

Ethique et systèmes d'armes autonomes : faire les bons choix

Par Jean-François Pelliard, directeur de l'Observatoire "Marine 2040", chercheur associé à la FMES.

Les débats éthiques sur les systèmes d'armes autonomes sont concentrés sur les armes létales, suffisamment "intelligentes" pour décider d'engager une cible tout en s'adaptant aux circonstances. Pourtant des questions éthiques sont posées dès les premiers niveaux d'autonomie. Parallèlement, l'intelligence artificielle qui procure cette autonomie permet, quand elle est utilisée à bon escient, de faciliter le respect de certains principes du droit humanitaire international. A cet égard, les rôles respectifs de l'homme et de la machine dans le processus décisionnel est déterminant. Ces débats méritent donc d'être élargis et les systèmes autonomes, sans en nier les risques, d'être considérés de manière moins manichéenne.

L'emploi de systèmes autonomes dans les opérations militaires soulève des questions techniques et juridiques, abordées dans deux précédents articles [voir : [L'intégration des drones navals dans une marine de combat](#) et [Le drone naval peut-il être un navire de guerre?](#)]. Il soulève également des questions éthiques, récurrentes pour les systèmes d'armes d'apparition récente et encore en devenir. Mais dans le cas des systèmes autonomes, recourant à l'intelligence artificielle (IA), apparaît la crainte spécifique que la technologie décide de l'emploi de la force armée à la place de l'homme.

Où fixer les limites de la technologie et des conditions d'emploi de ces systèmes et de l'IA pour se conformer à une éthique qui correspond aux valeurs des démocraties occidentales, sans pour autant faire preuve de naïveté et laisser ainsi des pays compétiteurs prendre une avance potentiellement décisive ? Aujourd'hui largement concentré sur les systèmes d'armes létales autonomes (SALA), ou robots tueurs, le débat mérite d'être élargi à d'autres systèmes en s'intéressant aux places respectives de l'homme et de la technologie dans la chaîne de décision. Ainsi le rapport entre autonomie des systèmes et respect des principes du droit international humanitaire pourrait bénéficier d'un regard moins manichéen.



Image d'un système d'arme létale autonome ou « robot tueur ». Source : [CICR](#).

Des questions éthiques soulevées dès les premiers niveaux d'autonomie

Dans un rapport d'information de 2020, les députés Claude de Ganay et Fabien Gouttefarde¹ s'intéressent aux SALA qu'ils définissent ainsi : « systèmes capables de choisir et d'engager seuls une cible, sans intervention humaine, dans un environnement changeant ». Chacun de ces mots a son importance pour caractériser les SALA tels que la technologie permet aujourd'hui de les concevoir². En effet une simple mine, arme du pauvre, est un système capable d'engager seul une cible dès lors que celle-ci, par son poids, sa masse magnétique ou le bruit qu'elle émet, et bien sûr par sa proximité, déclenche automatiquement le système de mise à feu. Des systèmes « plus intelligents » sont capables de se mouvoir vers une cible, de la choisir, selon sa signature acoustique ou thermique par exemple. Pour autant, ils ne sont pas des SALA car leur action est strictement encadrée par l'homme qui applique des règles d'engagement parfaitement définies. Ces règles doivent être comprises comme des conditions à réunir pour faire usage de la force. Elles découlent à la fois du cadre politique fixé à une intervention et de l'analyse des risques et des rapports de force sur le terrain. Dans la définition donnée des SALA, figure l'idée que le système analyse lui-même l'environnement pour déterminer si ces conditions sont réunies avant d'engager. Il est même intellectuellement concevable et sans doute technologiquement possible, à moyen terme, que le système précise les règles d'engagement à partir d'un cadre très général qui lui serait fixé. Exprimé autrement, les règles d'engagement fixées par l'homme pourraient alors devenir plus floues.

Cette capacité à s'adapter aux évolutions de l'environnement est conférée aux systèmes autonomes par l'IA. Sémantiquement, il est tentant de la rapprocher de l'intelligence humaine souvent définie comme une capacité à s'adapter aux situations nouvelles³. L'emploi trompeur du même substantif alimente un imaginaire collectif déjà abreuvé par des œuvres de science-fiction. Il existe donc une part de fantasme dans ce que seraient capables de réaliser les SALA. Pour autant, les questions éthiques sont légitimes car sans la rejoindre, la réalité se rapproche de la fiction. La militante américaine Jody Williams, prix Nobel de la paix en 1997 pour son rôle dans la campagne ayant conduit à l'interdiction des mines anti-personnel, s'engage aujourd'hui contre les SALA. Elle se place dans un champ moral en considérant que la prérogative humaine de tuer est une condition de la dignité humaine. Non définis juridiquement et donc non encadrés, les SALA soulèvent d'abord la question du respect des principes du droit international humanitaire (DIH). Parmi ces principes, les interprétations du devoir d'humanité dans l'action militaire peuvent diverger. Sur le plan philosophique, les SALA en sont bien sûr dépourvus.

