respect des règles de guerre et illustre les limites actuelles du DIH.

C) Le problème d'attribution de la responsabilité aux SALA

183. Lorsque l'agentivité humaine disparaît, une zone grise se crée sur le plan juridique. Qui est responsable lorsqu'un SALA commet une violation du DIH? En effet, «*d'un point de vue juridique, pour que le droit international humanitaire, qui régit les conflits armés, puisse s'appliquer, il faut que l'usage d'une arme puisse être rattaché à une*

²⁴⁰ Nathalie Nevejans, «la légalité des robots de guerre dans les conflits internationaux». *Op.cit*

responsabilité humaines.²⁴¹ Mais nous l'avons vu, un système d'arme entièrement autonome, qui prend lui seul la décision de tuer, pose des questions quant à la potentielle responsabilité humaine derrière cet acte. Or, le Statut de Rome est clair, sans auteur humain clairement identifié, aucun procès équitable n'est possible.²⁴²

184. Plusieurs hypothèses sont donc avancées, notamment dans la recherche d'une responsabilité dès la conception des SALA à l'individu qui l'a activé. Mais cette chaîne de responsabilité est fragmentée et diluée, Julien Ancelin parle même de responsabilité «fuyante» entre les décideurs politiques et militaires, industriels, programmeurs ou opérateurs.²⁴³

Ronald Arkin, quant à lui, suggère d'imposer un suivi technique afin d'identifier l'humain responsable de l'activation de la machine.²⁴⁴

185. Enfin, cette dilution de la responsabilité a des conséquences morales, en l'absence de coupable humain, aucune reconnaissance de faute, demande de pardon ni réconciliation ne peut avoir lieu. Cela rend impossible tout processus de justice post-conflit important dans le *jus post bellum*.²⁴⁵

²⁴¹ Ministère de l'Europe et des affaires étrangères «Systèmes d'armes létales autonomes, quelle est l'action de la France?», février 2020. Disponible sur: [**Le lien**].

²⁴² Dominique Lambert, p.84, *op.cit*

²⁴³ Julien Ancelin «Les systèmes d'armes létales autonomes (SALA): Enjeux juridiques de l'émergence d'un moyen de combat déshumanisé», *La Revue des droits de l'homme*, Actualités Droits-Libertés, mis en ligne le 25 octobre 2016, consulté le 20 novembre 2024. Disponible sur [**Le lien**].

²⁴⁴ Mdn Ruffo, *op.cit*.

²⁴⁵ Mdn Ruffo, *op.cit*.

186. Les SALA bouleversent les bases de la responsabilité pénale et du Droit International Humanitaire. En remplaçant l'humain dans la boucle de décision, ils rendent inopérants les fondements classiques que sont l'*actus reus*, la *mens rea* et l'agentivité humaine. Le droit est aujourd'hui dépassé par la technologie, incapable de désigner un responsable clair lorsque l'auteur matériel est une machine. Face à cette évolution, il devient indispensable de repenser les cadres juridiques et éthiques afin de garantir une responsabilité humaine inaliénable dans l'usage de la force. Car sans responsabilité, et «*sans humain pour combattre, il n'y aura pas d'homme pour faire la paix*».²⁴⁶

Section 2: Les moyens pour endiguer cette responsabilité

187. Face aux défis posés par la responsabilité des systèmes de ciblage humain intégrant de l'IA, plusieurs pistes peuvent être envisagées, notamment l'attribution d'une personnalité juridique aux robots (§1) ou le recours à des mécanismes alternatifs de responsabilisation (§2).

§ 1: *L'attribution de la personnalité juridique aux robots létaux autonomes: une solution illusoire ?*

188. Certains courants de pensée ont envisagé la reconnaissance d'une personnalité juridique pour les Systèmes d'Intelligence

²⁴⁶ *Ibid.*

Artificielle les plus perfectionnés et autonomes.²⁴⁷ L'idée serait que, à l'instar des personnes morales comme les sociétés ou les associations, ces entités pourraient se voir attribuer des obligations et une responsabilité propre. Dans le contexte militaire, cela pourrait potentiellement permettre d'imputer directement la responsabilité des actions dommageables commises par les SALA à ces entités elles-mêmes. En effet, il paraît problématique «*de conférer une autorité à des êtres, par ailleurs dotés d'une autonomie croissante, donc en mesure de prendre des décisions par eux-mêmes, sans leur confier en retour la moindre responsabilité*»²⁴⁸.

A) Autonomie décisionnelle des IA militaires?

1- Positions doctrinales

189. Sur le plan doctrinal, si Maître Alain Bensoussan plaide pour la création d'une personnalité juridique pour les robots²⁴⁹ en raison de leur autonomie et liberté de décision, cette position reste isolée au sein de la doctrine. Lawrence Solum, dans son article intitulé «*Legal Personhood for Artificial Intelligences*»,²⁵⁰ soulevait déjà la question de la personnalité juridique des IA, mais pointait le défaut d'intentionnalité comme une différence majeure avec l'intelligence humaine, ce qui est problématique notamment dans une optique d'imputation de responsabilité. Bien que le droit puisse parfois se

²⁴⁷ Alain Bensoussan dans sa Charte Droit des Robots de 2013 et la proposition de résolution du Parlement européen concernant des règles de droit civil sur la robotique de 2017.

²⁴⁸ Paul Isnard, *La cité et ses esclaves*, *Institutions, fictions, expériences*, Seuil, 2019, pp.55-74. Incise I, «Le robot est-il un esclave comme les autres?»

²⁴⁹ A. Bensoussan, «Les robots ont-ils une personnalité?», *Planète Robots* n°19, janv. 2013, disponible sur: **[Le lien]**

²⁵⁰ Lawrence B. Solum, "Legal Personhood for Artificial Intelligences", 70 *N.C.L. Rev.*1231 (1992), disponible sur: **[Le lien]**. Paragraphe c.

passer de l'intentionnalité, l'absence de libre arbitre chez un robot, qui suit sa programmation, constitue un obstacle fondamental à l'assimilation à une personne humaine dotée de volonté.

2- Proposition du Parlement

190. De plus, le Parlement européen, qui s'est saisi de la question en 2017, a proposé de créer «une personnalité juridique spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques responsables, tenues de réparer tout dommage causé à un tiers; il serait envisageable de conférer la personnalité électronique à tout robot qui prend des décisions autonomes ou qui interagit de manière indépendante avec des tiers».²⁵¹

Cette proposition montre un basculement rapide de l'idée d'une reconnaissance fonctionnelle vers un objectif de responsabilisation financière des robots en cas de dommages. Cependant, une lettre ouverte adressée à la Commission européenne a soulevé des inquiétudes quant aux critères, à l'utilité sociale et aux implications symboliques d'une telle reconnaissance et a insisté sur le fait que «d'un point de vue éthique et juridique, la création d'une personnalité juridique pour un robot est inappropriée quel que soit le modèle de statut juridique».²⁵² Sur la même lignée, la Commission mondiale d'éthique des connaissances scientifiques et des technologies de l'UNESCO a déclaré qu'«il serait absurde de les qualifier de «personnes» puisqu'ils sont dépourvus de certaines autres qualités généralement associées

²⁵¹ Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)).

²⁵² Open letter to the European Commission, Artificial Intelligence and Robotics. Disponible sur [\[Le lien\]](#).

*aux êtres humains comme le libre arbitre, l'intentionnalité, la conscience de soi, le sens moral et le sentiment de l'identité personnelle».*²⁵³

Finalement, le Comité économique et social européen a émis un avis défavorable,²⁵⁴ et le projet a été abandonné. La proposition de directive européenne du 28 septembre 2022 se concentrant sur la responsabilité du fait de l'IA, a, elle aussi, été abandonnée le 11 janvier 2025, ce qui montre la difficulté de cette question d'imputer la responsabilité à une machine.

B) Un patrimoine pour les robots autonomes militaires

191. La question du patrimoine des robots, nécessaire pour assumer une éventuelle responsabilité financière, soulève également des difficultés. Si Maître Bensoussan propose un capital fonction des risques encourus, derrière toute entité juridique, qu'elle soit morale ou envisagée pour les robots, se trouve la question des apports et des garanties, souvent assurées par des personnes physiques.²⁵⁵ En effet, la personne humaine reste bien derrière la prétendue personnalité du robot, en ce que Maître Bensoussan, prévoit à l'article 2 de sa Charte des droits des robots

²⁵³ Rapport de la COMEST de l'UNESCO sur l'éthique de la robotique, SHS/YES/COMEST-10/17/2 REV.Paris, 14 septembre 2017 Original: anglais. Disponible sur: [\[Le lien\]](#).

²⁵⁴ Comité Économique et Social Européen 526e Session Plénière du CESE des 31 mai et 1er juin 2017, Avis du Comité économique et social européen sur «L'intelligence artificielle: les retombées de l'intelligence artificielle pour le marché unique (numérique), la production, la consommation, l'emploi et la société» (avis d'initiative) (2017/C 288/01).

²⁵⁵ Diane Galbois-Lehalle, Cours de M1 «La notion de personne». Années 2023-2024.

que ces derniers disposent «*d'un représentant légal*».²⁵⁶ Dès lors, la création d'une personnalité juridique pour doter un SALA d'un patrimoine apparaît comme un détour complexe pour finalement revenir à des mécanismes de responsabilité impliquant des acteurs humains.

192. L'attribution d'une personnalité juridique aux robots létaux autonomes, se heurte à des obstacles conceptuels majeurs liés à l'absence d'intentionnalité et de libre arbitre, ainsi qu'à des difficultés pratiques concernant la mise en place d'un régime juridique et patrimonial autonome. L'abandon de cette idée par les institutions européennes témoigne de ces limites et de la nécessité d'explorer d'autres voies pour encadrer la responsabilité dans le domaine de l'IA militaire.

§ 2: *Vers des mécanismes alternatifs de responsabilisation pour les systèmes d'IA militaires*

193. Face aux limites de l'approche de la personnalité juridique, il est possible d'explorer des pistes alternatives pour endiguer la question de la responsabilité des systèmes d'intelligence artificielle, qui pourraient être adaptées au contexte militaire.

A) Un patrimoine d'affectation

194. Une première piste s'inspire du droit romain et du statut des esclaves. À Rome, l'esclave est juridiquement assimilé à une

²⁵⁶Alain Bensoussan, Charte des Droits des Robots, V.5.0, septembre 2015. Disponible sur [[Le lien](#)].

chose, il n'est pas sujet de droit, mais en constitue un objet.²⁵⁷ Pour autant, les esclaves pouvaient agir juridiquement pour le compte de leur maître, et le droit romain avait développé le concept de pécule, une forme de patrimoine d'affectation dont la valeur limitait la responsabilité du maître pour les actes de son esclave.²⁵⁸

195. Ce dispositif de patrimoine affecté se retrouvait aussi dans le statut de l'entreprise individuelle à responsabilité limitée où il était permis pour «*les personnes physiques exerçant une activité économique individuelle d'affecter, sur simple déclaration, un patrimoine à leur activité professionnelle, en permettant la séparation du patrimoine personnel du patrimoine professionnel*».²⁵⁹ L'un des avantages de ce dispositif est «*la limitation de responsabilité dont va bénéficier le commerçant*».²⁶⁰

196. Transposé au domaine des SALA, on pourrait envisager un mécanisme où un patrimoine serait spécifiquement affecté à la réparation des dommages causés par ces systèmes, sans pour autant leur conférer la personnalité juridique au contraire de la proposition de Maître Bensoussan (voir *supra* paragraphe n°191). La responsabilité resterait ancrée auprès des concepteurs, des fabricants, ou des autorités militaires déployant ces systèmes et

²⁵⁷ Thierry Garé et Anaïs Raynaud, *Droit des personnes et de la famille*, éd Spécial Droit, Col. Ellipses, 2^{ème} édition, 2018.

²⁵⁸ Diane Galbois-Lehalle, «L'attraction de la catégorie des personnes sur les entités numériques», acte de colloque, «la personnalité convoitée», 2023. Disponible sur [[Le lien](#)].

²⁵⁹ Fiche d'orientation Dalloz, «Entreprise individuelle à responsabilité limitée», juillet 2023. Publié par Dalloz dans son répertoire Fiches d'orientation.

²⁶⁰ Jean Derruppé et Thibault de Ravel d'Esclapon, «Fonds de commerce», octobre 2016, actualisation en janvier 2025, paragraphes 195-196. Répertoire de droit immobilier, *Dalloz Actualités*.

finalement des États, mais pourrait être limitée à la valeur de ce patrimoine affecté.

