Lena Trabucco

International Humanitarian Law and Lethal Autonomous Weapon Systems

Legal Considerations for Acquisition and Procurement

DJØF PUBLISHING
IN COOPERATION WITH THE
CENTRE FOR MILITARY STUDIES

International Humanitarian Law and Lethal Autonomous Weapon Systems

Legal Considerations for Acquisition and Procurement

Lena Trabucco

International Humanitarian Law and Lethal Autonomous Weapon Systems

Legal Considerations for Acquisition and Procurement

Djøf Publishing
In cooperation with the
Centre for Military Studies
2023

Lena Trabucco
International Humanitarian Law and
Lethal Autonomous Weapon Systems

Legal Considerations for
Acquisition and Procurement

*This publication is peer reviewed according to the standards
set by the Danish Ministry of Higher Education and Science.*

Cover: Kelly Chigozie Kjelsø Arazu
Print: Ecograf, Brabrand

Printed in Denmark 2023

ISBN 978-87-574-5585-4

Djøf Publishing
Gothersgade 137
1123 København K

Telefon: 39 13 55 00
e-mail: forlag@djoef.dk
www. djoef-forlag.dk

Editors' preface

The publications of this series present new research on defence and security policy of relevance to Danish and international decision-makers. This series is a continuation of the studies previously published as CMS Reports. It is a central dimension of the research-based services that the Centre for Military Studies provides for the Danish Ministry of Defence and the political parties behind the Danish defence agreement. The Centre for Military Studies and its partners are subject to the University of Copenhagen's guidelines for research-based services, including academic freedom and the arm's length principle. As they are the result of independent research, the studies do not express the views of the Danish Government, the Danish Armed Forces, or other authorities. Our studies aim to provide new knowledge that is both academically sound and practically actionable. All studies in the series have undergone external peer review. And all studies conclude with recommendations to Danish decision-makers. It is our hope that these publications will both inform and strengthen Danish and international policy formulation as well as the democratic debate on defence and security policy, in particular in Denmark.

The present publication is a result of the additional grant specifically aimed at research in the international legal challenges of the Danish Defence, which the parties to the Danish Defence Agreement have awarded to the Centre for Military Studies. The international legal research is conducted in collaboration with the Faculty of Law, University of Copenhagen, and the Royal Danish Defence College. Read more at: https://jura.ku.dk/icourts/research/intermil/

The Centre for Military Studies is a research centre at the Department of Political Science, University of Copenhagen. The centre conducts research into security and defence policy as well as military strategy. Read more about the centre, its activities, and other publications at: https://cms.polsci.ku.dk/english/

Copenhagen, May, 2023
Kristian Søby Kristensen, Kevin Jon Heller
and Astrid Kjeldgaard-Pedersen

Table of Contents

Abstract and Recommendations

This report aims to detail the legal challenges for Danish acquisition and procurement stakeholders in the context of lethal autonomous weapon systems (LAWS). The accelerated pace of artificial intelligence (AI) innovation, coupled with increasingly practical military applications, makes it likely that AI will play a prominent role in future warfare. The time is ripe for Danish policymakers to explore the opportunities available to the Danish Armed Forces as the promise and prominence of AI—and especially its weaponization—becomes central to discussions of future warfare. Unlike core Danish allies, Denmark has largely been silent regarding autonomous weapon systems and has not offered a policy position for future implementation or use.

As Denmark grapples with the future of autonomous military technology, this report examines the most pressing legal considerations for acquiring LAWS. I argue that Danish defense and legal stakeholders should consider two legal issues. The first is the nature of LAWS technology under the requirements of IHL. The analysis considers three major features of autonomous technology—AI transparency, machine predictability, and algorithmic bias—and details how they can complicate compliance with the IHL principles of distinction, proportionality, and precaution. The second legal issue examines how responsibility can be divided and attributed to the multiple actors across the lifecycle of the weapon system, including direct considerations for acquisition and procurement officials. This section details the international legal responsibility frameworks, including state responsibility, individual criminal responsibility, and corporate responsibility through the risks of strategic litigation for civilian firms as a potential barrier to defense collaboration and subsequent acquisition.

The AI landscape and the legal parameters for military applications are rapidly changing. If and when Denmark decides to acquire lethal autonomous weapons, these recommendations can help to guide future deliberations and planning.

1. **Formulate policy**. Formulating national policy regarding a Danish interpretation of LAWS can guide military decision-makers and legal advisers as AI continues to be important for discussions of future warfighting capabilities. Clear policy becomes all the more relevant as Denmark participates in international coalitions (e.g., the AI Partnership for Defense), but has largely remained silent on the LAWS issue. Danish allies (e.g., US, Australia, France) have already announced research and development (R&D) programs, and in some cases an accompanying ethical framework, toward responsible AI weapon systems development. Other Danish allies (e.g., Germany) have instead publicly opposed such weapons development. Without a formal policy, Denmark risks falling behind critical security partners on future dialogues regarding LAWS.
2. **Encourage inter-agency coordination**. Strengthened cooperation and coordination between the Danish Acquisition and Logistics Office (DALO) and the legal office of the Ministry of Defense can ensure greater IHL compliance. The legal challenges of incorporating LAWS into the Armed Forces will become critical to Danish warfighting capacity, and early stages of AI design should incorporate IHL. This will ensure that Denmark remains a competitive military ally with responsible, lawful autonomous systems.
3. **Restructure TEVV procurement**. Restructuring the Danish Testing, Evaluation, Validation, and Verification (TEVV) can help Denmark mitigate the inherent challenges to autonomous weapon systems; namely, algorithmic transparency, predictability, and bias. One option toward this end is to establish an iterative TEVV process offering accurate AI systems to Danish defense and legal stakeholders and ensuring that IHL standards are front-and-center of autonomous weapon development.
4. **Encourage joint acquisition/collaboration of LAWS through project development**. Project-based acquisition offers the greatest control and flexibility over LAWS design and development relative to off-the-shelf LAWS purchases. In order to ensure the maximum IHL compliance and sufficiently trained algorithms, Denmark may want to consider avenues for acquiring LAWS through project-based development. This report offers two starting points. First, Denmark can pursue joint acquisition participation in multi-national defense partnerships (e.g., NATO, the AI Partnership for Defense). Second,

the Danish Ministry of Defence (MoD) can create formal partnerships with Danish academic and industry experts at the cutting edge of AI design and development.

5. **Mitigate strategic litigation risks**. Global opposition to LAWS development increases the risks of defense and technology partners experiencing strategic litigation and reputational backlash for defense collaboration. The MoD can temper this issue through an actionable collaboration plan. Future LAWS/AI acquisition may want to mirror the Danish "triple helix" collaboration strategy for drone acquisition—combining research, industry, and the state to explore innovative acquisition potential—from the 2021 Danish National Defence Industrial Strategy. However, civilian collaboration will need to consider the risks of strategic litigation as a barrier for industry, especially smaller technology firms, with the potential to hinder an innovative Danish acquisition approach.

Resumé og anbefalinger

Denne rapport har til formål at analysere de juridiske udfordringer for danske beslutningstagere inden for anskaffelse af dødbringende autonome våbensystemer. Den accelererede innovation på området for kunstig intelligens (AI) kombineret med et stadig større potentiale for militær anvendelse gør det sandsynligt, at AI vil spille en fremtrædende rolle i fremtidens krigsførelse. Tiden er derfor inde til, at danske beslutningstagere undersøger de muligheder, AI giver det danske forsvar, efterhånden som teknologien – og især dens militære anvendelse – bliver mere central i diskussioner om fremtidens krig. I modsætning til sine kerneallierede har Danmark stort set været tavs med hensyn til spørgsmålet om autonome våbensystemer og har ikke fremlagt en politisk strategi om den fremtidige implementering og anvendelse af sådanne systemer.

