

Tadeusz ZIELIŃSKI¹

DYLEMATY UŻYTKOWANIA AUTONOMICZNYCH SYSTEMÓW BOJOWYCH W ODNIESIENIU DO PODSTAWOWYCH ZASAD MIĘDZYNARODOWEGO PRAWA HUMANITARNEGO

Wizja, w której maszyny będą podejmowały decyzje o życiu lub śmierci na polu walki staje się coraz bardziej realna. Symptomatyczne w tym kontekście wydaje się użytkowanie bezzałogowych statków powietrznych w konfliktach ostatniej dekady: w Iraku, Afganistanie, Libii, Syrii czy na Ukrainie. Zwiększająca się robotyzacja pola walki wywołuje wiele kontrowersji, szczególnie wśród organizacji pozarządowych, włącznie z postulatami całkowitego zakazu rozwoju i użytkowania autonomicznych systemów bojowych. Trwająca debata publiczna wyraźnie się zaostrza, a przyszłościowe użytkowanie przedmiotowych systemów ma zarówno swoich zwolenników jak i przeciwników. Kluczowe w tym kontekście, wydają się aspekty prawne wykorzystania autonomicznych systemów bojowych na polu walki.

Treści artykułu prezentują zasadnicze podejścia w definiowaniu autonomicznych systemów bojowych. Wskazano, że na gruncie prawa międzynarodowego brak jest jednoznacznej definicji tego rodzaju środka walki. Zwrócono uwagę na brak jednoznacznego rozgraniczenia pomiędzy systemami zautomatyzowanymi, a autonomicznymi. To z kolei powoduje problemy związane z właściwą interpretacją prawną użytkowania autonomicznych systemów bojowych. Zostały również opisane dylematy potencjalnego użytkowania autonomicznych systemów bojowych w odniesieniu do podstawowych zasad międzynarodowego prawa humanitarnego: rozróżniania, proporcjonalności oraz środków ostrożności. Zasadnicza konkluzja potwierdza tezę, że obecnie funkcjonujące normy prawa międzynarodowego nie wymagają zmian w kontekście użytkowania autonomicznych systemów bojowych. Jednocześnie, konieczna jest pogłębiona dyskusja mająca na celu rozstrzygnięcie pojawiających się wątpliwości. Wiążące ustalenia powinny zapaść na forum międzynarodowym przed rzeczywistym wprowadzeniem autonomicznych systemów bojowych do uzbrojenia sił zbrojnych.

Słowa kluczowe: automatyzacja, autonomiczne systemy bojowe, międzynarodowe prawo humanitarne, podejmowanie decyzji, robotyzacja pola walki.

¹ Płk dr hab. Tadeusz Zieliński¹, prof. ASzWoj, Wydział Zarządzania i Dowodzenia, Akademia Sztuki Wojennej, al. gen. Antoniego Chruściela „Montera” 103, 00-910 Warszawa; e-mail: www.akademia.mil.pl

Col. Tadeusz Zieliński, DSc, PhD, Faculty of Management and Command, War studies University; al. gen. Antoniego Chruściela „Montera” 103, 00-910 Warszawa; e-mail: www.akademia.mil.pl

1. WPROWADZENIE

Powszechne użycie bojowych bezzałogowych systemów powietrznych² w operacjach na terenie Iraku, Afganistanu, Syrii czy też przeciwko tzw. państwu islamskiemu wywołało głośną debatę publiczną na temat zastosowania nowych technologii w konfliktach zbrojnych. Nowe bronie, zawierające elementy automatyzacji lub autonomiczności wywołują wiele kontrowersji w dyskursie medialnym, włącznie z postulatami zakazu rozwoju i użytkowania w pełni autonomicznych systemów bojowych. Można zaryzykować stwierdzenie, że konwencjonalna technologia bezzałogowych systemów powietrznych zmieniła tradycyjne prowadzenie działań bojowych. Zdalne sterowanie bezzałogowymi statkami powietrznymi umożliwia zastosowanie uzbrojenia bez bezpośredniego zagrożenia dla personelu, który przeprowadza działania z odległego ośrodka dowodzenia (sterowania) i śledzi jego rezultaty na ekranie komputera. Eksperci uważają rozwój systemów autonomicznych za prawdziwie rewolucyjne zjawisko i wskazują, że stanowi to istotną zmianę w paradygmacie technologii wojskowej³. Jest to spowodowane tym, że w przeciwieństwie do obecnego stanu rzeczy, autonomiczne systemy bojowe nie tylko usprawniają techniczne możliwości stosowanych systemów, ale także spowodują, że w przyszłości człowiekowi pozostanie tylko podstawowa decyzja, związana z zastosowaniem danego systemu. Wszystkie dalsze, wynikowe decyzje dotyczące konkretnego użycia zostaną delegowane autonomicznemu systemowi⁴.

Obecnie na rynku nie funkcjonują w pełni autonomiczne systemy bojowe. Niektórzy inżynierowie robotyki wątpią nawet czy takie systemy będą mogły zostać kiedykolwiek zaprojektowane⁵. Z drugiej strony, większość ekspertów technicznych zakłada, że to tylko kwestia czasu, zanim takie systemy zostaną zaprojektowane i będą gotowe do wykorzystania. Potwierdzeniem może być fakt, że amerykański departament obrony opracował oficjalny plan wprowadzenia do służby systemów autonomicznych w sposób stopniowy, do 2038 roku⁶. W tym samym czasie, przedstawiciel Międzynarodowego Komitetu Czerwonego Krzyża (MKCK), w przemówieniu dla ONZ, słusznie zauważył, iż już dziś coraz więcej „krytycznych funkcji w systemach walki działa w sposób autonomiczny, tj. bez

² Bezzałogowy system powietrzny – system składający się z bezzałogowego statku powietrznego, systemu wsparcia oraz całości wyposażenia i personelu niezbędnego do kierowania bezzałogowym statkiem powietrznym, AAP-6 (2014) *Słownik terminów i definicji NATO*, NATO Standardization Agency, Bruksela 2014, s. 405.