B) Les SLA militaires: ni personne, ni chose

197. Une autre approche intéressante est celle des «centres d'intérêts» théorisée par Gérard Farjat qui dépasse la simple classification personne/chose.²⁶¹ L'auteur avance que la reconnaissance de ces centres d'intérêts, même dépourvus de personnalité juridique, permet de mieux refléter les réalités actuelles et de renforcer l'efficacité du système juridique. La notion de centre d'intérêts présente des fondements et des effets juridiques variés. Elle peut ouvrir droit à une protection juridique spécifique, supérieure à celle accordée aux choses, sans toutefois atteindre celle des personnes. Elle peut également impliquer «une charge particulière, une responsabilité élargie».²⁶² L'attribution de règles, de droits et d'obligations en constitue un aspect central. Bien que seules les personnes juridiques puissent être tenues responsables de la violation de ces règles, la reconnaissance d'un centre d'intérêts légitime leur application et permet d'étendre la responsabilité aux entités juridiques qui en assurent la gestion.

198. Appliquée aux systèmes d'arme de ciblage humain intégrant de PIA plus ou moins autonome, cette perspective permettrait d'envisager des formes spécifiques de responsabilité, en fonction de leur degré d'autonomie et des risques qu'ils présentent, sans les assimiler à des personnes.

²⁶¹ Gérard Farjat, «Entre les personnes et les choses, les centres d'intérêts», RTD Civ. 2002 p.221. Disponible en version PDF.

²⁶² *Ibid.*

199. Par ailleurs, le concept de «non-sujets de droit»²⁶³ du Doyen Carbonnier met en avant la possibilité de déconnecter la personnalité juridique de la qualité de sujet de droit. Ainsi, des entités ne possédant pas la personnalité juridique pourraient néanmoins être titulaires de droits et soumises à des obligations spécifiques. Dans le contexte militaire, cela pourrait se traduire par l'établissement de règles claires concernant le comportement attendu de ces systèmes, à savoir leur respect aux principes du DIH et les conséquences en cas de violation de ce droit, sans pour autant les considérer comme des personnes juridiques à part entière.

C) L'humain inévitablement au cœur de la responsabilité?

1- Une programmation humaine, une responsabilité humaine

200. Il est crucial de ne pas négliger la responsabilité des acteurs humains impliqués dans la conception, la programmation, le déploiement et le contrôle de ces systèmes. L'intégration des principes éthiques et les lois du DIH dans la programmation des IA militaires, ainsi que la mise en place de codes de conduite et de systèmes de contrôle humain effectifs, constituent des moyens essentiels pour prévenir les dommages et attribuer la responsabilité en cas d'incident.

201. À cet égard, les lois d'Asimov, initialement conçues comme un cadre éthique pour guider le comportement des robots dans leurs interactions avec les humains disposent que:

²⁶³ J. Carbonnier, Chapitre 3, «Être ou ne pas être. Sur les traces du non-sujet de droit», *Flexible droit*, L.G.D.J, 9^e ed. 1998.

- «1. Un robot ne peut porter atteinte à un être humain, ni, en restant passif, permettre qu'un être humain soit exposé au danger;
2. Un robot doit obéir aux ordres qui lui sont donnés par un être humain, sauf si de tels ordres entrent en conflit avec la première loi;
- 3.- Un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la première ou la deuxième loi.»

Une quatrième loi, la loi Zéro est induite des trois autres: «0. Un robot ne peut pas faire de mal à l'humanité, ni, par son inaction, permettre que l'humanité soit blessée». ²⁶⁴

S'inspirer de ces lois, «comme s'appliquant aux concepteurs, aux fabricants et aux opérateurs de robots, y compris de robots dotés d'autonomie et de capacités d'auto-apprentissage» ²⁶⁵ plutôt qu'aux robots eux-mêmes, souligne l'importance d'ancrer la responsabilité et le contrôle dans la sphère humaine.

202. De plus, Me Bensoussan dans sa Charte des droits des robots, évoque à l'article 6 que «L'utilisateur du robot est présumé responsable des agissements du robot sauf preuve contraire. La responsabilité des robots est soumise au droit commun de la responsabilité des personnes humaines selon l'ordre suivant:

- L'utilisateur du robot;
- Le fabricant de la plateforme d'intelligence artificielle;
- Les fabricants des composants corporels et incorporels du robot;

²⁶⁴ Isaac Asimov, exposées pour la première fois dans sa nouvelle *Cercle Vicieux* (Runaround), 1942.

²⁶⁵ Résolution du Parlement européen du 16 février 2017, *op.cit.*

- *Le propriétaire du robot.*».²⁶⁶

Cette gradation reflète l'idée selon laquelle les humains, à tous les niveaux de la chaîne, doivent rester responsables.

2- *La responsabilité du fait des robots*

203. Plusieurs auteurs suggèrent d'appliquer aux robots le régime de responsabilité du fait des choses. Ainsi, Patrick Lynn, Keith Abney et Georges Bekey estiment que «*la responsabilité morale pour les actions d'un robot échoit à la personne qui a accordé au robot la permission d'agir en son nom*».²⁶⁷ En ce sens, puisque le robot agit pour le compte d'un tiers, il ne peut être considéré comme responsable en lui-même.²⁶⁸ Dès lors, la théorisation de la responsabilité du fait des robots a été fait à partir de l'article 1243 du Code civil «*le propriétaire d'un robot, ou celui qui s'en sert, pendant qu'il est à son usage, est responsable du dommage que le robot a causé, soit que le robot fût sous sa garde, soit qu'il fût égaré ou échappé*».²⁶⁹

Ainsi, il n'est plus essentiel de proposer une personnalité juridique robotique ou électronique pour attribuer une quelconque responsabilité aux robots, puisque son propriétaire le sera. Dans le domaine militaire, «*le dépositaire de l'autorité militaire peut donc être tenu pour responsable des actes commis par un système autonome*».²⁷⁰

²⁶⁶ Alain Bensoussan, Charte des droits des robots, Publié par Lexing – Alain Bensoussan Avocats, V 5.0, septembre 2015

²⁶⁷ P. Lin, K. Abey, G.A. Bekey, *Robot Ethics: The Ethical and Social Implications of Robotics*, Cambridge, MIT Press, 2014, cite par Laure de Roucy-Rochegonde, «La guerre à l'ère de l'intelligence artificielle», p.138.

²⁶⁸ Laure de Roucy-Rochegonde, *op.cit.* p.138-139.

²⁶⁹ MdN Ruffo, *Itinéraire d'un robot tueur*, Edition Le Pommier/Humensis, 2018, p. 92.

²⁷⁰ Laure de Roucy-Rochegonde, *op.cit.*, p.139.

204. Néanmoins, une limite à ce régime peut être constatée. En effet, si l'autorité peut être reconnue responsable, cela ne le sera pas automatiquement et pas dans tous les cas. En effet, nous revenons au concept de *l'actus reus* et *mens rea*²⁷¹ mais cette fois-ci appliqué à l'autorité militaire en charge du SALA. Nous l'avons vu, le SALA agit en total autonomie, du ciblage, à la décision létale. Dès lors, *«il serait difficile de prouver l'élément intentionnel, qui suppose que l'autorité hiérarchique, ait connaissance des actes illicites avant qu'ils ne soient commis, qu'elle n'agisse pas pour les empêcher, ou qu'elle ne sanctionne pas a posteriori le subordonné à l'origine de ces actes»*.²⁷²

205. Ainsi, si un contrôle humain significatif est établi, c'est-à-dire une compréhension claire des actions du SALA et une connaissance préalable de ses comportements par l'autorité militaire en charge, alors le régime de responsabilité fondé sur le fait du robot apparaît comme une solution pertinente.

206. En conclusion, face aux défis de la responsabilité posés par les systèmes d'IA militaires autonomes, des mécanismes alternatifs à la reconnaissance de la personnalité juridique apparaissent plus prometteurs. Quoi qu'il en soit, *«pour que le droit international humanitaire, qui régit les conflits armés, puisse s'appliquer, il faut que l'usage d'une arme puisse être rattaché à une responsabilité humaine»*.²⁷³ S'agissant spécifiquement des SALA, l'un des onze principes directeurs

²⁷¹ Voir *supra* paragraphe n°180

²⁷² Laure de Roucy-Rochegonde, *op.cit.*, p.139.

²⁷³ Ministère de l'Europe et des affaires étrangères «Systèmes d'armes létales autonomes, quelle est l'action de la France?», février 2020. Disponible sur: [[Le lien](#)]

établis par le GEG souligne d'ailleurs que «*la responsabilité humaine dans les décisions relatives à l'emploi de systèmes d'armes doit être maintenue*».²⁷⁴

²⁷⁴ Assemblée nationale, Rapport d'information n°3248. *Op.cit.* Disponible sur [\[Le lien\]](#).

Chapitre 2: Vers un cadre juridique et éthique pour l'IA militaire

207. La militarisation de l'IA appelle à repenser son encadrement normatif, à travers la création d'un cadre juridique dédié (Section 1) et le rappel des limites incontournables imposées par le droit international humanitaire (Section 2).

Section 1: Un corpus juridique unique spécifique à l'IA militaire

208. L'élaboration d'un cadre juridique propre à l'IA militaire peut s'appuyer sur des principes existants comme la clause de Martens pour formuler un complément des lois de la guerre (§1), tout en nécessitant une rationalisation des normes applicables (§2).

§ 1: La clause de Martens ou un complément des lois de la guerre

209. La clause de Martens a fait sa première apparition dans le préambule de la Convention II de la Haye de 1899. Dans sa formulation originale, la Clause de Martens dispose:

«En attendant qu'un code plus complet des lois de la guerre puisse être édicté, les Hautes Parties contractantes jugent opportun de constater que, dans les cas non compris dans les dispositions réglementaires adoptées par elles, les populations et les belligérants restent sous la sauvegarde et sous l'empire des principes du droit des gens, tels qu'ils résultent des usages

*établis entre nations civilisées, des lois de l'humanité et des exigences de la conscience publique».*²⁷⁵

210. Cette clause est d'une importance majeure, dans la mesure où elle constitue l'«*unique disposition du droit international humanitaire qui établit une base de référence pour la protection des civils et des combattants lorsqu'une situation n'est couverte par aucun texte juridique*».²⁷⁶

211. Au fil des décennies et des conflits armés successifs, la clause de Martens a constamment prouvé son utilité, tout en voyant sa signification évoluer au gré des interprétations qui en ont été faites.

A) Les différentes interprétations de la Clause de Martens

1- Sens restreint

212. Dans son interprétation la plus restreinte, la clause de Martens rappelle que le droit international coutumier demeure applicable,²⁷⁷ notamment lorsque le droit conventionnel laisse certaines questions non codifiées ou ne les abordent pas

²⁷⁵ Convention (II) concernant les lois et coutumes de la guerre sur terre et son Annexe: Règlement concernant les lois et coutumes de la guerre sur terre. La Haye, 29 juillet 1899.

²⁷⁶ Human Rights Watch, «Répondre à l'appel. L'interdiction des robots tueurs, un impératif moral et juridique», 2018. Disponible sur [\[Le lien\]](#).

²⁷⁷ C. Greenwood, "Historical Development and Legal Basis", *The Handbook of Humanitarian Law in Armed Conflict*, Dieter Fleck (éd.), Oxford University Press, Oxford/New York, 1995, p.28 (par.129) cite par Rupert Ticehurst dans son article «La clause de Martens et le droit des conflits armés» publié dans la *Revue Internationale de la Croix-Rouge*, pp. 133-142.

explicitement.²⁷⁸ Cette interprétation est peu convaincante à deux égards. Tout d'abord parce que cette interprétation s'appuie uniquement sur la coutume et en oublie les lois de l'humanité et l'exigence de la conscience publique, rendant ces principes inefficaces.²⁷⁹ Enfin parce que, «*la clause elle-même ne peut pas servir de base pour interdire une arme, car une telle interdiction doit reposer sur une norme coutumière ou conventionnelle existantes*». ²⁸⁰ Étant donné qu'il n'existe pas à ce jour de norme coutumière sur les SALA, leur interdiction ne peut se faire à l'aune de cette Clause.

2- *Interprétation large*

213. Selon son interprétation la plus large, une action n'est pas automatiquement légale simplement parce qu'aucun traité ne l'interdit explicitement.²⁸¹ Cette vision est partagée par des États comme le Royaume-Uni qui, dans ses observations écrites lors d'un avis consultatif de la Cour Internationale de Justice sur la licéité de la menace ou de l'emploi d'armes nucléaires du 8 juillet 1996, que «*la clause de Martens stipule clairement que l'absence d'une interdiction*

²⁷⁸ Marcus Hickleton, "Shaky foundations: "killer robots" and the Martens Clause". *Perth International Law Journal*, vol. 4, janvier 2019, p.44. Disponible sur [Le lien].

