

Droits de l'homme, automatisation et déshumanisation des prises de décisions létales : les systèmes d'armement autonomes doivent-ils être interdits ?

Peter Asaro*

Philosophe des technologies, le professeur Peter Asaro a mené des travaux sur l'intelligence artificielle, les réseaux de neurones, le traitement du langage naturel et la vision robotique. Il est chercheur associé au Centre de recherche sur l'internet et la société de la faculté de droit de l'université de Stanford, aux États-Unis, et vice-président du Comité international pour le contrôle des armes robotisées, dont il est cofondateur. Il est également directeur des programmes de 2^e et 3^e cycles à l'École d'études des médias de la *New School For Public Engagement*, à New York City, États-Unis.

Résumé

Le présent article passe en revue la littérature consacrée récemment aux systèmes d'armement autonomes et à leur éventuelle interdiction internationale. Répondant aux commentateurs qui estiment que cette prohibition serait contestable à plusieurs titres, l'article montre qu'une telle interdiction trouverait son fondement théorique dans les normes relatives aux droits de l'homme et dans les principes humanitaires, non seulement moraux mais aussi juridiques. Ainsi, l'une des exigences implicites du droit international humanitaire, qui régit les conflits armés, est l'exercice d'un jugement humain. Cette exigence

* La version originale en anglais de cet article est publiée sous le titre « On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making », dans *International Review of the Red Cross*, Vol. 94, N° 886, été 2012, pp. 687-709.

figure implicitement dans trois principes cardinaux – distinction, proportionnalité et nécessité militaire – qui sont énoncés dans des traités internationaux tels que les Conventions de Genève de 1949 et qui sont fermement établis en droit international coutumier. Des principes similaires sont aussi implicites dans le droit international relatif aux droits de l'homme qui garantit, en toutes circonstances, certains droits fondamentaux à tous les êtres humains, indépendamment de leurs origines nationales ou de la législation locale. J'estime pour ma part qu'il existe une obligation spécifique qui s'applique à toute une gamme de systèmes automatisés et autonomes. Cette obligation découle de deux droits fondamentaux de la personne humaine – le droit à la vie et le droit à un procès équitable – et des conditions restreintes dans lesquelles il peut être dérogé à ces droits. Les individus et les États en temps de paix et en situation de conflit armé, les combattants, les organisations militaires et les États sont tenus de ne déléguer en aucun cas à des machines ou à des processus automatisés ni l'autorité ni la capacité d'employer la force létale si la légitimité morale et juridique d'un tel acte n'a pas été préalablement établie par un humain. Je soutiens qu'il serait bon que cette obligation soit établie comme norme internationale et que cela soit formalisé par un traité, avant que ne commencent à apparaître divers systèmes d'armes automatisées et d'armes autonomes susceptibles de menacer gravement les droits fondamentaux de tout individu.

Mots-clés : robots ; drones ; systèmes d'armement autonomes ; automatisation ; décisions létales ; droits de l'homme ; maîtrise des armements.



En septembre 2009, Jürgen Altmann, Noel Sharkey, Rob Sparrow et moi-même avons créé le Comité international pour le Contrôle des Armes robotisées (*International Committee for Robot Arms Control – ICRAC*)¹. Nous avons publié peu après une déclaration de mission qui contenait un appel au débat sur une éventuelle interdiction internationale des systèmes d'armement autonomes :

« Étant donné le rythme rapide du développement de la robotique militaire et les dangers pressants que ces avancées font peser sur la paix et la sécurité internationale, ainsi que sur les civils en temps de guerre, nous demandons à la communauté internationale d'entamer d'urgence un débat sur l'établissement d'un régime de maîtrise des armements visant à réduire la menace que ces systèmes représentent. Nous proposons que le débat porte sur l'interdiction de la mise au point, du déploiement et de l'emploi de systèmes d'armement autonomes télépilotés ; des machines ne devraient pas être autorisées à prendre la décision de tuer des êtres humains »².

1 Voir www.icrac.net [Sauf mention contraire, tous les sites internet ont été consultés en août 2013]

2 Jürgen Altmann, Peter Asaro, Noel Sharkey et Robert Sparrow, *Mission Statement of the International Committee for Robot Arms Control*, 2009, disponible sur : <http://icrac.net/statements/> [Traduction CICR].

Depuis lors, philosophes, juristes, officiers militaires, décideurs politiques, scientifiques et roboticiens se sont emparés de cette question. Au départ, les débats étaient surtout centrés sur l'incapacité, pour les systèmes d'armement autonomes existant déjà, de satisfaire aux exigences du droit international humanitaire (DIH) ; la possibilité que les futures technologies leur confèrent cette capacité a donné lieu à diverses conjectures. Deux questions retenaient alors particulièrement l'attention. D'une part, les systèmes armés autonomes sont-ils (ou seront-ils) capables de respecter les principes de distinction et de proportionnalité, comme l'exigent les Conventions de Genève ? D'autre part, sera-t-il possible d'imputer à quiconque la responsabilité des dommages ou des pertes que de tels systèmes pourraient causer de manière illicite ? À l'issue des discussions initiales, l'attention a commencé à se porter sur la question de savoir si le DIH a besoin d'être complété par un traité international interdisant explicitement ces technologies. Bien que la grande majorité du public et un certain nombre de spécialistes – universitaires, juristes, officiers militaires et ingénieurs – sont aujourd'hui d'avis que les systèmes létaux ne devraient pas être autonomes, certains commentateurs estiment qu'il pourrait être prématuré, inutile et même immoral d'établir une interdiction des systèmes d'armement autonomes au niveau international³. Je considère pour ma part que leur position est erronée et que nous devons agir sans tarder afin d'interdire ces systèmes. Je pense en effet que nous sommes tenus, pour des raisons d'ordre moral et juridique, d'empêcher toute délégation de l'autorité létale à des systèmes non humains et non supervisés. Il importe en outre que, dans nos travaux de recherche en science et en ingénierie, ainsi que dans l'affectation de nos ressources de développement, nous poursuivions l'objectif d'accroître la performance éthique des décideurs humains. Le présent article exposera donc un fondement théorique pour une interdiction internationale des systèmes d'armement autonomes reposant sur le droit international des droits de l'homme (DIDH) et sur le droit international humanitaire (DIH). Outre leur consécration et leur protection par un corpus juridique important (tant au plan international qu'interne), les droits humains ont également un statut moral, indépendant du droit en vigueur. Il est donc possible d'y puiser des conseils judicieux quant à l'extension du droit afin de répondre aux défis posés par les technologies émergentes. Je présenterai l'argument selon lequel une interdiction internationale des systèmes d'armement autonomes peut être fermement établie sur le principe que l'autorité de décider d'employer la force létale ne peut pas être légitimement déléguée à un processus automatisé ; cette autorité doit rester la responsabilité d'un agent humain qui est tenu de prendre une décision réfléchie et informée.

3 Ronald C. Arkin, *Governing Lethal Behavior in Autonomous Robots*, CRC Press, 2009 ; Gary Marchant, Braden Allenby, Ronald C. Arkin, Edward T. Barrett, Jason Borenstein, Lyn M. Gaudet, Orde F. Kittrie, Patrick Lin, George R. Lucas, Richard M. O'Meara et Jared Silberman, « International governance of autonomous military robots », dans *Columbia Science and Technology Law Review*, 30 décembre 2010, disponible sur : <http://ssrn.com/abstract=1778424> ; Kenneth Anderson et Matthew C. Waxman, « Law and ethics for robot soldiers », dans *Policy Review*, 28 avril 2012, disponible sur : <http://ssrn.com/abstract=2046375> .

Ce principe a des incidences sur plusieurs corpus juridiques, notamment le droit interne, le DIDH et le DIH. Étant donné que l'intérêt que suscite actuellement la mise au point de systèmes d'armement autonomes est principalement motivé par des applications militaires, je traiterai donc les incidences de ce principe sous l'angle du DIH. Ce même principe s'appliquerait toutefois à l'usage de systèmes d'armement autonomes par les États à des fins de police intérieure, de contrôle des mouvements de foule, de surveillance des frontières, de garde de prisonniers, ainsi que pour assurer la sécurité d'installations et du territoire, ou pour mener d'autres activités potentiellement meurtrières. Ces systèmes pourraient aussi être employés par des individus ou par des organisations pour toute une gamme d'utilisations relevant de la sécurité et impliquant l'emploi de la force. Je me focaliserai donc sur l'un des droits fondamentaux de la personne humaine – le droit à la vie – tout en sachant que des arguments similaires peuvent être invoqués à propos des décisions automatisées qui dérogent à d'autres droits humains, ou en empêchent l'exercice. Je pense notamment à l'automatisation de certaines activités (arrestation, détention et restriction de mouvement ; localisation, surveillance et poursuite ; déportation ; expulsion et saisie ; refus d'accès aux soins médicaux, interdiction de tenir des réunions publiques, suppression de la liberté de la presse et de la liberté d'expression, suppression des droits de vote ; et, enfin, suppression d'autres droits civils, politiques, économiques, sociaux et culturels)⁴.

Les systèmes d'armement autonomes

Un emploi accru de technologies hautement automatisées a été observé lors de récents conflits armés. Les drones armés pilotés à distance, utilisés dans un certain nombre de pays, notamment par les forces armées des États-Unis, en sont l'exemple le plus frappant. Ces aéronefs de combat sont capables d'effectuer de nombreuses procédures de vol automatisées très sophistiquées (par exemple, des opérations entièrement automatisées de décollage et d'atterrissage, des approches GPS autonomes ou le maintien en orbite à une altitude prédéfinie autour d'une cible localisée par GPS). Ces aéronefs possèdent aussi de nombreuses capacités en matière de collecte et de traitement automatisés des images. Bien qu'ils soient hautement *automatisés*, ces systèmes ne sont pas considérés comme *autonomes*, car ils sont encore placés sous la supervision et le contrôle direct d'un opérateur humain⁵. De plus, bien que ces systèmes soient porteurs d'armes dotées de cer-

4 Les droits de l'homme actuellement reconnus en droit international sont notamment, mais non exclusivement, les droits proclamés dans la Charte internationale des droits de l'homme des Nations Unies, composée de la Déclaration universelle des droits de l'homme, du Pacte international relatif aux droits civils et politiques et du Pacte international relatif aux droits économiques, sociaux et culturels.