³ Zob. P.W. Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century*, The Penguin Press, London 2009, s. 179–204.

⁴ Zob. *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, Christof Heyns, 26 April 2013, United Nations, General Assembly, A/HRC/23/47.

⁵ Zob. M.E. O’Connell, *Banning Autonomous Killing: The Legal and Ethical Requirement that Humans Make Near-Time Lethal Decisions* [w:] *The American Way of Bombing. Changing Ethical and Legal Norms, from Flying Fortresses to Drones*, red. M. Evangelista, H. Shue, Cornell University Press, New York 2014, s. 224, 226.

⁶ Zob. *Unmanned Systems Integrated Roadmap FY2013–2038*, US Department of Defense, Washington DC 2013, <https://www.defense.gov/Portals/1/Documents/pubs/DOD-USRM-2013.pdf> (dostęp: 6 listopada 2017 r.).

ingerencji człowieka”⁷. To z kolei generuje fundamentalne pytania na gruncie prawa międzynarodowego. Biorąc powyższe pod uwagę, celem artykułu jest identyfikacja autonomicznych systemów bojowych oraz wątpliwości związanych z ich użytkowaniem w kontekście podstawowych zasad międzynarodowego prawa humanitarnego.

2. AUTONOMICZNE SYSTEMY BOJOWE – KONTEKST DEFINICYJNY

W literaturze naukowej i oficjalnych dokumentach występuje wiele podejść do definiowania autonomicznych systemów bojowych. Obecnie nie istnieje żadna uniwersalna i akceptowalna na gruncie prawa międzynarodowego definicja. Jednakże, wspólną cechą w definiowaniu autonomicznych systemów bojowych jest to, że poziom zdolności w zakresie podejmowania decyzji przez zaprogramowane algorytmy, bez ingerencji człowieka, jest podstawowym kryterium decydującym o „autonomii”. Dlatego, autonomia sama w sobie, nie powinna być rozumiana w moralnie-filozoficznym sensie jako wolna wola osoby. Nawet autonomiczny robot może funkcjonować wyłącznie w ograniczeniach możliwości, które zostały w nim zaprogramowane za pomocą algorytmów. Z drugiej strony wyciąganie z tego wniosku, że nie może faktycznie istnieć żaden naprawdę autonomiczny system, ponieważ na którymś z etapów człowiek jest zawsze zaangażowany w macierz decyzyjną⁸, jest nieuzasadnionym upraszczaniem tych kwestii. Osoby, które używają tego argumentu nie widzą problemów związanych z funkcjonowaniem autonomicznych systemów bojowych. Według nich, ostatecznie wszystko pozostaje tak naprawdę bez zmian, ponieważ każdą decyzję podjętą przez autonomiczny system bojowy można odnieść do ludzkiej decyzji. Jednakże, niniejszy punkt widzenia nie uwzględnia poziomu autonomii, który może zostać zaimplementowany w dłuższej perspektywie w związku z wojskowymi badaniami nad nowymi technologiami. Nie trzeba dodawać, że wpływ człowieka nigdy nie zostanie całkowicie wykluczony. W konsekwencji, to ludzie decydują czy wdrożenie autonomicznego systemu bojowego powinno się odbyć, a w trakcie programowania decydują również o warunkach ramowych. Zatem – jak na razie – to ludzie podejmują decyzję o aktywacji robotów oraz o ich dezaktywacji. Jednakże, gdy decyzja o użyciu chociażby potencjalnie śmiertelnego uzbrojenia zostaje podjęta przed faktycznym wdrożeniem, nie można już mówić o jakiegokolwiek prawdziwej ludzkiej kontroli w kluczowym momencie. W bardzo skomplikowanych sytuacjach konfliktowych, scenariusze, z którymi zostaną skonfrontowane autonomiczne systemy bojowe wysłane na misję, nie mogą zostać w pełni przewidziane.

Nierozzerwalnie z autonomicznością związane są poziomy automatyzacji. W przeciwieństwie do autonomicznych systemów bojowych, systemy zautomatyzowane wykonują wyłącznie wcześniej zaprogramowane polecenia w przewidywalny sposób. Nie są w stanie samodzielnie reagować na nieprzewidziane zdarzenia. Takie systemy są zazwyczaj stosowane do przeprowadzenia w całości określonego, wcześniej ustalonego zadania (np. obrony przed pociskami przeciwokrętowymi) w ściśle określonym obszarze działań (np. 5 mil morskich wokół ochranianego okrętu) i przedziale czasowym. Miny również zaliczają się – jako

⁷ *Weapons: ICRC statement to the United Nations, General Assembly, 69th session, First Committee, 14 October 2014, New York*, <https://www.icrc.org/en/document/weapons-icrc-statement-united-nations-2014> (dostęp: 6 listopada 2017 r.).

⁸ Zob. M.N. Schmitt, J.S. Thurnher, *Out of the Loop: Autonomous Weapons Systems and the Law of Armed Conflict*, “Harvard National Security Journal”, Vol. 4, 2013, s. 231–280.

