



Killer robots: zegen of vloek? **Prof. Lode Lauwaert¹**

Techniekfilosoof, Hoger Instituut voor Wijsbegeerte, KU Leuven

Tijdens oorlogen of conflicten worden robots voor tal van zaken gebruikt: om mijnen onschadelijk te maken, vliegtuigen van brandstof te voorzien, munitie te dragen, etc. Het is niet uitgesloten dat die lange rij in de toekomst wordt uitgebreid. Want hoewel ze (wellicht) momenteel nergens worden gebruikt, is het mogelijk dat later een nieuwe technologie wordt geïntroduceerd: Lethal Autonomous Weapon Systems (LAWS) of killer robots, zoals ze doorgaans ook worden genoemd. Maar wat moeten we daaronder wel en niet verstaan? Waarom verzetten landen als Rusland en Israël zich tegen een (preventief) verbod op zulke technologie? Waarom precies wil men zo'n verbod? En hoewel sommige argumenten voor een verbod voor velen overtuigend zijn, is de vraag of dat zo is voor *alle* argumenten. De volgende paragrafen zijn aan deze vragen opgehangen. De tekst verdedigt dus geen standpunt in het debat, maar wil integendeel informeren en de discussie verdiepen.

1. Wat is een killer robot?

Killer robots zijn wapens. Het doel is dat ze tijdens een conflict of oorlog door een partij worden ingezet om te strijden tegen een andere partij. In dat opzicht verschillen ze van pakweg personenwagens, maar komen ze wel overeen met speren, zwaarden, pistolen, kanonnen, bommenwerpers en machinegeweren. Natuurlijk, ook personenwagens kunnen worden ingezet tijdens een conflict of oorlog, maar dat is niet de reden waarom ze bestaan. Wapens zijn echter met geen ander doel voor ogen gemaakt dan om te strijden.

Die omschrijving is te breed. Alle killer robots zijn wel wapens, maar niet alle wapens zijn killer robots. Een verdere precisering is dat LAWS zijn gemaakt om te *doden*, en niet om de tegenstander te verwonden of onschadelijk te maken door wapens te vernietigen. Daarop wijzen 'killer' en 'lethal'. Beide termen maken vanzelfsprekend ook duidelijk dat LAWS geen artefacten of anorganische materialen als target hebben. En ondanks het feit dat we ook dieren kunnen doden, verwijst 'killer robot' naar een wapen ontworpen om *mensen* te doden.

Volstaat het echter LAWS te omschrijven als dodelijke wapens? Neen, want ook een 9mm-pistool is een wapen gemaakt om mensen te doden. Het verschil tussen dat dodelijk wapen en een killer robot heeft alles te maken met het autonome karakter van de robot. In welke zin?

'Autonomie' kan meerdere zaken betekenen. In het geval van killer robots wil het zeggen dat die robots zonder menselijke tussenkomst kunnen functioneren. Zeker, killer robots zijn gemaakt door ingenieurs en programmeurs. En natuurlijk kunnen zij pas in werking treden nadat iemand 'op de knop drukte'. Maar eens dat laatste is gebeurd, kunnen zij op geheel zelfstandige basis doen waarvoor ze werden gemaakt (*human-out-the-loop*). Ze detecteren in de lucht, op het land of onder water de target, en

¹ De auteur schreef dit werk in eigen naam met de bedoeling om bij te dragen aan het debat rond killer robots. Deze paper geeft op geen enkele manier de visie weer van het Kenniscentrum Data & Maatschappij. Het Kenniscentrum probeert het debat met dergelijke papers enkel mee vorm te geven.

vervolgens beslissen ze om al dan niet te vuren. Killer robots zijn in dat opzicht het tegenovergestelde van 9mm-pistolen, die geen enkele autonomie bezitten.