5 Le terme « autonome » est utilisé en ingénierie pour qualifier un système qui fonctionne sans supervision ni contrôle direct d'un opérateur humain. Le terme « automatisé » s'applique à un système ou processus non supervisé, impliquant des opérations répétitives, structurées, de routine, effectuées avec peu de retour d'information (comme les lave-vaisselle), au contraire des systèmes ou processus

taines capacités automatisées (dans le cas, par exemple, des missiles guidés par laser et des bombes guidées par GPS), toutes les décisions de ciblage et de tir sont encore sous le contrôle direct d'un humain. Le présent article se développe autour des ramifications que peut avoir, sur les plans du droit et de l'éthique, l'automatisation des décisions de ciblage et de tir. Nous pouvons définir un « système d'armement autonome » comme étant tout système qui est capable de sélectionner sa cible et d'employer une force potentiellement meurtrière, sans aucune supervision ni aucune implication humaines directes dans la prise de décisions létales⁶. Si l'on s'en tient à cette définition, les aéronefs pilotés à distance utilisés aujourd'hui – tels que les drones de type *Predator* et *Reaper* – ne sont pas des systèmes d'armement autonomes. Néanmoins, il apparaît de plus en plus clairement que les activités qui, aujourd'hui, sont encore placées sous contrôle humain pourraient être automatisées dans un avenir proche, permettant ainsi l'élimination de toute intervention humaine directe pour sélectionner les cibles et décider d'employer la force létale contre elles. Il convient de relever que les aéronefs pilotés à distance ne sont pas le seul sujet de préoccupation. De nombreux systèmes déjà utilisés sur terre, en mer et sous la mer pourraient, eux aussi, être armés. De plus, certains systèmes défensifs fixes (tels que les tours de tir et les sentinelles) et divers modes opératoires des cyber-attaques pourraient être automatisés et devenir ainsi capables d'utiliser la force létale sans qu'aucun humain ne soit directement impliqué dans la sélection des cibles ou dans l'autorisation d'employer la force létale contre une cible donnée.

Certes, il existe divers exemples d'armes et de pratiques militaires qui se passent probablement d'intervention humaine directe dans la prise de décisions létales. Toutefois, la nouvelle vague de capacités technologiques a suscité, tant au sein de la communauté du droit international que parmi les militaires de carrière, de vives inquiétudes et interrogations quant à la légitimité morale et juridique de ces systèmes. Ainsi, comme l'a déclaré Jakob Kellenberger, alors président du Comité international de la Croix-Rouge (CICR), lors de la Conférence de San Remo (Italie) en septembre 2011 :

« robotisés » ou « autonomes » (comme les voitures sans conducteur, par exemple), qui opèrent, quant à eux, dans des environnements ouverts, dynamiques et non structurés, sur la base d'informations provenant de différents capteurs. Tous ces systèmes – malgré les distinctions ci-dessus et bien qu'ils restent tributaires des données imprédictibles fournies par les capteurs – obéissent, d'une part, à des instructions algorithmiques presque entièrement fixes et déterministes et, d'autre part, à des calculs de probabilité étroitement circonscrits (parfois utilisés à des fins d'apprentissage et de correction d'erreurs).

6 Je préfère utiliser le terme de « système d'armement autonome » plutôt que celui d'« arme autonome », de manière à signifier que chacun de ces systèmes peut être réparti entre divers éléments disparates qui, néanmoins, fonctionnent ensemble pour constituer un système d'armement autonome. Par exemple, un ordinateur se trouvant presque n'importe où dans le monde pourrait recevoir des données transmises par un drone de surveillance et utiliser ces informations pour lancer et diriger une frappe effectuée à l'aide d'un système d'armes téléguidées se trouvant dans un tout autre point du globe, tout cela sans intervention ni supervision humaines : cet ordinateur constituerait ainsi un « système d'armement autonome ». En d'autres termes, les différents éléments qui composent ces systèmes – à savoir, les capteurs, les moyens autonomes de ciblage et de prise de décision, et l'arme elle-même – n'ont pas besoin d'être directement attachés les uns aux autres ou de se trouver au même endroit : il suffit qu'ils soient connectés par des réseaux de communication.

Un système [d'armement] véritablement autonome serait doté d'une intelligence artificielle qui devrait être capable de mettre en œuvre le DIH. Bien que ce domaine suscite un grand intérêt et que la recherche soit largement financée, ces systèmes n'ont pas encore été adaptés aux armements. Développer de tels systèmes est un tel défi en termes de programmation que ce sera peut-être impossible. Il est clair que le déploiement de tels systèmes représenterait une véritable révolution conceptuelle et un changement qualitatif majeur dans la conduite des hostilités. Mais il soulèverait aussi tout un ensemble de problèmes fondamentaux du point de vue légal, éthique et sociétal, et ces problèmes doivent être pris en compte avant que ces systèmes ne soient développés ou déployés. Un robot pourrait être programmé de façon à se comporter de façon plus éthique et plus prudente qu'un être humain sur le champ de bataille. Mais que faire si, du point de vue technique, il est impossible de réaliser une programmation fiable d'un système d'armement autonome de façon à garantir qu'il respecterait le DIH sur le champ de bataille ? ... Cela étant, l'application de règles juridiques préexistantes à une technologie nouvelle soulève la question de savoir si ces règles sont suffisamment claires au vu des caractéristiques spécifiques – et peut-être sans précédent – de cette technologie, et également au vu de l'impact humanitaire qu'elle peut avoir dans un avenir prévisible. Dans certaines circonstances, les États choisiront, ou ont déjà choisi, d'adopter des règles plus spécifiques⁷.

Ainsi, comme Jakob Kellenberger l'a clairement indiqué, de graves interrogations subsistent quant à savoir si les technologies autonomes seront techniquement capables de se conformer aux dispositions actuelles du DIH. Nombreux sont les militaires qui reconnaissent que la technologie évolue dans le sens d'une autonomie accrue des systèmes d'armes létales ; toutefois, la plupart d'entre eux (y compris certains décideurs politiques du Bureau du Secrétaire à la Défense des États-Unis) expriment de graves préoccupations éthiques :

S'il est essentiel d'imposer des limites aux armes autonomes pour assurer des engagements éthiques, la tâche la plus difficile consiste à créer des armes autonomes « à sûreté intégrée ». L'environnement en temps de guerre, dans lequel les systèmes militaires opèrent, est à la fois confus et compliqué et, malgré cela, les systèmes autonomes doivent être capables de fonctionner de façon appropriée. L'adaptation de l'ennemi, la dégradation des communications, les risques pour l'environnement, la présence de civils sur le champ de bataille, les cyber-attaques, les dysfonctionnements, et les tensions en période de guerre sont autant d'éléments qui introduisent la possibilité que les systèmes autonomes soient confrontés à des situations non anticipées et

7 Jakob Kellenberger, « Discours d'ouverture », Le droit international humanitaire et les nouvelles technologies de l'armement, XXXIV^e table ronde sur les sujets actuels du droit international humanitaire, San Remo, Italie, 8-10 septembre 2011, pp. 5-6, disponible sur : <http://www.icrc.org/fre/resources/documents/statement/new-weapon-technologies-statement-2011-09-08.htm> .

que, par conséquent, ils agissent d'une façon qui n'a pas été voulue. Même les algorithmes relativement sophistiqués ne sont pas à l'abri d'une défaillance quand ils rencontrent une situation échappant aux paramètres intégrés lors de leur conception car, à la différence des êtres humains, ils ne possèdent pas l'intelligence contextuelle que l'on nomme parfois « bon sens » ou « sens commun ». La complexité des ordinateurs modernes vient encore compliquer le problème, car il est d'autant plus difficile d'anticiper tous les dysfonctionnements possibles ou tout comportement émergent, susceptible de survenir quand un système est mis en service⁸.

Même dotés d'intelligence artificielle, les systèmes autonomes doivent être préprogrammés et ils n'ont, au mieux, que des capacités très limitées en termes d'apprentissage et d'adaptation. Il sera donc difficile, voire impossible, de concevoir des systèmes capables de fonctionner malgré le « brouillard » et les tensions qui caractérisent le temps de guerre. Si l'on examine les incidences que cela peut avoir pour la protection des civils dans les conflits armés, diverses questions d'ordre éthique et juridique apparaissent. Tout d'abord, comment sera assuré le respect, exigé par le DIH, des principes de distinction, de proportionnalité et de nécessité militaire ? Ensuite, en cas d'emploi de la force létale, comment pourra-t-on établir les responsabilités et faire en sorte que les auteurs rendent compte de leurs actes ?

Maintes inquiétudes d'ordre éthique et sociétal sont suscitées par l'emploi de systèmes d'armement autonomes. Ces préoccupations sont notamment dues aux problèmes de la guerre asymétrique et de la redistribution des risques (qui pèsent non plus sur les combattants, mais sur les civils), ainsi qu'à la crainte de voir s'abaisser le seuil au-delà duquel les nations décident de livrer une guerre⁹. En effet, les systèmes d'armement autonomes tendent à éloigner de la zone de conflit les combattants qui les utilisent et à réduire les risques de pertes humaines pour ceux qui les possèdent. Ils ont donc pour effet de réduire les coûts et les risques politiques d'une guerre, ce qui pourrait se traduire par un abaissement général du seuil de l'entrée en guerre. Par ailleurs, les systèmes d'armement autonomes sont également susceptibles de provoquer de l'instabilité et de l'insécurité aux niveaux régional ou mondial, d'alimenter la course aux armements, de proliférer et tomber aux mains d'acteurs non étatiques, ou encore de provoquer l'escalade de certains conflits en l'absence d'intentions politiques humaines. La force létale pourrait être employée sans supervision humaine par des systèmes dotés de cette capacité, même dans des situations où les dirigeants politiques et les chefs militaires n'auraient pas jugé une telle action appropriée ; l'on verrait ainsi des conflits éclater ou escalader de manière involontaire, en dehors de tout contrôle humain direct¹⁰. Ces systèmes font donc peser une grave menace sur la

8 Paul Scharre, « Why unmanned », dans *Joint Force Quarterly*, N° 61, 2^e trimestre 2011, p. 92 [Traduction CICR].

9 Peter Asaro, « How just could a robot war be ? », dans Adam Briggles, Katinka Waelbers et Philip A. E. Brey (dir.), *Current issues in computing and philosophy*, IOS Press, Amsterdam, 2008, pp. 50-64, disponible sur : <http://peterasaro.org/writing/Asaro%20Just%20Robot%20War.pdf>.

10 L'on pourrait se référer par analogie au *Flash Crash* (« krach éclair ») du 6 mai 2010, au cours duquel les

stabilité internationale, ainsi que sur la capacité des instances internationales à gérer les conflits.

S'agissant de l'acceptabilité juridique des systèmes d'armement autonomes au regard du DIH existant¹¹, la question essentielle paraît être celle de savoir si ces systèmes seront ou non capables de respecter les principes de distinction et de proportionnalité¹². Tant la complexité de ces systèmes que notre incapacité de prévoir comment ils pourraient se comporter dans des contextes opérationnels complexes nous placent devant une difficulté supplémentaire : comment tester et vérifier qu'un système d'armement autonome venant d'être conçu satisfait aux prescriptions du DIH, tel qu'exigé par l'article 36 du Protocole additionnel I¹³ ? De manière plus générale, nous devrions également savoir comment « encadrer » les innovations technologiques, l'évolution étant de plus en plus rapide en ce qui concerne les nouvelles armes et les nouvelles tactiques¹⁴.

