JONATHAN KWIK, *LAWFULLY USING AUTONOMOUS WEAPON TECHNOLOGIES*, LA HAYE, TMC ASSER PRESS, 2024

Valentin Laugrand

Valentin Laugrand est candidat à la maîtrise en droit international à l'Université Laval, et en affaires internationales à l'Université Carleton. Il travaille actuellement à l'UNESCO sur les enjeux liés à l'éthique de l'intelligence artificielle. Ses recherches portent sur le droit international humanitaire et sur l'utilisation de systèmes d'armements autonomes en contexte urbain.

Depuis le début du XXIe siècle, le développement de drones intégrant l'intelligence artificielle (IA) s'est accéléré, permettant d'automatiser certaines fonctions de ciblage et déléguer des responsabilités auparavant réservées aux opérateurs humains[1]. Ces systèmes d'armements autonomes (SAA) alimentent d'importants débats entre les États, les juristes et les stratèges militaires, quant aux avantages opérationnels potentiels et la licéité de ces nouvelles armes[2]. Malgré la création d'un groupe d'experts gouvernementaux en 2016, les discussions diplomatiques visant à établir une définition commune et un cadre juridique permettant de superviser leur développement et leur utilisation n'avancent pas, en raison des divergences d'intérêts entre les États[3]. Aujourd'hui, ces discussions semblent d'autant plus pressantes alors que des rapports font état de l'usage de SAA ou de technologies de ciblage reposant sur l'IA dans les conflits en Libye, en Ukraine et au Moyen-Orient[4].

L'ouvrage *Lawfully Using Autonomous Weapon Technologies* apporte une analyse pratique et théorique de l'utilisation des SAA, et clarifie plusieurs enjeux juridiques clés liés à ces technologies. Jonathan Kwik, l'auteur de cette monographie, est chercheur au TMC Asser Institute de La Haye, et se spécialise en droit international humanitaire (DIH), en ciblage militaire et en IA[5]. Contrairement à de nombreux textes sur les SAA, cet ouvrage offre une analyse complexe et technique des enjeux relatifs à l'autonomisation et aux règles relatives à la conduite des hostilités. Il examine la façon dont les commandants militaires peuvent gérer les particularités technologiques de l'IA, tout

1. Altab Hossin et al, « Integrating artificial intelligence in unmanned vehicles: navigating uncertainties, risks, and the path forward for the fourth industrial revolution » (2025) 12:312 Humanities & Soc Sciences Communications ; Valeri Modebadze, « The importance of drones in modern warfare and armed conflicts » (2021) 1:2 J Soc Sciences & Arts 89.
2. Mariarosaria Taddeo et Alexander Blanchard, « A Comparative Analysis of the Definitions of Autonomous Weapons Systems » (2022) 28:37 Science & Engineering Ethics.
3. Plusieurs puissances militaires telles que la Chine, la Corée du Sud, les États-Unis, la France, l'Inde, Israël, le Royaume-Uni, et la Russie ont des programmes visant à développer des systèmes d'armes présentant différents degrés d'autonomie. Certains États freinent ainsi la mise en place d'une régulation internationale, afin de faire avancer leurs programmes militaires respectifs.
4. *Lettre datée du 8 mars 2021, adressée à la Présidente du Conseil de sécurité par le Groupe d'experts sur la Libye créé par la résolution 1973 (2011) du Conseil de sécurité*, CS NU, 76e année, Doc NU S/2021/229 (2021) ; Jean-Marc Rickli et Federico Mantellassi, « The War in Ukraine: Reality Check for Emerging Technologies and the Future of Warfare » (2024) Geneva Paper no 34/24 ; Amira Mhadhbi, « Quelles sont les nouvelles armes israéliennes de haute technologie et comment l'intelligence artificielle est-elle impliquée? » *BBC News Afrique* (11 janvier 2024), en ligne : <bbc.com> [perma.cc/E22T-NFHA].
5. Jonathan Kwik, « About Dr Jonathan Kwik » (2025), en ligne : <jonathankwik.com> [perma.cc/V8W9-4EPZ].

en respectant leurs obligations en matière de DIH. Kwik identifie les facteurs environnementaux, technologiques, opérationnels et juridiques qui doivent être considérés avant de déployer des SAA. Il précise également les conditions pouvant mener à l'attribution d'une responsabilité pénale individuelle.