Mais il serait dangereux de limiter la réflexion éthique à ces drones tueurs totalement autonomes. Comme les véhicules, les systèmes d'armes peuvent être catégorisés selon leur niveau d'autonomie. Par exemple, les systèmes téléopérés, capables de se mouvoir et de naviguer dans la plupart des situations sans intervention humaine majeure, requièrent un opérateur pour l'exécution de leur mission. C'est le cas du drone MALE *Reaper* et de la quasi-totalité des drones navals de surface actuels. Or la distanciation entre l'opérateur et le champ de bataille soulève des questions éthiques. Elle limite la

¹ Assemblée Nationale - Rapport d'information N° 3248 sur les systèmes d'armes létaux autonomes - enregistré le 22 juillet 2020

https://www.assemblee-nationale.fr/dyn/15/rapports/cion_def/l15b3248_rapport-information.pdf

² En 2020, les parlementaires considéraient que de tels systèmes n'existaient pas encore. Il est difficile de l'affirmer aujourd'hui.

³ Définition du « Dictionnaire de la philosophie » de Larousse : ensemble des fonctions psychologiques d'adaptation pratique aux situations nouvelles qui mobilisent des compétences abstraites dans le traitement des problèmes.

perception de la réalité de la mission. La distanciation physique induit une distanciation psychologique : elle atténue d'autant plus l'empathie pour le combattant adverse que les conditions du combat, les risques pris, ne sont pas symétriques. L'opérateur vit la scène sans en réaliser toutes les dimensions dans l'instant. Il déclenche le tir de missiles sur des silhouettes dépersonnalisées dans le confort relatif d'une salle de contrôle. L'expérience américaine de l'emploi des *Reaper* en Afghanistan et en Irak a cependant révélé une prise de conscience a posteriori avec des conséquences importantes sur la santé mentale des anciens opérateurs. Certains d'entre eux s'interrogent même sur le respect des principes de distinction et de discrimination du DIH. Au-delà de l'éthique du combattant, l'éthique du donneur d'ordres vis-à-vis de ses soldats est alors en question, même si ces situations éloignent les soldats d'un risque physique.

Plus généralement, de nombreux systèmes à divers degrés d'autonomie recourent à l'IA. Contrairement aux algorithmes classiques, dits déterministes, ceux utilisés par l'IA ne permettent pas de garantir un résultat. C'est la différence majeure entre l'automatisme et l'autonomie de décision. Ce résultat est par ailleurs dépendant de la qualité des données utilisées pour l'apprentissage de la machine et de la représentativité de ces données pour l'environnement particulier de mise en œuvre réelle du système : le périmètre de validité des modèles auxquels elles aboutissent est limité. Un système optronique entraîné à identifier des cibles menaçantes aux abords du port de Toulon, rade bordée par des reliefs et pratiquée par des navires militaires et de plaisance, ne sera pas adapté au port d'Abu Dhabi, proche du désert et fréquenté par les navires traditionnels des Émirats. Dans ces conditions, la confiance accordée à un système pour qu'il respecte le cadre qui lui a été fixé ne peut être absolue. Cette incertitude combine des dimensions d'efficacité et d'éthique. Elles sont évidentes quand l'IA contribue directement à une action létale. La nouvelle torpille française F21 recourt à l'IA dans le difficile travail de trouver la cible qui lui a été assignée et de l'engager sans être auparavant neutralisée. Ce faisant, peut-elle se tromper de cible ? L'environnement et l'éventuelle présence de tiers conditionnent la décision de tir qui reste humaine. Mais dans d'autres cas, l'IA influence la décision humaine d'engager la force létale. Face au flux de plus en plus important d'informations arrivant au décideur, elle filtre, hiérarchise et propose des options. L'éthique autant que l'efficacité imposent que l'IA donne au décideur, autant que des conseils, des indices de confiance sur la qualité de ce qu'elle produit. Elles imposent également au décideur une certaine prise de recul. Malgré ces interrogations légitimes, il convient de reconnaître aux systèmes autonomes des atouts que n'ont pas les armements plus classiques.