Un autre motif de préoccupation tient au fait que l'opérateur de ces systèmes d'armement autonomes peut ne pas être identifiable. En d'autres termes, d'une part, aucun être humain ne pourrait être tenu responsable des actes commis dans une situation donnée par un système d'armement autonome et, d'autre part, un tel système pourrait agir de manière si imprévisible qu'il serait injuste de tenir l'opérateur responsable des actes commis¹⁵. Ces systèmes pourraient donc exclure la possibilité d'établir toute responsabilité pénale individuelle, celle-ci

systèmes de courtage automatique à haute fréquence se sont emballés et ont provoqué une baisse de 1000 points de l'indice des cours (Dow Jones) : cette baisse de 9 % en moyenne des valeurs cotées a été la plus violente que la Bourse de New York ait connue jusqu'alors. Voir sur *Wikipedia*, l'article consacré au *Flash Crash* du 6 mai 2010, disponible sur : http://fr.wikipedia.org/wiki/2010_aux_%C3%89tats-Unis.

- 11 Noel Sharkey, « Death strikes from the sky : the calculus of proportionality », dans *IEEE Technology and Society Magazine*, Vol. 28, No. 1, 2009, pp. 16-19 ; Noel Sharkey, « Saying 'No!' to lethal autonomous targeting », dans *Journal of Military Ethics*, Vol. 9, N° 4, 2010, pp. 369-383 ; Markus Wagner, « Taking humans out the loop : implications for the international humanitarian law », dans *Journal of Law Information and Science*, Vol. 21, 2011, disponible sur : http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1874039 ; Matthew Bolton, Thomas Nash et Richard Moyes, « Ban autonomous armed robots », 5 mars 2012, Article36.org, 5 mars 2012, disponible sur : <http://www.article36.org/statements/ban-autonomous-armed-robots>.
- 12 Voir, en particulier, les articles 51 et 57 du Protocole additionnel I. *Protocole additionnel aux Conventions de Genève du 12 août 1949 relatif à la protection des victimes des conflits armés internationaux (Protocole I)*, 8 juin 1977, 1125 UNTS 3, entré en vigueur le 7 décembre 1978. Voir <http://www.icrc.org/applic/ihl/dih.nsf/Article.xsp?action=openDocument&documentId=1C05636226FBEE14C12563BD002C24B0> (article 51 – Protection de la population civile) et <http://www.icrc.org/applic/ihl/dih.nsf/Article.xsp?action=openDocument&documentId=9FBD35598315A3E8C12563BD002C25A3> (article 57 – Précautions dans l'attaque).
- 13 Le texte intégral de l'article 36 du Protocole additionnel I (Armes nouvelles) est le suivant : « 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 ». Voir : <http://www.icrc.org/applic/ihl/dih.nsf/Article.xsp?action=openDocument&documentId=39C6B917F5E974F5C12563BD002C22BE>.
- 14 Richard M. O'Meara, « Contemporary governance architecture regarding robotics technologies : an assessment », dans Patrick Lin, Keith Abney et George Bekey, *Robot Ethics*, MIT Press, Cambridge, MA, 2011, pp. 159-168.
- 15 Robert Sparrow, « Killer Robots », dans *Journal of Applied Philosophy*, Vol. 24, N° 1, 2007, pp. 62-77.

exigeant la capacité de porter des jugements moraux ainsi que la détermination d'une intention criminelle (*mens rea*)¹⁶. Des systèmes d'armement autonomes placés sous la supervision ou le commandement d'opérateurs humains pourraient commettre des atrocités ou causer une tragédie. Deux notions fondamentales – la responsabilité du commandant et l'obligation de superviser les personnes placées sous ses ordres – se trouveraient alors gravement sapées. Les opérateurs humains seraient à l'abri de poursuites alors que, dans d'autres circonstances, les actes commis auraient été considérés comme des crimes de guerre. Il apparaît donc de plus en plus important que, d'une part, les États soient tenus responsables de la conception et de l'emploi des systèmes d'armement autonomes et que, d'autre part, des règles soient imposées au niveau international.

Nous sommes aujourd'hui à la croisée des chemins. Il nous faut décider quelle attitude nous entendons adopter, en tant que communauté internationale, vis-à-vis de ces systèmes. Allons-nous les traiter comme de nouvelles extensions de technologies anciennes, ou comme un changement qualitatif vers un nouveau genre de technologie? Tels qu'ils sont aujourd'hui, le DIH et le DIDH permettent-ils de répondre aux défis posés par les technologies létales des systèmes autonomes ou ont-ils besoin d'être modifiés et de faire l'objet d'extensions mineures ou de révisions majeures? Une interdiction des systèmes d'armement autonomes est-elle souhaitable ou risque-t-elle de venir perturber la mise au point d'armes nouvelles qui seraient mieux à même de respecter les normes morales et juridiques?

J'estime pour ma part que les systèmes d'armement autonomes représentent un changement qualitatif de la technologie militaire, précisément parce qu'ils suppriment l'exercice du jugement humain dans l'engagement de la force létale. Ces systèmes menacent de saper les droits de l'homme du fait de l'absence de tout jugement et examen critique de la part d'un opérateur humain. Il existe donc de bonnes raisons de clarifier le DIH et le DIDH en codifiant explicitement l'interdiction de l'emploi de systèmes d'armement autonomes. Ces raisons résistent d'ailleurs à toutes les critiques formulées jusqu'ici à leur encontre. Une telle clarification et une telle codification permettraient notamment :

- 1) d'éviter diverses « pentes glissantes » menant vers des systèmes d'armement autonomes, en imposant une limite de principe qui départage ce qui peut être automatisé et ce qui ne peut pas l'être ;
- 2) d'orienter les futurs investissements dans le développement technologique vers des conceptions davantage centrées sur l'humain et capables d'améliorer le respect des normes éthiques et juridiques dans les conflits armés ;
- 3) d'éliminer le risque que les nouvelles technologies entraînent une désstabilisation plus radicale des normes éthiques et juridiques régissant les conflits armés et, enfin ;
- 4) d'établir le principe juridique selon lequel les processus automatisés ne satisfont pas aux exigences morales à respecter quand une vie humaine est en jeu.

16 M. Wagner, *op. cit.*, note 11, p. 5.

Il serait donc souhaitable que la communauté internationale prenne des mesures visant à établir une interdiction internationale des systèmes d'armement autonomes en se fondant sur les normes de protection des droits de l'homme ainsi que sur d'autres normes qui protègent l'individu.

La prise de décisions létales

Afin de développer l'argument selon lequel leur emploi est moralement et juridiquement inadmissible, nous devons clarifier en quoi les systèmes d'armement autonomes ne remplissent pas les conditions nécessaires et suffisantes dans lesquelles il est permis de tuer pendant un conflit armé. La notion de « système d'armement autonome » doit également être affinée. Il suffit toutefois, à ce stade, de définir que la catégorie des systèmes d'armement autonomes regroupe l'ensemble des systèmes automatisés qui peuvent utiliser la force létale en l'absence de toute décision spécifique, consciente et délibérée d'un opérateur, contrôleur ou superviseur humain.

Certes, ces systèmes ne sont pas sans précédents. Des précurseurs de divers types ont déjà été utilisés dans des conflits armés, notamment les mines et autres dispositifs déclenchés par les victimes, ainsi que certains missiles guidés et quelques systèmes automatiques de défense. En un certain sens, ces systèmes sont eux-mêmes moins des « armes » que des systèmes automatisés qui sont dotés d'armes ou qui contrôlent des armes. Ils constituent donc un défi pour les modes de pensée traditionnels dans le domaine des armes et de la maîtrise des armements, qui ont tendance à se concentrer essentiellement soit sur l'arme en tant qu'outil ou instrument, soit sur les effets destructeurs de cette arme. Au contraire, les systèmes d'armement autonomes nous forcent à penser en termes de « systèmes » qui pourraient réunir toute une variété de configurations de capteurs, de traitement de l'information et de déploiement d'armements, et ils nous obligent à nous concentrer sur la façon dont est prise la décision d'employer la force létale¹⁷.

Il existe, au sein des forces armées des États-Unis, une ligne politique qui consiste à suivre un modèle prévoyant « un homme dans la boucle » (*human-in-the-loop*) en cas d'emploi de la force létale. Cette expression est utilisée dans le domaine de l'ingénierie des facteurs humains pour indiquer qu'un être humain fait partie intégrante du système. S'agissant de l'utilisation de la force létale, le système déterminant est celui qui inclut le cycle de prise de décisions au cours duquel intervient la décision de recourir à la force létale. Dans le jargon militaire, ce cycle de prise de décisions est appelé « chaîne de frappe » et, selon l'armée de l'air des États-Unis, comporte six phases : *find*, *fix*, *track*, *target*, *engage*, *assess*

17 Au sens de l'article 36 du Protocole additionnel I aux Conventions de Genève, les systèmes d'armement autonomes doivent faire l'objet d'un examen de licéité. L'obligation porte en effet sur « toute nouvelle arme, nouveau moyen ou nouvelle méthode de guerre ». Le fait de doter une arme existante (déjà autorisée) de moyens autonomes de ciblage ou de tir constitue une nouvelle manière d'utiliser cette arme.

(trouver-fixer-suivre-cibler-engager-évaluer)¹⁸. Un débat s'est engagé récemment sur l'opportunité d'adopter un modèle intégrant « un homme dans la boucle », l'opérateur humain supervisant un ou plusieurs systèmes qui automatisent un grand nombre de tâches des six phases de ce cycle. Ce changement de paradigme paraît créer une position intermédiaire entre le contrôle humain direct du modèle « un homme dans la boucle » et un système d'armement entièrement autonome. Néanmoins, l'élément décisif qui détermine si un système donné est, ou n'est pas, un système d'armement autonome réside dans sa capacité à automatiser soit l'étape « cibler » soit l'étape « engager » en l'absence de tout contrôle humain direct. Nous pouvons donc dire que tout système qui est capable de sélectionner des cibles et d'employer la force potentiellement létale en l'absence de volonté délibérée et d'examen spécifique de la part d'un agent humain entre dans la catégorie des « systèmes d'armement autonomes ».

Cette définition reconnaît que, sur les plans de l'éthique et du droit, le problème fondamental consiste à établir soit la présence d'un lien de causalité entre la prise de décision automatisée et l'emploi d'une arme ou de la force létale, soit, inversement, l'absence de rapport causal entre la décision humaine et le fait qu'un système automatisé contrôle directement le recours à la force létale. Ainsi, le problème central, sur les plans moral et juridique, réside dans le fait que des humains puissent abandonner leurs propres responsabilités dans la prise de décisions pour les déléguer à des systèmes autonomes conçus pour tuer.