możliwie najprostszy przykład – do niniejszej koncepcyjnej definicji zautomatyzowanych systemów bojowych. Im bardziej skomplikowana dziedzina funkcji oraz im większy lub bardziej zróżnicowany obszar wdrożenia, tym bardziej rozmyte staje się rozróżnienie pomiędzy systemami autonomicznymi i zautomatyzowanymi. Nie można z całkowitą pewnością powiedzieć, gdzie dokładnie znajduje się granica w odniesieniu do tego kontinuum, a precyzyjne rozróżnienie pomiędzy systemami autonomicznymi i automatycznymi nie zawsze jest możliwe. Jednakże prawdopodobnie nie jest to nawet konieczne dla wyjaśnienia istotnych kwestii prawnych. Eksperti MKCK aktualnie nie stosują rozróżnienia pomiędzy systemami autonomicznymi i automatycznymi. Według nich, tym co ma decydujące znaczenie jest to czy krytyczne funkcje – czyli niezależne podejmowanie decyzji o życiu i śmierci – są powierzone systemowi⁹. Niniejsze podejście jest bardzo przekonujące. Niezależnie od definicji technicznych, tym co ma decydujące znaczenie z etycznego i prawnego (międzynarodowego) punktu widzenia jest to, gdzie wyznaczona została granica pomiędzy delegowanymi decyzjami oraz decyzjami, które muszą pozostać w bezpośredniej kontroli człowieka. Podejście MKCK daje również jasno do zrozumienia, że funkcjonowanie autonomicznych systemów bojowych nie w każdym przypadku musi być problematyczne. Autonomicznie funkcjonujące łodzie podwodne służące do wyszukiwania min lub autonomiczne systemy ofensywne wykorzystywane wyłącznie w przestrzeni kosmicznej do neutralizowania wrogich satelitów wojskowych nie wzbudzają takich samych pilnych kwestii etycznych i prawnych, co robot bojowy zaprogramowany do walki w mieście, który bezpośrednio napotyka ludzi, zarówno wrogich bojowników (żołnierzy), jak i osoby niewalczące (ludność cywilną).

Wnioski z analizy dostępnej literatury przedmiotu wskazują, że w kwestie definicyjne odnoszące się do autonomicznych systemów oraz ich identyfikacja i kategoryzacja opierają się na kilku podejściach, które dominują w publicznej debacie.

Specjalny Sprawozdawca ONZ, Christof Heyns, opisuje autonomiczne systemy bojowe jako roboty gromadzące dane na temat swojego otoczenia za pomocą sensorów, które następnie są przetwarzane w celu podjęcia decyzji, ostatecznie realizowanej przez zainstalowane komponenty (np. uzbrojenie lub środki transportu)¹⁰. Zgodnie z (szeroką) definicją departamentu brony USA, systemy bojowe są autonomiczne, jeżeli po ich aktywacji są zdolne do niezależnego, czyli bez dalszej interwencji człowieka wyboru celu i jego zaatakowania. Autonomia występuje również wtedy, gdy ludzie podejmujący decyzje są w stanie anulować konkretne decyzje dotyczące ataku, nawet po aktywacji¹¹. Natomiast MKCK za decydujący czynnik w definicji uznaje również niezależność zainstalowanych krytycznych systemów związanych z wyborem celu, jego rozpoznaniem, śledzeniem, i atakowaniem¹². We wnioskach z badań odnoszących się do autonomicznych systemów bojowych, organizacja pozarządowa *Human Rights Watch* przedstawiła podejście, które jest bardzo często

⁹ Zob. *Weapons: ICRC statement to the United Nations...*

¹⁰ *Report of the Special...*, s. 7–8.

¹¹ *Autonomy in Weapons Systems. Directive 3000.09*, U.S. Department of Defense, 21 November 2012, s. 13, <http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf> (dostęp: 12 listopada 2017 r.).

¹² *Report of the Expert Meeting on Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*, 26–28 March 2014, Geneva, 9 May 2014, s. 1, <https://shop.icrc.org/expert-meeting-autonomous-weapon-systems-technical-military-legal-and-humanitarian-aspects-26-to-28-march-2014.html> (dostęp: 12 listopada 2017 r.).

powielane w literaturze¹³. Kryterium wyjściowym definicji jest poziom ludzkiego uczestnictwa w podejmowaniu decyzji przez system w odniesieniu do wykonywania przez niego poszczególnych funkcji: obserwacji, orientacji, podejmowania decyzji i działań w odpowiedzi na daną sytuację lub środowisko. W tym kontekście wyróżnia się systemy semi-autonomiczne (ang. *human in the loop*), w których istnieje możliwość automatycznego działania do pewnego momentu, a następnie konieczna jest ingerencja człowieka. Przykładowo, do tej kategorii należą bezzałogowe systemy powietrzne wykorzystywane przez USA w Pakistanie, Afganistanie i Jemenie, które są kontrolowane zdalnie z odległego ośrodka dowodzenia. Grupa ta obejmuje również uzbrojenie, w przypadku którego wybór celu jest dokonywany przez człowieka, lecz namierzanie celu i jego zniszczenie odbywa się już niezależnie. Druga grupa obejmuje systemy nadzorowane (ang. *human on the loop*), w których istnieje możliwość nieprzerwanego działania autonomicznego, ale z możliwością w dowolnej chwili ingerencji człowieka. Ludzie wyłącznie monitorują działania systemu i mogą interweniować jeżeli jest to konieczne, aby odwołać decyzje maszyny. Ten wariant określany jest również jako monitorowana autonomia, jednakże Specjalny Sprawozdawca ONZ, Heyns, wskazuje, że w rzeczywistości opcja interwencji może być bardzo ograniczona, jeśli robot podejmuje decyzje w ciągu kilku milisekund. W tym przypadku, znacząca kontrola nie jest *de facto* możliwa¹⁴. Trzecia grupa, to systemy w pełni autonomiczne (ang. *human out of the loop*), bez możliwości ingerencji człowieka. Maszyna działa niezależnie w odniesieniu do wszystkich niezbędnych kroków w przeprowadzaniu ataku. Ponadto, oczekuje się, że przyszłe generacje robotów będą miały możliwość uczenia się „z własnych doświadczeń” i w ten sposób będą niezależnie dostosowywać swoje funkcjonowanie, a może nawet je rozszerzać¹⁵.

Warto również wspomnieć o propozycji profesora Noela Sharkey’a zajmującego się sztuczną inteligencją i robotyką. Jego zdaniem, należy rozróżniać pięć poziomów ludzkiej kontroli nad systemami uzbrojenia: a) człowieka sam rozpatruje cele przed zainicjowaniem jakiegokolwiek ataku; b) program komputerowy dostarcza listę celów i człowiek wybiera, który z nich zaatakować; c) program wybiera cel, a człowiek musi dać zgodę przed przeprowadzeniem ataku; d) program wybiera cel, a człowiek ma ograniczony czas, w którym może odwołać atak; e) program wybiera cel i inicjuje atak bez udziału człowieka¹⁶.