Maar nuance is nodig, want autonomie bestaat in gradaties. Er zijn wapens die geen killer robots zijn, maar die wel in een *bepaalde mate* autonoom zijn. Sommige drones kunnen bijvoorbeeld enkel opstijgen, vliegen en vuren indien zij door een menselijke operator worden bestuurd, hoewel ze wel zelfstandig omgevingsbeelden maken en verwerken. Of neem de SGR-A1, de robot gebruikt door Zuid-Korea om te verhinderen dat soldaten uit Noord-Korea dichterbij komen. Die is niet enkel uitgerust met sensoren die iedere beweging van mensen aan de grens op eigen houtje detecteren, maar ook met machinegeweren. Toch is de SGR-A1 geen killer robot. Dat komt omdat de robot enkel zal vuren indien het commando van een mens afkomstig is (*human-in-the-loop*).

Moeten we dan besluiten dat ook antipersoonsmijnen LAWS zijn? Want zulke mijnen zijn toch ontworpen om te doden, en ook zij veronderstellen toch geen menselijke interventie? Het belangrijkste verschil is dat killer robots zijn uitgerust met artificiële intelligentie (niet enkel met de *Good-Old-Fashioned AI*). Dat houdt in dat we de verwerking van de omgevingsinformatie die de robot via sensoren ontvangt mogelijk niet kunnen begrijpen, en dat er geen volledige zekerheid is over wat de robot zal beslissen en doen. Natuurlijk, ook antipersoonsmijnen ontvangen input die leidt tot output. Mijnen zijn (meestal) echter geen intelligente systemen, wat onder meer inhoudt dat je zeker bent van de uitkomst. Indien een sensor signaleert dat een mens de mijn aanraakt of in de nabijheid is, dan weet je zeker dat de mijn ontploft.

Tot slot: hoewel met 'killer robots' doorgaans wordt verwezen naar wapensystemen om aan te vallen, is dat niet noodzakelijk zo. LAWS kunnen in zowel defensief als offensief opzicht worden gebruikt. Hetzelfde geldt voor het mobiele karakter van killer robots. Meestal heeft men het over robots die zich in de lucht, onder water of op het land bewegen. Nochtans zijn ook een *statische* autonome wapensystemen die zijn ontworpen om mensen te doden LAWS.

2. Campaign to stop killer robots

De vele pleidooien tegen LAWS zijn deel van een geschiedenis van verzet tegen het gebruik van wapens en de regulering ervan. De Verklaring van Sint-Petersburg in 1868 wordt doorgaans gezien als het begin van die geschiedenis. Ze verbood het gebruik van explosieve kogels, dat wil zeggen: kogels die in het lichaam ontploften zodra ze erin binnendrongen. In de decennia daarna volgden tal van verdragen. De internationale vredesconferentie in Den Haag in 1907 bande de zogenaamde dumdkogels (kogels die versplinterden in het lichaam), de VN-conventie in 1980 het gebruik van boobytraps, terwijl op de Ottawa-conventie in 1997 zo'n veertigtal landen besloten om geen landmijnen meer te produceren en te gebruiken.

Het wereldwijde verzet tegen killer robots startte officieel in oktober 2012. In New York werd toen beslist om de Campaign to Stop Killer Robots op te starten, waaraan intussen tientallen niet-gouvernementele organisaties uit meer dan 30 landen deelnemen. Enkele maanden later, in mei 2013, spraken tal van landen op een bijeenkomst van de Verenigde Naties in Genève over LAWS. Pakistan riep er als eerste land ooit op voor een verbod op de productie en het gebruik van autonome wapens. De voorbije jaren volgden nog 29 landen, met Namibië als het meest recente in oktober 2019. Een andere belangrijke datum is 27 februari 2014. Die dag keurde het Europese Parlement een niet-bindende resolutie goed voor een verbod op killer robots.

Ook tal van wetenschappers, ondernemers en publieke figuren mengden zich de voorbije jaren in de strijd tegen LAWS. In juli 2015 werd bijvoorbeeld een open brief verspreid die als volgt eindigt: "Een militaire AI wapenwedloop starten is een slecht idee, en moet worden voorkomen door een verbod op offensieve autonome wapens zonder betekenisvolle menselijke controle." De brief werd

voorgesteld aan het begin van de International Joint Conference on Artificial Intelligence in Buenos Aires. Tot de ondertekenaars behoren onder anderen Stephen Hawking, Elon Musk en Steve Wozniak, de mede-oprichter van Apple.