Il convient cependant de relever que l'inclusion d'un humain dans le processus de prise de décisions létales constitue une exigence nécessaire, mais non pas suffisante. Pour être légitime, ce processus doit remplir trois autres conditions : d'une part, l'opérateur humain (qui prend la décision d'employer la force létale après avoir vérifié que la cible est légitime) doit disposer de suffisamment de temps pour réfléchir. D'autre part, cet opérateur doit avoir reçu une formation appropriée et être bien informé. Enfin, il doit être tenu responsable et répondre de ses actes. Certes, il serait facile de placer des personnes insuffisamment formées devant un écran sur lequel défile une liste de cibles désignées et de demander à ces « opérateurs » de vérifier les cibles avant d'appuyer sur un bouton pour donner l'autorisation d'employer la force létale contre ces cibles. Ces « opérateurs » ne feraient pas mieux que des automates s'ils étaient contraints de prendre rapidement des décisions, sans avoir le temps de réfléchir, ou sans avoir accès à des informations pertinentes et suffisantes sur la base desquelles ils pourraient prendre une décision valable, ou s'ils étaient soumis à des tensions extrêmes sur les plans physique et émotionnel. Nous tenons généralement compte de ce genre de facteurs quand nous évaluons la validité d'une décision prise par un individu. Si un opérateur est placé dans de telles conditions, nous sommes moins enclins à le tenir responsable tant des décisions prises que de toute conséquence fortuite pouvant en avoir résulté, même s'il devra tout de même répondre de ses

18 Julian C. Cheater, « Accelerating the kill chain via future unmanned aircraft », Blue Horizons Paper, Center for Strategy and Technology, Air War College, avril 2007, p. 5, disponible sur : http://www.au.af.mil/au/awc/awcgate/cst/bh_cheater.pdf.

actes. Puisque ces facteurs diminuent la responsabilité de l'individu qui prend la décision, la conception et l'emploi de systèmes qui augmentent la probabilité que la prise de décisions se déroule de cette façon constituent en eux-mêmes des actes irresponsables. À mon avis, si l'on se place du point de vue de l'éthique de l'ingénierie et de la conception, le fait de concevoir délibérément des systèmes qui excluent l'intervention d'agents qui puissent répondre de leurs actes et être tenus responsables est en lui-même non éthique, irresponsable et immoral. Au moment d'établir les critères à l'aune desquels nous évaluons la prise de décisions létales, nous devrions éviter toute confusion entre, d'une part, les « circonstances atténuantes » que nous accordons aux humains qui agissent dans des conditions difficiles et, d'autre part, les idéaux auxquels nous rattachons les critères utilisés. De plus, le fait que lors de la prise de telles décisions, la performance d'un humain puisse être ramenée au niveau de celle d'un système autonome ne signifie pas que nous devions aussi abaisser les critères à l'aune desquels nous jugerons ces décisions.

Certes, le libellé détaillé de la définition des systèmes d'armement autonomes devant figurer dans un traité international sera nécessairement choisi au terme d'un processus de négociations. Néanmoins, la pierre angulaire d'un tel traité devrait être l'établissement du principe selon lequel des vies humaines ne peuvent pas être supprimées sans une décision informée et réfléchie, prise par un humain, quant à chacune de ces vies, ce principe s'appliquant dans chaque cas, sans exception, d'emploi de la force létale. Dès lors, tout système automatisé qui exclut les humains du processus de prise de décisions létales ne respecte pas ce principe et doit être interdit. Une telle proposition a un caractère novateur dans le domaine du contrôle des armes, car elle ne vise pas une arme particulière, mais plutôt la manière dont est prise la décision d'utiliser cette arme. Les traités de limitation des armements précédents ont essentiellement porté sur des armes spécifiques et sur leurs effets, ou sur le caractère intrinsèquement indiscriminé d'une arme donnée. Une interdiction des systèmes d'armement autonomes devra, au contraire, se concentrer sur la délégation de l'autorité d'employer la force létale à un processus automatisé qui n'est placé sous aucune supervision humaine directe et qui échappe à tout contrôle discrétionnaire.

« Meurtre légal » et nécessité d'un jugement humain

Pour qu'il soit légal de tuer pendant un conflit armé, cet acte doit être en accord avec les règles du DIH. En particulier, les parties à un conflit armé sont tenues de respecter les principes de distinction et de proportionnalité. La capacité des systèmes d'armement autonomes de respecter ces principes suscite de nombreux débats. La conjecture la plus ambitieuse consiste à dire que nous serons peut-être capables de programmer les systèmes d'armement autonomes de manière telle que, demain, ils seront capables de se conformer aux dispositions du DIH, ainsi qu'aux règles d'engagement et aux ordres du commandant

s'appliquant spécifiquement à une mission donnée¹⁹. Les promoteurs de cette idée (qui s'appuie sur la tradition dite de la « programmation par contraintes ») affirment que le DIH pourrait être traduit en règles de programmation définissant strictement quelles actions sont prohibées dans une situation donnée. Ainsi, un hypothétique « dispositif de contrôle éthique » pourrait empêcher un système d'armement autonome de mener une action qui, selon ce dispositif, est explicitement prohibée par le DIH. Ronald C. Arkin estime que les systèmes d'armement autonomes pourraient choisir de se sacrifier dans des situations où nous ne nous attendrions pas à ce que des humains le fassent. Ces systèmes pourraient dès lors commettre beaucoup moins d'erreurs et de manquements que des soldats humains et seraient donc plus aptes à respecter les règles du DIH.

En première analyse, cette proposition paraît assez convaincante et même J. Kellenberger lui a reconnu un certain attrait :

À l'occasion du débat sur ces nouvelles technologies, il nous faut voir également quels sont les avantages qu'elles pourraient apporter si elles contribueraient à une meilleure protection. Respecter les principes de distinction et de proportionnalité signifie qu'il faut prendre certaines précautions dans l'attaque, comme indiqué à l'article 57 du Protocole additionnel I. Cet article prévoit notamment l'obligation pour un attaquant de prendre toutes les précautions pratiquement possibles quant au choix des moyens et méthodes d'attaque en vue d'éviter et, en tout cas, de réduire au minimum les pertes en vies humaines dans la population civile, les blessures aux personnes civiles et les dommages aux biens de caractère civil qui pourraient être causés incidemment. Dans certains cas, les cyberopérations ou le déploiement d'armes télécommandées ou de robots pourraient faire incidemment moins de victimes civiles et causer moins de dommages aux biens de caractère civil que l'emploi d'armes classiques. Des précautions accrues devraient également être possibles dans la pratique, du fait simplement que ces armes sont déployées depuis suffisamment loin et souvent, avec suffisamment de temps pour que la cible soit choisie avec soin et que le moment de l'attaque soit décidé de façon à minimiser l'impact sur la population civile et les biens de caractère civil. On pourrait considérer que dans de telles circonstances, l'application de cette règle voudrait qu'un commandant évalue s'il peut obtenir le même avantage militaire en utilisant ces moyens et méthodes de guerre, s'ils sont applicables²⁰.

Assurément, le fait de renforcer la protection des personnes et des biens civils dans les futurs conflits armés serait avantageux. Nous devons toutefois éviter de tirer trop vite des conclusions qui conduiraient à autoriser l'emploi de systèmes d'armement autonomes. Cet argument apparemment simple repose sur

19 R. C. Arkin, *op. cit.*, note 3, pp. 71-91.

20 J. Kellenberger, *op. cit.*, note 7, p. 6.

un grand nombre d'hypothèses, et nous risquerions de nous laisser induire en erreur et d'en oublier la raison d'être et la signification du DIH.

Pendant un conflit armé, le but ultime du DIH est de protéger les personnes qui ne participent pas, ou ne participent plus, directement aux hostilités et de restreindre le recours à certains moyens et méthodes de guerre. Il est tentant de penser que de tels objectifs peuvent être objectivement et facilement mesurés. Nous aimerions pouvoir croire que le principe de distinction est semblable à une règle de tri qui permettrait de scinder le monde en deux catégories : les civils d'un côté, les combattants de l'autre. Nous voudrions croire en l'existence d'une règle, aussi complexe fût-elle, qui permettrait d'attribuer définitivement chaque individu à l'une ou l'autre de ces deux catégories²¹, mais les choses sont bien plus compliquées que cela. Je prendrai ici pour exemple la difficulté de définir le sens de l'expression « civil participant aux hostilités ». Le CICR a établi un ensemble de lignes directrices, soigneusement formulées, visant à définir ce qui constitue « un acte de participation directe aux hostilités », en conséquence duquel un civil perd les protections que le DIH octroie normalement aux civils²². Ces lignes directrices précisent dans quelles conditions il est possible de conclure qu'un civil constitue une cible légitime. Ces conditions sont les suivantes : 1) seuil de nuisance, 2) rapport direct de cause à effet et 3) lien de belligérance. Chacune de ces trois conditions est explicitée dans le Guide interprétatif du CICR mais, pour les fins du présent article, il suffira de les résumer brièvement :

Pour qu'un acte spécifique atteigne le niveau de nuisance requis pour constituer une participation directe aux hostilités, il doit être de nature à nuire aux opérations militaires ou à la capacité militaire d'une partie à un conflit armé. En l'absence d'effets nuisibles sur le plan militaire, le seuil peut également être atteint si un acte est susceptible d'infliger des pertes en vies humaines, des blessures ou des destructions à des personnes ou à des biens protégés contre les attaques directes. Dans ces deux cas, les actes atteignant le niveau de nuisance requis ne peuvent constituer une participation directe aux hostilités que s'ils satisfont en outre aux exigences de relation directe de causalité et de lien de belligérance.

L'exigence de causation directe est satisfaite lorsque l'on peut raisonnablement attendre de l'acte spécifique en question (ou d'une opération militaire concrète et coordonnée dont cet acte fait partie intégrante) qu'il cause directement – en une seule étape causale – des effets nuisibles atteignant le seuil requis. Néanmoins, des actes satisfaisant à l'exigence de causation directe et atteignant le seuil de nuisance requis ne constitueront une par-

21 Il existe en effet, dans la littérature sur les armes autonomes, une tendance à se référer à la notion de « discrimination » plutôt qu'au principe de distinction (ce qui renvoie à la notion de « tâche de discrimination » dans les domaines de la psychologie cognitive et de l'intelligence artificielle). Voir, à ce sujet, le point de vue de Noel Sharkey dans cette édition.

22 Nils Mezler, *Guide interprétatif sur la notion de participation directe aux hostilités en droit international humanitaire*, CICR, Genève, 2010, disponible sur : http://www.cicr.org/fre/assets/files/other/icrc_001_0990.pdf.

ticipation directe aux hostilités que si le troisième critère – celui du lien de belligérance – est également rempli.

Afin de satisfaire à l'exigence du lien de belligérance, un acte doit être spécifiquement destiné à causer directement des effets nuisibles atteignant le seuil requis, à l'avantage d'une partie à un conflit armé et au détriment d'une autre. En règle générale, les effets nuisibles causés n'ont pas le lien de belligérance requis pour constituer une participation directe aux hostilités dans les cas suivants: a) dans le cadre de la légitime défense individuelle ou de la défense d'autrui contre les actes de violence interdits par le DIH; b) dans l'exercice du pouvoir ou de l'autorité sur des personnes ou sur un territoire; c) en tant qu'élément des troubles civils contre une telle autorité; d) lors de situations de violence entre civils.