L'analyse de l'auteur se concentre sur le rôle des commandants militaires opérants sur le terrain qui sont responsables du déploiement des SAA, qu'il nomme *Deployer*. Kwik avance qu'en raison de leur proximité avec le champ de bataille, ces derniers demeurent les mieux placés pour exercer un contrôle suffisant sur les SAA, et par le fait même, en mesure de jouer un rôle décisif dans le respect des règles de DIH[6]. Son analyse s'articule en six parties et se déploie en quatorze chapitres. Elle couvre à la fois les dimensions techniques et technologiques des SAA, incluant les indicateurs de performance, les causes potentielles de défaillances et les contre-mesures ennemies, mais également le contenu des règles de DIH relatives au ciblage et à la conduite des hostilités.

À travers une analyse technique poussée des capacités technologiques des SAA – description de leurs capteurs, de leur programme interne d'IA, de leurs données, de leur programmation et de l'environnement dans lequel ils sont déployés, l'auteur examine les considérations et les comportements que les commandants opérationnels (*Deployer*) doivent adopter afin d'utiliser des SAA de façon licite[7].

La première partie contient un chapitre unique. Intitulée « Framing the study », elle introduit les principaux concepts abordés dans l'ouvrage. L'auteur y définit les SAA, en adoptant une définition assez large : « *weapon system utilizing (partial) data-driven AI models* »[8]. Contrairement à de nombreux auteurs qui utilisent habituellement les critères de l'autonomie ou le niveau d'implication d'un opérateur humain pour définir ces technologies, l'auteur distingue les SAA des autres systèmes qui comportent des fonctions autonomes, par la dépendance des SAA aux données : « *data-driven machine learning is what truly separates contemporary AI from older systems that exhibit automatic functionalities* »[9]. Les notions centrales de l'ouvrage sont définies, comme les termes *Deployer*, *Operator* ou encore *Providing Entity*, et un aperçu complet de l'analyse prodiguée dans l'ouvrage est présenté.

6. Jonathan Kwik, *Lawfully Using Autonomous Weapon Technologies*, La Haye, TMC Asser Press, 2024 à la p 7 [Kwik, « Lawfully Using Technologies »].
7. *Ibid* à la p 9.
8. *Ibid* à la p 11.
9. *Ibid* à la p 10.

La deuxième partie, « Controling AWS », aborde la notion de « contrôle humain significatif », aussi appelé *Meaningful Human Control* (MHC). Cette notion a souvent été mobilisée dans les discussions politiques et juridiques pour encadrer le niveau d'autonomie laissée aux SAA, mais plusieurs déplorent son inutilité sur le plan opérationnel, car son contenu concret reste flou[10]. Cette partie, qui contient un chapitre unique, présente les résultats d'une étude empirique sur les différentes façons dont le MHC a été conceptualisé dans la littérature, et identifie les moments clés du processus de ciblage durant lesquels les commandants peuvent exercer une influence décisive sur les SAA[11]. Le schéma présenté comporte cinq leviers : la conscience de la situation (*awareness*), le choix de l'arme (*weapon selection*), les contrôles contextuels (*context control*), les capacités de prédiction (*prediction*) et les mécanismes de responsabilité (*accountability*)[12]. Ces éléments permettent d'assurer une influence suffisante sur les SAA, pour être dit « en contrôle ». Ce schéma s'avère un élément central puisque l'auteur y fait référence tout au long de l'ouvrage, afin d'identifier et de comprendre l'interaction entre le commandant opérationnel (*Deployer*), le SAA, avec ses indicateurs techniques et son IA, l'environnement et le contexte d'utilisation, et la responsabilité pénale individuelle lié aux décisions prises dans ce cadre.