In België werd het verzet tegen killer robots op politiek vlak vooral op gang getrokken door Groen en Ecolo. In 2016 bijvoorbeeld diende volksvertegenwoordiger Wouter De Vriendt een resolutievoorstel in dat pleit voor een (inter)nationaal verbod. Mede daardoor werd het thema sindsdien al meermaals besproken in de Kamer. In juli 2018 werd uiteindelijk een resolutie goedgekeurd die uitging van de meerderheidspartijen. Dat zou tegemoet komen aan de wil van de meerderheid van de Belgische bevolking. Uit een peiling van YouGov blijkt immers dat 60% van de Belgen meent dat de regering zich moet verzetten tegen de ontwikkeling en het gebruik van LAWS. In het licht daarvan is het opmerkelijk dat ons land op een VN-conferentie in augustus 2018 in Genève niet ijverde voor een internationaal verbod.

3. Argumenten pro

Wie pleit tegen killer robots moet zich bewust zijn van een aantal voordelen van zulke technologie. Doorgaans onderscheidt men er een viertal. Tijdens debatten leggen de voorstanders van killer robots meestal minstens één van die voordelen in de schaal.

Een eerste argument pro killer robots is dat ze op een aantal vlakken technisch erg geavanceerd zijn. LAWS werken bijzonder snel: ze hebben slechts een aantal milliseconden nodig om hun doelwit te detecteren en uit te schakelen. Bovendien kunnen ze grote hoeveelheden informatie tegelijk en snel verwerken, waardoor ze een goede inschatting kunnen maken van de situatie. Ook kunnen killer robots met een hoge mate van precisie te werk gaan. Wanneer het doel is om *deze* target uit te schakelen, dan zullen ze (wellicht) weinig of geen *collateral damage* veroorzaken. Killer robots zijn in dat opzicht superieur aan bijvoorbeeld bommenwerpers, die doorgaans wel (veel) bijkomende schade veroorzaken. Wie dat laatste technische voordeel aanhaalt, moet wel weten dat niet enkel killer robots dat voordeel hebben. Ook drones bestuurd door mensen hebben de capaciteit om de schadelijke neveneffecten van aanvallen te beperken.

Ten tweede bieden LAWS een financieel voordeel. Dat komt omdat killer robots autonoom functioneren. Terwijl militaire operaties doorgaans de betrokkenheid van veel personeelsleden vereisen, veronderstelt een killer robot geen menselijke interventie (behalve 'de druk op de knop'). Maar zelfs indien voor een militaire ingreep slechts één persoon nodig is (om een halfautomatische drone te besturen bijvoorbeeld), zijn killer robots financieel voordeliger. Die technologie is niet bijzonder duur om te maken, terwijl militair personeel een grote kost met zich meebrengt. Concreet: in de VS hangt aan de loopbaan van een soldaat een kostenplaatje van vier miljoen dollar vast, terwijl LAWS minder tien procent van dat bedrag zouden kosten.

Een derde argument dat in het voordeel van killer robots pleit, is dat het functioneren van zulke technologie, anders dan dat van soldaten, niet wordt gehinderd door emotionele, mentale of lichamelijke factoren. Militairen kunnen bijvoorbeeld niet werken zonder te eten en te rusten, en functioneren niet (goed) onder bepaalde weersomstandigheden zoals regen en hitte. Daarnaast is het mogelijk dat hun geheugen hen in de steek laat of dat ze onder druk van de omstandigheden in de war raken. En tot slot hebben soldaten natuurlijk ook emoties die de goede afloop van een interventie kunnen bemoeilijken of onmogelijk maken. Denk bijvoorbeeld aan angst en medelijden. Sterker nog, bepaalde emoties kunnen zelfs aan de basis liggen van geweld dat het doel van de operatie voorbij schiet. Hoeveel onnodig, excessief lijden vloeit niet voort uit pakweg haat, woede of wraak?