Appliqués conjointement, les trois critères requis – seuil de nuisance, causation directe et lien de belligérance – permettent d'établir une distinction fiable entre, d'une part, les activités constituant une participation directe aux hostilités et, d'autre part, les activités qui, bien qu'elles se produisent dans le contexte d'un conflit armé, n'entrent pas dans le cadre de la conduite des hostilités et, par conséquent, n'entraînent pas, pour les personnes civiles, la perte de protection contre les attaques directes. Toutefois, même quand un acte spécifique constitue une participation directe aux hostilités, le type et le degré de la force utilisée pour le réprimer doivent être en conformité avec les règles et les principes du DIH et de toute autre branche applicable du droit international²³.

Ces lignes directrices tentent de définir la voie à suivre pour déterminer qui est une cible légitime et qui ne l'est pas. Pourtant, elles ne sont pas même appelées « règles » : ce ne sont que des lignes directrices destinées à aider un agent moral à franchir de multiples étapes d'interprétation et de jugement. Pour être à même de déterminer si un individu spécifique, dans des circonstances spécifiques, remplit chacune des trois conditions énoncées, cet agent moral doit être parvenu à un degré élevé de compréhension d'une situation complexe. Il doit notamment avoir apprécié les implications, aux niveaux tactique et stratégique, du préjudice potentiel, ainsi que le statut d'autres individus potentiellement menacés, la nature des structures causales, ainsi que les relations et implications causales directes des actions d'un individu en particulier et, enfin, la situation socioculturelle et l'état psychologique de l'individu pour que ses intentions et actions puissent être qualifiées d'« actions militaires » et ne pas être considérées, par exemple, comme relevant de l'exercice de l'autorité publique ou de la légitime défense.

Que veulent vraiment dire ceux qui déclarent que nous pouvons transcrire les règles du DIH dans un programme informatique ? S'agit-il simplement de prendre des règles juridiques qui ont été élaborées pour régir des actions humaines et de les traduire en codes programmés pour imposer des contraintes

23 N. Mezler, *op. cit.*, note 22, pp. 48-67.

aux actions d'une machine ? Le prochain Protocole additionnel aux Conventions de Genève devrait-il être écrit directement en code informatique ? Le DIH ne comporte-t-il pas des éléments qui ne peuvent pas être programmés ? Il est tentant d'aborder ce problème sous son aspect technique, de voir les décisions et les actions d'un combattant comme une « boîte noire », de comparer le soldat humain au soldat robot, et de prétendre que celui des deux qui commet le moins d'erreurs au regard du DIH est le soldat « le plus éthique ». Cette stratégie d'argumentation a souvent été utilisée dans l'histoire de l'intelligence artificielle.

Deux questions se posent en fait ici. La première est de caractère empirique : quand il en va de la vie ou de la mort de personnes humaines, peut-on attendre d'une machine, d'un ordinateur ou d'un processus automatisé qu'ils prennent chacune de ces décisions à un niveau de « performance » jugé acceptable ? La seconde question est d'ordre moral : peut-on admettre que de telles décisions de vie ou de mort soient prises par une machine, un ordinateur ou un processus automatisé ? Si nous ne parvenons pas à démontrer qu'une machine ne devrait pas prendre de telles décisions, nous n'aurons plus qu'à nous demander si (et quand) des programmeurs de talent seront à même de concevoir un jour un système informatique doté de ces capacités et, au minimum, quand allons-nous permettre à des machines de prendre de telles décisions.

L'histoire de l'intelligence artificielle est instructive à cet égard. Elle nous indique en effet que, de manière générale, un tel problème ne peut pas être résolu par le calcul numérique, mais que si nous arrivons à le définir très précisément et à le simplifier, nous avons quelques chances de succès. Cela dit, nous pourrions aussi établir une comparaison entre, d'une part, le genre de problèmes que l'intelligence artificielle a su résoudre (dans le jeu d'échecs, par exemple) et, d'autre part, le genre de problèmes rencontrés dans l'application du DIH. Certes, les exigences posées par le DIH constituent en un certain sens des « règles », mais elles sont bien différentes des règles des échecs dans la mesure où, pour être appliquées de façon appropriée en toute situation donnée, elles requièrent une grande part de jugement interprétatif. De plus, le contexte dans lequel s'appliquent les règles du DIH, la nature et la qualité des informations disponibles, ainsi que les diverses interprétations pouvant se trouver en concurrence ou en conflit peuvent grandement varier, tant tout au long d'un conflit donné que d'un jour à l'autre, voire d'une heure à l'autre.

Nous pourrions vouloir ajouter que l'intelligence est exclusivement humaine, mais si quelqu'un est un jour capable de la définir de manière suffisamment spécifique, ou de la réduire à l'exécution de telle ou telle tâche concrète, alors il sera peut-être possible de programmer un ordinateur pour qu'il remplisse « mieux » cette même tâche. Ce faisant, nous modifions nécessairement la définition de l'intelligence, qui n'est plus considérée comme une aptitude complexe mais comme la capacité d'accomplir une tâche spécifique. Peut-être n'est-il pas si important (malgré les incidences sur les plans sociétal et culturel) de redéfinir l'intelligence pour tenir compte des dernières avancées en informatique ? Par contre, quand il s'agit de moralité et de suppression de vies humaines, voulons-nous vraiment redéfinir le sens du qualificatif « moral » de manière à le rendre applicable aux

systèmes d'armement autonomes? Quels lendemains préparerions-nous en donnant à des systèmes automatisés l'autorité de décider de la vie ou de la mort d'êtres humains? En l'absence de tout jugement humain, comment pouvons-nous avoir l'assurance que ces décisions ne seront pas prises de manière arbitraire?

L'automatisation des règles du DIH déprécierait probablement le rôle qu'elles jouent en réglementant le comportement éthique des divers protagonistes. Les risques liés à une telle évolution expliquent peut-être aussi pourquoi les développeurs s'efforcent de conserver « dans la boucle » des humains à qui il incombe de lever les ambiguïtés et de passer les décisions au crible d'une appréciation morale. Comme l'a déclaré Sir Brian Burridge, qui a commandé les forces aériennes du Royaume-Uni (RAF) en Irak de 2003 à 2005 :

Au regard du droit des conflits armés, il subsiste l'exigence d'évaluer la proportionnalité et, dans ce cadre, une attente demeure, celle que l'humain se trouvant au bout de la chaîne de transmission accomplisse la dernière évaluation en appréciant la situation sur la base d'un jugement rationnel. Les conflits postmodernes nous mettent face ... à des champs de bataille ambigus, non linéaires. Dès lors, nous ne pouvons pas « sortir de la boucle » l'opérateur humain, le commandant, l'analyste – tous ceux qui luttent contre l'ambiguïté. Le débat sur le maintien de « l'homme dans la boucle » doit aller bien au-delà de cela²⁴.

De par sa nature même, le DIH – dont le but est de réguler le comportement des êtres humains et des organisations humaines pendant les conflits armés – présume que les combattants sont des êtres humains. En ce sens, le DIH est anthropocentrique. Malgré tous les efforts déployés par ses auteurs pour que le DIH soit clair et précis, son application ne peut pas être effective dans une situation donnée sans passer par des niveaux d'interprétation multiples. Le DIH complète ses règles par des orientations heuristiques auxquelles les agents humains doivent se conformer ; il demande explicitement aux combattants d'examiner de façon raisonnée les implications de leurs actions ; enfin, en faisant explicitement appel à leur humanité, le DIH demande aux combattants de faire preuve de compassion et d'exercer leur jugement. Ce faisant, le droit n'impose pas d'effectuer un calcul spécifique, mais il impose aux combattants le devoir d'évaluer de manière réfléchie le coût potentiel en vies humaines et en biens des diverses modalités d'action dont ils disposent.

La justice ne peut pas être automatisée

De par sa nature même, le droit est incomplet. Il reste soumis à l'interprétation et peut toujours faire l'objet d'un examen ultérieur. Aussi bien intentionnées

24 Brian Burridge, « UAVs and the dawn of post-modern warfare: a perspective on recent operations », dans *RUSI Journal*, Vol. 148, N° 5, octobre 2003, pp. 18-23 [Traduction CICR].

que puissent être les lois et les règles de droit, et même si leur élaboration a été approfondie et réfléchie, aucun régime juridique n'est, et ne peut être, parfait. En évolution constante, ce système est conçu de telle manière que la charge d'en gérer l'application dans le monde des affaires humaines incombe à des institutions créées par l'homme. Afin de permettre le bon fonctionnement du système judiciaire, un certain nombre d'agents humains – juges, procureurs, avocats, témoins, jurés – s'engagent dans des processus complexes d'interprétation et de jugement. Les uns et les autres participent activement à l'évaluation des concordances pouvant exister entre un ensemble de règles abstrait et une situation concrète donnée. Le droit à une procédure régulière porte essentiellement sur l'obligation, pour ceux qui « disent le droit », de rendre compte publiquement de cette démarche délibérative.

Certes, il pourrait exister un jour un programme informatique destiné à remplacer les agents humains et à automatiser leurs décisions. J'estime toutefois que cela porterait fondamentalement atteinte au droit à une procédure régulière, qui consiste essentiellement à permettre, d'une part, de mettre en cause les règles et la pertinence de leur application dans une circonstance donnée et, d'autre part, de faire appel à la rationalité et à la compréhension humaines suffisamment informées. Arrive-t-il que des humains occupant de telles fonctions commettent des fautes ? Oui, bien sûr, cela arrive. Cependant, la compréhension, la rationalité et le jugement humains valent davantage que tout ensemble concevable de règles fixées ou que tout système informatique. De plus, les arguments avancés dans un cas donné, la possibilité d'interjeter appel afin de renverser les décisions judiciaires, ainsi que les diverses façons dont les opinions et la jurisprudence renseignent l'interprétation des lois, font apparaître que tout jugement juridique exige la mise en regard de perspectives différentes, incompatibles et même parfois contradictoires, dont il importe ensuite de tirer les enseignements. Les systèmes informatiques ou algorithmiques actuels ne possèdent pas une telle « compétence », et rien ne dit qu'ils la posséderont un jour.

Plus important encore, le jugement humain est un élément constitutif du système judiciaire. En d'autres termes, un système de justice, quel qu'il soit, ne peut s'appliquer à des êtres humains que s'il est fondé sur la raison humaine. La justice, en tant que telle, ne peut pas être déléguée à des processus automatisés. Bien sûr, dans le cadre des procédures administratives et juridiques, l'automatisation de diverses tâches est de nature à aider les humains qui, ainsi, seront mieux à même de poser leurs jugements, ou de le faire avec une plus grande efficacité. Par contre, l'automatisation ne peut pas exonérer les humains de l'obligation d'examiner les éléments de preuve, de délibérer sur des interprétations alternatives et de se former une opinion basée sur des informations suffisantes. De manière générale, les efforts visant à automatiser la justice administrative n'ont pas amélioré la performance humaine – ils l'ont même fortement dégradée²⁵. Automatiser ces aspects essentiels du jugement humain dans les procé-

25 Danielle Keats Citron, « Technological due process », dans *Washington University Law Review*, Vol. 85, 2008, pp. 1249-1292.

dures judiciaires équivaldrait à déshumaniser la justice, et une telle perspective devrait être rejetée par principe.