Mens Danmark overvejer sin position med hensyn til fremtidig anvendelse af autonom militærteknologi, undersøger denne rapport de mest relevante juridiske udfordringer i forbindelse med anskaffelsen af dødbringende autonome våbensystemer. Rapporten argumenterer for, at danske beslutningstagere bør overveje to juridiske spørgsmål. Det første er forholdet mellem autonome våbensystemers teknologi og kravene i den humanitære folkeret. Analysen gennemgår tre centrale kendetegn ved dødbringende autonome våbensystemers teknologi – AI-gennemsigtighed, maskinforudsigelighed og algoritmisk bias – og undersøger, hvordan de udfordrer den humanitære folkerets principper om distinktion, proportionalitet og forsigtighed. Det andet juridiske spørgsmål er, hvordan ansvar kan opdeles mellem og henføres til flere aktører på tværs af våbensystemets livscyklus, herunder overvejelser med direkte relevans for beslutningstagere involveret i anskaffelse. Dette afsnit beskriver de folkeretlige rammer for ansvar, herunder statsansvar, individuelt strafferetligt ansvar og virksomhedsansvar, og kommer ind på virksomheders risiko for at blive mål for strategiske retssager som en potentiel hindring for forsvarssamarbejde og efterfølgende anskaffelse.

AI-teknologi og de juridiske rammer for dens militære anvendelse udvikler sig hurtigt. Hvis og når Danmark beslutter at anskaffe dødbrin-

gende autonome våbensystemer, kan disse anbefalinger bidrage til at informere fremtidige drøftelser og fremtidig planlægning.

1. **Formulere politik**. Udformning af en dansk politik vedrørende dødbringende autonome våbensystemer kan vejlede militære beslutningstagere og juridiske rådgivere, efterhånden som AI bliver stadig mere centralt i diskussionen om fremtidige militære kapaciteter. En klar politik er især relevant, eftersom Danmark deltager i internationale koalitioner (f.eks. AI Partnership for Defense), men stort set har forholdt sig tavs med hensyn til spørgsmålet om dødbringende autonome våbensystemer. Danske allierede (f.eks. USA, Australien og Frankrig) har allerede annonceret forsknings- og udviklingsprogrammer og i nogle tilfælde en medfølgende etisk ramme for ansvarlig udvikling af AI-våbensystemer. Andre danske allierede (f.eks. Tyskland) har i stedet offentligt modsat sig en sådan våbenudvikling. Uden en formel politik risikerer Danmark at komme bagud i forhold til kritiske sikkerhedspartnere i fremtidige dialoger om dødbringende autonome våbensystemer.
2. **Tilskynde til koordinering mellem myndigheder**. Styrket samarbejde mellem Forsvarsministeriets Materiel- og Indkøbsstyrelse og det juridiske kontor i Forsvarsministeriets departement kan sikre overholdelsen af den humanitære folkeret. De juridiske udfordringer ved at integrere dødbringende autonome våbensystemer i forsvaret vil blive afgørende for Danmarks fremtidige militære kapacitet, og udvikling og inkorporering af AI-teknologi bør fra et tidligt stadie tage højde for den humanitære folkeret. Det vil sikre, at Danmark forbliver en konkurrencedygtig militær allieret med ansvarlige autonome våbensystemer, der handler indenfor den humanitære folkeret.
3. **Omstrukturere TEVV-anskaffelse**. Omstrukturering af den danske test-, evaluerings-, validerings- og verifikationsproces (TEVV-proces) kan hjælpe Danmark med at afbøde de iboende udfordringer knyttet til autonome våbensystemer; især algoritmisk gennemsigtighed, forudsigelighed og bias. En mulighed i den henseende er at etablere en iterativ TEVV-proces, der tilbyder danske beslutningstagere præcise AI-systemer, samtidig med at den humanitære folkerets standarder bliver centrale for udvikling af autonome våbensystemer.
4. **Tilskynde til fælles anskaffelse af og samarbejde om dødbringende autonome våbensystemer gennem projektudvikling**. Pro-

jektbaseret anskaffelse giver den største kontrol og fleksibilitet i forbindelse med design og udvikling af dødbringende autonome våbensystemer sammenlignet med anskaffelse af færdigproducerede hyldevarer. For at sikre den bedst mulige overholdelse af den humanitære folkeret og tilstrækkeligt trænede algoritmer kan Danmark med fordel overveje at deltage i internationale udviklingsprojekter for så vidt angår dødbringende autonome våbensystemer. Denne rapport giver to udgangspunkter: For det første kan Danmark deltage i fælles anskaffelsesprogrammer i multinationale forsvarssamarbejder (f.eks. NATO og AI Partnership for Defense). For det andet kan Forsvarsministeriet skabe formelle partnerskaber med danske akademiske og industrielle eksperter, som besidder den nyeste viden inden for AI-design og -udvikling.

5. **Afbøde risikoen for strategiske retssager**. Global modstand mod udviklingen af dødbringende autonome våbensystemer øger risikoen for, at forsvarsindustrielle og forsvarsteknologiske virksomheder kan blive mål for strategiske retssager og kampagner med negative konsekvenser for deres omdømme. Forsvarsministeriet kan mindske dette problem gennem en klar handlingsplan for samarbejdet. Fremtidig anskaffelse af dødbringende autonome våbensystemer og AI-teknologi kan med fordel følge den såkaldte "Triple Helix-model" formuleret i regeringens strategi for dansk forsvarsindustri fra 2021, der samler aktører fra forskning, industri og stat i en undersøgelse af potentialet for innovative droneanskaffelser. Et sådant samarbejde, der inkluderer den civile sektor, bør dog afveje risikoen for, at strategiske retssager kan blive en hindring for industriens – især mindre teknologivirksomheders – deltagelse, hvilket risikerer at bremse udviklingen af en innovativ dansk tilgang til anskaffelse af autonome våbensystemer.

ACKNOWLEDGEMENTS

The author is grateful for valuable input from an anonymous peer reviewer as well as invaluable feedback and support from colleagues at CMS. A very special thanks to Kevin Jon Heller for reading many iterations of this report and his unwavering support.

1

Introduction

The undisputed advantages of artificial intelligence (AI), particularly lethal autonomous weapon systems (LAWS), coupled with the accelerating pace of AI innovation, make it likely that AI will be a prominent feature in future warfare. As armed forces and warfighters across the globe envision and prepare for future hostilities and operations, it is becoming clear that AI will be an imperative tool for militaries to maintain a warfighter advantage. But questions remain as states pivot from understanding military applications in the abstract toward visions of responsibly acquiring, procuring, and implementing future autonomous weapon systems. LAWS are autonomous systems that are able to respond independently to their environment without human intervention or supervision. These systems remain under development, and the technology is still brittle; nevertheless, the complexity of these systems requires that states like Denmark begin exploring the legal uncertainties.

It is vital to consider the acquisition and procurement phase of LAWS for two reasons.[1] First, the nature of the systems does not fit in traditional acquisition models; and established processes that work for conventional weapons do not necessarily account for the complexity of LAWS. This report deals exclusively with one aspect of the acquisition process: the international legal regulation of new weapon systems and the substantial legal uncertainty that exists for defense officials in the acquisition of LAWS. For example, David van Weel, NATO Assistant

1. There is an important distinction to acknowledge between acquisition and procurement. Acquisition is the broader process that includes the development of new systems, which will be a recurring theme in this report. Procurement, by contract, is the purchasing of new systems.

Secretary General for Emerging Security Challenges, was recently questioned about NATO's interpretation of responsibility when a LAWS engages in unlawful targeting conduct, and whether state actors or LAWS manufacturers should incur this responsibility. NATO does not currently have an answer, but the question requires closer examination of the responsibility from multiple angles, especially the enhanced risk for state responsibility.[2] Second, the acquisition and procurement phase contains high-stake decision-making in the context of LAWS, which is discussed at length. It is important to recognize that new considerations will factor into acquisition decisions, such as the inherent nature of the technology, and these considerations will have major implications for the Danish Armed Forces.