Przedstawione powyżej definicje i podejścia do autonomicznych systemów bojowych należy traktować jako schematyczne i uproszczone. W rzeczywistości, granice są często niewyraźne. Można sobie wyobrazić prawdopodobne scenariusze, w których żołnierze pozostają decyzyjni wyłącznie formalnie, a ważne aspekty procesów decyzyjnych są przekazywane maszynie, np. wybór celu zgodnie z trzecim poziomem według Sharkey’a. Nawet jeśli podejmowanie decyzji w sprawie rzeczywistego zastosowania autonomicznego systemu bojowego pozostaje w rękach człowieka, to trzeba przyjąć, że w sytuacjach stresowych i pod presją czasu – typowe cechy każdej misji bojowej – zaangażowane osoby będą

¹³ *Losing Humanity: The Case against Killer Robots*, Human Rights Watch 2012, s. 2, https://www.hrw.org/sites/default/files/reports/arms1112_ForUpload.pdf (dostęp: 12 listopada 2017 r.).

¹⁴ *Report of the Special...*, s. 8.

¹⁵ M.N. Schmitt, J.S. Thurnher, *Out of the Loop: Autonomous Weapons...*, s. 240.

¹⁶ N. Sharkey, *Towards a Principle for the Human Supervisory Control of Robot Weapons*, 2014, s. 11, [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/2002471923EBF52AC1257CCC0047C791/\\$file/Article_Sharkey_PrincipleforHumanSupervisory.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/2002471923EBF52AC1257CCC0047C791/$file/Article_Sharkey_PrincipleforHumanSupervisory.pdf) (dostęp: 12 listopada 2017 r.).

w razie wątpliwości powierzać zadanie maszynie. Takie zachowanie, z tendencją do ufania zautomatyzowanej maszynie, nawet jeśli występują istotne wskazania, że jest zawodna lub robi błędy w niektórych przypadkach – jest znane pod nazwą „inklinacja ku automatyce”¹⁷. W takim przypadku, ludzka kontrola, chociaż technicznie możliwa, w rzeczywistości jest bez znaczenia.

3. ZASADY MIĘDZYNARODOWEGO PRAWA HUMANITARNEGO W ODNIESIENIU DO AUTONOMICZNYCH SYSTEMÓW BOJOWYCH

Należy zauważyć, że w odniesieniu do najważniejszych kwestii prawnych, autonomia sama w sobie nie wzbudza znaczących kontrowersji. Niezależnie działające poszukiwacze min lub roboty przeznaczone do rozbrajania bomb w prosty sposób uzasadniają to stwierdzenie. Kontrowersje może wzbudzać delegowanie maszynom krytycznych decyzji generujących potencjalne powstanie problemów etycznych i prawnych. Szczególnie krytyczne jest przekazanie decyzji odnoszących się do życia ludzkiego. Przykłady pojazdów bez kierowcy oraz robotów medycznych wskazują, że niniejsze zagadnienie nie ogranicza się jedynie do wojskowych systemów bojowych. Jednakże, w odniesieniu do tych ostatnich, wyzwania prawne i etyczne wysuwają się na pierwszy plan w związku z faktem, że są one projektowane i używane do działań, które mogą powodować straty ludzkie i materialne. W tym kontekście zasadnym jest sformułowanie pytania: czy autonomiczne systemy bojowe mogą być rozwijane i stosowane w zgodzie z podstawowymi zasadami międzynarodowego prawa humanitarnego (MPH)?

W debacie na temat autonomicznych systemów bojowych, do tej pory skupiano się na kwestii czy takie systemy byłyby zdolne do przestrzegania zasad międzynarodowego prawa humanitarnego. Są to przepisy prawa międzynarodowego, które mają zastosowanie w czasie trwania konfliktu zbrojnego. Modyfikują one kluczowe przepisy dotyczące praw człowieka, w szczególności prawa do życia. Zatem, w trakcie konfliktu zbrojnego prawo do życia jest przyznawane wyłącznie według międzynarodowego prawa humanitarnego. Oznacza to między innymi, że wroddzy żołnierze i bojownicy mogą być atakowani w zasadzie w każdej chwili, jako uzasadnione cele wojskowe, nawet jeśli nie stanowią bezpośredniego zagrożenia dla innych stron konfliktu w danym momencie. W odniesieniu do międzynarodowego prawa humanitarnego zasadnym jest przeanalizowanie podstawowych zasad tego prawa w kontekście autonomicznych systemów bojowych.

Po pierwsze, przestrzeganie zasady rozróżniania. Wiąże się to bezpośrednio z odpowiedzią na pytanie czy możliwe jest zaprogramowanie autonomicznych systemów bojowych w taki sposób, aby były one w stanie rozróżnić ochraniających cywilów od wrogich bojowników – nawet w skomplikowanych scenariuszach sytuacyjnych. Ludność cywilna nigdy nie jest uzasadnionym celem, nawet w konfliktach zbrojnych. Na tym założeniu opiera się zasada rozróżniania, która została w szczególności określona w Artykule 51 I Protokołu Dodatkowego do Konwencji Genewskiej, mająca także zastosowanie na mocy prawa zwyczajowego wszystkich państw. Z ustępu 2 tego przepisu jasno wynika, że ani populacje

¹⁷ P. M. Asaro, *Modelling the Moral User*, “IEEE Technology and Society Magazine”, Vol. 28, Issue 1, Spring 2009, s. 22, <http://peterasaro.org/writing/Asaro%20Modeling%20Moral%20User.pdf> (dostęp: 20 listopada 2017 r.).