Si j'affirme que l'on devrait rejeter par principe l'automatisation du raisonnement humain dans les procédures judiciaires, c'est parce qu'il n'existe encore, à mon avis, aucun système automatisé (quel que soit son niveau de performance) que nous pourrions accepter en tant que substitut d'un être humain. En somme, quand un système judiciaire, un État ou ses agents sont appelés à prendre des décisions touchant aux droits fondamentaux de la personne humaine, les agents et les fonctionnaires de l'État à qui appartient la décision finale doivent être eux-mêmes humains. Des motifs non seulement moraux mais également juridiques peuvent être avancés à l'appui de ce principe. En effet, indépendamment de son statut sur le plan éthique, ce principe est un élément constitutif et essentiel du système de justice lui-même.

Au sein de chaque armée, du commandant en chef au simple soldat, il existe de nombreux niveaux de délégation de l'autorité. Toutefois, à chacun de ces niveaux se trouve un humain à qui incombent à la fois l'autorité et la responsabilité du recours à la force. La responsabilité du supérieur hiérarchique ne permet pas d'éviter l'obligation morale et juridique de déterminer si l'emploi de la force est approprié dans une situation donnée. Cette obligation peut être transférée à un autre agent humain responsable, le supérieur qui délègue son autorité étant alors tenu de superviser la conduite de son subordonné. Dans la mesure où les systèmes d'armement autonomes ne sont pas des agents humains responsables, cette autorité ne peut pas leur être déléguée.

En ce sens, le respect du principe de distinction peut être vu non seulement comme l'obéissance à une règle qui demande de distinguer les combattants des civils, mais aussi comme la prise en compte des vies humaines qui risqueraient d'être perdues en cas d'emploi de la force létale. Il est donc nécessaire qu'un être humain prenne une décision basée sur des informations suffisantes avant d'ôter la vie d'autrui. Cette exigence apparaît plus nettement quand il s'agit de respecter le principe de proportionnalité. Il faut en effet, dans ce cas, mettre en balance la valeur de vies humaines (qu'il s'agisse de civils ou de combattants) et la valeur d'objectifs militaires. Or, aucune de ces valeurs n'est fixe et, d'une certaine façon, elles découlent précisément des jugements éthiques qui entrent dans les calculs de proportionnalité.

Voilà donc pourquoi l'on ne saurait prétendre qu'un système d'armement autonome serait moralement supérieur à un soldat humain, au motif qu'il pourrait être technologiquement capable de commettre moins d'erreurs en accomplissant une tâche de discrimination, ou de trouver des moyens de neutraliser des cibles militaires réduisant de façon optimale le risque de provoquer des pertes et des dommages disproportionnés. Cela ne signifie nullement qu'il n'est pas désirable de poursuivre de tels buts. S'il existait vraiment des technologies qui soient capables de faire la différence entre civils et combattants mieux que tout être humain, ou mieux que le combattant « moyen », ces technologies devraient alors être déployées non pas pour pouvoir se passer de tout jugement humain, mais pour aider les combattants à respecter le principe de distinction. De même, s'il

existait une technologie capable de déterminer la marche à suivre pour détruire un objectif militaire en causant un minimum de dommages collatéraux et en réduisant au maximum les pertes et les dommages disproportionnés, cette technologie pourrait être employée par des combattants humains, tenus de prendre en connaissance de cause la décision d'employer la force létale dans une situation donnée.

Tout processus automatisé – aussi performant soit-il, et même si sa performance excède notablement celle d'un humain – devrait être soumis au contrôle d'un opérateur humain avant de pouvoir légitimement utiliser la force létale. Cette exigence est requise sur le plan technologique pour ce qui est de l'avenir prévisible car, pendant quelque temps encore, les robots n'atteindront pas le niveau de performance des êtres humains. Surtout, c'est une exigence morale, et dans de nombreux cas importants, juridique. J'affirme donc que, de manière générale, nous avons le devoir de ne pas autoriser des systèmes d'armement autonomes à faire usage de la force létale si aucune supervision et aucun contrôle direct ne sont exercés par un humain.

Deux stratégies sont utilisées aujourd'hui pour défendre l'argument selon lequel les systèmes d'armement autonomes pourraient constituer un moyen de guerre moralement ou juridiquement supérieur à ceux qui sont aujourd'hui déployés dans les conflits armés. Cet argument se présente sous de nombreuses variantes, que je range pour ma part en deux catégories : 1) les arguments pragmatiques, axés sur les défaillances du processus de prise de décisions létales pendant les conflits armés, et insistant sur les possibles/hypothétiques améliorations technologiques à attendre de l'automatisation de ces décisions²⁶ ; 2) les arguments insistant sur le fait que si de tels systèmes permettent de réduire de manière générale les risques encourus par les combattants et/ou les civils (la diminution du nombre de victimes venant le confirmer), il existe un impératif moral enjoignant de les utiliser. Ces mêmes arguments ont été avancés dans le passé pour promouvoir les armes de précision²⁷ ainsi que, plus récemment, pour promouvoir les drones de type *Predator* et certaines armes létales commandées à distance²⁸.

Les armes plus précises sont-elles plus « morales » que les armes de moindre précision ? Imaginons que nous nous trouvons devant le choix suivant : attaquer une cible militaire en utilisant des munitions à guidage de précision, et donc avec un faible risque de dommages collatéraux, ou attaquer la même cible en lançant un « tapis de bombes », avec un niveau élevé de probabilité (sinon de certitude) que les dommages collatéraux seront importants. Il est assez facile de défendre l'argument selon lequel les munitions à guidage de précision seraient préférées par tout individu placé devant un tel choix. Voilà précisément, toutes choses étant égales par ailleurs, le type de décision morale et légale que nous

26 Ronald C. Arkin, « Governing lethal behavior: embedding ethics in a hybrid deliberative/reactive robot architecture », Georgia Institute of Technology, Technical Report GUT-GVU-07-11, 2007, p. 11.

27 Human Rights Watch, « International humanitarian law issues in the possible U.S. invasion of Iraq », dans *The Lancet*, 20 février 2003.

28 Bradley Jay Strawser, « Moral Predators: the duty to employ uninhabited aerial vehicles », dans *Journal of Military Ethics*, Vol. 9, N° 4, 2010, pp. 342-368.

devons prendre aujourd'hui. Bien sûr, l'expression « toutes choses étant égales par ailleurs » peut être un grand fourre-tout dans lequel chacun glissera ce qu'il voudra. Certes, au moment de décider de la manière d'engager une cible, l'opérateur devrait préférer l'arme la plus précise. Fort bien, mais l'arme choisie n'est pas éthiquement étrangère à cette décision. En fin de compte, c'est l'opérateur humain qui choisit d'utiliser (ou non) l'arme qu'il juge la plus « morale » ; même la plus précise des armes peut être utilisée de manière illégale et immorale. Tout ce qu'apporte la précision, c'est la possibilité, pour l'utilisateur de l'arme, d'opter pour une façon d'agir plus éthique : elle ne détermine ni ne garantit l'adoption d'un tel comportement.

L'argument ci-dessus peut paraître purement sémantique, mais la distinction qui est faite est cruciale. Nous ne mettons pas fin à nos responsabilités morales en recourant à des technologies plus précises. Par contre, comme dans le cas d'autres systèmes automatisés (tels que le régulateur de vitesse ou le pilotage automatique d'un véhicule), nous continuons à tenir l'opérateur responsable tant du système qu'il gère que de la décision ultime d'enclencher ou de désenclencher le système automatisé ; l'opérateur est ultimement responsable de la pertinence de ses choix. En effet, la plupart du temps, comme nous l'avons vu dans l'emploi des munitions à guidage de précision et des drones de combat, ces technologies augmentent en fait le fardeau moral qui pèse sur l'opérateur, à savoir : faire en sorte que les cibles soient correctement sélectionnées et que les civils soient épargnés. Dès lors, à mesure que nos technologies gagnent en sophistication, nous devrions les concevoir de telle manière que leur perfectionnement améliore notre conduite morale.

Il y a quelque chose de profondément étrange dans l'idée que l'on pourrait accroître la moralité de la conduite de la guerre en automatisant celle-ci (à tel point que l'humain n'y jouerait plus aucun rôle) ou, tout au moins, en automatisant les décisions de recours à la force létale. La stratégie rhétorique visant à rendre ces arguments plus convaincants consiste à mettre l'accent sur les défaillances morales des humains en temps de guerre. Les actes dictés par la peur ou le désespoir et les erreurs commises sous l'effet du stress ou de la contrainte et dans le « brouillard de la guerre » sont montrés du doigt. L'étape suivante consisterait donc à proposer une solution technologique susceptible d'éliminer de telles défaillances. L'idée pourrait être séduisante, mais aucune technologie de ce type n'existe encore. Par ailleurs, deux points cruciaux n'ont pas été relevés à propos des nouveaux types de technologies automatisées que nous voyons apparaître. Premièrement, en éloignant les soldats des risques immédiats de la guerre (ce que font les systèmes opérés à distance, sans pour autant automatiser les décisions létales), nous pouvons aussi leur éviter de subir une grande part des pressions psychologiques évoquées plus haut, et donc de commettre les erreurs qui s'ensuivent. Deuxièmement, si un système automatisé était capable de meilleures performances que les humains dans les tâches de discrimination ou dans les calculs de proportionnalité, il pourrait tout aussi facilement remplir une fonction consultative auprès des décideurs humains. De tels systèmes automatisés assisteraient et renseigneraient les opérateurs

sans qu'il soit nécessaire de leur transférer l'autorité de recourir à la force létale en l'absence de toute décision (basée sur des informations suffisantes) prise par un humain²⁹.

Arguments contre l'interdiction des systèmes d'armement autonomes

Dans une note d'orientation publiée récemment, Anderson et Waxman ont présenté un examen critique des propositions d'interdiction des systèmes d'armement autonomes³⁰. Ils concluent qu'il est important d'établir des normes internationales relatives à l'emploi des systèmes d'armement autonomes, mais qu'une interdiction ne constitue pas la meilleure option. Tant l'argumentation que beaucoup de conclusions présentées par ces auteurs soulèvent néanmoins de nombreux problèmes. Leur argumentation repose essentiellement sur deux postulats :

Dans la recherche de solutions aux dilemmes juridiques et éthiques qui accompagnent ces technologies émergentes, il est capital de reconnaître que l'évolution de ces technologies est inéluctable et qu'elle se fera de manière progressive. La politique des États-Unis visant à résoudre ces dilemmes devrait être élaborée en partant de ces postulats. Étant donné le développement et le déploiement (certains, mais graduels) de ces systèmes, ainsi que les avantages humanitaires liés à la précision que présentent quelques-uns d'entre eux, certaines mesures proposées – telles que des traités d'interdiction – sont à la fois impossibles à implanter et contestables sur le plan de l'éthique³¹.