This report explores the legal challenges for the acquisition and procurement of LAWS. The legal challenges are one aspect of broader acquisition uncertainties, some of which are discussed in this report. But Danish acquisition officials will find that the legal uncertainties of LAWS are deeper than the standard weapons legal review process; an issue that Iben Yde discusses comprehensively in the 2021 report on the Danish weapons review and autonomous weapons.[3] Rather, this report investigates two legal challenges that are broader in nature: the nature of LAWS technology under international humanitarian law (IHL), and who can be held responsible when LAWS engage in unlawful conduct. These are two legal considerations with which the Danish Ministry of Defence (MoD) will likely require extensive coordination with the MoD legal office. First, qualities that are embedded in AI that are discussed below (transparency, predictability, bias) will likely be problematic for demonstrating compliance with IHL. These are qualities which were not at issue in conventional weapons procurement, because lethal decision-making was ultimately left to human judgement. In contrast, the autonomous nature of LAWS requires a deeper understanding of the

2. Sebastian Sprenger, "NATO Tees up Negotiations on Artificial Intelligence in Weapons," April 27, 2021, https://www.c4isrnet.com/artificial-intelligence/2021/04/27/nato-tees-up-negotiations-on-artificial-intelligence-in-weapons/.
3. Iben Yde, "Autonome våbensystemer i danske våbenscreeninger—Nye udfordringer og krav til implementeringen af den folkeretlige våbenscreeningsforpligtelse," Djøf Publishing, Copenhagen, June 2021.

technology, especially machine learning systems, in order to identify and demonstrate IHL compliance. This report offers a step in that direction.

Second, it is necessary to reconsider legal frameworks of responsibility and accountability in the event of unlawful machine conduct. This is especially pertinent as LAWS development requires substantial cooperation with the civilian defense industry, including the technology sector, which is not a traditional partner for defense procurement. This report explores the question of responsibility from multiple international legal frameworks which account for the range of actors involved in the lifecycle of LAWS development, including state responsibility, individual criminal responsibility, command responsibility, and corporate responsibility.

These issues are especially relevant for Danish defense stakeholders due to the Danish participation in the US-led Partnership for Artificial Intelligence for Defense. The first of its kind, this partnership was announced in late 2020, and it fosters cooperation in AI development for defense purposes and supports partnership collaboration and the sharing of development and data. Danish participation acknowledges the significance of AI and development of LAWS as a strategic asset and priority in the partnership, especially with the United States. This kind of partnership therefore holds implications for both partnership military cooperation and Danish defense/technology industry. As such, it is vital for Danish defense stakeholders to explore the legal issues presented here and to consider the policy recommendations in order to implement responsible and reliable AI technology and acquire weapon systems that meet international legal regulations.

The report concludes with five policy recommendations for Danish defense stakeholders regarding these legal uncertainties for future autonomous weapon system acquisition and procurement.

The AI landscape and legal parameters for military applications are changing rapidly. If and when Denmark decides to acquire lethal autonomous weapons, these recommendations can contribute to guiding future deliberations and planning.

1. **Formulate a policy**. Formulating a national policy regarding a Danish interpretation of LAWS can guide military decision-makers and legal advisers, as AI continues to be important for discussions of future warfighting capabilities. Clear policy becomes all the more

relevant as Denmark participates in international coalitions (e.g., the AI Partnership for Defense), but has largely remained silent on the LAWS issue. Danish allies (e.g., US, Australia, France) have already announced R&D programs, and in some cases an accompanying ethical framework, toward responsible AI weapon systems development. Other Danish allies (e.g., Germany) have instead publicly opposed such weapons development. Without a stated policy, Denmark risks falling behind critical security partners on future dialogue regarding LAWS.

2. **Encourage inter-agency coordination**. Strengthened cooperation and coordination between the Danish Acquisition and Logistics Office (DALO) and the legal office of the Ministry of Defense can ensure greater IHL compliance. The legal challenges of incorporating LAWS into the Armed Forces will become critical to the Danish warfighting capacity, and early stages of AI design should incorporate IHL. This will ensure that Denmark remains a competitive military ally with responsible, lawful autonomous systems.
3. **Restructure TEVV procurement**. Restructuring the Danish Testing, Evaluation, Validation, and Verification (TEVV) can help Denmark to mitigate the inherent challenges to autonomous weapon systems—namely algorithmic transparency, predictability, and bias. One option toward this end is to establish an iterative TEVV process offering accurate AI systems to Danish defense and legal stakeholder and ensuring that IHL standards are front-and-center of autonomous weapon development.
4. **Encourage joint acquisition/collaboration of LAWS through project development.** Project-based acquisition offers the greatest control and flexibility over LAWS design and development relative to purchasing LAWS off the shelf. In order to ensure the maximum IHL compliance and sufficiently trained algorithms, Denmark may want to consider avenues for acquiring LAWS through project-based development. This report offers two starting points. First, Denmark can pursue joint acquisition participation in multinational defense partnerships (e.g., NATO, the AI Partnership for Defense). Second, the Danish MoD can create formal partnerships with Danish academic and industry experts at the cutting edge of AI design and development.

5. **Mitigate strategic litigation risks**. Global opposition to LAWS development increases the risks of defense and technology partners experiencing strategic litigation and reputational backlash for defense collaboration. The MoD can temper this issue through an actionable collaboration plan. Future LAWS/AI acquisition may want to mirror the Danish "triple helix" collaboration strategy for drone acquisition—which combines research, industry, and the state to explore innovative acquisition potential—from the 2021 Danish National Defence Industrial Strategy. However, civilian collaboration will need to consider the risks of strategic litigation as a barrier for industry, especially smaller technology firms, with the potential to hinder an innovative Danish acquisition approach.

2

Artificial Intelligence and Weapon Systems

AI has been described less as a single technology and more a "system of systems," or general purpose technology with dual-use capabilities that is more akin to electricity than a single weapon, like a stealth bomber.[4] AI's range of applications and capacity for change is therefore extensive. Risa Brooks describes the range potential for AI in the security sphere:

> *AI is an enabling technology that can be used in diverse domains of military activity—everything from weapons systems, intelligence, logistics, and training to the learning tools employed in professional military education. When combined with robots, AI will increase the ability of machines to operate autonomously. With advances in robots, computing and neuroscience, military personnel will be able to compensate for their cognitive and physical limitations with biotechnology and implantable devices. These technologies will fundamentally reshape the character of war, if not—as some have speculated—its very nature.*[5]

4. See National Security Commission on Artificial Intelligence (NSCAI), February 2021, p. 7 [hereafter NSCAI 2021]. For a thoughtful discussion comparing AI to other military technological revolutions and a useful comparison of AI to electricity, see Michael Horowitz, Gregory C. Allen, Elsa B. Kania and Paul Scharre, "Strategic Competition in an Era of Artificial Intelligence," Center for New American Security, 2018.
5. Risa Brooks, "Technology and Future War Will Test US Civil-Military Relations," *War on the Rocks*, November 26, 2018, https://warontherocks.com/2018/11/technology-and-future-war-will-test-u-s-civil-military-relations/.