Pour revenir aux termes de la définition des SALA, leur capacité à choisir une cible, à discriminer conformément à un principe du DIH, peut être considérée comme favorable à l'éthique. Au contraire, quoi de plus aléatoire qu'une mine antipersonnel incapable de distinguer le soldat du non combattant ? Qu'un système dormant capable de cibler précisément un adversaire mais persistant sans être capable de s'adapter à l'environnement ? Avec une dose d'incertitude qu'il sera toujours difficile de mesurer, l'IA conçue selon certains critères et les systèmes d'armes qui y recourent ont donc également une capacité à minimiser certains risques éthiques. Cet aspect est développé plus loin.

Ainsi l'autonomie que les technologies peuvent conférer aux systèmes d'armes soulève effectivement des questions éthiques, mais elles peuvent être relativisées par rapport à celles posées par des générations d'armes plus anciennes. C'est bien l'usage qui en est fait et la manière dont les systèmes sont intégrés dans la chaîne de mise en œuvre qui restent déterminants.

Le critère essentiel des places respectives de l'homme et des technologies dans la chaîne de décision

Le rapport parlementaire cité plus haut constate justement que tout système d'armes, dont l'acception est étendue ici à l'ensemble des systèmes mis en œuvre dans une manœuvre militaire, qu'ils soient ou non létaux, appartient à une chaîne de commandement. Celle-ci s'exerce de la définition de la mission jusqu'à son exécution, et même dans son exploitation ultérieure. Or, sauf à se projeter dans le monde

futur de Terminator⁴ où l'humanité mène une guerre existentielle contre les machines, le rôle de l'homme dans cette chaîne reste dominant. Une guerre entre drones exclusivement n'aurait pas grand sens car finalement, la guerre est un choc des volontés humaines. Le rôle de l'homme est dominant aux échelons politique et stratégique. Il l'est également à l'échelon tactique dans la boucle OODA (Observer – Orienter – Décider – Agir) dont l'accélération est tout à la fois permise par nos technologies et rendue nécessaire par les technologies identiques ou concurrentes utilisées par l'adversaire.

Le contrôle de l'homme sur un système d'armes plus ou moins autonome s'exerce donc d'abord dans ce cadre. Le système doit accomplir une mission définie par l'homme. Pour cela, il convient de réunir au moins deux parmi les trois critères essentiels suivants⁵. Le premier est de correctement encadrer la mission, d'en fixer les limites. C'est ce que l'industriel Naval Group appelle justement « l'autonomie décisionnelle contrôlée » (ADC®), qu'il développe pour ces systèmes de drones navals. Le cadrage comporte bien sûr des directives sur l'objectif et la manière de réaliser la mission. Mais il peut également très simplement contraindre l'action dans le temps et dans l'espace, évitant ainsi, par exemple, tout risque dans une zone de forte activité civile. Ces principes autorisent par ailleurs un fonctionnement autonome d'un drone normalement téléopéré lorsque la liaison avec son contrôleur est perdue. Le deuxième critère, déjà évoqué, est celui de la confiance dans la capacité du système à respecter ce cadrage et plus largement à réaliser ce qui est attendu de lui dans les conditions qui lui ont été fixées. Plus la confiance sera élevée, plus l'homme acceptera de confier au système des missions complexes et risquées. L'éthique du militaire est ainsi d'une certaine manière déléguée à la machine. Et comme tout chef militaire sait que la délégation ne dissout pas la responsabilité, la confiance dans la machine ne saurait être aveugle. L'éthique passe donc par une conscience éclairée des limites de la machine. C'est ici qu'intervient le troisième critère, selon les cas alternatif ou complémentaire du précédent, qui permet à l'homme de contrôler la mission : le veto. Malgré la confiance, être en mesure d'empêcher le drone d'agir est la meilleure garantie que son action restera dans le cadre éthique défini. Cela suppose cependant une liaison permettant de percevoir son comportement et son environnement, et de transmettre l'ordre de veto. La permanence de cette liaison ne peut malheureusement pas toujours être assurée, en particulier dans le milieu sous-marin, mais aussi dans des environnements contestés, notamment en cas de brouillage.

Aucun de ces trois critères n'est absolu et ne peut être strictement garanti. Le risque éthique des systèmes autonomes ne peut donc être écarté. Pour autant, faut-il les interdire ou en limiter le développement ? La France aujourd'hui, comme toutes les démocraties occidentales, exclut la disparition complète de l'homme dans la boucle de mise en œuvre de la force létale. Qu'en sera-t-il de nos principaux compétiteurs ? La Chine, très avancée, reste discrète sur ses intentions et d'éventuels projets. L'administration Trump ne s'est pas clairement prononcée sur la question. Que vaudrait une interdiction même partielle dans un monde où le respect du droit international est à géométrie variable ? Sans nier les risques représentés par ces systèmes, la question pourrait être inversée : peuvent-ils au contraire contribuer au respect des principes du DIH⁶ ?