Nous avons ici plusieurs arguments contre la proposition d'un traité international d'interdiction. Les auteurs insistent tout d'abord sur les deux postulats de départ : l'apparition de ces technologies est inévitable, et leur développement sera progressif. Ils n'apportent cependant aucun élément de preuve ni aucun argument à l'appui de l'une et l'autre de ces hypothèses, alors qu'il existe de solides raisons de les rejeter. Ils avancent ensuite l'argument selon lequel certains de ces systèmes pourraient présenter des avantages sur le plan humanitaire et que, par conséquent, les mesures visant à les interdire sont à la fois « impossibles à implanter » et « contestables sur le plan de l'éthique ». Je viens d'expliquer pourquoi il n'est pas « contestable sur le plan de l'éthique » d'affirmer que même les plus précis des systèmes d'armement autonomes mettent en péril les droits de l'homme. Je m'intéresserai donc maintenant aux deux postulats préliminaires d'Anderson et Waxman, ainsi qu'à leur éventuelle incidence

29 Peter Asaro, « Modeling the moral user: designing ethical interfaces for tele-operation, dans *IEEE Technology & Society*, Vol. 28, N° 1, 2009, pp. 20-24, disponible sur : <http://peterasaro.org/writing/Asaro%20Modeling%20Moral%20User.pdf>.

30 K. Anderson et M. C. Waxman, *op. cit.*, note 3, p. 13. [Traduction CICR]

31 *Idem*, p. 2.

sur la mise en œuvre d'une interdiction internationale des systèmes d'armement autonomes.

Les systèmes d'armement autonomes sont-ils inévitables ?

Pourquoi devrions-nous partir du principe que les systèmes d'armement autonomes sont inévitables ? Qu'est-ce que cela pourrait réellement signifier ? En tant que philosophe et historien des sciences et des technologies, je rencontre souvent des déclarations sur le caractère « inéluctable » des découvertes scientifiques ou des innovations technologiques. La popularité de cette assertion est due en grande partie au caractère rétrospectif de toute démarche historique et au fait que, pour réfléchir à la possible évolution des technologies, nous nous basons sur ce que nous avons pu comprendre dans le passé. En d'autres termes, il paraît facile aujourd'hui pour nous d'affirmer que l'invention de l'ampoule électrique, ou du téléphone, ou de n'importe laquelle des technologies était inéluctable, puisqu'elle a bel et bien eu lieu... Il est difficile d'imaginer ce que serait le monde sans ces inventions. Pourtant, un examen attentif des détails historiques révèle que, le plus souvent, le succès d'une technologie a été extrêmement tributaire de toute une série de facteurs. Dans la plupart des cas, l'adoption de la technologie n'a pas été garantie par le succès de l'innovation. De fait, les moyens et les manières de son éventuelle utilisation dépendent toujours d'une grande variété de forces qui s'exercent aux niveaux sociétal et culturel. Assurément, il suffit de se pencher sur les nombreux grands « couacs » de la technologie et, par exemple, sur la succession d'échecs enregistrés par la commercialisation de l'ampoule électrique avant son premier succès, pour se rendre compte que les technologies pouvant être qualifiées d'« inévitables » ou d'« inéluctables » sont très rares, voire inexistantes. Même le succès de l'ampoule électrique doit beaucoup à l'innovation et au développement des services d'électricité, et son adoption généralisée s'explique par la mise au point d'une foule d'autres appareils électriques (les grille-pain, par exemple). Désormais bien plus rapide, l'évolution des technologies reste un phénomène toujours aussi dynamique et imprédictible.

Peut-être Anderson et Waxman veulent-ils dire que le développement de ces technologies aura fort probablement lieu ? Voilà qui est plus plausible. En effet, des systèmes rudimentaires peuvent déjà exécuter les fonctions essentielles d'un système d'armement autonome ; ils seraient néanmoins incapables de satisfaire aux normes juridiques internationales en vigueur exigeant le respect des principes de distinction et de proportionnalité³². Mais même en ignorant les limitations juridiques existantes, le fait que nous soyons capables de construire des technologies létales autonomes ne signifie pas nécessairement que nous les utiliserons. D'aucuns pourraient prétendre que divers types de systèmes d'armement autonomes peuvent déjà être fabriqués et que, dès lors, c'est leur

32 M. Wagner, *op. cit.*, note 11, pp. 5-9.

adoption qui est « inéluctable ». Un tel argument équivaldrait cependant à dissimuler d'importants éléments qui différencient l'invention d'une technologie et son adoption généralisée au sein de la société. Il existe certainement des motivations très solides pour adopter ces technologies, y compris le désir de réduire les risques courus par les militaires et de diminuer les frais et le personnel mobilisé dans diverses opérations et capacités militaires.

Ou alors, peut-être Anderson et Waxman veulent-ils dire que nous devrions partir du postulat qu'il est inévitable que soient mis au point des systèmes d'armement autonomes capables de respecter les règles exigeant un certain degré de discrimination et de proportionnalité. Ce postulat n'est cependant qu'une affirmation empirique relative aux capacités de technologies encore non existantes, et qui sont mesurées à l'aune de critères encore non définis. Sur un plan purement empirique, nous ignorons si ces technologies verront ou non le jour et si nous parviendrons à nous mettre d'accord sur des critères acceptables pour en évaluer les performances. Dès lors, pourquoi devrions-nous croire que le développement de ces technologies est inéluctable³³ ?

La question cruciale est donc la suivante: ces nouvelles technologies peuvent-elles déjà, ou pourront-elles, satisfaire aux exigences du droit international? Je répondrai que cela est loin d'être certain. Les arguments affirmant la supériorité éthique des soldats robots ressemblent étrangement à ceux que l'on a entendus aux premiers temps de l'intelligence artificielle selon lesquels les ordinateurs finiraient par battre nos grands maîtres humains au jeu d'échecs. Et voilà qu'avec un retard de quarante ans par rapport aux prédictions initiales, l'ordinateur *Deep Blue* d'IBM a réussi à battre Gary Kasparov. Entre le jeu d'échecs et le DIH, il existe toutefois d'importantes différences qui méritent d'être relevées. Les échecs sont un jeu assez bien défini, basé sur un ensemble de règles et qui se prête à l'analyse informatique. Il n'est question dans ce jeu ni d'interprétation, ni de normes sociales. Bien qu'il ait lui aussi des règles, le droit international est autre chose que le jeu d'échecs. Pour pouvoir être appliqué aux situations du monde réel, le droit, quel qu'il soit, nécessite interprétation et jugement. Certes, les précédents historiques et les normes établies facilitent ces activités d'interprétation et de jugement mais ils ne les déterminent pas de manière stricte. Le corpus constitué par la jurisprudence, les procédures, les plaidoiries et les appels est apte à défendre de vieux principes ou à établir de nouveaux précédents. Ce faisant, il établit des normes et des principes, même si le sens de ces normes et principes continue d'évoluer au fil du temps.

Prétendre avec insistance que les systèmes d'armement autonomes sont inévitables est donc un argument assez pernicieux. D'une part, un tel postulat ferait apparaître l'établissement d'une interdiction comme automatiquement

33 Prenons, à titre de comparaison, le cas des voitures électriques, une technologie qui existe déjà depuis une centaine d'années. Malgré la récente popularité des voitures hybrides (essence-électricité) et les très bonnes performances de certaines voitures électriques, bien peu de gens seraient prêts à soutenir l'argument selon lequel notre transition vers la voiture électrique est inévitable. Or, la technologie dont nous parlons est non seulement « possible », mais elle existe déjà!

impossible à respecter ou à mettre en œuvre. En d'autres termes, si l'on admet que les systèmes prohibés existeront et seront utilisés, à quoi bon les interdire ? Bien sûr, ces systèmes n'existent pas et ne sont pas utilisés aujourd'hui. De fait, même s'ils étaient déjà utilisés, une interdiction pourrait empêcher leur usage ultérieur. Loin d'être impossible à respecter ou à mettre en œuvre, une interdiction pourrait contribuer de manière assez efficace à infléchir les trajectoires de l'innovation vers des systèmes plus utiles et véritablement conformes à l'éthique. Il semble évident que la catégorie des systèmes d'armement autonomes peut être définie de façon suffisamment claire. Nous pourrions donc débattre ensuite de la manière dont un traité s'appliquerait (ou ne s'appliquerait pas) à certains cas « limites » (blindages réactifs, systèmes de défense antimissile balistique ou systèmes de surveillance, par exemple). Il est peu probable qu'une interdiction vienne prohiber tout recours, sans exception, à l'automatisation dans les conflits armés. Elle devrait, par contre, établir une norme internationale selon laquelle il est illicite d'employer des systèmes qui prennent des décisions létales de manière automatisée. Les traités internationaux d'interdiction des mines terrestres et des armes à sous-munitions n'ont peut-être pas fait entièrement disparaître les mines terrestres et les armes à sous-munitions, ni supprimé leur emploi dans les conflits armés. Toutefois, depuis que ces traités existent, il est plus difficile, pour les fabricants, de produire ces armes de façon profitable et, pour les militaires, de les employer sans répercussions au sein de la communauté internationale.

Il apparaît en outre que si l'on prend pour postulat de départ le caractère inévitable des systèmes d'armement autonomes, leur acceptabilité est nécessairement tenue pour acquise. Pourtant, en définitive, ce qui compte, c'est de savoir quels critères internationaux seront utilisés pour déterminer l'acceptabilité de ces systèmes. En d'autres termes, quelles sont les normes de conduite qui seront jugées acceptables par la communauté internationale ? Accepter de postuler que le développement et l'emploi de ces technologies sont inéluctables empêche tout nouveau débat sur le bien-fondé et la pertinence de la démarche qui consiste à étudier, développer et employer ces technologies. Bref, comme dans le cas de toute autre technologie, ni la mise au point ni l'emploi de systèmes d'armement autonomes ne sont inévitables. Certes, de tels systèmes sont réalisables ; s'ils ne l'étaient pas, il n'y aurait aucun besoin de les interdire. Néanmoins, leur développement nécessite encore d'importants investissements. Même si nous ne parvenons pas à empêcher la création de certaines technologies, nous pourrions toujours défendre notre position quant à l'acceptabilité morale et juridique de leur emploi. Il ne suffit pas qu'une technologie existe pour que son emploi soit acceptable.

Les conséquences d'un développement progressif des systèmes d'armement autonomes

Permettez-moi de revenir au second postulat d'Anderson et Waxman, à savoir que le développement des systèmes d'armement autonomes ne sera pas brusque

mais progressif. Quelle contribution ce postulat est-il censé apporter à l'argumentation de ces auteurs ? Toutes les technologies connaissent, d'une manière ou d'une autre, ce type de développement. Pourquoi cela devrait-il modifier notre façon d'appréhender les implications éthiques et juridiques de telles innovations ? Anderson et Waxman tentent-ils simplement de dissiper la crainte d'un grand bond technologique à l'issue duquel des robots remplaceraient les soldats humains ? À mesure que ces auteurs développent leur argumentation, nous comprenons mieux ce qu'ils veulent dire : la transition vers les systèmes d'armement autonomes ne sera pas brutale ; elle se fera « à petits pas », chacun de ces pas étant soigneusement examiné. Il s'agit là d'une inversion assez habile de l'argument de la « pente glissante ». Anderson et Waxman ne disent pas que ces technologies sont dangereuses parce qu'elles nous encouragent à déléguer toujours davantage d'autorité aux systèmes automatisés (ce qui, à terme, nous amènerait à doter les robots de l'autorité – illégitime mais réelle – de cibler et de tuer des êtres humains). Ces systèmes seront légitimes parce que chaque étape de leur mise au point nous aura paru acceptable. En quelque sorte, nous devrions accepter la conclusion de cette ligne de raisonnement au motif que nous avons pu l'atteindre grâce à une série d'ajustements moraux, dont aucun n'est apparu, en lui-même, trop détestable.