²⁷⁹ *Ibid.*

²⁸⁰ T. McFarland, préc., note 101 à la page 104. Cité dans le mémoire de recherche *Les systèmes d'armes autonomes et le droit international humanitaire: Une légalité contestée*, p.45, par Julien Grenon, Université de Montréal, juillet 2022.

²⁸¹ Y. Sandoz, C. Swinarski, B. Zimmermann (éd.), *Commentaire des Protocoles additionnels du 8 juin 1977 aux Conventions de Genève du 12 août 1949*, CICR/Martinus Nijhoff, Genève, 1986, pp.38-39 (par.55) ; N. Singh et E. McWhinney, *Nuclear Weapons and Contemporary International Law*, 2^e éd. Martinus Nijhoff, Dordrecht, 1989, pp. 46-47 cité par Rupert Ticehurst dans son article "La clause de Martens et le droit des conflits armés", publié dans la *Revue Internationale de la Croix-Rouge*, pp. 133-142.

*conventionnelle spécifique sur l'emploi des armes nucléaires ne signifie pas pour autant que leur usage soit autorisé.*²⁸² La clause imposerait donc un seuil minimal de protection pour les civils et les combattants, fondé non seulement sur les règles écrites mais aussi sur des obligations morales et humanitaires implicites, destinées à éviter des conséquences inacceptables du point de vue de l'humanité.²⁸³

3- *Sens moderne*

214. L'interprétation modérée de la clause de Martens, interprétation généralement retenue par le droit international,²⁸⁴ considère que cette disposition ne constitue pas une source autonome de droit, mais qu'elle joue un rôle important en tant qu'outil d'interprétation des normes existantes du droit international humanitaire.²⁸⁵ Dans les situations où les règles conventionnelles ou coutumières sont ambiguës, incomplètes ou sujettes à plusieurs interprétations, la clause invite les juges à privilégier la lecture la plus conforme aux principes d'humanité et

²⁸² Lettre du 16 juin 1995 du conseiller juridique du Foreign and Commonwealth Office du Royaume-Uni de Grande-Bretagne et d'Irlande du nord, accompagnée des observations écrites du Gouvernement du Royaume-Unis P.48, disponible sur: [**Le lien**].

²⁸³ CCW, working paper submitted by Switzerland, "A "Compliance-Based" Approach to Autonomous Weapon Systems", CCW/GGE.1/2017/WP.9, ONU, 10 novembre 2017, paragraphe 18. Disponible en version PDF sur [**Le lien**].

²⁸⁴ Marcus Hickleton, "Shaky foundations: "killer robots" and the Martens Clause". p.48, *op. cit.*

²⁸⁵ Marcus Hickleton, "Shaky foundations: "killer robots" and the Martens Clause". p.45, *op. cit.*

aux exigences de la conscience publique,²⁸⁶ entendues comme «*la conception qu'une population a du bien et du mal*».²⁸⁷ Elle ne permet donc pas d'interdire à elle seule une arme ou une conduite en l'absence de réglementation claire, mais peut renforcer ou orienter l'application de règles déjà existantes en cas de doute sur leur interprétation.

De plus, les exigences de la conscience publique ne permettent pas à elles seules d'interdire les SALA puisqu'il est difficile de «*déterminer l'opinion des citoyens dans tous les pays du monde*»²⁸⁸ et surtout puisqu'elles sont fluctuantes en fonction du contexte. Sur ce point, une enquête menée par Michael Horowitz montre que «*les citoyens interrogés se révélaient également beaucoup plus enclins à soutenir le recours aux armes autonomes lorsqu'il était avancé que leur utilisation permettrait de protéger les forces américaines*».²⁸⁹

215. Dès lors, il apparaît peu probable que la Clause de Martens puisse interdire l'utilisation des SALA à l'aune des exigences de la conscience publique.

B) *Un nouveau cadre juridique exclusif*

216. Même si la clause de Martens ne crée pas directement de nouvelles lois, elle est précieuse pour interpréter et faire évoluer le

²⁸⁶ Antonio Cassese, "The Martens Clause: Half a Loaf or simply Pie in the Sky?" *European Journal of International Law*, Volume 11, Issue 1, 2000, Pages 187–216, p. 212. Disponible en version PDF sur [Le lien].

²⁸⁷ Laure de Roucy-Rochegonde, p.191. *Op. cit.*

²⁸⁸ Laure de Roucy-Rochegonde, p.192. *Op. cit.*

²⁸⁹ M.C. Horowitz, "Public Opinion and the Politics of the Killer Robots Debate", *Research & Politics*, 3/1, 2016, p.1-8. Cité par Laure de Roucy-Rochegonde, *op.cit.*, p.193.

droit international humanitaire, surtout dans des domaines où les nouvelles technologies posent des défis éthiques ou juridiques.²⁹⁰

217. Ainsi, au-delà de l'interprétation juridique de la clause de Martens, il convient de s'attarder sur sa première phrase: «*en attendant qu'un code plus complet des lois de la guerre puisse être édicté [...]*». Cette mention suggère que la clause n'a jamais eu vocation à constituer un socle définitif, mais une transition en l'absence de normes plus précises. Par conséquent, au lieu de se contenter de l'appliquer par extension aux systèmes d'armes autonomes intégrant de l'intelligence artificielle, létaux ou non, il paraît aujourd'hui nécessaire de s'atteler à la rédaction d'un véritable corpus juridique adapté aux réalités technologiques contemporaines. En effet, il semble peu pertinent de continuer à invoquer une disposition issue de la Conférence de La Haye de 1899 pour encadrer les usages d'outils technologiques aussi complexes et récents que les systèmes d'armes autonomes intégrant de l'IA.

218. A la lumière de cette réflexion, une alternative consisterait à élaborer un cadre juridique autonome, spécifiquement conçu pour réguler l'emploi militaire de l'intelligence artificielle et des SALA.

²⁹⁰ Joanne Kirkham, *Les systèmes d'armes létaux autonomes: une étude de droit international*, thèse de doctorat en droit public soutenue le 10 décembre 2022, École doctorale de droit international droit européen, relations internationales et droit comparé, Université Paris Panthéon-Assas. Paragraphe 99. *Op.cit.* Disponible sur le [Le lien].

219. Dans ce cadre, une approche pertinente consisterait à adapter au domaine militaire certaines régulations civiles existantes, notamment celles issues du droit européen. Ainsi, le règlement européen sur l'intelligence artificielle du 13 juin 2024 qui ne s'applique pas au domaine militaire,²⁹¹ pourrait néanmoins constituer une base de travail. Bien qu'il ne s'agisse pas ici de transposer mécaniquement une régulation civile dans un contexte de conflit armé, les exigences prévues pour les systèmes à «haut risque» pourraient inspirer un encadrement rigoureux des systèmes d'IA militaires. Cette démarche permettrait un compromis entre l'innovation technologique et le respect des principes fondamentaux du droit international humanitaire, sans nécessairement aller jusqu'à l'interdiction pure et simple des SALA.

220. Parmi les exigences clés du règlement européen, figure notamment à l'article 4 du règlement, la «maîtrise de l'IA», qui impose aux fournisseurs et utilisateurs de systèmes à haut risque de garantir un niveau de compétence suffisant à leur personnel, en fonction de leur formation, et du contexte opérationnel.²⁹² Une telle exigence répondrait aux attentes exprimées par le Groupe d'Experts Gouvernementaux réunis dans le cadre de la CCAC, notamment en matière d'explicabilité des systèmes. L'article 14 du rè-

²⁹¹ Bensamoun, Alexandra et Loiseau, Grégoire, *Droit de l'intelligence artificielle*, 2e édition. - Paris La Défense : LGDJ, un savoir-faire de Lextenso, DL 2022. - 1 vol. (594 p.).

²⁹² 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) (Texte présentant de l'intérêt pour l'EEE). Disponible sur [**Le lien**].

glement prévoit également une surveillance humaine accrue, rejoignant les principes défendus dans le projet de résolution du Conseil de l'Europe, selon lesquels les SALA doivent rester prévisibles afin de satisfaire au principe de précaution.²⁹³ Ces principes permettraient également de rendre l'humain réellement responsable.

221. D'autres dispositions du RIA telles que la gestion des risques (article 9), la documentation technique en faveur de l'explicabilité (article 11), ou encore les garanties de précision, de robustesse et de cybersécurité (article 15), pourraient aussi être transposées, ou à tout le moins adaptées, aux spécificités des systèmes militaires. L'objectif serait alors de garantir un contrôle humain significatif tout en autorisant un usage encadré, maîtrisé et conforme aux impératifs opérationnels contemporains.

222. Une telle proposition pourrait constituer un point de convergence entre les États partisans d'un encadrement juridique strict, sans pour autant interdire le recours aux SALA, et ceux qui défendent leur développement dans un cadre éthique et responsable. En conciliant exigence juridique et flexibilité technologique, cette approche offrirait une réponse pragmatique aux défis posés par l'intégration croissante de l'IA dans les conflits armés.

²⁹³ Conseil de l'Europe, Commission des questions juridiques et des droits de l'homme, «L'Émergence des systèmes d'armes létales autonomes (SALA) et leur nécessaire appréhension par le droit européen des droits humains», Projet de résolution adopté à l'unanimité par la Commission le 14 novembre 2022. Disponible en version PDF sur [\[Le lien\]](#).

§ 2 Les moyens de rationaliser

A) Les instruments existants de régulation des SALA

223. La régulation des systèmes d'armes létales autonomes repose actuellement sur une pluralité d'initiatives institutionnelles et normatives. Les premières alertes ont été lancées par les concepteurs eux-mêmes, avant d'être relayées par des organisations non gouvernementales telles qu'Amnesty International, appelant à leur interdiction. Sur le plan étatique, des démarches ont également été entreprises pour porter cette question à l'agenda international, notamment par la France dès 2013 lors de la Conférence du désarmement. Le cadre principal demeure la Convention sur certaines armes classiques, décrite comme «*l'instrument clé du droit international humanitaire*»²⁹⁴ en raison de sa vocation à «*établir un équilibre entre les nécessités militaires et les considérations humanitaires*». ²⁹⁵ À l'occasion d'un webinaire organisé par le Bureau des affaires de désarmement des Nations Unies, il a été rappelé, dans les remarques liminaires de Mme Lindholdt Hove Rietveld, que l'Union européenne a décidé en 2021 de renforcer et de soutenir la Convention sur certaines armes classiques. Il y a également été affirmé que le Groupe d'Experts Gouvernementaux constitue, selon cette perspective, le cadre le plus approprié pour mener les discussions relatives aux systèmes d'armes létales autonomes.²⁹⁶

²⁹⁴ GGE SALA - Déclaration générale: Intervention de l'Ambassadrice Camille Petit (Genève, 26 août 2024). Disponible sur [[Le lien](#)].

²⁹⁵ *Ibid.*

²⁹⁶ Webinar of United Nations, Office for Disarmament Affairs, "Experts Reflect on Novel Approaches to "Meaningful Human Control", 7 mars 2023. Disponible sur : [[Le lien](#)].

224. Cette convention, par sa structure souple, permettrait une certaine adaptabilité aux «technologies émergentes dans le domaine des SALA».²⁹⁷ En son sein, le Groupe d'Experts Gouvernementaux a adopté en 2019 onze principes directeurs réaffirmant la nécessité du respect du droit international humanitaire et de la responsabilité humaine. Il a également été souligné que «*le concept de contrôle humain significatif est une question centrale*»²⁹⁸ dans ces discussions. Toutefois, ces principes ne revêtent pas de valeur contraignante, ce qui limite leur portée normative.²⁹⁹ Par ailleurs, certains États comme la Chine ont exprimé leur soutien à un instrument juridique, à condition qu'il soit «*uniquement au sein de la Convention sur certaines armes classiques*»,³⁰⁰ ce qui illustre la persistance d'une approche conservatrice au sein de cette enceinte.

B) Une efficacité remise en question

225. Malgré l'existence de ces instruments, leur efficacité demeure limitée face à l'accélération du développement des SALA. La CCAC, en raison de son mode de fonctionnement basé sur le consensus, se heurte à de nombreux blocages.³⁰¹ Il est ainsi observé que «face au blocage de la CCAC, la question des armes autonomes

²⁹⁷ Nathalia A. Smuha, *The Cambridge Handbook of the law, ethics and policy of artificial intelligence*, Cambridge University Press, Avril 2025, Chapitre 20 "Artificial Intelligence and Armed Conflicts", Katerina Yordanova, pp. 411-428.