La troisième partie, intitulée « Understanding AWS », comporte quatre chapitres et analyse en profondeur les dimensions techniques et technologiques des SAA. Elle constitue une base essentielle permettant l'interprétation des règles de DIH dans les chapitres suivants. Cette partie s'intéresse plus précisément à la phase préparatoire d'une attaque, durant laquelle le responsable du déploiement doit évaluer les systèmes à sa disposition, anticiper les effets escomptés de l'attaque, et étudier l'environnement opérationnel[13].

Le troisième chapitre explore les propriétés fondamentales de l'IA, ses avantages et ses limites, ainsi que la manière dont ces éléments interagissent avec les réalités complexes du champ de bataille. L'auteur détaille le processus décisionnel de l'IA sur le terrain à travers l'analyse approfondie du modèle

10. Michael C Horowitz et Paul Scharre, « Meaningful Human Control in Weapon Systems: A Primer » (2015) Center for a New American Security, Document de travail n° 031315 à la p 6 ; Amanda Musco Eklund, « Meaningful Human Control of Autonomous Weapon Systems: Definitions and Key Elements in the Light of International Humanitarian Law and International Human Rights Law » (2020) Swedish Defence Research Agency, Document de travail n° FOI-R-4928-SE à la p 27.
11. Kwik, « Lawfully Using Technologies », *supra* note 6 à la p 27.
12. *Ibid* aux pp 33–40.
13. *Ibid* à la p 51.

« *acquisition-analysis-decision-action* »[14]. Il distingue le *symbolic AI* du *machine learning*, en décrivant le processus d'acquisition des données nécessaires à l'apprentissage et au développement de l'IA[15]. Il expose ensuite le processus décisionnel : de l'acquisition des données par les capteurs, passant par l'analyse, et la prise de décision, pour aboutir à la prise d'action. L'auteur explique que le comportement et les performances des SAA dépendent autant de leurs propriétés techniques intrinsèques, que de l'environnement dans lequel ils sont déployés[16]. Il introduit à ce titre le concept d'*Intended Operational Environment* (IOE), désignant le contexte et l'environnement précis qui sont identifiés en fonction des capacités concrètes des SAA, pour un usage optimal et conforme aux règles de DIH[17].

Le quatrième chapitre est consacré à une analyse approfondie des principaux indicateurs de performance de l'IA : la précision (*accuracy*), la robustesse (*robustness*), la fiabilité (*reliability*) et l'intelligibilité (*understandability*)[18]. Les différents paramètres métriques, quantitatifs et qualitatifs, propres à chacun d'eux sont examinés. Les connaissances en IA et en robotique de l'auteur permettent de pousser la réflexion au-delà des analyses juridiques habituelles consacrées à ce sujet.

Le cinquième chapitre aborde les causes spécifiques pouvant potentiellement mener à une défaillance des SAA. Ces causes incluent les problèmes liés aux données d'entrainement (*training data*), aux erreurs humaines, aux dysfonctionnements techniques des composantes ou du système (*component and system failures*), ou encore à une utilisation en dehors du cadre opérationnel prévu (IOE)[19].

Enfin, le sixième chapitre examine les contre-mesures susceptibles d'être utilisées par un adversaire pour perturber ou neutraliser les SAA. L'auteur distingue la manipulation des données mobilisées par l'IA pour prendre ses décisions (*adversarial inputs*), de l'empoisonnement des données dès la phase de conception et d'apprentissage (*poisoning*), et de l'utilisation de portes dérobées (*backdoors*) — une combinaison particulièrement efficace des deux méthodes précédentes pour manipuler les SAA une fois déployés[20].

14. *Ibid.*
15. *Ibid* aux pp 56–57.
16. *Ibid* aux pp 68–71.
17. *Ibid.*
18. *Ibid* à la p 79.
19. *Ibid* à la p 105.
20. *Ibid* à la p 129.