cywilne, ani pojedyncze osoby cywilne nie mogą być przedmiotem ataku¹⁸. W odniesieniu do autonomicznych systemów bojowych wynika z tego kilka problemów. Przede wszystkim musi zostać wyjaśnione czy sensory takich systemów będą w stanie dokonać niezbędnego rozróżnienia z wystarczającą wiarygodnością. Już teraz jest to dużym wyzwaniem technologicznym i niektórzy eksperci robotyki uważają, iż jest to niewykonalne¹⁹. Jednakże, poza czysto fizyczną zdolnością do rozróżniania, należy wziąć pod uwagę fakt, że przestrzeganie imperatywu rozróżniania wymaga wysoce skomplikowanych procesów oceny. W krytycznych sytuacjach konfliktu zbrojnego zawsze pojawią się problematyczne osądy moralne. Nawet zakładając znaczne postępy w technologii sensorów, pozostaje pytanie czy ten aspekt będzie kiedykolwiek obsługiwany przez algorytmy. Odnosi się to w szczególności do typowych sytuacji we współczesnych konfliktach zbrojnych, które charakteryzują się coraz większą niejasnością i złożonością. Pod takimi sloganami jak „asymetryczne działania bojowe” oraz „działania bojowe w terenie zurbanizowanym” kryją się ogromne trudności związane z jednej strony z rozróżnianiem pomiędzy nieregularnymi bojownikami i innymi uzasadnionymi celami stron konfliktu oraz ludnością cywilną – która musi być chroniona – z drugiej strony. Nie jest to tylko kwestia identyfikacji przez sensory poszczególnych typów uzbrojenia lub umundurowania, lecz przede wszystkim interpretacji ludzkich zachowań. Nie jest jasne czy systemy komputerowe kiedykolwiek będą w stanie spełnić nakaz rozróżniania w złożonych sytuacjach konfliktowych. Warto zauważyć w tym kontekście, że wytyczne armii Stanów Zjednoczonych dotyczące etycznego postępowania na misjach, wyraźnie obejmują kryterium „przeczcucia”. Żołnierz, przed użyciem broni ma zdecydować czy dane działanie jest moralnie właściwe. Nawet zwolennicy autonomicznych systemów bojowych przyznają, że takie rozważania nie mogą zostać osiągnięte za pomocą algorytmów²⁰. Szczególnie problematyczne wydaje się zaprogramowanie zasad zachowania w niejednoznacznych sytuacjach bojowych. Stąd, niektórzy eksperci co najmniej kwestionują pogląd, że jest możliwe wyposażenie autonomicznych systemów bojowych w zdolność do „wątpliwości”, która mogłaby skutecznie zapobiec kontynuowaniu ataku w takiej sytuacji²¹. Z drugiej strony, żołnierze mają już teraz do czynienia z taką samą trudnością. Popołniają błędy i mogą świadomie lub nieświadomie naruszać zasadę rozróżniania. Stres, gniew czy strach to czynniki, które mogą wywołać przestępstwa prawne lub czynią je bardziej prawdopodobnymi. Natomiast maszyny nie muszą radzić sobie z takimi emocjami. Jest to podnoszone jako główny argument, dzięki któremu autonomiczne systemy bojowe miałyby znacznie większe możliwości niż ludzie w odniesieniu do przestrzegania zasady

¹⁸ Artykuł 51 (2): Celem ataków nie mogą być ani ludność cywilna jako taka, ani osoby cywilne. Zabronione są akty i groźby przemocy, których głównym celem jest zastraszenie ludności cywilnej, Protokół dodatkowy do konwencji genewskich z 12 sierpnia 1949 r., dotyczący ochrony ofiar międzynarodowych konfliktów zbrojnych, Genewa, 8 czerwca 1977 r. (Dz.U. z 1992 r., nr 41, poz. 175, załącznik).

¹⁹ Zob. N.E. Sharkey, *The Evitability of Autonomous Robot Warfare*, „International Review of the Red Cross”, Vol. 94, 2012, s. 787, 788.

²⁰ R.C. Arkin, *Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture*, Technical Report GIT-GVU-07-11, s. 51, <http://www.cc.gatech.edu/ai/robotlab/online-publications/formalizationv35.pdf> (dostęp: 20 listopada 2017 r.).

²¹ Zob. M.N. Schmitt, *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, „Harvard National Security Journal Features”, 2013, s. 16, <http://harvardnsj.org/wp-content/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf> (dostęp: 23 listopada 2017 r.).

rozdzielania w złożonych sytuacjach²². Robotowi, który nie musi obawiać się o życie, znacznie łatwiej jest przestrzegać założeń wymaganego przez międzynarodowe prawo, że każdy jest cywilem, który musi być chroniony, aż do momentu, gdy faktycznie wyciągnie broń. Żołnierz w trosce o własne życie ma nadrzędny interes, aby odwrócić niniejsze założenie. Jednocześnie powstaje jednak pytanie, dlaczego roboty powinny mieć prawo do zabijania wrogich bojowników i działania na takich samych zasadach jak żołnierze. Jeżeli roboty są rzeczywiście lepsze od człowieka, jak często się twierdzi, i biorąc pod uwagę, że (pancerne) roboty mogłyby walczyć bez znacznego zagrożenia dla siebie, powinny one być również zobowiązane do odpowiednio wyższych, zgodnie z aktualnymi technologiami, standardów bezpieczeństwa.

Należy zauważyć, że w kontekście zasady rozróżniania, do rozważenia pozostaje więcej aspektów niż tylko sama ochrona ludności cywilnej. Żołnierze lub bojownicy przeciwnika, którzy poddali się lub są ranni, zalicza się do *hors de combat*²³ i przysługuje im ochrona zgodnie z normami prawa międzynarodowego. Jest co najmniej wątpliwe czy autonomiczne systemy bojowe będą w stanie niezawodnie rozpoznawać niniejsze kryteria rozróżniania. Aby było to możliwe, musiałyby one rozpoznawać gesty, mimikę twarzy oraz emocje i prawidłowo je oceniać. Wydaje się to prawie niemożliwe²⁴.