²⁹⁸ *Ibid.*

²⁹⁹ *Ibid.*

³⁰⁰ Nations Unies, «Première Commission: les pays soulignent l'importance d'une prise en compte des dimensions éthique et humanitaire de certaines armes classiques», AG/DSI/3748 24 octobre 2024. Disponible sur [\[Le lien\]](#).

³⁰¹ Laure de Roucy-Rochegonde, «Usage militaire et éthique de l'IA & Droit international», Cours de M2 Droit de l'intelligence artificielle

est discutée dans d'autres enceintes des Nations Unies»,³⁰² notamment à l'Assemblée générale, qui a adopté en 2023³⁰³ et 2024³⁰⁴ des résolutions largement soutenues par les États. Ces résolutions, bien qu'elles ne lancent pas encore de négociations formelles, ont permis de déplacer le débat dans un cadre où «les décisions se prennent à la majorité»,³⁰⁵ permettant ainsi d'éviter les obstructions systématiques. En décembre 2024, une résolution a été adoptée à 166 voix, prévoyant des discussions informelles en 2025³⁰⁶ et «*décide d'inscrire à l'ordre du jour provisoire de sa quatre-vingtième session la question intitulée «Système d'armes létaux autonomes»*».³⁰⁷

226. À la lumière de précédents historiques, comme les traités d'Ottawa de 1997 et celui d'Oslo en 2008, ayant permis de réduire voire d'interdire significativement la prolifération des mines antipersonnel et des armes à sous-munitions,³⁰⁸ un nouveau traité sur les SALA serait «*l'unique moyen réaliste d'aboutir à l'adoption de normes juridiques solides dans un avenir proche*».³⁰⁹ A ce titre, le Comité

³⁰² Aymeric Elluin, «L'urgence d'un traité encadrant les systèmes d'armes autonomes», 13 janvier 2025. Disponible sur [[Le lien](#)].

³⁰³ Assemblée Générale des Nations Unies, «Résolution adoptée par l'Assemblée générale le 22 décembre 2023», A/RES/78/241. Disponible en version PDF sur [[Le lien](#)].

³⁰⁴ Assemblée Générale des Nations Unies, «Résolution adoptée par l'Assemblée générale le 2 décembre 2024», A/RES/79/62. Disponible en version PDF sur [[Le lien](#)].

³⁰⁵ Aymeric Elluin, «L'urgence d'un traité encadrant les systèmes d'armes autonomes», *op.cit.*

³⁰⁶ Assemblée Générale des Nations Unies, 2 décembre 2024. *Op.cit.* p.3, par. 7.

³⁰⁷ *Ibid.*, p.4, par. 12.

³⁰⁸ Aymeric Elluin, «L'urgence d'un traité encadrant les systèmes d'armes autonomes» *op.cit.*

³⁰⁹ *Ibid.*

International de la Croix-Rouge et le Secrétaire Général des Nations Unies appellent à «*négozier avant 2026 de nouveaux accords sur ces armes*».³¹⁰

Section 2: Les limites absolues du DIH

227. Toutefois, bien qu'il existe des mécanismes pour adapter le DIH à l'ère de l'IA militaire, les limites du droit international humanitaire tiennent autant aux impératifs stratégiques et de souveraineté des États (§1) qu'à leur volonté politique de s'y conformer ou non (§2).

§ 1: Les enjeux stratégiques et de souveraineté

A) L'IA militaire comme levier de supériorité technologique et dissuasive

228. L'intelligence artificielle appliquée aux technologies militaires s'impose progressivement comme un facteur clé de supériorité stratégique. Intégrée aux SALA, l'IA permet non seulement de renforcer la précision et la rapidité des opérations,³¹¹ tout en compensant certaines contraintes structurelles, telles que la réduction des effectifs militaires dû à une difficulté de recrutement et «*la réticence de plus en plus grande des nations occidentales à risquer la vie*

³¹⁰ Nations Unies, «L'ONU et la Croix-Rouge appellent à un accord international inédit sur les armes autonomes», 5 octobre 2023. Disponible sur [**Le lien**].

³¹¹ Bibliothèque du Parlement du Canada, Ariel Shapiro, «Systèmes d'armes autonomes: quelques répercussions sur la sécurité internationale et sur le Canada», 20/12/2019. Disponible sur [**Le lien**].

de ses soldats sur des théâtres d'opérations extérieures». ³¹² Le développement des SALA suscite ainsi un intérêt militaire soutenu, à tel point que certains observateurs y voient les prémices d'une nouvelle révolution des techniques de la guerre. ³¹³ Dans un contexte où l'évolution accélérée des technologies permet la conception de systèmes pouvant suppléer, voire remplacer, l'humain au combat, les «robots tueurs» apparaissent comme une réponse pragmatique aux exigences de modernisation. ³¹⁴ À ce titre, un collectif de spécialistes influents en intelligence artificielle et en robotique avait déjà alerté en 2015 sur le fait que «*la technologie de l'intelligence artificielle a atteint un niveau tel que le déploiement de ces systèmes est – pratiquement, sinon légalement – envisageable dès les prochaines années, et non les prochaines décennies*», ³¹⁵ soulignant que ces armes pourraient constituer «*la troisième révolution des techniques de guerre, après la poudre à canon et les armes nucléaires*». ³¹⁶ L'IA militaire ne se réduit donc pas à un simple outil d'efficacité: elle devient un levier de puissance, un facteur de dissuasion, un symbole d'avance technologique, et «*se présente comme la voie principale de la supériorité tactique*». ³¹⁷ Les États qui choisiraient d'écarter, par principe, le développement de ces technologies, s'exposeraient dès lors à un

³¹² Dominique Lambert, p.38, *op.cit.*

³¹³ Laure de Roucy-Rochegonde, «Systèmes d'armes létales autonomes: y aura-t-il un Terminator tricolore ?», 225/09/2020. Disponible sur [Le lien].

³¹⁴ *Ibid.*

³¹⁵ Autonomous Weapons Open Letter: AI & Robotics Researchers, 9 février 2016. Disponible sur [Le lien].

³¹⁶ *Ibid.*

³¹⁷ Jean-Christophe Noël, «Intelligence artificielle: vers une nouvelle révolution militaire?», *op.cit.*

déclassement stratégique³¹⁸ et ce refus pourrait être interprété comme une vulnérabilité, à l'heure où la maîtrise de l'IA militaire est progressivement intégrée dans les doctrines de sécurité nationale de plusieurs puissances.

B) *Tensions entre refus d'ingérence juridique et respect des obligations du DIH*

229. En devenant parties aux conventions de Genève, l'article 1^{er} commun à toutes les Conventions de Genève dispose que «*Les Hautes Parties contractantes s'engagent à respecter et à faire respecter la présente Convention en toutes circonstances*».³¹⁹ Cette obligation découle d'une règle fondamentale du droit international, *pacta sunt servanda*³²⁰ prévu par l'article 26 de la Convention de Vienne du 23 mai 1969, et implique d'appliquer de bonne foi leurs engagements conventionnels. Elle ne saurait être vidée de son sens par des réserves en faveur de la sécurité de l'État ou des nécessités militaires.³²¹ D'autre part, malgré un appel conjoint du Secrétaire général des Nations Unies et du CIRC exhortant «*les dirigeants politiques à «établir d'urgence des interdictions et des limitations spécifiques pour les systèmes d'armes autonomes, afin de protéger l'humanité*»³²² avant 2026, les principales puissances militaires telles que les Etats-Unis, la Chine ou encore la Russie ont indiqué qu'elles se réservaient le

³¹⁸ Laure de Roucy-Rochegonde, «Systèmes d'armes létales autonomes: y aura-t-il un Terminator tricolore?» *op.cit.*

³¹⁹ *Convention de Genève pour l'amélioration du sort des blessés et des malades dans les forces armées en campagne du 12 août 1949*, I, Article 1.

³²⁰ *Convention de Vienne sur le droit des traités*, Vienne, 23 mai 1969. Article 26.

³²¹ Abdelwahad Biad, *Droit international humanitaire*, ed. Ellipses, coll. «Mise au point», 2^{ème} édition, 2006, p. 82.

³²² Nations Unies, «L'ONU et la Croix-Rouge appellent à un accord international inédit sur les armes autonomes», 5 octobre 2023, *op.cit.*

droit de développer de tels systèmes. Ces prises de position traduisent en réalité une absence d'intention, de la part de ces acteurs dominants, de s'engager dans un cadre juridique contraignant limitant l'usage de l'IA militaire ou de ralentir leurs investissements dans ce domaine stratégique.³²³ Toutefois, cette réalité entre en tension avec les initiatives de l'ONU et du CICR et s'explique notamment par la stratégie militaire, fondée sur la logique de rivalité géopolitique et la dynamique concurrentielle qui caractérise le système international actuel.³²⁴ En définitive, l'effectivité du droit international humanitaire se heurte aux logiques de puissance et aux impératifs de souveraineté technologique, révélant une tension structurelle entre les obligations juridiques des États et les dynamiques géopolitiques contemporaines fondées sur la compétition stratégique.

§ 2. *Volonté des États de respecter le DIH ou non*

A) *Les facteurs déterminants du respect du DIH par les États*
230. Hormis les obligations conventionnelles à respecter par les États, la promotion du respect du DIH peut reposer sur d'autres mécanismes tels que la transparence, la redevabilité et la pression internationale. Le respect du principe de transparence dans la conduite des opérations militaires, ainsi que la communication d'informations pertinentes, renforcent la légitimité des acteurs armés souhaitant être perçus comme des interlocuteurs ou

³²³ Thierry Berthier, «L'ONU veut réguler l'intelligence artificielle militaire mais se heurte au principe de réalité concurrentielle», *Revue Conflits*, publié le 13 août 2023. Disponible sur [[Le lien](#)].

³²⁴ Thierry Berthier, *Revue Conflits*, *op.cit.*

représentants légitimes. Dans cette perspective, le respect du droit des conflits armés devient un levier stratégique pour asseoir cette reconnaissance.³²⁵ La redevabilité, quant à elle, implique que les violations du DIH soient effectivement sanctionnées. En effet, «*les violations graves du DIH, sont constitutives de crime international, imprescriptible et inamnistiable, donnant lieu à réparation et à répression*»³²⁶. Dès lors les sanctions dû aux violations du DIH sont, a minima, un moyen de le faire respecter.

231. Par ailleurs, l'influence de l'opinion publique nationale et internationale, joue un rôle crucial dans l'évaluation de la conduite des États. Les États se soucient de leur image et de leurs choix politiques aux yeux de la communauté internationale et de leur propre population, ce qui peut les pousser à agir conformément au DIH pour «*gagner le soutien de la population*».³²⁷

En somme, ces mécanismes interconnectés constituent des leviers essentiels pour encourager les États à respecter le DIH, en renforçant la confiance, en assurant la justice et en mobilisant la communauté internationale autour des principes humanitaires fondamentaux.

³²⁵ Madalena Vasconcelos Rosa et Juliane Garcia Ravel, «DIH en action: sept facteurs permettant d'influencer les comportements dans l'objectif de renforcer le respect du droit international humanitaire», 19 novembre 2020. Disponible sur [[Le lien](#)].

³²⁶ David Cumin, *Manuel de Droit de la Guerre*, p.160, *op.cit.*

³²⁷ Madalena Vasconcelos Rosa et Juliane Garcia Ravel, *op.cit.*

B) *Le respect sélectif du DIH selon les intérêts politiques et opérationnels*

232. À l'inverse, le respect du droit international humanitaire est souvent loin d'être absolu et peut être sélectivement appliqué ou interprété en fonction des intérêts politiques et opérationnels des États et des belligérants.³²⁸ Les belligérants tendent parfois à privilégier une logique de transactions ou à invoquer l'exceptionnalisme plutôt que la stricte conformité aux règles, déclarant respecter le droit uniquement si la partie adverse fait de même, ou affirmant que l'adversaire est si nuisible que les dispositions limitant l'usage de la force peuvent être ignorées.³²⁹ Cette tendance se manifeste également face aux nouvelles technologies de guerre, comme le débat autour des systèmes d'armes létales autonomes l'illustre. Pour qu'une régulation internationale des SALA soit efficace, elle nécessite l'adhésion de toutes les puissances militaires concernées.³³⁰ Cependant, comme l'expérience l'a montré avec d'autres tentatives de régulation, telles que les conventions d'Ottawa sur les mines antipersonnel ou d'Oslo sur les armes à sous-munitions, les principales puissances

³²⁸ Duncan McLean et Michiel Hofman, «Droit international humanitaire, souveraineté des États et érosion du consensus humanitaire: la fin de l'humanitarisme?», publié dans Numéro 23 – Droit international humanitaire: le grand retour... en arrière? *Revue Alternatives Humanitaires*. Traduit de l'anglais par Sophie Jeangeorges. Disponible sur [\[Le lien\]](#).