There is no officially accepted definition of lethal autonomous weapons, but the United States Department of Defense (DoD) definition is a useful starting point. According to the DoD, LAWS are "weapon system(s) that, once activated, can select and engage targets without further intervention by a human operator."[6] Lethal autonomous weapons systems identify and engage a target without human guidance or direction, or without a human "on the loop."[7] These systems offer a decisive advantage by propelling warfighting capabilities to machine speed that is beyond the capacity of human cognition.[8] As such, this could mean the speed of decision-making and action will outpace any human opponents—and certainly any chain of command—in such scenarios.[9] The operational benefits and strategic significance are potentially enormous. Nevertheless, some experts caution against the increasing hype around LAWS and remind us that the technology is still far from providing these capabilities. Some experts maintain, "It remains unclear when, whether, and in what contexts greater degrees of autonomy will provide clear advantages,"[10] and strong geopolitical and strategic incentives remain to pursue LAWS R&D to mitigate a large strategic competitor advantage, and states are steadily increasing their investments in LAWS development.[11]

There are two types of AI systems to consider for greater understanding of the capacity of LAWS. The first AI system type covers "hand coded" systems in which the weapon system decision-making is limited to

6. See United States Department of Defense, Directive 3000.09, November 21, 2012. Incorporating change May 8, 2017, https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.
7. Gary Schaub Jr. and Jens Wenzel Kristoffersen, "In, on, or out of the Loop? Denmark and Autonomous Weapon Systems," CMS Report, February 2017.
8. Michael C. Horowitz, "When Speed Kills: Lethal Autonomous Weapon Systems, Deterrence and Stability," *Journal of Strategic Studies* 42, no. 6 (2019); Elsa Kania, "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power," Center for New American Security, November 2017.
9. Kania, "Battlefield Singularity."
10. Kania, "Battlefield Singularity," 38.
11. Horowitz, "When Speed Kills;" for broader analysis on AI and geopolitical contexts, see Horowitz, Allen, Kania, and Scharre, "Strategic Competition;" Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," Texas National Security Review 1, no. 3 (2018); Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt, 2018); Samuel Bendett, "Should the U.S. Army Fear Russia's Killer Robots?" The National Interest, November 2017.

the initial data parameters.[12] There are pros and cons to this approach: On the one hand, hand-coded systems lead to more predictable LAWS and enhance human (or operator) trust in the system. On the other hand, the system will be limited in response options and not necessarily prepared for changing or evolving circumstances on the ground. These issues will be discussed in greater detail below.

The second type of system is a machine learning system, which unlike hand-coded systems is able to make decisions beyond the initial programming parameters. Essentially, these are systems designed to mimic human learning processes. They have the ability to learn and advance their own processes. Machine learning systems operate by "discovering correlations between variables in a dataset, often to make predictions or estimates of some outcome."[13] These systems are not "programmed" in the traditional sense. Instead, programmers create a structure allowing machine learning systems flexibility to learn and adapt to changing conditions and environments. As with hand-coded systems, there are pros and cons. Machine learning offers a more adaptive and potentially more accurate system that can learn from changing battlefield dynamics. But machine learning inevitably raises concerns over the control of lethal decision-making and the degree of risk of system errors.

Three legal concerns dominate the discourse surrounding the development and implementation of LAWS. The first is the generally applicability of IHL to LAWS and whether these systems challenge fundamental rules on the conduct of hostilities. The second issue relates to LAWS capacity for situational awareness and behavior regarding machine learning. The rules of warfare require responding to constantly changing dynamics on the ground, and there are legitimate concerns regarding machine performance toward such environmental changes. Third, and finally, substantial legal uncertainty remains regarding accountability for miscalculated decisions in the system, as well as system overrides or failures that may result in conflict escalation and unnecessary risks to civilians.

12. Tobias Vestner and Altea Rossi, "Legal Reviews of War Algorithms," *97 International Legal Studies 509* (2021).
13. David Lehr and Paul Ohm, "Playing with the Data: What Legal Scholars Should Learn about Machine Learning," *UC Davis Law Review 51* (2017): 671; quoted in Vestner and Rossi, "Legal Reviews," 537.

This report considers each of these claims and offers potential solutions for the procurement and acquisition process which may help mitigate some of these legal concerns.

3

Methodology

This report has two research objectives aimed at addressing the main concerns at the center of this emerging technology: to evaluate the technological qualities of LAWS against the legal backdrop of IHL and to evaluate the frameworks of responsibility for actors involved throughout the design-to-deployment weapon system lifecycle.

To conduct this analysis, this report employs a doctrinal legal method drawing on relevant international legal instruments, particularly on IHL sources and other instruments relevant to IHL and state responsibility. This offers an assessment of new technology on extant legal requirements as relevant to procurement officials. Additionally, this report takes the next step to situate the legal analysis within a broader security context. Research on AI is still emerging and it is vital to include legal analyses within larger security contexts for a comprehensive view of the challenges related to the procurement and implementation of autonomous weapons.

4

AI Acquisition and Procurement: The Challenges Ahead

It is not hard to imagine AI as a revolutionary tool for the international security environment by virtue of its potential to redefine the range of military capabilities. And with this change comes the redefinition of military-civilian partnerships for AI development. Militaries have long partnered with the civilian defense industry for the development and manufacturing of weapons, but acquiring LAWS opens a new sector for defense collaboration, and the technology sector has different incentives than those of traditional defense contractors.

Historically, traditional procurement processes have posed problems for fast-moving technological development because of the many restrictions, regulations, or lack of resources that may be necessary to stay at the cutting edge of technology development.[14] The incredibly fast-paced environment of AI innovation, coupled with the imperative of working with nontraditional defense partners, places procurement programs in a particularly challenging spot in acquiring LAWS.

There are two AI procurement challenges that Denmark and others will face in the future. First, and historically the most problematic, is the speed of procurement processes, which are often too slow to match the accelerating pace of AI innovation. In some contexts, the pace of gov-

14. For a thorough investigation into acquisition and new technology, see Philip S. Anton et al., "Strategies for Acquisition Agility: Approaches for Speeding Delivery of Defense Capabilities," RAND Corporation (2020), https://www.rand.org/pubs/research_reports/RR4193.html.

ernment procurement has deterred critical AI innovators or technology stakeholders from accepting military contracts.[15] In the US, for example,

> *...the main barrier to doing business with the Defense Department isn't ethics but bureaucracy. Except for a small number of defense contractors, most companies see the Pentagon as too small and low-margin a market compared to US civilian consumers—let alone Chinese ones—to justify the effort of complying with its complex regulations and congressionally imposed restrictions. Of the top 100 AI companies in the world, only two do business with DOD... "The two worst words you can have in a business plan, when you're raising money from tier one venture investors in Silicon Valley, are 'government customers'...it is the kiss of death."*[16]

The slow process of acquisition and working with new private sectors that are unfamiliar with government regulations and restrictions is clearly a hurdle. Lengthy testing standards and iterative maintenance with external providers leaves some acquisition experts to conclude that defense and military AI might need to be a government rather than an industry responsibility.[17] But maintaining military AI as a strictly government responsibility is unlikely—the tech industry is at the forefront of innovation and the government cannot compete with salaries and benefits of the tech industry in order to recruit and maintain AI talent.[18]

15. Experts have offered numerous reasons technology companies would be hesitant to cooperate with the military, including: ideological beliefs (Maaike Verbruggen, "The Role of Civilian Innovation in the Development of lethal Autonomous Weapon Systems," *Global Policy 10*, no. 2 [2019]), lengthy security clearance processes (Trevor Taylor, "Artificial Intelligence in Defence: When AI Meets Defence Acquisition Processes and Behaviours," *RUSI Journal* 164, nos 5/6 [2019]), loss of intellectual property rights (Renaud Ballais and Renelle Guichard, "Defense Innovation, Technology Transfers and Public Policy," *Defence and Peace Economics 17*, no. 3 [2006]) and profit margins (Verbruggen, "Civilian Innovation"). For example, profit margins for major US information technology companies are around 30-40%, and the Pentagon caps profit margins for defense companies at 15% (Loren Thompson, "Five Reasons Why Silicon Valley Won't Partner with The Pentagon," *Forbes*, April 27, 2015, https://www.forbes.com/sites/lorenthompson/2015/04/27/five-reasons-why-silicon-valley-wont-partner-with-the-pentagon/?sh=37e3ca8f4de9).
16. Sydney J. Freedberg, Jr., "Google Helps Chinese Military, Why Not US?" *Breaking Defense*, Quoted in Taylor (2019).
17. Taylor, "Artificial Intelligence," 80.
18. For insight into how the US can tackle this AI talent gap, see James Ryseff, "How to (Actually) Recruit Talent for the AI Challenge" *War on the Rocks*, February 2020.