Kolejna z zasad międzynarodowego prawa humanitarnego – zasada proporcjonalności – koresponduje w kontekście autonomicznych systemów bojowych z pytaniem czy możliwe jest zaprogramowanie autonomicznych systemów bojowych w taki sposób, aby nadmierne straty i zniszczenia niezamierzone (ang. *collateral damage*) zostały bezpiecznie wykluczone. Poza zasadą rozróżniania, zasada proporcjonalności jest kolejnym podstawowym filarem międzynarodowego prawa humanitarnego. Pomimo tego, że zasada rozróżniania restrykcyjnie zabrania bezpośrednich ataków na ludność cywilną, dopuszczalne są w stopniu proporcjonalnym niezamierzone straty wśród ludności cywilnej i szkody materialne, na przykład, jeżeli zostanie zaatakowany bezpośrednio cel wojskowy i przypadkowo zostaną zabite lub ranne osoby cywilne. Niniejsza zasada, która ma również zastosowanie do wszystkich państw na mocy międzynarodowego prawa zwyczajowego, znajduje swój pisemny wyraz w art. 51 (5)(b)²⁵ oraz art. 57 (2)(a)(iii)²⁶ I Protokołu Dodatkowego do Konwencji Genewskich. Rozważania na temat powiązań pomiędzy przewidywaną przewagą

²² R.C. Arkin, *Governing Lethal Behavior...*, s. 6.

²³ Francuski termin używany w dyplomacji i prawie międzynarodowym, jako określenie żołnierza, który jest niezdolny do pełnienia swoich wojskowych funkcji [przyp. aut.].

²⁴ M. Sassòli, *Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to Be Clarified*, "International Law Studies", Vol. 90, 2014, s. 308, 322.

²⁵ Następujące rodzaje ataków będą między innymi traktowane jako wykonane bez rozróżnienia: ataki, co do których można przypuszczać, że wywołają również straty w życiu ludzkim wśród ludności cywilnej, ranienia osób cywilnych, szkody w dobrach o charakterze cywilnym lub połączenie tych strat i szkód, jeśli byłyby one nadmierne w porównaniu z oczekiwaną konkretną i bezpośrednią korzyścią wojskową, *Protokół dodatkowy do konwencji...*, art. 51 (5) b.

²⁶ Planujący lub decydujący o podjęciu ataku powinni: powstrzymać się od podjęcia ataku, który mógłby spowodować niezamierzone straty w życiu ludzkim wśród ludności cywilnej, ranienia osób cywilnych, szkody w dobrach o charakterze cywilnym lub takie straty i szkody łącznie, których rozmiary byłyby nadmierne do oczekiwanej konkretnej i bezpośredniej korzyści wojskowej, tamże, art. 57 (2)(a)(iii).

wojskową i możliwymi stratami wśród ludności cywilnej wymagają złożonego procesu podejmowania decyzji opartego na wartościach i konkretnych przypadkach, w których dane okoliczności muszą zostać uwzględnione w kontekście całości. Ponownie pojawia się pytanie czy może to zostać przeprowadzone za pomocą obliczeń z wykorzystaniem algorytmów. Czy autonomiczny system bojowy mógłby w rzeczywistości właściwie ocenić zalety wojskowej przewagi w operacji? Jest wątpliwe, aby autonomiczne systemy bojowe, przynajmniej w najbliższej przyszłości, mogły zostać zaprogramowane do dokonania takiej oceny. Jeżeli przedmiotowa zasada w odniesieniu do autonomicznych systemów bojowych nie mogłaby zostać zaimplementowana lub systemy te nie będą mogły z całkowitą pewnością dokonać rozróżnienia pomiędzy ochranianą ludnością cywilną, a żołnierzami i bojownikami przeciwnika, to mogłyby być stosowane wyłącznie w sytuacjach, w których ludność cywilna nie jest obecna w działaniach bojowych. Przeciwnicy tego argumentu dowodzą, że wymagane programowanie, przypuszczalnie jest możliwe do osiągnięcia w stopniu, w jakim byłoby to równoważne z ludzką oceną w tej samej sytuacji²⁷. Naukowiec zajmujący się tematyką robotyki, Ronald Arkin idzie nawet o krok dalej i wskazuje na brak instynktu przetrwania w przypadku autonomicznych systemów bojowych. Według niego, autonomiczne systemy bojowe nie mogą zostać wprowadzone w błąd i użyć nadmiernej siły zbrojnej, co oznacza, że przestrzeganie zasady proporcjonalności mogłoby zostać zapewnione w znacznie bardziej niezawodny sposób²⁸. Argument ten nie przekonuje osób, które już teraz wątpią czy autonomiczne systemy mogłyby zostać zaprojektowane w taki sposób, aby odpowiednio stosować zasadę proporcjonalności i właściwie jej przestrzegać.

Trzecia z zasad międzynarodowego prawa humanitarnego – środki ostrożności – wiąże się z koniecznością odpowiedzi na pytanie czy możliwe jest zaprogramowanie autonomicznych systemów bojowych w taki sposób, aby były w stanie przestrzegać zobowiązań wynikających z prawa humanitarnego odnoszących się do podejmowania odpowiednich środków ostrożności w kontekście ataku. Zasada ta nawiązuje bezpośrednio do nakazu rozróżniania oraz zasady proporcjonalności i zgodnie z art. 57(1) I Protokołu Dodatkowego wymaga, aby „w toku prowadzenia operacji wojskowych stale troszczyć się o oszczędzanie ludności cywilnej, osób cywilnych oraz dóbr o charakterze cywilnym”²⁹. W związku z tym, należy stale podejmować środki mające na celu zapobieganie powstawaniu strat wśród ludności cywilnej. Obowiązek zadbania o niniejszą kwestię ma zastosowanie dla sił zbrojnych i dotyczy wszystkich osób biorących udział w przygotowaniach: dowódców, ale także producentów i programistów danych systemów uzbrojenia. Stąd, powołujący się na tę zasadę wskazują na konieczność ciągłego nadzoru ludzkiego (żołnierzy) w celu ewentualnego reagowania na zmieniające się okoliczności³⁰. Inni wyrażają wątpliwości co do powyższego argumentu, ponieważ komputerowo sterowane systemy bojowe przetwarzają informacje znacznie szybciej niż ludzie, a zatem mogą szybciej reagować. Wątpliwe jest, w jakim stopniu można zakładać, że żołnierze nadzorujący systemy bojowe w sytuacjach ekstremalnych są w stanie interweniować, jeśli autonomiczny system bojowy przygotowuje się do naru-

²⁷ M.N. Schmitt, *Autonomous Weapon Systems...*, s. 19.