³²⁹ Allocution de Peter Maurer, Président du CICR, Centre d'études stratégiques et internationales, Washington DC, 10 mai 2019. Traduction de la transcription. «Les règles de la guerre appartiennent-elles au passé?», 10/05/2019. Disponible sur [\[Le lien\]](#).

³³⁰ Assemblée nationale, rapport d'information n°3248. *Op.cit.* Disponible sur [\[Le lien\]](#).

militaires concernées n'y ont souvent pas adhéré.³³¹ Cela souligne comment les intérêts nationaux et militaires peuvent primer sur les efforts de régulation universelle.³³² De plus, une interdiction préventive visant uniquement les États pourrait ne pas empêcher les groupes terroristes, qui ne sont pas liés par de telles obligations, de recourir aux SALA, ce qui démotiverait les puissances militaires à s'engager dans un processus réglementaire.³³³ Ces dynamiques reflètent les défis rencontrés pour garantir la pertinence et le respect du DIH dans un monde où la compétition entre puissances s'accroît et où les conflits sont de plus en plus fragmentés et complexes.³³⁴

233. Conclusion du Titre 2. Finalement, l'intégration de l'IA militaire, et en particulier les SALA, met à rude épreuve le DIH en raison de l'érosion du contrôle et de la responsabilité humaine. Si des pistes juridiques alternatives à la personnalité juridique pour les machines existent pour encadrer la responsabilité, et si un nouveau cadre normatif inspiré des régulations civiles est envisageable, la compatibilité réelle de ces systèmes avec le DIH dépendra *in fine* de la capacité des États à maintenir un contrôle humain significatif, à garantir la responsabilité humaine en cas de violation, malgré les défis techniques, éthiques et les fortes pressions géopolitiques et surtout leur capacité à vouloir à respecter le DIH.

³³¹ *Ibid.*

³³² *Ibid.*

³³³ *Ibid.*

³³⁴ Allocution de Peter Maurer. *Op.cit.*

Conclusion

234. L'essor de l'Intelligence Artificielle dans le domaine militaire, en particulier pour le ciblage humain, soulève de profonds défis techniques, juridiques et géopolitiques dans le cadre du Droit International Humanitaire. La question centrale est de savoir si ce cadre, issu d'une tradition ancienne, peut encore encadrer efficacement ces nouvelles technologies. Dans un contexte de course à l'armement, les divergences entre États freinent l'émergence d'un consensus sur la régulation. Face à ces limites, l'adaptation du DIH semble inévitable, qu'elle passe par un cadre juridique propre à l'IA militaire ou par l'ajustement de normes existantes. À défaut d'accord rapide au sein de la CCAC, des alternatives comme un traité contraignant sur les SALA ou des mécanismes de transparence et de responsabilisation apparaissent essentiels pour préserver les principes humanitaires fondamentaux.

235. En fin de compte, la compatibilité des SIA de ciblage humain avec le DIH n'est pas acquise et dépendra de la capacité de la communauté internationale à adapter le cadre normatif. Les limites techniques de l'IA actuelle, l'érosion du contrôle humain induite par l'accélération des opérations et les biais cognitifs, ainsi que la difficulté à attribuer la responsabilité juridique sont des obstacles majeurs au respect des principes fondamentaux du DIH. Bien que certains pensent qu'il est possible de programmer les SIA conformément au DIH, l'incapacité de l'IA à réaliser des jugements éthiques et contextuels rend cette tâche extrêmement délicate, voire impossible pour une autonomie complète. Néanmoins, la nécessité de maintenir un contrôle humain effectif sur les décisions

létales et de garantir la responsabilité en cas de violation du DIH est une ligne rouge éthique et juridique que les États doivent impérativement maintenir. L'avenir du DIH à l'ère de l'IA militaire repose sur un équilibre délicat entre l'innovation technologique et l'impératif de protéger l'humain dans les conflits armés. La conformité des SIA au DIH est incertaine, mais non impossible si les États œuvrent collectivement à l'adaptation du droit et à la mise en place d'un cadre éthique solide, en garantissant que l'humain reste maître de la décision létale.

236. Bien que la question de la régulation des SALA ne soit pas nouvelle, elle prend aujourd'hui une dimension accrue avec l'émergence de nouvelles formes d'IA. Alors que certains soutenaient jusqu'alors l'inexistence concrète des SALA, l'apparition de l'IA agentic dans laquelle, *«l'intelligence artificielle passe du statut d'assistant à celui d'acteur. Elle devient capable de réfléchir, de planifier et d'agir pour atteindre un objectif»*,³³⁵ remet en question cette perspective. Plus qu'une adaptation du DIH, il faudra certainement repenser les codes de la guerre, non pas pour les humains mais pour les IA militaires.

³³⁵ Julien Perez, professeur au Laboratoire de Recherche de l'EPITA (école pour l'informatique et les techniques avancées). Actualités EPITA, «Les agents IA, un tournant majeur pour l'intelligence artificielle», 05/05/2025. Disponible sur [[Le lien](#)].

Postface

Rédigée le 27 mai 2026

La publication récente de la lettre encyclique *Magnifica Humanitas*³³⁶ du pape Léon XIV, et plus particulièrement ses développements concernant les *Armes et l'IA*, offre un éclairage précieux sur les enjeux éthiques liés à l'essor de l'intelligence artificielle dans le domaine militaire. Parue un an après la rédaction de ce mémoire, elle en confirme la pertinence et montre que les questions juridiques et éthiques soulevées par l'usage de l'IA en conflit armé demeurent au cœur des préoccupations internationales. Cette prise de position du Saint-Siège prolonge la réflexion entreprise dans ce mémoire, en rappelant que les technologies ne sont jamais neutres et que leur intégration dans la conduite de la guerre traduit une certaine vision de l'humanité. En cela, l'encyclique constitue un repère moral essentiel dans un débat où les considérations techniques et stratégiques tendent parfois à occulter la dimension profondément humaine de l'usage militaire de l'IA, notamment dans les systèmes de ciblage humain.

L'analyse développée dans ce mémoire rejoint largement les préoccupations exprimées par le Pape Léon XIV, notamment quant aux risques d'érosion du contrôle humain, de

³³⁶ LETTRE ENCYCLIQUE MAGNIFICA HUMANITAS DU SAINT-PÈRE LÉON XIV SUR LA PROTECTION DE LA PERSONNE HUMAINE À L'ÈRE DE L'INTELLIGENCE ARTIFICIELLE, Rome, près de Saint-Pierre, le 15 mai de l'année 2026. Disponible sur [\[Le lien\]](#)

déréponsabilisation, et de déshumanisation du conflit armé. L'encyclique souligne que l'IA peut rendre la guerre «plus accessible» et abaisser «le seuil du recours à la violence», en transformant les victimes en «simples données». Cette critique rejoint les constats formulés dans ce mémoire sur les limites possibles des systèmes d'armes autonomes et des outils de ciblage fondés sur l'IA, dont l'opacité, la vitesse d'exécution et la logique d'optimisation peuvent fragiliser les principes fondamentaux du DIH, en particulier les principes de distinction, de précaution et de proportionnalité. Ainsi, l'encyclique et ce mémoire s'accordent pour dénoncer les dangers d'une automatisation qui risque d'effacer la dimension morale du jugement humain.

Face à ces dérives, la lettre du Saint-Père insiste sur la nécessité de préserver un contrôle humain effectif, conscient et responsable, en refusant toute délégation de la décision létale à des processus opaques ou automatisés. Cette exigence fait écho aux propositions développées dans ce mémoire, articulées autour du contrôle humain significatif et de critères tels que la transparence, la traçabilité, la robustesse ou encore la supervision humaine, eux-mêmes inspirés du cadre de l'AI Act et transposés aux réalités militaires. L'humain doit demeurer l'acteur responsable, celui qui décide en dernier ressort et protège la dignité de l'autre, même au cœur du conflit. L'IA peut éclairer et assister, mais elle ne peut ni discerner ni porter la charge morale de la violence armée.

En définitive, la rencontre entre *Magnifica Humanitas* et ce mémoire révèle une conviction commune: l'humain doit rester au cœur de toute décision. Dans ce travail de recherche, cette exigence se traduit par l'adaptation du droit international humanitaire à l'usage de l'IA en contexte de conflit armé. Le DIH peut encore répondre

aux défis posés par ces technologies, à condition d'être renforcé, précisé et réaffirmé. L'un et l'autre appellent ainsi à une vigilance éthique et juridique constante, afin que la puissance des machines ne supplante jamais la conscience humaine. *Magnifica Humanitas* vient dès lors conforter la conclusion de ce mémoire: même face à l'automatisation croissante de la guerre, la responsabilité et le jugement moral doivent demeurer fondamentalement humains.

Bibliographie

I. Ouvrages généraux: traités, manuels, et cours

BENSAMOUN, Alexandra et LOISEAU, Grégoire, *Droit de l'intelligence artificielle*, 2^{ème} édition, - Paris La Défense: LGDJ, un savoir-faire de Lextenso, DL 2022. - 1 vol. (594 p.)

CUMIN, David, *Manuel du droit de la guerre*, Bruxelles: Bruylant, Collection Paradigme, 2^{ème} ed., 2020, 550 p.

DE ROUCY-ROCHEGONDE, Laure, «Usage militaire et éthique de l'IA et Droit international,» Cours du Master 2 Droit de l'IA. Année 2024-2025.

DECHAUX, Raphaël, «Réglementation éthique de l'IA dans les organisations internationales,» Cours de M1 Droit de l'intelligence artificielle. Année 2023-2024.

GALBOIS-LEHALLE, Diane, «La notion de personne,» Cours de M1 Droit de l'intelligence artificielle. Année 2023-2024

LEHUGER, Auguste, «Applications récentes de l'apprentissage machine,» Cours de M2 Droit de l'intelligence artificielle. Année 2024-2025.

MINISTÈRE DES ARMÉES, *Manuel de droit des opérations militaires*, édition 2022. Disponible sur [[Le lien](#)].

RUFFO, Marie des Neiges, «Éthique militaire et IA», Cours du Master 2 Droit de l'IA. Année 2024-2025.

STANDFORD ENCYCLOPEDIA OF PHILOSOPHY, *Agency*, First published Mon Aug 10, 2015; substantive revision Mon Oct 28, 2019. Disponible sur [[Le lien](#)].

TABOUY, Laure, «Transhumanisme», Cours de M2 Droit de l'intelligence artificielle. Année 2024-2025.

UZAN, Pierre, «Philosophie de l'esprit», Cours du Master 2 Droit de l'IA. Année 2024-2025.

II. Thèses

DE LA PANOUSE – DE FRANCE, Ysens, *Le droit à l'épreuve de l'automatisation des systèmes militaires robotisés terrestres*, 2019. HAL Id: tel-03823506. Disponible sur [[Le lien](#)].

KIRKHAM, Joanne, *Les systèmes d'armes létaux autonomes: une étude de droit international*, thèse de doctorat en droit public soutenue le 10 décembre 2022, École doctorale de droit international droit européen, relations internationales et droit comparé, Université Paris Panthéon-Assas soutenue le 10 décembre 2022. Disponible sur le [[Le lien](#)].

III. Ouvrages spécialisés

A SMUHA, Nathalie, *The Cambridge Handbook of the law, ethics and policy of artificial intelligence*, Cambridge University Press, Avril 2025

ASIMOV, Isaac, *Cercle Vicioux*, (Runaround), 1942

BIAD, Abdelwahad, *Droit internationale humanitaire*, édition Ellipses, coll. «Mise au point», 2^{ème} édition, 2006

DE ROUCY-ROCHEGONDE, Laure, *La guerre à l'ère de l'intelligence artificielle, quand les machines prennent les armes*, aux presses universitaires de France, édition PUF, 1^{ère} édition, édition 2024

DUNANT, Heny, *Un souvenir de Solférino*, 1962

ELLUL, Jacques, *La technique ou l'enjeu du siècle*, 1954, ed. Economica, 1990, coll. classiques des sciences sociales dirigée par Hervé Coutau-Bégarie

GARÉ, Thierry et RAYNAUD, Anaïs, *Droit des personnes et de la famille*, éd Spécial Droit, Col. Ellipses, 2^{ème} édition, 2018

ISMARD, Paul, *La cité et ses esclaves*, Institutions, fictions, expériences, Seuil, 2019

LAMBERT, Dominique, *La robotique et l'intelligence artificielle*,
Collection Que penser de...? Namur, Paris: Fidélité: Éditions
Jésuites, DL 2019. - 1 vol. (138 p.)