Some states have already acted to adjust acquisition processes to better accommodate the accelerating pace of AI innovation. For example, the United States implemented two programs designed to streamline acquisition within modern technological development—the Better Buying Power (BBP) initiative and Tradewind. BBP was originally launched in 2010 to implement modified acquisition principles that foster better working relationships with civilian industry.[19] The goal was to emphasize that defense structures should, to the greatest extent possible, be tailored to the content of the product being acquired.[20] To increase flexibility in cooperation with civilian industry, there are multiple acquisition models to best fit the product and industry requirements—rather than a single model which is typical of traditional acquisition protocol. The process is also designed to adapt as the technology and needs of the DoD shift. The most recent iteration of this program, BBP 3.0, incorporates acquisition principles to focus on technical excellence and innovation, including continuous improvements to acquisition models, achieving dominant capabilities, and incentivizing innovation and competition.

The US Tradewind initiative is specifically for AI acquisition and officially launched in summer 2021.[21] The goal of Tradewind is to simplify and accelerate the AI acquisition and implementation into US warfighting capabilities. It encourages a streamlined business acquisition model, which allows the DoD to collaborate more easily with nontraditional partners. In recognizing the challenges inherent to AI acquisition, the DoD recognized the

> *...need(s) to accelerate the adoption of AI-enabled capabilities—essential to strengthening our military, increasing the effectiveness and efficiency of our operations, and enhancing the security of the Nation. [Tradewind]...will replace the current cumbersome procurement model, making the process more collaborative, research-based and effective.*[22]

19. Department of Defense, "Better Buying Power: Acquisition, Technology, Logistics," https://www.ustranscom.mil/dbw/docs/BBP_Fact_Sheet.pdf.
20. See Department of Defense Memorandum, 5000.02, January 2015, https://dod.defense.gov/Portals/1/Documents/pubs/Getting-Acquisition-Right-Jan2017.pdf.
21. Department of Defense, Joint AI Center, https://tradewindfaq.org/.
22. Ibid.

The UK has similarly streamlined processes for technological acquisition, especially regarding AI. A recent independent review of defense acquisition noted the UK MoD acquisition and procurement structures were unsustainable for future technology development.[23] The report advocated for the parliament to conduct routine strategic defense reviews as a mechanism for periodic "resetting" of MoD plans and procedures. Finally, the report recommends specific changes for equipment programs to mitigate wasteful spending on acquiring new technology with traditional procurement measures.[24]

Conversely, some experts have cautioned against the US and UK approaches to speedy technological procurement, warning that attempting to compensate for the pace of innovation could lead to irresponsible or dangerous concessions in other important areas, such as program management, sustainment, or other areas.[25]

The second procurement challenge has received less attention than the challenges of speed—and that is the global nature of private sector development. The wide accessibility and wide-scale attention on advanced technology has shone a public spotlight on the implications for warfare. The global discourse that the nature or character of warfare may be changed with the introduction of LAWS has taken away the exclusivity of new military technology. Accordingly, "the loss of exclusivity means the likelihood of technological surprise is far higher."[26] The pressure to stay ahead of competitors can pressure procurement officials to find fast, tenable solutions to quickly gain a competitive edge. The United States has acknowledged that military innovation will not lead the way in developing LAWS, but rather the best solution is to be "fast adapters—as opposed to sole developers—of technology, helping to integrate advanced commercial capability for strategic advantage."[27] This "fast follower" approach is the overarching framework for the US programs aimed at streamlining the AI acquisition process.

23. Bernard Gray, Review of Acquisition for the Secretary of State for Defence, October 2009, https://delta.bipsolutions.com/docstore/ReviewAcquisitionGrayreport.pdf.
24. Ibid.
25. Jonathan P. Wong "Bad Idea: Overly Focusing on Development and Acquisition Speed," RAND Commentary, December 16, 2020, https://www.rand.org/blog/2020/12/bad-idea-overly-focusing-on-development-and-acquisition.html.
26. Defense Innovation Unit Annual Report (2017), p. 2.
27. Ibid., p. 2.

Denmark has thus far opted not to follow the lead of the US and UK to create a streamlined acquisition for AI systems, and this type of program innovation is not necessarily in the best interest of Denmark. In general, not a lot is publicly available about the Danish acquisition process. Nevertheless, some key observations are relevant for understanding future approaches to LAWS and the legal context surrounding it. Traditionally, Denmark has preferred a procurement strategy of purchasing systems "off the shelf" (OTS) for numerous reasons (e.g., acquisition efficiency, saving money). The majority of the procurement process occurs in the Danish Acquisition and Logistics Office (DALO). The office has a mandate to implement the national procurement plan and collaborate with industry partners that will fulfill aims outlined in the Danish defense agreement and encourage international cooperation and Danish interoperability with military partners.

But there may be reasons for Denmark to be cautious of purchasing OTS LAWS systems, as the nature of LAWS does not easily translate into OTS solutions. Certainly, there are non-lethal AI systems that would prove unproblematic for OTS procurement; but lethal weapon systems would pose significant legal questions as to the transparency, predictability, and bias of an OTS system. The best way for Denmark to tackle these qualities of LAWS optimally would be to seek project-based solutions for Denmark to maintain control over the development and legal parameters of the system.

Project-based acquisition is the best path toward implementing the greatest adherence to Danish legal obligations because of the inherent nature of autonomous systems, as discussed in the next section. Denmark has two starting points toward this end. The first is joint acquisition programs with allies and partners through extant defense AI partnerships. Denmark's participation in the AI Partnership for Defense may be a useful platform to acknowledge and undertake solutions for LAWS acquisition challenges. While the AI Partnership for Defense does not explicitly recognize acquisition as a goal of the partnership, the contributing partners are still navigating the scope and potential of the partnership, and there is a common interest in advancing AI acquisition and adoption swiftly and responsibly. Additionally, Danish officials can ex-

plore NATO acquisition or procurement standards among the Allies.[28] As a standing multi-national partnership, NATO is uniquely situated to establish lawful and ethical standards for future adoption of LAWS, and Denmark can benefit from NATO's stewardship.

A second starting point is formal partnerships with Danish industry and academic experts at the cutting edge of AI development.[29] Defense stakeholders can address many of the challenges detailed in this report with a strengthened collaboration with experts in industry and academia to innovate acquisition processes and policies to adapt to the realities of AI and LAWS. Future acquisition planning can utilize this community to ensure that the legal framework and ethical considerations are deeply embedded in Danish acquisitions of LAWS.

This will be expanded in the pages below. But first, we need to develop the requirements of IHL, the foundations of AI technology, and the legal questions posed by a lethal autonomous system.

28. For more on NATO and AI standards, see Zoe Stanley-Lockman and Lena Trabucco, "NATO's Role in Responsible AI Governance in Military Affairs" forthcoming in *Oxford Handbook on AI Governance*, Available at SSRN, https://ssrn.com/abstract=3939769, August 2021.
29. The United States Air Force (USAF) created a similar research partnership with the Massachusetts Institute of Technology (MIT) creating the Artificial Intelligence Accelerator. This is a multidisciplinary team of researchers and military to conduct research enabling rapid prototyping, risk reduction, and ethical considerations for the USAF. Part of this partnership includes acquisition officers to learn how to manage AI programs and develop policies, processes, and lessons for future acquisition. See "Artificial Intelligence Acquisition Guidebook" Department of the Air Force and Massachusetts Institute of Technology, February 2022, https://aia.mit.edu/wp-content/uploads/2022/02/AI-Acquisition-Guidebook_CAO-14-Feb-2022.pdf.

5

Artificial Intelligence, Lethality, and International Law

International humanitarian law (IHL), sometimes also referred to as the law of armed conflict or laws of war, stipulates important restrictions in warfare. In order to comply with IHL regulations, there are limits to the types of weapons that can be used in an armed conflict. Article 36 of Additional Protocol I (1977) requires a legal review process to assess the compatibility of new means and methods of warfare (including new weapons) with IHL rules. A brief description of the IHL rules which govern the use of means of methods of warfare is included in Box 1.