²⁸ R.C. Arkin, *Governing Lethal Behavior...*, s. 58.

²⁹ *Protokół dodatkowy do konwencji...*, art. 57 (1).

³⁰ *Academy Briefing No. 8: Autonomous Weapon Systems under International Law*, Geneva Academy of International Humanitarian Law and Human Rights, November 2014, s. 16, <https://www.geneva-academy.ch/our-projects/publications?start=10> (dostęp: 30 listopada 2017 r.).

szenia zasady międzynarodowego prawa humanitarnego³¹. Można zatem założyć, że pełne przestrzeganie przez autonomiczne systemy bojowe zasady odnoszącej się do środków ostrożności będzie możliwe w konfliktach bez udziału ludności cywilnej. Rzeczywista decyzja w odniesieniu do zasad międzynarodowego prawa humanitarnego zostanie w takim przypadku podjęta przez dowódcę, który wysyła maszynę w rejon działań bojowych. Jednakże, pozostaje pytanie, jak bardzo realny jest taki scenariusz (bez udziału ludności cywilnej), biorąc pod uwagę formy współczesnych konfliktów. Po tym, jak autonomiczne systemy bojowe zostaną stworzone i będą gotowe do użycia trudno będzie nie ulec pokusie ich stosowania, nawet w złożonych i nieprzewidywalnych sytuacjach.

4. PODSUMOWANIE

Należy zauważyć, że wiele prezentowanych argumentów odnoszących się do przyszłościowego funkcjonowania autonomicznych systemów bojowych w zgodzie z zasadami międzynarodowego prawa humanitarnego ma charakter spekulacyjny. Widoczne są raczej przemawiające zarówno za zwolennikami, jak i oponentami stosowania tego rodzaju środków walki. Międzynarodowe prawo humanitarne nakłada surowe wymagania na ocenę i zdolność do interpretacji najbardziej złożonych sytuacji krytycznych w ramach konfliktu zbrojnego. Wynika to z faktu, że współczesne konflikty charakteryzują się wysokim poziomem nieprzewidywalności i złożoności. Prawdopodobnie, nawet najbardziej złożone systemy autonomiczne nigdy nie będą posiadać inteligencji kontekstowej obejmującej wszystkie aspekty ich funkcjonowania. Jest co najmniej wątpliwe czy autonomiczne systemy bojowe będą w stanie właściwie wypełniać swoje zadania w przypadku nieplanowej zmiany misji. Z drugiej strony, jeśli rzeczywiście okaże się, że są one w stanie lepiej przestrzegać nakazu rozróżniania, zasady proporcjonalności i środków ostrożności w porównaniu z żołnierzami, to można argumentować, że dowódca wojskowy jest zobowiązany (przynajmniej moralnie) do zastosowania tego rodzaju środków walki zamiast ludzi w celu wykonania misji, w której spodziewana jest obecność ludności cywilnej. Tylko w taki sposób będzie zdolny do zapewnienia ochrony ludności cywilnej w możliwie najlepszy sposób.

W debacie nad autonomicznymi systemami bojowymi podkreśla się również, że współcześnie najsłabszym ogniwem, w zakresie wojskowych procedur podejmowania decyzji, jest człowiek. Bardzo często, zbrodnie wojenne są wynikiem emocjonalnego przeciążenia żołnierzy w obliczu złożoności wydarzeń i strachu na polu walki. Żołnierze powinni stosować zasady, które ustanowiła społeczność międzynarodowa na podstawie powszechnie obowiązujących norm. Właśnie wtedy, gdy stosowanie ustalonych zasad jest sporne, algorytmy są zazwyczaj bardziej zdolne do właściwego rozwiązania danego problemu, stąd możliwość angażowania autonomicznych systemów bojowych.

Wydaje się również, że cały system międzynarodowego prawa humanitarnego, mimo wszystko, opiera się na domniemanym założeniu, że jego zasady dotyczą i mają zastosowanie do człowieka podejmującego decyzje w odniesieniu do konfliktów zbrojnych. Jeśli adresatami ustalonych obowiązków i zakazów są ludzie, których emocjonalność, podatność na błędy i instynkt przetrwania są już uwzględnione w architekturze prawnej, to pytanie czy autonomiczne systemy bojowe mogą być zgodne z międzynarodowym prawem

³¹ P. Alston, *Lethal Robotic Technologies: The Implications for Human Rights and International Humanitarian Law*, "Journal of Law, Information & Science", Vol. 21, 2011/2012, s. 36, 55.

humanitarnym jest błędne. Bardziej decydująca może być kwestia odnosząca się do tego czy zasady zawarte w obowiązującym systemie prawnym nadal są właściwymi przepisami w przypadku zaangażowania autonomicznych systemów bojowych do działań wojennych. Z niniejszych rozważań można wywnioskować, że takie systemy – jeśli w ogóle – powinny spełniać znacznie wyższe standardy.

Reasumując, istnieje potrzeba pogłębionej dyskusji na temat przyszłościowego użytkowania autonomicznych systemów bojowych, obejmującej kwestie: techniczne, moralne oraz prawne. Uczestnikami tej dyskusji powinny być zarówno gremia decyzyjne (w wymiarze narodowym i międzynarodowym) oraz uznane autorytety zajmujące się problematyką etyczno-moralną i prawną w kontekście nowoczesnych technologii.