Wellicht het meest aangehaalde argument in het voordeel van killer robots, ten slotte, hangt opnieuw samen met het volkomen autonome karakter van LAWS. Omdat men die wapens kan inzetten zonder

menselijke tussenkomst, kan het gebruik van die technologie ertoe leiden dat het aantal dodelijke slachtoffers in eigen rangen sterk afneemt. Wie zich op dat argument beroept, moet wel beseffen dat ook dat voordeel niet uniek is voor killer robots. Ook de bestuurder van gewapende drones loopt immers geen dodelijk gevaar (al leert onderzoek wel dat drone-operatoren een niet gering risico op posttraumatische stressstoornissen lopen).

4. Argumenten contra

Hoewel de meeste betrokken partijen in het debat de voordelen van LAWS erkennen, meent lang niet iedereen dat ze ook beslissend zijn. Volgens sommigen wegen ze niet op tegen minstens één van de volgende twee *principiële* tegenargumenten.

Een eerste argument contra killer robots heeft te maken met morele verantwoordelijkheid. Stel, een killer robot doodt een burger in plaats van een terrorist. Bepaalde stemmen opperen dat niemand daarvoor verantwoordelijk is. Is het immers fair iemand verantwoordelijk te houden voor een overlijden dat voortvloeit uit de beslissingsprocedure van de robot en waarover niemand controle heeft? Hoewel het aanduiden van een verantwoordelijke belangrijk is, vindt men dat argument in de literatuur doorgaans niet overtuigend. Zolang de dood van de burger *voorzien* is en men wist dat dat kon gebeuren, kun je diegenen die hebben besloten om de killer robot te gebruiken moreel verantwoordelijk houden voor de dood van de burger.

Een tweede vaak aangehaald argument luidt dat het doden van een soldaat door een killer robot een miskennis van de waardigheid van de soldaat inhoudt. Dat tegenargument veronderstelt een duidelijk begrip van menselijke waardigheid. Maar wat bedoelt men daar wel en niet mee? Kun je dat glashelder uitleggen? Stel echter dat je er wel een goed zicht op hebt. Waarom vereist respect voor de waardigheid van het slachtoffer een interventie door mensen? Waarom wordt met andere woorden de waardigheid van een soldaat *niet* geschonden indien die wordt gedood door een andere mens, en *wel* indien die door een robot wordt gedood? Bovendien luidt een centraal principe in de militaire ethiek dat geweld niet meer schade dan nodig mag veroorzaken. Volstaat het uit hoofde van het respect voor de waardigheid van de ander dan niet om geen excessief geweld te gebruiken?

Daarnaast beroept men zich ook op niet-principiële argumenten. Die focussen op de negatieve gevolgen die zouden kunnen voortvloeien uit het gebruik van dodelijke autonome wapensystemen. Zie hier de twee die het vaakst worden aangehaald.

Een derde argument is dat de mogelijkheid om die wapens te gebruiken de drempel zou verlagen om een oorlog te beginnen. Dat argument is gebaseerd op de al eerder vermelde voordelen van killer robots. Precies omdat het gebruik van die wapens inhoudt dat er op het terrein minder soldaten zullen sterven, zal men minder aarzelen om een oorlog te starten. Hoewel het niet onwaarschijnlijk is dat het gebruik van LAWS tot meer oorlogen zal leiden, is dat toch ook niet zeker. Vooreerst kunnen ook nu wapens worden gebruikt die op het terrein niet leiden tot slachtoffers in het eigen kamp. Denk aan langeafstandsraketten of drones die vanop grote afstand worden bestuurd. Heeft de mogelijkheid om zulke wapens te gebruiken geleid tot meer oorlog? Bovendien is het risico op slachtoffers niet de enige overweging om al dan niet een oorlog te starten. Er spelen ook economische motieven of het risico op beschadiging van de eigen infrastructuur.

Ten vierde wordt ook gewezen op dé morele wet tijdens conflicten en oorlogen, namelijk het discriminatieprincipe. Dat luidt dat je een onderscheid moet maken tussen soldaten en burgers. Het is gerechtvaardigd dat de eerste groep een target is, terwijl het verboden is om individuen die niet strijden intentioneel aan te vallen, te schaden of te doden.