Box 1. Key international humanitarian law rules governing the use of means and methods of warfare

Distinction

Parties to armed conflicts must at all times distinguish between civilians and combatants, and between civilian objects and military objectives. Attacks may only be directed against combatants and military objectives, never against civilians or civilian objects.[30] Lawful targets include combatants and civilians directly participating in hostilities, and objects that constitute military objectives.[31]

30. Protocol Additional to the Geneva Conventions of August 12, 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), adopted June 8, 1977, 1125 UNTS 3, entered into force December 7, 1978, Articles 48, 51(2) and 52(1); International Committee of the Red Cross, Customary IHL Study 1 (Customary IHL), Rules 1 and 7.
31. Protocol I, Articles 48 API, 51(2) and (3), 52(1) and (2); Customary IHL, Rules 1 and 7.

Prohibition of indiscriminate attacks

Indiscriminate attacks are attacks of a nature to strike military objectives and civilians and civilian objects without distinction, either because the attacks are not directed at a specific military objective, they employ a method or means of combat that cannot be directed at a specific military objective, or they employ a method or means of combat the effects of which cannot be limited as required by international humanitarian law (IHL).[32]

Prohibition of disproportionate attacks

The rule of proportionality prohibits attacks which, although directed at a military objective, are expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects or a combination thereof, that would be excessive in relation to the concrete and direct military advantage anticipated.[33]

Precautions in attack

In the conduct of hostilities, IHL requires parties to armed conflicts to take constant care to spare the civilian population, civilians, and civilian objects. The obligation to take precautions in attack requires persons who plan, decide on, and carry out attacks to:

- do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects, and are not subject to special protection but are military objectives
- take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event minimizing, incidental loss of civilian life, injury to civilians, and damage to civilian objects
- refrain from deciding to launch an attack if doing so may be expected to cause disproportionate civilian harm, and

  cancel or suspend an attack if it becomes apparent that the object is not a military one, or is subject to special protection, or that the attack may be expected to cause disproportionate civilian harm.[34]

Source: "Limits on Autonomy in Weapon Systems" ICRC & SIPRI (italics added)

32. Protocol I, Article 51(4); Customary IHL, Rules 11-13.
33. Protocol I, Article 51(5)(b); Customary IHL, Rule 14 ICRC Customary IHL Study.
34. Protocol I, Article 57(2)(a) and (b); Customary IHL, Rules 15-19.

New technology must be able to demonstrate compliance with four key IHL principles. Firstly, the principle of distinction requires parties of a conflict to, at all times, distinguish between civilians and combatants as well as civilian objects and military objectives.[35] Second, there is a prohibition of indiscriminate attacks, which are attacks that are not directed at a military objective, or employ a means or method of warfare that by nature cannot be directed at military objectives, or an attack that employs a means or method of warfare that causes effects which cannot be limited as required under Additional Protocol I (API).[36] Third, states are prohibited from engaging in attacks which are not proportionate to the military advantage gained from the attack; this requires that attacks on military objects which are expected to result in loss of civilian life or damage to civilian objects (or a combination thereof) must not be excessive in relation to the concrete and direct military advantage that is anticipated.[37] Finally, there is a requirement to take precautions in attacks to spare civilian life and objects.[38]

Not all of the IHL principles are problematized in the context of autonomous weapons. For example, the second principle, prohibition of indiscriminate attacks, does not necessarily pose a significant challenge for hyper-precise technologies like LAWS. Precision technology is largely able to meet the legal threshold for directing attacks toward legitimate targets. However, the final three principles (distinction, proportionality, and precaution) merit deeper assessment in the context of lethal autonomous weapons. The rest of this report will discuss the principles of distinction, proportionality, and precaution in the context of LAWS.

5.1. The Principle of Distinction

The principle of distinction requires that the parties to a conflict distinguish between lawful targets (combatants, military objectives, and civil-

35. Additional Protocol to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), adopted 8 June 1977, entered into force 7 Dec. 1978, Articles 48, 52(1); ICRC Customary IHL Study 1 (Customary IHL), Rules 1 and 7.
36. Protocol I, Article 51(4); Customary IHL, Rules 11-13.
37. Protocol I, Article 51(5)(b); Customary IHL Rule 14.
38. Protocol I, Article 57(2)(a) and (b); Customary IHL, Rules 15-19.

ians directly participating in hostilities) and unlawful targets (civilians, civilian objects, and persons *hors de combat*).[39]

Certain conditions, notably urban spaces, can complicate compliance with the distinction requirement because there is not always certainty or clarity as to who is a legitimate combatant or what a legitimate military object is. Critics of LAWS argue that the technology is particularly inept to make such an assessment effectively.[40] According to critics, the assessment of information that is necessary to comply with the distinction requirement is too complicated for a machine. For example, distinctions between civilians protected under IHL and civilians that are not because they are directly participating in hostilities (DPH) is a highly complex and situational assessment. Would a machine be able to distinguish an armed civilian who is currently DPH from an armed civilian police officer, who is protected under IHL? Strict object recognition (e.g., whether a person is holding a gun) does not in itself capture the environment in which LAWS must be able to navigate and which they must interpret.

The AI technology is not yet capable of making such decisions. The object recognition technology is still easy to manipulate, such as through hacking or deliberately misleading the recognition software, as discussed below. Nevertheless, this capability in itself (i.e., autonomously recognizing images and making data-driven calculations) is not necessarily inherently incompatible with IHL regulations. For example, LAWS could lawfully be employed in spaces where it is unnecessary to distinguish between civilian and military objects. As Michael Schmitt usefully acknowledged,

> *Not every battlespace contains civilians or civilian objects. When they do not, a system devoid of any capacity to distinguish protected persons and objects from lawful military targets can be used without endangering the former...The inability of weapon systems to distinguish bears on the*

39. Protocol I, Article 48.
40. See, e.g., Noel E. Sharkey, "The Inevitability of Autonomous Robot Warfare," *International Red Cross Review* (2012).

> *legality of their use in particular circumstances...but not their lawfulness, per se.*[41]

Regardless, the state of current technology has a long way to go in enhancing object recognition before it is anywhere close to the distinction standards required under IHL. This is particularly important for distinction as, currently, object recognition algorithms are easy to trick and manipulate. Small changes to an object can result in misclassifications and expose a high risk of intentional manipulation of the surrounding context by adversaries, or "adversarial attacks."[42] In one recent example, researchers at OpenAI, a US-based AI research laboratory, found that simply writing "iPad" on a piece of paper and placing it on any object will lead to the classification of an iPad.[43] In an armed conflict where there are substantial legal limits to targeting, simply writing the word "civilian" or "school" on combatants or objects could intentionally mislead LAWS or other autonomous systems and carry significant legal consequences.

Certainly, military AI researchers recognize the vulnerabilities for object recognition and improvements are happening rapidly, but the state of the technology is still too brittle to meet a standard of performance compliant with the distinction requirements under IHL. As discussed below, acquisition officials must reconsider the nature and qualities of machine learning systems in order to properly meet the complexities of LAWS and distinction.

5.2. The Principle of Proportionality

The second challenge for LAWS lies in the difficulties in complying with the proportionality requirement. The principle of proportionality

41. Michael N. Schmitt, "Autonomous Weapon Systems and International Humanitarian Law: A Reply to Critics," *Harvard National Security Journal* (2013). Quoted in Rebecca Crootof, "Killer Robots Are Here: Legal and Policy Implications," *Cardozo Law Review* 36 (2015): 1874.
42. Vestner and Rossi, "Legal Reviews;" AI scientists at MIT have shown the vulnerability in the technology with object recognition and simple images such as a turtle with a rifle led to significant misclassifications by the algorithms. See Will Knight, "Military Artificial Intelligence Can Be Easily and Dangerously Fooled," *MIT Technology Review* (2019).
43. Open AI Blog https://openai.com/blog/multimodal-neurons/, March 4, 2021.

prohibits attacks that are expected to result in any loss of civilian life or damage to civilian objects that exceeds the relative or concrete military advantage that is anticipated as a result of the attack.[44] This standard requires commanders to conduct complex proportionality assessments using the available knowledge to weigh elements that are extremely subjective and potentially constantly changing on the ground.[45]

The complexity of proportionality assessments presents a particularly difficult challenge for AI developers because they require dynamic, multivariate assessments that could, potentially, introduce new information or data instantly. The complexity at play in these assessments has led some to conclude that the proportionality assessment requires a "distinctively human judgement."[46] But this does not mean that AI systems cannot offer significant support for proportionality assessments.