LITERATURA

1. AAP-6 (2014) *Słownik terminów i definicji NATO*, NATO Standardization Agency, Bruksela 2014.
2. *Academy Briefing No. 8: Autonomous Weapon Systems under International Law*, Geneva Academy of International Humanitarian Law and Human Rights, November 2014, <https://www.geneva-academy.ch/our-projects/publications?start=10> (dostęp: 30 listopada 2017 r.).
3. Alston P., *Lethal Robotic Technologies: The Implications for Human Rights and International Humanitarian Law*, "Journal of Law, Information & Science", Vol. 21, 2011/2012.
4. Arkin R.C., *Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture*, Technical Report GIT-GVU-07-11 <http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf> (dostęp: 20 listopada 2017 r.).
5. Asaro P. M., *Modelling the Moral User*, "IEEE Technology and Society Magazine", Vol. 28, Issue 1, Spring 2009, <http://peterasaro.org/writing/Asaro%20Modeling%20Moral%20User.pdf> (dostęp: 20 listopada 2017 r.).
6. *Autonomy in Weapons Systems. Directive 3000.09*, U.S. Department of Defense, 21 November 2012, <http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf> (dostęp: 12 listopada 2017 r.).
7. *Losing Humanity: The Case against Killer Robots*, Human Rights Watch 2012, https://www.hrw.org/sites/default/files/reports/arms1112_ForUpload.pdf (dostęp: 12 listopada 2017 r.).
8. O'Connell M.E., *Banning Autonomous Killing: The Legal and Ethical Requirement that Humans Make Near-Time Lethal Decisions [w:] The American Way of Bombing. Changing Ethical and Legal Norms, from Flying Fortresses to Drones*, red. M. Evangelista, H. Shue, Cornell University Press, New York 2014.
9. Protokół dodatkowy do konwencji genewskich z 12 sierpnia 1949 r., dotyczący ochrony ofiar międzynarodowych konfliktów zbrojnych, Genewa, 8 czerwca 1977 r. (Dz. U. z 1992 r., nr 41, poz. 175, załącznik).
10. *Report of the Expert Meeting on Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*, 26–28 March 2014, Geneva, 9 May 2014, <https://shop.icrc.org/expert-meeting-autonomous-weapon-systems-technical-military-legal-and-humanitarian-aspects-26-to-28-march-2014.html> (dostęp: 12 listopada 2017 r.).
11. *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, Christof Heyns, 26 April 2013, United Nations, General Assembly, A/HRC/23/47.

12. Sassòli M., *Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to Be Clarified*, "International Law Studies", Vol. 90, 2014.
13. Schmitt M.N., *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*, "Harvard National Security Journal Features", 2013, <http://harvardnsj.org/wp-content/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf> (dostęp: 23 listopada 2017 r.).
14. Schmitt M.N., Thurnher J.S., *Out of the Loop: Autonomous Weapons Systems and the Law of Armed Conflict*, "Harvard National Security Journal", Vol. 4, 2013.
15. Sharkey N.E., *The Evitability of Autonomous Robot Warfare*, "International Review of the Red Cross", Vol. 94, 2012.
16. Sharkey N., *Towards a Principle for the Human Supervisory Control of Robot Weapons*, 2014, [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/2002471923EBF52AC1257CCC0047C791/\\$file/Article_Sharkey_PrincipleforHumanSupervisory.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/2002471923EBF52AC1257CCC0047C791/$file/Article_Sharkey_PrincipleforHumanSupervisory.pdf) (dostęp: 12 listopada 2017 r.).
17. Singer P. W., *Wired for War: The Robotics Revolution and Conflict in the 21st Century*, The Penguin Press, London 2009.
18. *Unmanned Systems Integrated Roadmap FY2013–2038*, US Department of Defense, Washington DC 2013, <https://www.defense.gov/Portals/1/Documents/pubs/DOD-USRM-2013.pdf> (dostęp: 6 listopada 2017 r.).
19. *Weapons: ICRC statement to the United Nations*, General Assembly, 69th session, First Committee, 14 October 2014, New York, <https://www.icrc.org/en/document/weapons-icrc-statement-united-nations-2014> (dostęp: 6 listopada 2017 r.).

DILEMMAS OF UTILIZATION OF LETHAL AUTONOMOUS WEAPON SYSTEMS AS REGARD TO BASIC PRINCIPLES OF INTERNATIONAL HUMANITARIAN LAW

A vision where machines will make decisions on life or death on the battlefield becomes more and more expressive. Symptomatic in this context seems to be the use of unmanned aircraft during conflicts of the last decade: Iraq, Afghanistan, Libya, Syria or Ukraine. The increasing robotization of the battlefield causes a lot of controversy, especially among non-governmental organizations, including postulates of a total ban on the development and use of lethal autonomous weapon systems. The ongoing public debate is clearly sharpening and the future use of these systems has both its supporters and opponents. In this context, the legal aspects of the application of lethal autonomous weapon systems on the battlefield seem to be crucial.

The paper presents the basic approaches in defining lethal autonomous weapon systems. It has been pointed out that there is no unambiguous definition of this type of combat tool inside international law. Attention is drawn to the lack of a clear separation between automated and autonomous systems. This, in turn, causes problems related to the proper legal interpretation of the utilization of lethal autonomous weapon systems. The dilemmas of the potential use of lethal autonomous weapon systems in relation to the basic principles of international humanitarian law: distinction, proportionality and precautionary measures have also been articulated. It was pointed out that currently operating norms of international law do not require changes in the context of the use of lethal autonomous weapon systems. At the same time,

an in-depth discussion is necessary to resolve emerging doubts. Binding arrangements should be made on the international forum before the actual introduction of lethal autonomous weapon systems to equip the armed forces.

Keywords: automatization, decision-making, international humanitarian law, lethal autonomous weapon systems, robotization of the battlefield.

DOI: 10.7862/rz.2018.hss.16

Przesłano do redakcji: styczeń 2018 r.

Przyjęto do druku: kwiecień 2018 r.