LE MAHABHARATA, *Mort de Bhûrisravas*, cité par
CHALIAND, Gérard, *Anthologie mondiale de la stratégie, des
origines au nucléaire*. 1990, Paris, Laffont

RICOEUR, Paul, *Soi-même comme un autre*, éditions du Seuil, 1990

RUFFO, Marie des Neiges, *Itinéraire d'un robot tueur*, Edition Le
Pommier/Humensis, 2018

SCHARRE, Paul, *Army of None: Autonomous Weapons and the Future
of War*, New York, W. W. Norton and Company, 2018

SCHARRE, Paul, *Four battlegrounds: power in the age of artificial
intelligence*, publié par W.W Norton & Company. Extrait du
passage disponible sur [**Le lien**].

VON CLAUSEWITZ, Carl, *De la guerre*, 1832, édition abrégée et
présentée par Gérard Chaliand, collection tempus, éditions
Perrin

IV, Rapports, textes législatifs et autres

ANSSI, «Recommandations de sécurité pour un système d'IA
générative», ANSSI-PA-102, 29/04/2024. Disponible sur [**Le
lien**]

ASSEMBLÉE GÉNÉRALE DES NATIONS UNIES,
«Résolution adoptée par l'Assemblée générale le 22 décembre
2023», A/RES/78/241. Disponible en version PDF sur [**Le
lien**]

ASSEMBLÉE GÉNÉRALE DES NATIONS UNIES,
«Résolution adoptée par l'Assemblée générale le 2 décembre
2024», A/RES/79/62. Disponible en version PDF sur [**Le
lien**]

ASSEMBLÉE GÉNÉRALE DES NATIONS UNIES,
Rapport du Rapporteur spécial sur les exécutions
extrajudiciaires, sommaires ou arbitraires, Christof Heyns, 9
avril 2013. Document A/HRC/23/47. Disponible sur [**Le
lien**].

ASSEMBLÉE GÉNÉRALE DES NATIONS UNIES,
Soixante-dix-neuvième session Première Commission Point
98 ss) de l'ordre du jour Désarmement général et complet:
systèmes d'armes létaux autonomes, A/C.1/79/L.77, 18
octobre 2024. Disponible sur [**Le lien**]

ASSEMBLÉE NATIONALE, Rapport d'information n°3248
en conclusion des travaux d'une mission d'information sur les
systèmes d'armes létaux autonomes, 22 juillet 2020.
Disponible sur [**Le lien**].

CICR, Rapport, «Le droit internationale humanitaire et les défis
posés par les conflits armés contemporains: engagement

renouvelé en faveur de la protection dans les conflits armés à l'occasion du 70^e anniversaire des Conventions de Genève», 2020. Disponible sur [[Le lien](#)].

CODE LIEBER, 1863, disponible sur: [[Le lien](#)].

COMEST (Commission mondiale d'éthique des connaissances scientifiques et des technologies) de l'UNESCO, Rapport sur l'éthique de la robotique, SHS/YES/COMEST-10/17/2 REV.Paris, 14 septembre 2017 Original: anglais. Disponible sur: [[Le lien](#)].

COMITÉ ÉCONOMIQUE ET SOCIAL EUROPÉEN, 526^e session plénière du CESE des 31 mai et 1er juin 2017, Avis du Comité économique et social européen sur «L'intelligence artificielle: les retombées de l'intelligence artificielle pour le marché unique (numérique), la production, la consommation, l'emploi et la société» (avis d'initiative) (2017/C 288/01).

COMITÉ ÉTHIQUE DE LA DÉFENSE, «Avis sur l'usage des technologies d'intelligence artificielle par les forces armées», 14 janvier 2025. Disponible sur [[Le lien](#)].

CONSEIL DE L'EUROPE, Commission des questions juridiques et des droits de l'homme, «L'Émergence des systèmes d'armes létales autonomes (SALA) et leur nécessaire appréhension par le droit européen des droits humains», Projet de résolution adopté à l'unanimité par la Commission le 14 novembre 2022. Disponible en version PDF sur [[Le lien](#)]

CONSEIL DE L'EUROPE, Commission des questions juridiques et des droits de l'homme, «L'Émergence des systèmes d'armes létales autonomes (SALA) et leur nécessaire appréhension par le droit européen des droits humains», Projet de résolution adopté à l'unanimité par la Commission le 14 novembre 2022. Disponible en version PDF sur [**Le lien**]

CONVENTION (II) DE LA HAYE concernant les lois et coutumes de la guerre sur terre et son Annexe: Règlement concernant les lois et coutumes de la guerre sur terre. La Haye, 29 juillet 1899. Disponible sur [**Le lien**].

CONVENTION DE GENÈVE pour l'amélioration du sort des militaires blessés dans les armées en campagne. Genève, 22 août 1864. Disponible sur [**Le lien**].

CONVENTION DE GENÈVE pour l'amélioration du sort des blessés et malades dans les armées en campagne. Genève, 6 juillet 1906. Disponible sur [**Le lien**]

CONVENTION DE GENÈVE pour l'amélioration du sort des blessés et des malades dans les forces armées en campagne du 12 août 1949 . Disponible sur [**Le lien**].

CONVENTION DE LA HAYE pour le règlement pacifique des conflits internationaux, 1899. Disponible sur [**Le lien**].

CONVENTION DE VIENNE sur le droit des traités. Faite à Vienne le 23 mai 1969. Entrée en vigueur le 27 janvier 1980. Nations Unies, Recueil des Traités, vol. 1155, p. 331. Disponible sur [[Le lien](#)].

CONVENTION ON CERTAIN CONVENTIONAL WEAPONS, «Rapport du Groupe d’experts gouvernementaux sur les technologies émergentes dans le domaine des systèmes d’armes létaux autonomes sur sa session de 2021», 22 février 2022. CCW/GGE.1/2021/3. Disponible sur [[Le lien](#)].

GOVERNEMENT DES ÉTATS-UNIS, Déclaration au Groupe d’experts gouvernementaux sur les systèmes d’armes létaux autonomes, “Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the CCW”, Genève, 10 avril 2018. Disponible en version PDF sur: [[Le lien](#)].

GOVERNEMENT DU ROYAUME-UNI, Lettre du 16 juin 1995 du conseiller juridique du Foreign and Commonwealth Office du Royaume-Uni de Grande-Bretagne et d’Irlande du nord, accompagnée des observations écrites du Gouvernement du Royaume-Unis P.48, disponible sur: [[Le lien](#)]

ICRC POSITION PAPER, “Artificial intelligence and machine learning in armed conflict: A human-centred approach”, mars 2021. Disponible sur [[Le lien](#)].

JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE, n°0285 du 9 décembre 2018, Commission d'enrichissement de la langue française, «Vocabulaire de l'intelligence artificielle (liste de termes, expressions et définitions adoptés)». Disponible sur [[Le lien](#)].

MINISTÈRE DE L'EUROPE ET DES AFFAIRES ÉTRANGÈRES, «11 principes sur les systèmes d'armes létaux autonomes». Disponible sur [[Le lien](#)].

MINISTÈRE DE L'EUROPE ET DES AFFAIRES ÉTRANGÈRES, «Systèmes d'armes létales autonomes, quelle est la position de la France?». Disponible sur [[Le lien](#)].

MINISTÈRE DES AFFAIRES ÉTRANGÈRES DE LA RÉPUBLIQUE POPULAIRE DE CHINE, «Position Paper of the People's Republic of China on Regulating Military Applications of Artificial Intelligence (AI)», 14 décembre 2021. Disponible sur: [[Le lien](#)].

NATIONS UNIES, «Appel conjoint du Secrétaire général de l'ONU et de la Présidente du CICR demandant aux États d'énoncer de nouvelles interdictions et limitations pour les systèmes d'armes autonomes», 5 octobre 2023. Disponible sur [[Le lien](#)].

NATIONS UNIES, «Première Commission: les pays soulignent l'importance d'une prise en compte des dimensions éthique et humanitaire de certaines armes classiques», AG/DSI/3748 24 octobre 2024. Disponible sur [[Le lien](#)].

NORME ISO/IEC TS 5723, Technical specification, “Trustworthiness – Vocabulary”, First edition, 2022-07. Disponible en version PDF sur [**Le lien**].

PARIS DECLARATION on “Maintaining Human Control in AI enabled Weapon Systems,” Publié le 11 février 2025. Disponible sur [**Le lien**].

PARLEMENT EUROPEEN, «Résolution du Parlement européen du 12 septembre 2018 sur les systèmes d’armes autonomes». Disponible sur [**Le lien**].

PROTOCOLE ADDITIONNEL (I) aux Conventions de Genève du 12 août 1949 relatif à la protection des victimes des conflits armés internationaux, du 8 juin 1977. Disponible sur [**Le lien**].

RAPPORT DE LA TASK FORCE, «L’intelligence artificielle au service de la défense», 2019. Disponible sur [**Le lien**].

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). Disponible sur [**Le lien**].

RÉSOLUTION DU PARLEMENT EUROPÉEN DU 16 FEVRIER 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)). Disponible sur [**Le lien**].

U.S. DEPARTMENT OF STATE, “Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy”, 9 novembre 2023. Disponible sur [**Le lien**]

UNITED NATIONS INSTITUTE FOR DISARMAMENT RESEARCH, “Table-Top Exercises on the Human Element and Autonomous Weapons Systems, Summary Report”, Giacomo Persi Paoli, Alice Spazian, Alisha Anand, This report summarises the findings of UNIDIR’s series of regional Table-Top Exercises conducted between September 2020 and June 2021. Disponible en version PDF sur [**Le lien**]

VILLANI, Cédric, « Donner un sens à l’intelligence artificielle », rapport du 28 mars 2018. Rapport officiel. Mission confiée par le Premier Ministre Édouard Philippe. Mission parlementaire du 8 septembre 2017 au 8 mars 2018. Disponible sur [**Le lien**].

V. Articles et publications juridiques

ANCELIN, Julien, « Les systèmes d’armes létaux autonomes (SALA): Enjeux juridiques de l’émergence d’un moyen de combat déshumanisé », *La Revue des Droits de l’Homme*, Actualités Droits-Libertés, mis en ligne le 25 octobre 2016, consulté le 20 novembre 2024. Disponible sur [**Le lien**].

ANDERSON, Kenneth et C. WAXMAN, Matthew, “Law and ethics for autonomous weapons systems: why a ban won’t work and how laws of war can”, *American University Washington College of Law Research Paper No. 2013-11*, 2013. Disponible sur [[Le lien](#)].

ARKIN, Ronald, *Governing Lethal Behaviour in Autonomous Robots* (CRC Press, 2009);

ANDERSON, Kenneth et C. WAXMAN, Matthew, «Law and ethics for robot soldiers», *Policy Review*, no 176 (2012). Cité par Christof Heyns, rapporteur spécial sur les exécutions extrajudiciaires, sommaires ou arbitraire. Assemblée Générale des Nations Unies, A/HRC/23/47, 9 avril 2013. Disponible sur [[Le lien](#)].

B. SOLUM, Lawrence, «Legal Personhood for Artificial Intelligences”, 70 N.C.L. Rev.1231 (1992), disponible sur: [[Le lien](#)].

BENSOUSSAN, Alain, «Les robots ont-ils une personnalité?», *Planète Robots* n°19, janv. 2013, disponible sur: [[Le lien](#)]

BENSOUSSAN, Alain, «Charte des Droits des Robots,» V.5.0, septembre 2015. Disponible sur [[Le lien](#)]

CARBONNIER, Jean, Chapitre 3, «Être ou ne pas être. Sur les traces du non-sujet de droit», *Flexible droit*, L.G.D.J, 9^e ed. 1998

CASSESE, Antonio, “The Martens Clause: Half a Loaf or simply Pie in the Sky?” *European Journal of International Law*, Volume 11, Issue 1, 2000, Pages 187–216, p.212. Disponible en version PDF sur [[Le lien](#)]

DERRUPPÉ, Jean et DE RAVEL D’ESCLAPON, Thibault, «Fonds de commerce», octobre 2016, actualisation en janvier 2025, paragraphes 195-196. Répertoire de droit immobilier, Dalloz Actualités

DURHIN, Nathalie, «Systèmes d’armes létaux autonomes: ne pas mélanger juridique et éthique», *Revue Défense Nationale*, 2018/HS1 n°Hors-série, pp. 167 à 176. Disponible sur [[Le lien](#)]. Consulté le 14/01/2025

ELLUIN, Aymeric, «L’urgence d’un traité encadrant les systèmes d’armes autonomes», 13 janvier 2025. Article publié sur le site Ritimo. Disponible sur [[Le lien](#)]

FARJAT, Gérard, «Entre les personnes et les choses, les centres d’intérêts», *RTD Civ.* 2002 p.221. Disponible en version PDF sur [[Le lien](#)].