Il serait plus sensé de considérer ce processus comme une pente glissante qui nous mène à un résultat qui nous paraît inacceptable. Nous devrions, dès lors, rechercher avec davantage de soin un principe sous-jacent que nous pourrions invoquer afin de stopper notre périlleuse glissade vers une conclusion indésirable. Or, une limite fondée sur des principes existe, et nous pouvons l'imposer aux systèmes d'armement autonomes. Cette prescription peut être énoncée ainsi : un être humain doit être impliqué de manière significative chaque fois que la décision d'employer la force létale doit être prise. Nous pourrions, certes, assouplir cette règle en faisant intervenir divers systèmes technologiques de contrôle partagé et de contrôle hiérarchique. Nous pourrions aussi concevoir les systèmes d'armement autonomes de manière telle que la règle imposée soit au contraire plus claire, et que les décisions létales soient mieux informées³⁴.

De la nécessité d'établir des normes

Les conclusions d'Anderson et Waxman sont erronées en ce qui concerne les implications d'une interdiction des systèmes d'armement autonomes. Par contre, ces auteurs ont raison sur deux points, à savoir la signification de l'établissement de normes visant à régir l'emploi de ces systèmes et la nécessité d'imposer certaines contraintes :

Les États-Unis doivent néanmoins agir avant que les attentes internationales relatives à ces technologies se durcissent, sous l'influence de ceux qui souhai-

34 P. Asaro, *op. cit.*, note 29, pp. 20-24 .

metaient imposer des interdctions irréalistes, ineffectives ou dangereuses, ou de ceux qui préféreraient peu de contraintes, ou pas de contraintes du tout³⁵.

Anderson et Waxman prennent donc acte ici du fait que ces technologies ouvrent effectivement un nouvel espace moral. Nul ne sait à ce stade ce que la communauté internationale acceptera en tant que nouvelles normes applicables à la conduite des hostilités à l'ère de la robotique et de l'automatisation. Ces auteurs ont aussi raison d'affirmer que les États-Unis – à la fois superpuissance et chef de file dans le développement de ces nouvelles technologies – sont mieux placés que quiconque pour créer des précédents et établir les normes qui façonneront l'avenir des conflits armés. Par contre, Anderson et Waxman n'ont pas montré en quoi il serait irréaliste d'interdire les systèmes d'armement autonomes; ils n'ont pas non plus apporté la moindre preuve qu'une telle interdiction serait inefficace ou immorale. Examinons donc tour à tour chacune de ces thèses.

Comment devons-nous comprendre l'argument selon lequel une interdiction des systèmes d'armement autonomes serait irréaliste? Cela signifie-t-il qu'une telle interdiction serait difficile à implanter dans la pratique? Tous les traités de maîtrise des armements présentent des défis de mise en œuvre. L'interdiction des systèmes d'armement autonomes ne devrait pas être exceptionnellement plus difficile que d'autres à mettre en application – elle n'est donc pas irréaliste. Devons-nous plutôt comprendre qu'il serait politiquement difficile de rallier un soutien suffisant pour une telle interdiction? D'après mon expérience personnelle, je pense qu'une telle interdiction trouverait un soutien auprès de nombreux individus (en particulier parmi les officiers militaires et les responsables politiques, mais aussi parmi les ingénieurs et les dirigeants de l'industrie de la défense). De par ma pratique du dialogue avec le public, je sais aussi que les systèmes d'armement automatisés, au même titre que les risques potentiels qui leur sont liés, suscitent des appréhensions morales, fortes et largement répandues. Je dirai donc *a minima* qu'une interdiction n'est pas irréaliste dans la mesure où elle est susceptible d'obtenir un large soutien de la part du public et des autorités.

En fait, la seule raison de considérer qu'une telle interdiction est irréaliste serait d'accepter le premier postulat (non démontré) d'Anderson et Waxman proclamant le caractère inéluctable des systèmes d'armement autonomes. Il est vrai que si nous acceptons comme une fatalité le développement de ces systèmes, toute tentative visant à empêcher que l'inévitable se produise paraît effectivement irréaliste. Il n'y a toutefois rien d'inévitable dans le développement d'une technologie émergente, dont les capacités n'existent pas encore, et au sujet de laquelle aucune norme n'a encore été établie.

Anderson et Waxman anticipent également un argument contre les systèmes d'armement autonomes, fondé sur un principe moral:

35 K. Anderson et M. C. Waxman, *op. cit.*, note 3, p. 2. [Traduction CICR]

Une deuxième objection, d'ordre moral, consiste à dire que le retrait de tout agent moral humain de la « boucle de tir » constitue simplement une erreur en soi. Une machine, aussi performante fût-elle, ne peut pas totalement remplacer la présence d'un véritable agent moral, c'est-à-dire d'un être humain possédant une conscience et la faculté de jugement moral (malgré les imperfections de la nature humaine). D'ailleurs, le titre du présent document est délibérément provocateur puisqu'il assimile « robot » à « soldat », alors que c'est précisément ce que l'on ne devrait jamais tenter de faire.

Il s'agit là d'un argument auquel il est difficile de répondre, car il s'arrête sur un principe moral que l'on ne peut qu'accepter ou refuser³⁶.

L'objection dont parlent Anderson et Waxman s'appuie sur ce qui est censé être une sorte de principe en soi³⁷. Ces auteurs supposent dès lors que la seule raison d'accepter ce principe serait de se fier à ses propres intuitions morales. À mon avis, les arguments présentés plus haut dans le présent article ont clairement démontré que le principe moral au nom duquel nous nous opposons aux systèmes d'armement autonomes est en fait contenu de manière implicite dans le DIH et s'exprime à travers ses diverses formulations et exigences anthropocentriques. Ce principe est en outre implicite dans la structure même des lois et de l'administration de la justice (droit à une procédure régulière, notamment). Si la présence d'un humain est requise, c'est en tant qu'« agent juridique », indépendamment de la nécessité qu'il soit un « agent moral ».

Ce n'est pas uniquement parce que la décision de tuer est lourde à prendre (bien qu'elle le soit, assurément). La décision de tuer un humain ne peut être légitime que si elle est non arbitraire. Or, en l'absence de contrôle, de supervision et de responsabilité exercés par un humain, il n'y a aucun moyen de garantir le caractère non arbitraire de l'emploi de la force. Tuer sans que s'exercent la raison, le jugement et la compassion d'un humain est donc immoral – et devrait être illégal.

Conclusions

Afin de préserver la moralité humaine, la dignité, la justice et le droit, nous ne pouvons pas accepter qu'un système automatisé prenne la décision de tuer un humain. Nous devrions, afin de respecter ce principe, prohiber les systèmes d'armement autonomes. Quand il est question de vie ou de mort, chaque cas mérite d'être examiné avec attention et considération par un humain, étant donné le poids moral que porte en soi la suppression d'une vie humaine.

À mesure qu'elle progresse, la technologie permet à l'humanité d'exercer sur le monde un contrôle toujours plus étendu. Ce nouveau contrôle s'accompagne d'une responsabilité accrue. Bien sûr, cela semble évident

³⁶ *Idem*, p. 11.

³⁷ M. Bolton, T. Nash et R. Moyes, *op. cit.*, note 11.

dans le cas des technologies qui influencent le bien-être de la population et la protection de l'environnement, mais cela vaut également pour les technologies militaires. Le développement de technologies militaires avancées n'implique pas nécessairement que celles-ci peuvent être utilisées (et seront utilisées) avec davantage de prudence et de manière plus éthique, même si une telle possibilité existe. Mais les nouvelles capacités s'accompagnent d'un risque de régression sur les plans de l'éthique et de la moralité, plutôt que d'un potentiel de progrès. En définitive, la technologie déterminera la nature des progrès que nous ferons sur le plan moral dans la manière de conduire les hostilités. L'impact sera plus profond que le simple fait de donner aux combattants la possibilité de conduire des guerres peu coûteuses en vies humaines. Cela va au-delà des exigences du DIH et du DIDH. En choisissant les armes et les tactiques avec lesquelles nous nous engageons dans un conflit armé, nous opérons également un choix moral quant au monde dans lequel nous souhaitons vivre et pour lequel nous voulons nous battre, et quant aux conditions légitimes dans lesquelles nous pouvons faire advenir ce monde. Au moment de ces choix, nous devons résister aux arguments voulant que « la fin justifie les moyens ». Nous devons aussi reconnaître que les moyens mobilisés pour opérer des changements dans le monde, ou pour résister à de tels changements, deviennent ainsi un aspect de ce monde. Si nous souhaitons vraiment construire un monde dans lequel les conflits armés seront à la fois inutiles et inacceptables, nous devons, pour y parvenir, adopter un processus permettant que chaque nouvelle innovation technologique vienne élever, et non abaisser, nos principes moraux.

La communauté internationale devrait entamer des discussions sur l'élaboration d'un traité d'interdiction des systèmes d'armement autonomes. Dans la mesure où de tels systèmes n'existent pas encore, le fait de les interdire contribuerait à infléchir le développement des technologies militaires de demain en l'axant non pas sur ce que l'on nomme les « systèmes éthiques », mais sur des systèmes qui soient capables de réellement améliorer le comportement éthique des humains dans les conflits armés. Ceux qui s'opposent à une telle interdiction basent leur argumentation tant sur des affirmations infondées de l'inévitabilité de ces technologies que sur des allégations mensongères concernant des technologies éthiquement évoluées. Aussi longtemps que les capacités potentielles de ces technologies resteront incertaines, leur émergence aura lieu dans un champ de normes éthiques et juridiques évoluant rapidement. Bien sûr, nous aimerions pouvoir faire confiance à la promesse de guerres plus éthiques, rendues possibles par ces hypothétiques systèmes d'armement autonomes. La réalité est malheureusement différente : ces systèmes peuvent aussi dégrader nos conceptions et nos critères en matière de conduite éthique ; en nous incitant à poursuivre la mise au point d'une technologie improbable qui menace nos droits humains au niveau fondamental, ils risquent en outre de nous amener à ne plus chercher à améliorer le raisonnement moral humain par le biais des technologies. Ces systèmes risquent aussi de nous dissuader de continuer à renforcer et à améliorer le DIH et le DIDH afin qu'ils soient à même de relever de façon appropriée et moralement acceptable les défis posés par ces nouvelles technologies.