La quatrième partie, la plus dense de l'ouvrage, « AWS in the Legal-Operational Context », est divisée en quatre chapitres et consacrée à l'application concrète des connaissances techniques acquises précédemment pour l'analyse juridique et opérationnelle. Elle explore comment le responsable du déploiement (*Deployer*) peut déterminer si l'utilisation d'un SAA est à la fois conforme aux exigences du DIH et justifiée d'un point de vue opérationnel.

L'analyse débute en présentant les caractéristiques techniques des SAA qui influencent leur conformité aux règles du DIH relatives au ciblage. Les principes de distinction, de proportionnalité et de précaution figurent au cœur de l'analyse juridique, de même que le contenu des articles 51 et 57 du *Protocole additionnel I aux Conventions de Genève*[21].

Le huitième chapitre occupe une place centrale dans l'ouvrage, car l'auteur dépasse les formulations normatives vagues souvent rencontrées dans les études sur les SAA[22]. Alors que la littérature sur le sujet se contente habituellement d'évoquer la nécessité pour ces systèmes d'être suffisamment fiables, cohérents, sécurisés, etc., sans définir concrètement ces éléments, l'auteur identifie quant à lui les variables clés à considérer pour l'évaluation juridique d'un SAA, particulièrement au regard des règles de ciblage[23]. Il examine les propriétés intrinsèques de ces systèmes : leur capacité à distinguer les cibles militaires des civils ou leur degré d'opacité ; les particularités de l'environnement opérationnel, telles que la densité d'entités civiles présentes sur le terrain ou les conditions météorologiques ; ainsi que les paramètres propres à la mission, incluant les contraintes temporelles et spatiales, et les catégories spécifiques de cibles[24]. Il explique que les limites de certaines variables peuvent

21. *Protocole additionnel aux Conventions de Genève du 12 août 1949 relatif à la protection des victimes des conflits armés internationaux*, 8 juin 1977, 1125 RTNU 3, arts 51, 57.
22. L'auteur fait référence aux recherches suivantes, mais ce problème est assez généralisé dans la doctrine : Comité international de la Croix-Rouge, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, Versoix, ICRC, 2016 à la p 81 ; Geoffrey S Corn, « Autonomous Weapons Systems: Managing the Inevitability of "Taking the Man out of the Loop" » dans Nehal C Bhuta et al, dir, *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge, Cambridge University Press, 2016 à la p 226 ; Ian S Henderson, Patrick Keane et Joshua Liddy, « Remote and Autonomous Warfare Systems: Precautions in Attack and Individual Accountability » dans Jens David Ohlin, dir, *Research Handbook on Remote Warfare*, Cheltenham, Edward Elgar, 2017, 335 à la p 361 ; William H Boothby, « Control in Weapons Law » dans Rogier Bartels et al, dir, *Military Operations and the Notion of Control Under International Law*, La Haye, TMC Asser Press, 2021, 369 aux pp 384–85.
23. Kwik, « Lawfully Using Technologies », *supra* note 6 à la p 168.
24. *Ibid* aux pp 172–208.

être compensées en ajustant les paramètres opérationnels[25]. Par exemple, si un SAA ne parvient pas à distinguer un tramway d'un char d'assaut, la délimitation préalable du champ d'action par le commandant, à une zone d'action excluant la présence de tramways, permet de compenser les limites du système[26]. Le chapitre approfondit par ailleurs les enjeux liés aux indicateurs de performance quantitatifs et qualitatifs, notamment en ce qui concerne le taux d'erreur dans l'identification des cibles. Il insiste sur l'importance de contextualiser ces indicateurs, rappelant que les performances des SAA en matière de reconnaissance dépendent étroitement du contexte, du moment de leur engagement, ainsi que de la nature de l'environnement opérationnel.

Le neuvième chapitre constitue un autre élément central de l'analyse. Il offre une étude approfondie de cinq règles du DIH en matière de ciblage, à savoir l'interdiction des attaques indiscriminées, le principe de proportionnalité, l'obligation de minimiser les dommages collatéraux, l'exigence d'émettre des avertissements avant une attaque, ainsi que celle de suspendre ou d'annuler une attaque. Ce chapitre met l'accent sur la manière dont les spécificités techniques des SAA influencent le respect de ces règles. L'auteur précise notamment la portée exacte de la notion d'« attaque » dans le contexte des SAA, pour mieux définir les limites juridiques associées à leur emploi[27].