FICHE D’ORIENTATION DALLOZ, «Entreprise individuelle à responsabilité limitée», juillet 2023.

GALBOIS-LEHALLE, Diane, «L’attraction de la catégorie des personnes sur les entités numériques», acte de colloque, «la personnalité convoitée», 2023. Disponible sur [[Le lien](#)].

GREENWOOD, C., “Historical Development and Legal Basis”, *The Handbook of Humanitarian Law in Armed Conflict*, Dieter Fleck (éd.), Oxford University Press, Oxford/New York, 1995, p.28 (par.129) cite par Rupert Ticehurst dans son article “La clause de Martens et le droit des conflits armés” publié dans la *Revue Internationale de la Croix-Rouge*, pp. 133-142.

HICKLETON, Marcus, “Shaky foundations: “killer robots” and the Martens Clause,” *Perth International Law Journal*, vol. 4, Janvier 2019, p.44. Disponible sur [**Le lien**].

MCFARLAND, T., préc., note 101 à la page 104. Cité dans le mémoire de recherche *Les systèmes d'armes autonomes et le droit international humanitaire: Une légalité contestée*, p.45, par Julien Grenon, Université de Montréal, juillet 2022.

MCLEAN, Duncan et HOFMAN, Michel, «Droit international humanitaire, souveraineté des États et érosion du consensus humanitaire: la fin de l’humanitarisme?», publié dans Numéro 23 – Droit international humanitaire: le grand retour... en arrière?. *Revue Alternatives Humanitaires*. Traduit de l’anglais par Sophie Jeangeorges. Disponible sur [**Le lien**].

NATIONS UNIES, communiqué de presse, «Première Commission: les pays soulignent l’importance d’une prise en compte des dimensions éthiques et humanitaires de certaines armes classiques», AG/DSI/3748, 24 octobre 2024. Disponible sur [**Le lien**].

NATO Communications and Information Agency, “NATO acquires AI-enabled warfighting system”, 14 avril 2025. Disponible sur: [**Le lien**].

NEVEJANS, Nathalie, «La légalité des robots de guerre dans les conflits internationaux», *Recueil Dalloz* 2016, p.1273.

PETIT, Camille, Représentation permanente de la France auprès de la Conférence du désarmement à Genève, GGE SALA - Déclaration générale: Intervention de l’Ambassadrice Camille Petit (Genève, 26 août 2024). Disponible sur [**Le lien**].

RAYNE, Serge, *Répertoire de droit pénal et de procédure pénal, Armes-Définition*, Chapitre 2-Définition (45-68), février 2023.

SHARKEY, Noël, «Weapons of Indiscriminate Lethality”, in Forum des informaticiens pour la paix et la responsabilité sociétale (FIF Kommunikation), n°1/09, pp. 26-29. Cité par J. Ancelin: «Les systèmes d’armes létaux autonomes (SALA): Enjeux juridiques de l’émergence d’un moyen de combat déshumanisé», 2016. Parargaphe.3. Disponible en version PDF.

SMARANDESCU, Ioana-Alexandra et VARZARU, Nina-Maria, “The new means of waging war. Artificial intelligence in the field of international humanitarian law”, *Romanian Journal of International Law*, RJIL No.30/2024, P. 35-54.

STIP COMPASS, “Joint Artificial Intelligence Center,” last updated: June 3, 2024. Disponible sur [**Le lien**].

VASCONCELOS ROSA, Madalena et GARCIA RAVEL, Juliane, «DIH en action: sept facteurs permettant d’influencer les comportements dans l’objectif de renforcer le respect du droit international humanitaire», 19 novembre 2020. Disponible sur [**Le lien**].

Y. SANDOZ, C. SWINARSKI, B. ZIMMERMANN (éd.), Commentaire des Protocoles additionnels du 8 juin 1977 aux Conventions de Genève du 12 août 1949, CICR/ Martinus Nijhoff, Genève, 1986, pp.38-39 (par.55) ; N. Singh et E. McWhinney, *Nuclear Weapons and Contemporary International Law*, 2^e éd. Martinus Nijhoff, Dordrecht, 1989, pp. 46-47 cité par Rupert Ticehurst dans son article “La clause de Martens et le droit des conflits armés”, publié dans la *Revue Internationale de la Croix-Rouge*, pp. 133-142.

VI. Jurisprudence

TPIY, Trial Chamber II, *The prosecutor vs. Clément Kayishema and Obed Ruzindana*, 21 mai 1999, §198. Disponible sur [**Le lien**].

VII. Références non juridiques

ABRAHAM, Yuval, “Lavander: The AI machine directing Israel’s bombing spree in Gaza”, *+972 Magazin*, 3 avril 2024. Disponible sur [**Le lien**].

ACTU IA, «Etats-Unis: les dernières nouvelles autour de l'IA du Pentagone et du Joint Artificial Intelligence Center». Disponible sur: [[Le lien](#)].

AMNESTY INTERNATIONAL, «Quelles armes sont interdites par le droit international humanitaire?» disponible sur [[Le lien](#)].

ARKIN, Ronald, *Lethal Autonomous Systems and the Plight of the Non-combatant*, No137, juillet 2013. Publié dans le livre *The Political Economy of Robots: Prospects for Prosperity and Peace in the Automated 21st Century* (International Political Economy Series), 1st ed. 2018 Edition, par Ryan Kiggins(Editeur). Disponible sur [[Le lien](#)].

B. KANIA, Elsa, “The Pla’s trajectory from informatized to “intelligentized” warfare”, 8 juin 2017. Disponible sur [[Le lien](#)].

BARBA-KAY, Antón, “There Is No Ethical Automation: Sanislav Petrov’s Ordeal by Protocol”, *Journal of military ethics*, 2024, Vol 23, n°3-4, pp. 277-288. Disponible en version PDF sur [[Le lien](#)].

BARRIER, Félicité, «Les systèmes armés létaux autonomes (Sala): vers une nouvelle course à l’armement?» pp. 19 à 24 de la *Revue Défense Nationale*, 2018/5, n°810. Disponible sur [[Le lien](#)].

BERTHIER, Thierry, «L'ONU veut réguler l'intelligence artificielle militaire mais se heurte au principe de réalité concurrentielle», *Revue Conflits*, publié le 13 août 2023. Disponible sur [[Le lien](#)].

BERTHIER, Thierry, «Systèmes armés semi-autonomes: que peut apporter l'autonomie?», *Revue Défense Nationale*, 2019/5, n°820, pp. 74- 80. Disponible sur [[Le lien](#)].

BIBLIOTHÈQUE DU PARLEMENT DU CANADA, Ariel Shapiro, «Systèmes d'armes autonomes: quelques répercussions sur la sécurité internationale et sur le Canada», 20/12/2019. Disponible sur [[Le lien](#)].

BOUTHERIN, Grégory, «Maîtrise des armements non conventionnels: le salut viendra-t-il du soft disarmament?», *AFDI*, LIII, 2007, p. 243. Disponible en version PDF sur [[Le lien](#)].

BRANDAO, Caroline, «L'analyse des nouveaux systèmes d'armes robotisés sous le prisme de l'article 36 du Protocole additionnel I», publié dans la *Revue Défense Nationale*, 2018/HS1 n°Hors-Série, pp. 135 à 143. Disponible sur [[Le lien](#)].

BROADHURST, Clea, «L'IA, le nouvel outil économique et géopolitique de la Chine», Article publié le 06/02/2025 dans le journal numérique rfi. Disponible sur [[Le lien](#)].

BUGNION, François, «Droit international humanitaire coutumier», 24 novembre 2007, version PDF disponible sur: [**Le lien**], consulté le 25/04/2024.

CAMPAIGN TO STOP KILLER ROBOT, Note d'information, «Maintenir le contrôle sur les systèmes d'armes». Disponible en version PDF sur [**Le lien**].

CAMPAIGN TO STOP KILLER ROBOT, “Report on outreach on the UN report on “lethal autonomous robotics””, 31 juillet 2013. Disponible en version PDF sur [**Le lien**].

CARDIA, Anthony, «Les conflits armés à l'époque de l'intelligence artificielle», 26 novembre 2024. Disponible sur [**Le lien**].

CASIER, Frédéric, «L'appel d'urgence à de nouvelles règles internationales sur les systèmes d'armes autonomes face aux risques humanitaires croissants», 11/01/2024. Disponible sur [**Le lien**].

CICR, «Droit international humanitaire, réponses à vos questions», 0703/001 05.2015 15000. Disponible en version PDF sur [**Le lien**].

CICR, «Moyens et méthodes de guerre» disponible sur: [**Le lien**]

CICR, «Un nouveau pas vers l'imposition de limites à l'autonomie des systèmes d'armes», 12/04/2018. Disponible sur [**Le lien**].

CICR, «Réunion d'experts sur les systèmes d'armes létales autonomes», 15/11/2017. Disponible sur [**Le lien**].

COMMISSION NATIONALE INFORMATIQUE ET LIBERTES (CNIL), Définition de robustesse d'une intelligence artificielle par la CNIL. Disponible sur [**Le lien**].

COMANDRÉ, Mathias et MEVELLEC, Mattéo, «Les défis éthiques et les risques de l'intelligence artificielle militaire», Institut National des Affaires Stratégiques et Politiques. Disponible sur [**Le lien**].

DARTHMOUTH COLLEGE, “A proposal for the Dartmouth summer research project on artificial intelligence”, August 31, 1955. Disponible sur [**Le lien**]

DATASULTING, Article de blog, «Algorithme, Intelligence Artificielle, Machine Learning, Deep Learning: quelles différences?», 22 juin 2023. Disponible sur [**Le lien**].

DE LUCA, Anne, «Le calife, le soldat et le paysan: l'art de la guerre en islam», Les Cahiers de l'Orient, 2017/1 n°125, Article de revue disponible sur: [**Le lien**], consulté le 25/04/2025.

DE MATTEIS, Ludovic, JANNY, Steeven, NATHAN, Solal et SHU-QUARTIER, Wenqi, “Introduction à l'apprentissage automatique», École normale supérieure, Paris-Saclay, édité le 24/05/2022. Disponible sur [**Le lien**].

DE ROUCY-ROCHEGONDE, Laure et FÉREY, Amélie, «Bombardement israéliens à Gaza: comment l'intelligence artificielle annihile les préventions éthiques», 07/05/2024. Disponible sur [[Le lien](#)].

DE ROUCY-ROCHEGONDE, Laure et FÉREY, Amélie, «De l'Ukraine à Gaza: l'intelligence artificielle en guerre», article publié dans la revue *Politique étrangère*, vol.89, n°3, automne 2024. Disponible en version PDF sur [[Le lien](#)].

DE ROUCY-ROCHEGONDE, Laure, «Systèmes d'armes létales autonomes: y aura-t-il un Terminator tricolore?», 225/09/2020. Disponible sur [[Le lien](#)].

FERRARA, Emilio, «Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies», 26 décembre 2023. Disponible sur [[Le lien](#)].

FORKLIN, Julien, «L'avenir de l'IA dans la guerre: concepts clés et implications». Disponible sur [[Le lien](#)].

FOURNIER, Paul, «Interdiction de l'usage de l'arc et de l'arbalète par le IIe concile de Latran en 1139», *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres*, 1916, 60-3/pp. 295-296. Disponible sur: [[Le lien](#)].

FUTURE OF LIFE INSTITUTE, «Autonomous Weapons Open Letter: AI & Robotics Researchers», 9 février 2016. Disponible sur [[Le lien](#)].

FUTURE OF LIFE INSTITUTE, “A Diplomat’s Guide to Autonomous Weapons Systems”, 2024. Disponible en version PDF sur [[Le lien](#)].

General R. ALLEN, John and HUSAIN, Amir, “On hyperwar”, U.S Naval Institute, Juillet 2017. Disponible sur [[Le lien](#)].

GOMEZ, Éric, «Le ciblage dans le domaine militaire», Ouvrage *Les guerres de l’information à l’ère numérique*, Sous la direction de Céline Marangé et Maud Quessard, Hors collection, 2021, Presses Universitaires de France, 456 pages. Disponible sur [[Le lien](#)].

GOUTEFANGEA, Patrick, «Alan Turing et l’intelligence artificielle: le «jeu de l’imitation» et «l’IA forte», 24 juin 2017, Hal open science. Disponible sur: [[Le lien](#)].