Proportionality assessments have two distinct elements. The first is the prediction of civilian damage, and the second is the anticipated direct or relative military advantage to be gained from the attack, and these must be weighed against each other.[47] The first element—prediction of civilian damage—requires a calculation that is particularly suited to the strengths of AI; that is, speedy data analysis, such as blast radius and risks to civilian population. This kind of calculation is an advantage that AI has over human cognitive limitations and can perform this data analysis at an accelerated pace.

The second element is the controversial application of AI, which, given the current state of technology, exposes the limits of AI in the proportionality assessment. Human Rights Watch, a consistent critic of the development of LAWS, claimed that proportionality assessments, "require more than a balancing of quantitative data, and a robot [LAWS] could not be programmed to duplicate the psychological processes in human judgment that are necessary to assess proportionality."[48] This is

44. Protocol I, art. 51.
45. Crootof, "Killer Robots."
46. Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Human Rights Council, U.N. Doc. A/HRC/23/47, April 2013, at 14.
47. For an expansion of this argument, see Schmitt, "Autonomous Weapon Systems," especially pp. 18-23.
48. Human Rights Watch, "Losing Humanity: the Case against Killer Robots" (November 2012): 30, https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots.

widely acknowledged as the limitation of LAWS and proportionality assessments.[49] This assessment, especially when taking account of the inherent limitations of the technology discussed in the next section, will be a central issue if LAWS enter the acquisition procedures.

For these reasons, some critics claim that AI technology will never be able to conduct in-theater proportionality assessments because the constant re-calibration of the information and the dynamic environment requires a high threshold of qualitative analysis that goes beyond quantitative, data-driven analysis.[50] Critics also acknowledge the inherent difficulties in determining responsibility for miscalculations in proportionality assessments—it is unclear where the responsibility lies for algorithm-generated calculations, especially if the environment changes significantly.[51] This question of responsibility is revisited in section seven. Nevertheless, it should also be recognized that, as AI innovation accelerates and the capabilities of LAWS, in theory, become more powerful, it remains to be seen whether machines are able to overcome the legal challenges that section six expands on AI innovation and the state of the technology itself.

5.3. Principle of Precaution

Article 57 of API requires state parties to take constant care to protect the civilian population, civilians, and civilian objects.[52] This includes an obligation for the persons who plan, decide on, and carry out attacks to

49. For example, Michael Schmitt, "Autonomous Weapon Systems," in direct response to the Human Rights Watch report "Losing Humanity" accepts that the quantification of military advantage is a challenge in utilizing this weapon, though acknowledges that it is not impossible to overcome this challenge. See Schmitt, "Autonomous Weapon Systems," 21.
50. Lieutenant Colonel Alan L. Schuller, "Artificial Intelligence Effecting Human Decisions to Kill: The Challenge of Linking Numerically Quantifiable Goals to IHL Compliance," *Journal of Law and Policy for the Information Society 15*, nos 1-2 (2019); on AI and Proportionality more generally, see Jeroen van den Boogaard, "Proportionality and Autonomous Weapons Systems," *Journal of Humanitarian Legal Studies* (2015).
51. Crootof, "Killer Robots;" see also Human Rights Watch and International Human Rights Clinic, Harvard Law School, "Advancing the Debate on Killer Robots: 12 Key Arguments for a Preemptive Ban on Fully Autonomous Weapons" (May 2014), https://www.hrw.org/sites/default/files/related_material/Advancing%20the%20Debate_8May2014_Final.pdf.
52. Protocol I, Article 57(1).

take precautions when doing so. The precaution principle is unique from the distinction and proportionality principles previously discussed, because it demands a certain conduct that it is not related to the outcome of the attack, as with distinction and proportionality.[53]

The rule requires the attacker to have "constant care" to protect civilian populations and civilian objects from the effects of an attack. To meet this obligation, Article 57 requires planners or decision-makers for attacks to: (1) "do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects and are not subject to special protection but are military objectives," (2) cancel an attack if it becomes apparent that the rules of distinction or proportionality will be breached, (3) provide "effective advance warning" of an attack if it may affect the civilian population, "unless circumstances do not permit," (4) "when a choice is possible between several military objectives for obtaining a similar advantage, select the attack on which may be expected to cause the least danger to civilian lives and to civilian objects," and (5) "take all precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects."[54]

The procurement or acquisition of any autonomous weapon system must consider how the technology satisfies the rules of precaution. Some of the above requirements will likely be satisfied if the LAWS satisfies the standards for distinction and proportionality; for example, a LAWS with the advanced capabilities to distinguish civilians/combatants and civilian objects/military objectives, and conduct a proportionality assessment, will likely be able to cancel an attack if these requirements cannot be met. However, it is the requirements that require an extra step of conduct not inherently tied to other IHL principles that merit further exploration.[55]

53. Kimberley Trapp, "A Framework of Analysis for Assessing Compliance of LAWS with IHL (API) Precautionary Measures," Convention on Certain Conventional Weapons (CCW) Informal Meeting of Experts (2016), https://discovery.ucl.ac.uk/id/eprint/1513936/1/Trapp_CCW%20Informal%20Meeting%20of%20Experts%20(2016).pdf.
54. Protocol I, Article 57(2-3); see also Jeffrey S. Thurnher, "Feasible Precautions in Attack and Autonomous Weapons," in *Dehumanization of Warfare: Legal Implications of New Weapon Technologies* (eds Wolff Heintschel von Heinegg, Robert Frau, Tassilo Singer) (2018).
55. Eric Talbot Jensen, "Autonomy and Precautions in the Law of Armed Conflict," *International Law Studies 96* (2020).

The first requirement obliges decision-makers to "do everything feasible to verify" that the objects of the attack are legitimate military objectives. Essentially, this precaution reflects the distinction principle. A LAWS that satisfies the distinction requirement will likely also be able to satisfy this requirement. However, it is important to note that the "do everything feasible" requirement requires high transparency of the LAWS. Transparency is discussed in depth below, but it is important to flag that the precaution standards will require a high degree of transparency, or explainability, to understand how the LAWS settled on its conclusion and determine if it did everything feasible to make the assessment.

The third requirement, "provide effective advance warning" to civilian populations potentially impacted by the attack, also warrants further scrutiny. As already discussed, the main operational benefit of employing LAWS is the rapid decision-making and accelerated pace of the decision to attack when executing the attack. With the pace of these systems and the ability to analyze large amounts of data (e.g., determining if a civilian population will be affected), fulfilling the advance warning requirement is not likely problematic. Implementing a simple loudspeaker into the LAWS that can warn nearby civilian populations is sufficient to meet this criterion.[56] The exception to the advanced warning requirement includes military factors; for example, advanced warning would not be necessary if it compromised the operational need for a surprise attack.