⁴ Un futur aujourd'hui proche car le scénario du film, sorti aux États-Unis en 1984, place cette guerre en 2029.

⁵ Dans ses travaux consacrés aux SALA, Jean-Baptiste Jeangène Vilmer propose un code de conduite plus limitatif dans l'emploi de ces systèmes, code dont certaines règles se rapprochent du veto mentionné ici. Mais surtout, il recommande d'éviter l'emploi de ces armes létales dans certains contextes, par exemple en milieu urbain, et qu'elles ne s'en prennent qu'aux objets et non aux humains.

⁶ Ces principes dont certains ont été évoqués plus haut, peuvent être résumés ainsi : la distinction de la nécessité militaire, conjuguée à la proportionnalité des moyens employés, la discrimination entre combattants et non combattants, et malgré les conflits une certaine humanité qui interdit de causer des maux superflus.

L'autonomie des systèmes d'armes peut favoriser le respect des principes du DIH

Dans le cas des drones, systèmes d'armes autonomes mobiles, la capacité à observer au plus près le terrain, avec une meilleure persistance, est objectivement un facteur d'amélioration de l'information. Par là-même, elle facilite la discrimination des situations et la distinction entre combattants et non combattants dès lors que la volonté de le faire est présente. Une vidéo prise à quelques centaines de mètres ou moins, transmise en temps réel, est beaucoup plus facilement interprétable qu'un ensemble d'informations synthétiques issues de radars ou d'équipements de guerre électronique. Même quand l'information est partielle, complexe à interpréter, qu'elle provienne ou non de systèmes autonomes, l'intelligence artificielle améliore la perception de la réalité et là encore permet de mieux discriminer. Le chef militaire peut alors adapter le niveau d'emploi de la force au juste besoin pour produire les effets recherchés. Le compromis entre sécurité éthique et sécurité de la mission n'en devient que plus facile à établir. Cette perception optimiste reste malgré tout aujourd'hui assez théorique. Elle peut cependant orienter le progrès technologique.

Enfin il est permis de s'interroger sur les faillibilités respectives de l'homme et du système autonome. La responsabilité humaine dans la destruction d'un Airbus iranien par l'USS Vincennes, en 1988, a été reconnue par les États-Unis. Une IA serait sans doute capable aujourd'hui, dans un cas similaire, d'alerter l'opérateur radar qui à l'époque n'avait pas correctement interprété la cinématique de l'avion, erreur déterminante ayant conduit à l'ordre de tir. L'IA peut donc limiter le risque d'erreur humaine. A l'avenir, des systèmes autonomes seront sans doute statistiquement plus fiables que l'homme dans l'engagement du feu. Le prouver restera une gageure. Et vaut-il mieux en cas d'erreur être la victime d'une machine ou d'un homme ? Derrière la statistique réside un risque difficile à faire accepter par le politique et par l'opinion publique. Et n'existent que des cas particuliers impliquant des individus. La question est aujourd'hui purement philosophique et chacun peut avoir sa réponse. Elle sera posée demain de manière plus concrète et peut-être plus clivante, opposant les tenants de la technologie aux sceptiques, tous pouvant se réclamer de valeurs humanistes.

Même s'ils en concentrent les enjeux, les SALA ne sont pas les seuls systèmes autonomes soulevant des questions éthiques. Ces questions doivent être considérées sans oublier que l'autonomie et l'IA peuvent aussi contribuer au respect des principes du DIH.

Croire en la supériorité de l'intelligence humaine sur la machine peut aboutir à lui confier la technologie la plus évoluée, en lui laissant le soin d'en faire le meilleur usage. L'Histoire récente n'autorise pas cette naïveté. Alors qu'une mise à jour de la Convention de 1980 sur certaines armes classiques, intégrant les SALA, est souhaitée par certains pays et organisations, il conviendra de trouver le bon compromis pour protéger sans être trop restrictif, sous peine de voir cette Convention respectée uniquement par les pays qui l'accepteraient. En l'absence prolongée de cadre international, les démocraties occidentales ne devront pas elles non plus tomber dans une certaine naïveté qui les laisserait sans défense. La bataille pour arriver à un consensus contraignant sur le sujet de l'illégitimité ou non des armes (nucléaires, biologiques, mines antipersonnel, ... et SALA) doit être menée avec la lucidité et le sens des responsabilités nécessaires dans un monde qui se brutalise chaque jour un peu plus.