HUMAN RIGHTS WATCH, «Gaza: Les outils numériques de l’armée israélienne mettent en danger les civils», 10 septembre 2024. Disponible sur [[Le lien](#)].

HUMAN RIGHTS WATCH, «Répondre à l’appel. L’interdiction des robots tueurs, un impératif moral et juridique», 2018. Disponible en version PDF sur [[Le lien](#)].

HUMAN RIGHTS WATCH, «Robots tueurs», disponible sur [[Le lien](#)]

HUMAN RIGHTS WATCH, «Stopper les robots tueurs, positions des pays sur l'interdiction des armes entièrement autonomes et sur le maintien du contrôle humain», 26 août 2020. Disponible sur [[Le lien](#)].

HUMAN RIGHTS WATCH, «Un danger pour les droits humains, systèmes d'armes autonomes et processus décisionnels numériques», résumé et recommandations, 28 avril 2025. Disponible en version PDF sur [[Le lien](#)].

IBM, «Qu'est-ce que la gestion des vulnérabilités?», Guide 2025 de la cybersécurité (actualisé en 2026). Disponible sur [[Le lien](#)]

INSTITUT NATIONAL DES AFFAIRES STRATÉGIQUES ET POLITIQUES, «Implications géopolitiques et compétition internationale: la nouvelle course à l'IA militaire», disponible sur [[Le lien](#)].

ISMAILOVIC, Muzen, «Ciblage algorithmique: le rôle de l'intelligence artificielle dans les frappes israéliennes à Gaza et ses implications éthiques», 20 février 2025. Disponible en version PDF sur [[Le lien](#)].

JEANGENE VILMER, Jean-Baptiste, «Robots létaux autonomes: amorce d'un débat», décembre 2012. Disponible en version PDF sur [[Le lien](#)].

KENNITH, Payne, «Artificial Intelligence and the Nature of War», in: Sweijts Tim et Michaels Jeffrey H. (dir.), *Beyond Ukraine: Debating the Future of War*, Oxford University Press,

2024, p. 223-240. Cité par la *Revue Nationale de Défense*, n°879, avril 2025.

LA TOUPIE, dictionnaire de la politique, «responsabilité», disponible sur [**Le lien**]

LINDLOFF, Chiara and SIEGERT, Ingo, “Defining bias in AI-systems: Biased models are fair models”, 25 février 2025. Disponible en version PDF sur [**Le lien**].

MAURER, Peter, Président du CICR, Allocution, Centre d'études stratégiques et internationales, Washington DC, 10 mai 2019. Traduction de la transcription. «Les règles de la guerre appartiennent-elles au passé?». 10/05/2019. Disponible sur [**Le lien**].

MICHAEL C. Horowitz, «Public Opinion and the Politics of the Killer Robots Debate», *Research & Politics*, 3/1, 2016, p.1-8. Cité par Laure de Roucy-Rochegonde, *op.cit.*, p.193.

MINISTÈRE DES ARMÉES, Un accord entre l'AMIAD et Mistral IA, «IA de défense: Sébastien Lecornu inaugure le pôle recherche de l'Amiad», 5 mars 2025. Disponible sur [**Le lien**].

MINSTÈRE DES ARMÉES, Biblioville, «Les systèmes d'armes létales autonomes (SALA)», mars 2021. Disponible en version PDF sur [**Le lien**].

MOSCONI, Jean, «Sur quelques capacités et incapacités des machines», *Bulletin de la Société Française de Philosophie*, 3, juillet-septembre 1991, p. 86. Cité par Patrick Goutefangea.

MUCCI, Tim et STRYKER, Cole, «Qu'est-ce que la superintelligence artificielle», IBM, 18 décembre 2023. Disponible sur [**Le lien**]

MUST PARIS, «MQ-1 Predator – Rôle et évolution du drone de combat US», 1 décembre 2024, disponible sur: [**Le lien**], consulté le 27/04/2024

NATIONS UNIES, «L'ONU et la Croix-Rouge appellent à un accord international inédit sur les armes autonomes», 5 octobre 2023. Disponible sur [**Le lien**].

NOËL, Jean-Christophe, «Intelligence artificielle: vers une nouvelle révolution militaire?», Octobre 2018. Disponible en version PDF sur [**Le lien**].

NOËL, Jean-Christophe, «L'IA a-telle un avenir pour les armées?», article paru dans la *Revue Défense Nationale*, 2019/N°820. Disponible sur [**Le lien**].

NOIZET, Thomas, «Les essais de drones, graal ou chimère?», *Revue Défense Nationale* 2023/HS13 N° Hors-série, pp. 239-256. Disponible sur [**Le lien**].

OPEN LETTER TO THE EUROPEAN COMMISSION, Artificial Intelligence and Robotics, 2017. Disponible sur [**Le lien**].

P. LIN, K. ABEY, G.A. BEKEY, *Robot Ethics: The Ethical and Social Implications of Robotics*, Cambridge, MIT Press, 2014, cité par Laure DE ROUCY-ROCHEGONDE, «La guerre à l'ère de l'intelligence artificielle», p.138.

PÉGNY, Maël et ISSAM IBNOUHSEIN, Mohamed, «Quelle transparence pour les algorithmes d'apprentissage machine?», mai 2018, pp.7-8. Disponible en version PDF sur [**Le lien**]

PÉRÉMARTY, Léa, «Les armes létales autonomes: entre mythes et confusions», 26 juillet 2023. Disponible sur [**Le lien**]

PEREZ, Julien, professeur au Laboratoire de Recherche de l'EPITA (école pour l'informatique et les techniques avancées). Actualités EPITA, «Les agents IA, un tournant majeur pour l'intelligence artificielle», 05/05/2025. Disponible sur [**Le lien**].

PICTET, Jean (sous la direction de), *Les Conventions de Genève du 12 août 1949: Commentaire, Vol. IV, Convention de Genève relative à la protection des personnes civiles en temps de guerre* (ci-après Commentaire GC IV), CICR, Genève, 1958, p. 219. Cité par Cordula Droegge Original anglais, publié sous le titre “In Truth the *leitmotiv*: the prohibition of torture and other forms of ill-treatment in international humanitarian law”, *International*

Review of the Red Cross, Vol. 89, No 867, septembre 2007, pp.515-541. Disponible sur [[Le lien](#)]

REVUE DÉFENSE NATIONALE, 2019/5 n°820, Avant-propos – Intelligence artificielle et défense, par Florence Parly, pp. 9 à 17. Disponible sur [[Le lien](#)].

REVUE INTERNATIONALE DE LA CROIX-ROUGE, «Position du Comité international de la Croix-Rouge sur les systèmes d'armes autonomes et documentations de référence», IRCC No.915, janvier 2022. Disponible sur [[Le lien](#)].

RUFFO, Marie des Neiges, «Avons-nous le choix d'utiliser l'IA en temps de guerre?», 47^e numéro des *Cahiers de la sécurité et de la justice de la Revue trimestrielle de l'IHEMI*, 12 mars 2020. Disponible sur [[Le lien](#)].

SAUNOIS, Laurent, «Enjeux et défis de l'Intelligence artificielle (IA) pour l'OTAN», *Revue Défense Nationale*, n°879, Avril 2025. Disponible sur [[Le lien](#)].

SAUVETON, Pierre, «IA militaire: l'OTAN choisit Palantir, la France répond avec Artemis, GenIAI... et Comand AI», 23 avril 2025. Disponible sur [[Le lien](#)].

SCHARRE, Paul, *Four Battlegrounds—Power in the Age of Artificial Intelligence*, W.W. Norton & Company, 2023, 496 pages. Cité par la Revue Nationale de Défense, n°879, avril 2025.

Site WATSON, «Le Russe qui a sauvé le monde», 14/04/2024.
Disponible sur [**Le lien**].

SOETAERT, Sammy, «Une IA explicable est-elle la clé d'un développement ultérieur, publié le 11/03/2022. Disponible sur [**Le lien**].

STEWART, Ruben et HINDS Georgia, «Les algorithmes de la guerre: le recours à l'intelligence artificielle pour la prise de décision dans les conflits armés», 22 février 2024. Disponible sur [**Le lien**].

SVEN, Clement, «NATO and AI: Navigating the Challenges and Opportunities», *Special Report for Science and Technology Committee of NATO Parliamentary Assembly*, 24 novembre 2024. Disponible sur [**Le lien**].

TANGUY, Vincent, «Poutine pense dominer le monde en maîtrisant l'intelligence artificielle», article publié dans *Sciences et Avenir*, 05/09/2017. Disponible sur [**Le lien**].

THALES GROUP, «La DSTA et Thales annoncent la création d'un laboratoire commun axé sur l'Intelligence Artificielle pour renforcer les systèmes de défense de Singapour», publié le 04 avril 2025. Disponible sur [**Le lien**].

TICEHURST, Rupert, «La clause de Martens et le droit des conflits armés», *Revue Internationale de la Croix-Rouge*. Volume 79, Issue 824, April 1997, pp.133-142. Disponible en version PDF sur [**Le lien**].

TUAL, Morgane, «A peine lancée, une intelligence artificielle de Microsoft Dérape sur Twitter» article publié dans *Le Monde*, 24 mars 2016. Disponible sur [**Le lien**].

TUHIN, Muhammad, «7 examples of How Deep Learning Is Used Today», *Science News Today*, 30 avril 2025. Disponible sur [**Le lien**].

UNITED NATIONS, Webinar of United Nations, Office for Disarmament Affairs, “Experts Reflect on Novel Approaches to “Meaningful Human Control””, 7 mars 2023. Disponible sur: [**Le lien**].

Dr VALANTIN, Jean- Michel, «Militarisation de l’Intelligence artificielle en Chine», The Red Team Analysis Society, 23 avril 2018. Disponible sur [**Le lien**].

YANN LECUN, Définition de l’IA par Yann Lecun. Article issu de la revue DéfiS, n°8, 12 novembre 2019, écrit et publié sous la direction de l’Institut National des Hautes Études de la Sécurité et de la Justice (INHESJ) qui a rejoint l’Institut des Hautes Études du Ministère de l’Intérieur (IHEMI). Disponible sur [**Le lien**].

全军军事术语管理委员会 [All-Military Military Terminology Management Committee], 中国人民解放军军语 [People’s Liberation Army Military Terminology], (Beijing: 军事科学出版社 [Military Science Press], 2011). Cité par Elsa B. Kania,

“ “ AI weapons” in China’s Military Innovation”, Avril 2020.
Disponible en version PDF sur [**Le lien**].

[人大代表刘国治：人工智能将加速军事变革进程], cité par
Dr Michel Valantin, «Militarisation de l’Intelligence artificielle
en Chine», The Red Team Analysis Society, 23 avril 2018.

Ce mémoire de recherche étudie les défis que pose l'essor de l'intelligence artificielle (IA) au droit international humanitaire (DIH), avec une approche spécifique sur les systèmes d'IA appliqués au ciblage humain. Adoptant une approche croisant les dimensions juridiques, techniques et éthiques, il analyse la compatibilité des systèmes d'armes létales autonomes et des outils d'aide à la décision sous supervision humaine avec les principes fondamentaux du DIH, tels que la distinction, la proportionnalité et la précaution.

L'étude met en évidence la façon dont ces technologies, susceptibles d'assister voire de remplacer l'intervention humaine, remettent en cause les fondements du droit des conflits armés, historiquement pensés pour une mise en œuvre humaine. Elle souligne les difficultés de leur intégration dans le cadre juridique existant, liées à leurs limites techniques, opacité des algorithmes, fiabilité incertaine, biais algorithmique, mais aussi aux enjeux juridiques et aux tensions géopolitiques qu'elles engendrent.

Dans ce contexte, l'étude questionne la capacité du DIH à encadrer efficacement ces nouvelles technologies utilisées en contexte opérationnel. Elle insiste sur la nécessité de préserver un contrôle humain significatif et interroge les mécanismes de responsabilité en cas de défaillance. Face aux divergences entre États et à l'absence de consensus international, elle met en avant l'urgence d'une adaptation du cadre normatif.

Ce travail propose ainsi une réflexion approfondie sur l'avenir du DIH à l'ère de l'intelligence artificielle, au croisement des exigences juridiques, des contraintes opérationnelles et des enjeux éthiques.

Ce mémoire a reçu le premier prix de l'an 2026 dans le cadre du concours annuel 'Manfred Rosenberger' d'EuroISME pour le meilleur mémoire d'étudiant (Master of Arts). Pour plus d'informations sur le concours, veuillez consulter le site www.euroisme.eu