En ce qui concerne l'interdiction des attaques indiscriminées, l'auteur soutient que le déploiement d'un SAA en dehors des scénarios pour lesquels il a été spécifiquement conçu, ou sans validation préalable des cibles par un opérateur humain, devrait constituer en soi une violation de ce principe[28]. Par ailleurs, il insiste sur le fait que l'ensemble des mesures de contrôle contextuel identifiées au huitième chapitre doivent être considérées comme des obligations nécessaires à la réduction des risques, conformément au principe de précaution.

L'auteur approfondit au passage le principe de proportionnalité, en discutant notamment du moment précis auquel le calcul relatif aux dommages collatéraux potentiels doit être réalisé. Il traite de l'obligation faite aux commandants (*Deployer*) de suspendre ou d'annuler une attaque lorsque l'objectif militaire visé devient illicite ou lorsque les pertes humaines et matérielles apparaissent excessives par rapport à l'avantage militaire attendu[29].

25. *Ibid* à la p 167.
26. *Ibid*.
27. *Ibid* à la p 229.
28. *Ibid* à la p 228.
29. *Ibid* à la p 244.

Enfin, Jonathan Kwik rappelle que même si les SAA peuvent contribuer au respect de certains principes de DIH, c'est principalement au commandant humain (*Deployer*) qu'incombe la responsabilité juridique ultime de s'assurer que ces règles soient pleinement respectées tout au long du processus de ciblage[30].

Le dixième chapitre aborde la question de l'utilité opérationnelle des SAA. Il souligne que de nombreux avantages attribués à ces systèmes ne leur sont pas nécessairement exclusifs, et qu'en conséquence, les commandants militaires doivent impérativement balancer les avantages potentiels, avec les risques liés à leur utilisation[31].

La cinquième partie, « Criminal Liability », traite des défis spécifiques que posent les SAA en matière de responsabilité pénale individuelle. Le onzième chapitre introduit brièvement la problématique du « déficit de responsabilité » qui caractérise fréquemment les débats sur les SAA. Le douzième chapitre approfondit cet examen en expliquant comment les caractéristiques techniques des SAA compliquent l'attribution de responsabilité aux individus chargés de leur déploiement. Selon l'auteur, les propriétés des SAA compliquent la démonstration des éléments constitutifs nécessaires à l'attribution d'une responsabilité pénale. Dans le cadre des SAA, l'*actus reus* (l'acte matériel) devrait correspondre à la décision humaine initiale d'activer le système dans des circonstances précises, plutôt qu'aux actions effectuées par le système sur le champ de bataille.

Le cœur du problème réside néanmoins dans la difficulté d'établir la *mens rea* (l'élément mental), indispensable pour engager une responsabilité pénale. La complexité technique et l'opacité des modèles d'IA réduisent considérablement la capacité des commandants à percevoir ou anticiper les risques liés à leur utilisation. Ce problème varie toutefois en fonction du niveau d'intention du *Deployer*. Certaines formes de *mens rea*, notamment celles correspondant à des actes intentionnels ou réalisés en connaissance de cause (*purposely* ou *knowingly*), sont compatibles avec l'emploi des SAA et ne posent généralement pas de problème pour l'attribution de responsabilité. Les situations de négligence ne génèrent pas non plus d'obstacles significatifs. Le véritable défi est lié aux degrés intermédiaires d'intention liés à la prise de risque, où les caractéristiques techniques de l'IA rendent difficile, voire impossible, d'évaluer précisément la nature et l'étendue des risques associés au déploiement des SAA. L'auteur qualifie ces risques de « génériques » ou « d'inconnus »[32].

30. *Ibid* à la p 267.
31. *Ibid* à la p 275.
32. *Ibid* à la p 315.