Gehoorzaamheid aan dat principe is niet vanzelfsprekend. Het veronderstelt dat je subtiele verschillen tussen handelingen, voorwerpen, gezichtsuitdrukkingen en kleding kunt onderscheiden; het vereist ook dat je de complexiteit van situaties kunt zien en interpreteren. Inschatten of robots kunnen gehoorzamen aan het discriminatieprincipe houdt dus in dat je je moet beroepen op het oordeel van ingenieurs, informatici, robotici en computerwetenschappers. En laten zij het nu precies daarover volkomen eens zijn: killer robots zijn niet of onvoldoende in staat om burgers en soldaten te onderscheiden (en zullen dat vermoedelijk nooit zijn). In een open brief uit 2013 van 272 experts op het vlak van artificiële intelligentie lezen we bijvoorbeeld: “De technologie van autonome robotwapens zit vol beperkingen en onbekende toekomstige risico’s. Er is geen duidelijk wetenschappelijk bewijs dat zo’n wapens in staat zijn – of in de voorziene toekomst in staat zullen zijn – tot een accurate identificering van doelwitten, voldoende situationeel bewustzijn hebben (...)” De technische onmogelijkheid om voldoende te gehoorzamen aan het discriminatieprincipe is voor velen dan ook een doorslaggevend argument om te pleiten tegen killer robots.

5. Meer debat is nodig!

De discussie over killer robots vereist niet alleen dat verder wordt ingezoomd op een aantal problemen met argumenten. Daarnaast zijn er nog minstens twee andere moeilijkheden die moeten worden weggewerkt (of die minstens moeten worden aangekaart).

Hoewel, ten eerste, de meeste betrokkenen de definitie van hierboven hanteren, volgt niet iedereen die omschrijving. Het International Committee of the Red Cross bijvoorbeeld, dat nochtans aan hetzelfde debat deelneemt, vertrekt van een bredere definitie. Het focust niet op *Lethal Autonomous Weapon Systems* maar heeft het over autonome wapens in het algemeen. Die verscheidenheid aan definities heeft niet enkel als onwenselijk gevolg dat niet iedere partij over hetzelfde debatteert. Naast spraakverwarring zijn er ook juridische gevolgen. Dewelke?

Wie pleit voor een preventief verbod op autonome wapens in het algemeen (en niet louter op *Lethal Autonomous Weapon Systems*) argumenteert tegen méér zaken dan wie uitsluitend pleit tegen killer robots. Concreet: wie autonome wapens wil verbieden is in principe ook gekant tegen de Iron Dom in Israël, terwijl een verbod op killer robots niet inhoudt dat de Iron Dom wordt verboden. Dat technologisch systeem is immers ontworpen om het Israëlische grondgebied te beschermen tegen *raketten* die worden afgevuurd door de Palestijnen in de Gazastrook, en niet tegen mensen. Kortom, een internationaal gedeelde definitie dringt zich op!

Ten tweede is er niet enkel onenigheid over de definitie, maar ook over de reikwijdte van het verzet tegen LAWS. Neem bijvoorbeeld het argument dat verband houdt met het discriminatieprincipe. De voorspelling van experts dat robots zich niet of bijzonder moeilijk aan dat principe kunnen houden, is volgens sommigen een reden voor een verbod op het gebruik van LAWS *waar dan ook*. Anderen argumenteren dan weer dat het volstaat dat killer robots worden verboden op plaatsen waar zich burgers bevinden. Killer robots zouden volgens hen wel mogen worden gebruikt in gebieden waar zich uitsluitend strijders bevinden.

Een andere vraag waarop niet eensgezind wordt geantwoord is of enkel het *gebruik* van killer robots moet worden verboden. Of betreft het verbod ook productie, handel en onderzoek? En indien je ook onderzoek aan banden wil leggen, waarover heb je het dan precies? Bepaalde stemmen menen dat je uitsluitend het onderzoek gericht op het maken van wapens moet verbieden. Anderen menen dat daaronder ook de wetenschappelijke studies moeten vallen die mogelijk gebruikt kunnen worden voor productie, hoewel dat gebruik zelf niet het doel van het onderzoek is. Er moet met andere woorden niet enkel dringend eensgezindheid komen over het begrip van killer robots, er moet ook eensgezindheid komen over wat men nu precies wel en niet wil verbieden.