Enfin, le treizième chapitre souligne une situation particulièrement délicate, celle de l'ignorance délibérée (*manufactured ignorance*). Cette dernière se manifeste lorsque les commandants pourraient obtenir une connaissance plus précise des risques associés à un SAA, notamment en analysant les résultats des déploiements antérieurs, mais négligent volontairement cet effort pour des raisons opérationnelles ou pratiques[33].

Finalement, la sixième partie, « Conclusions », rassemble les constats et les recommandations issues des chapitres précédents, et propose deux contributions distinctes. La première synthétise les principaux enseignements dégagés de chaque thème abordé, et identifie des axes prioritaires pour de futures recherches. L'auteur conclut que les *Deployers* doivent assumer un rôle central tout au long du processus de ciblage mené par les SAA, afin de garantir le maintien d'une forme de contrôle tout au long de l'attaque – depuis la phase préparatoire, jusqu'à l'évaluation post-déploiement – permettant ainsi de respecter les obligations prévues par le DIH. La seconde contribution, d'ordre plus appliqué, prend la forme d'un guide opérationnel destiné aux *Deployers*, en leur fournissant un guide pratique applicable sur le terrain[34].

L'ouvrage de Jonathan Kwik se distingue par la profondeur technique de son examen, en combinant les aspects technologiques des SAA (les capteurs, les indicateurs de performance ou encore le processus décisionnel), à une décomposition des règles de ciblage du DIH. Cette approche permet de produire une réflexion complète sur la compatibilité entre des éléments techniques complexes et les exigences juridiques en matière de conduite des hostilités. Toutefois, cette technicité constitue à la fois la force et la faiblesse de cet ouvrage. D'une part, la complexité de l'analyse réduit l'accessibilité de celui-ci. Les concepts abordés et le niveau de détail exigent des connaissances préalables, tant en au niveau du DIH que de l'IA ou des technologies militaires autonomes. Cet ouvrage est ainsi destiné à des spécialistes ou à des lecteurs déjà familiers avec ces éléments. D'autre part, l'accent mis sur les conditions techniques d'emploi des SAA tend à reléguer au second plan l'analyse juridique. Les règles de DIH sont parfois survolées ou réduites à quelques références, notamment liées à la notion de « MHC ».

L'une des grandes forces du travail de Kwik réside dans son attention particulière à l'environnement opérationnel, élément trop souvent négligé par la

33. *Ibid* à la p 339.
34. *Ibid* à la p 389.

doctrine[35]. En intégrant cette dimension dans l'évaluation de la licéité des SAA, l'auteur enrichit considérablement le débat et offre un cadre d'analyse plus complet et plus ancré dans la réalité des opérations militaires contemporaines.

35. Tim McFarland, *Autonomous Weapon Systems and the Law of Armed Conflict: Compatibility with International Humanitarian Law*, Cambridge, Cambridge University Press, 2020 ; Afonso Seixas-Nunes, *The Legality and Accountability of Autonomous Weapon Systems: A Humanitarian Law Perspective*, Cambridge, Cambridge University Press, 2022 ; Jai Galliott, Duncan Macintosh et Jens David Ohlin, dir, *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, Oxford, Oxford University Press, 2021 à la p 202 ; Edward Hunter Christie et al, « Regulating lethal autonomous weapon systems : exploring the challenges of explainability and traceability » (2024) 4 AI & Ethics 229 ; Afia Kada, Bettayeb Miloud et Merah Ahmed, « AI-Powered Autonomous Weapons and IHL » (2025) 8 :1 J Leg & Econs Research 1070 ; Laura Bruun, Marta Bo et Netta Goussac, *Compliance with International Humanitarian Law in the Development and Use of Autonomous Weapon Systems : What Does IHL Permit, Prohibit and Require?*, Stockholm International Peace Research Institute, 2023 à la p 8 ; Vincent Boulanin, Laura Bruun et Netta Goussac, *Autonomous Weapon Systems and International Humanitarian Law : Identifying Limits and the Required Type and Degree of Human–Machine Interaction*, Stockholm International Peace Research Institute, 2021. Netta Goussac, *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human–Machine Interaction,* Stockholm International Peace Research Institute, 2021.