The fourth precaution rule requires an attack to select the object of attack expected to result in the least harm to civilian populations or objects. As some experts have also argued, this is a difficult assessment for a LAWS for the same reasons the proportionality assessment would be difficult.[57] The assessment requires a complex analysis of military advantage and other variable military factors that could be dynamic and evolving in

56. Thurnher, "Feasible Precautions;" Thurnher explains in greater depth: "There are no established standard forms for the warnings. The warnings can be a general message delivered either to the leadership of the enemy nation or directly to the civilian population. An attacker is not required to explicitly detail the particular time or place of the planned attack, but the warning should provide as much detail as the circumstances allow. The warning must be sufficient to provide the civilian population the opportunity to take measures to avoid the dangers. The delivery method of the warning can vary and can range from methods such as leaflets or media broadcasts," p. 111.
57. Thurnher, "Feasible Precautions."

the course of a situation. As argued above in the case of proportionality, however, this is not necessarily an impossible task for LAWS. The data analysis capacity vastly exceeds human capacity, and LAWS will likely be able to evaluate all of the available data necessary to make this assessment faster than human counterparts will. Nevertheless, the issue is the conclusion that the machine will draw from that data analysis, including whether it is the correct one. Certainly, the current state of AI and machine learning technology would be unable to carry out this assessment to a standard that will satisfy the military; but that does not mean it will be impossible in the future.

The final precaution rule requires attackers to choose the means and methods most likely to minimize harm to civilians. States can satisfy this requirement in a number of ways. If LAWS were to have multiple munition options, it would have to be able to decide which method would accomplish its mission while minimizing danger to civilians.[58] Additionally, LAWS will likely be able to satisfy alternative means of warfare to minimize civilian damage. For example, LAWS are able to surveille a target and use metadata to determine the moment the target is farthest away from civilian populations. This is an assessment that LAWS are arguably better designed to carry out than human operators.

In sum, there are concerns and challenges regarding precautionary measures and LAWS. The procurement of weapon systems ought to ensure that the system can satisfy the steps necessary for precaution; as the previous discussion demonstrated, however, many of the concerns can be resolved by ensuring distinction and proportionality compliance. Nevertheless, certain conditions (e.g., the advance warning capability) will be necessary to ensure.

5.4. Conclusions

The above analysis offers a deeper discussion of IHL obligations and the main considerations necessary for LAWS procurement that may be unique or distinct from conventional weapon systems with human control. In order to assess the potential challenges for LAWS compliance,

58. Thurnher, "Feasible Precautions."

however, it is also important to look more closely at the technology itself and the qualities or features of AI that procurement stakeholders will need to consider. In particular, the next section tackles three qualities of AI that experts generally consider the most problematic to operationalize in a weapon system: transparency, predictability, and algorithmic bias.

6

Decoding the Tech: Transparency, Predictability, and Algorithmic Bias

Evaluating LAWS compliance with IHL regulations requires a deeper understanding of the technology. AI experts have extensively considered the inherent risks of LAWS in a range of scenarios, from inadvertent crisis escalation to system failure likelihood; but the legal debate about the qualities of technology as problematic (or not) for IHL warrants further investigation.[59] IHL experts have discussed LAWS compliance broadly to determine whether IHL can apply to LAWS as an emerging weapon system. But answering that question requires a deeper engagement with IHL and the parameters and determinates of machine learning and behavior.

This section examines three qualities of AI (transparency, predictability, and bias) under the distinction, proportionality, and precaution requirements. Ultimately, for LAWS to be in compliance with IHL, the testing and evaluation (T&E) processes must monitor the evolution of these qualities, ideally with the cooperation of legal experts. Incorporating IHL standards into the design and development processes can mitigate potential legal challenges down the road, which is best accomplished by re-imagining existing testing, evaluation, validation and verification (TEVV) protocols.

59. For a brief description of AI technology and challenges for defense, see Stanley-Lockman and Trabucco, "NATO's Role."

6.1. Algorithmic Transparency

Algorithmic transparency (sometimes also called explainability, traceability, or the "black box effect") is the understanding of why and how a machine produced a particular outcome. Transparent AI is typically included in national ethical frameworks because it is an important component of trustworthy AI and can simultaneously improve AI accuracy.[60] If we can understand the process leading to the output, then we can better identify where improvements can be made to optimize that output. Understanding LAWS and the risks of failure in the system increases our knowledge and understandings of IHL compliance.

Establishing a degree of transparency may also be vital to demonstrate LAWS compliance with IHL. Transparency allows operators to understand machine distinction determinations, which optimizes the likelihood of high compliance with IHL distinction requirements. This allows operators to improve the system, including the eradication of inconsistencies or other failures in the LAWS data analysis. Transparent systems offer a higher guarantee of machine optimization and legal compliance.

Transparency in LAWS is significant when addressing system failure, or the failure to operate as intended, which clearly has implications for IHL legal compliance. Transparency in the systems uncovers system vulnerabilities, limitations, and risk of failure.
Put another way,

> *(L)AWS would generally fail differently to how human soldiers would fail; humans may fail to adequately perform some task for which they are, in theory, "programmed" (trained) due to inattention, fatigue and a host of other human factors...but when a human encounters situations outside their circle of expertise, they are able to apply some common sense, imagination and other abilities to lessen the negative impact. Computers...do not typically fail to follow their programming but, being bound*

60. See also Cynthia Rudin and Joanna Radin, "Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson from an Explainable AI Competition," *Harvard Data Science Review* 1, no. 2 (2019).

> *to stay strictly within the limits of that programming, tend to "fail" more suddenly when faced with unanticipated circumstances.*[61]

Also important is that transparency enhances trust from human operators and can reduce the likelihood of misperceptions and miscalculations.[62] Studies repeatedly reveal that transparency about a system's failures or errors is important to establish trust from human operators, including trust in the data output—which is particularly important in operations involving human-machine teaming.[63]

6.2. Algorithmic Predictability

Another key quality of LAWS necessary for IHL compliance is machine predictability. Algorithmic predictability means that the machine behavior is expected and iterative given a set of environmental conditions. A LAWS must have a level of performance predictability in order to meet the legal obligations under IHL. In May 2021, the International Committee of the Red Cross (ICRC) argued that "unpredictable autonomous weapons should be ruled out, notably because of their indiscriminate effects, and that this would be best achieved through a prohibition of unpredictable autonomous weapons."[64] Operators must be able to trust the system to perform consistently replicating outcomes from training data in hostile conditions.

A number of factors make LAWS performance predictability—and by extension IHL compliance—challenging. The first is the increasing complexity of technology. As the systems are being designed to conduct

61. Tim McFarland, *Autonomous Weapons Systems and the Law of Armed Conflict: Compatibility with International Humanitarian Law*, Cambridge University Press, 2020, p. 63.
62. Department of Defense, Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity (2018), https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF; see also Deloitte, Transparency and Responsibility in Artificial Intelligence: A Call for Explainable AI, https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/innovatie/deloitte-nl-innovation-bringing-transparency-and-ethics-into-ai.pdf.
63. Margarita Konaev, Tina Huang, Husanjot Chahal, "Trusted Partners: Human-Machine Teaming and the Future of Military AI," CSET *Issue Brief*, February 2021.
64. Peter Maurer, ICRC Statement, May 12, 2021, https://www.icrc.org/en/document/peter-maurer-role-autonomous-weapons-armed-conflict.

more sophisticated tasks in complex environments, the technology requires more command input combinations and environmental stimuli.[65] Increasing the scale and complexity of the software naturally increases the rate of errors and predictability, and it should be monitored as systems develop.[66] Another challenge will be the sustainability of performance predictability. These systems are expected to perform over a longer period of time than other weapon systems.[67] On the one hand, these systems are designed for sustained efforts, and they yield "potentially unlimited persistent capabilities without degradation due to fatigue or lack of attention."[68] On the other hand, these benefits are conditioned on the system's ability to perform predictably over lengthy periods of time.[69]

The predictability challenge depends partly on the type of AI initially programmed. A recent article by Tobias Vestner and Altea Rossi argues that the difference between a lethal AI system that is restricted to the initial parameters established by the programmer (called "hand-coded programming") and machine learning systems (which can operate beyond the initial parameters established by programmers) underpins the legal requirement of predictability.[70] In their assessment, AI systems operating within hand-coded programming will be less problematic for legal compliance because the machine is limited to the initial inputs, and thus more predictable. But hand-coded systems also have disadvantages, as they are unlikely to have the dynamic response capacity to changing combat conditions. Machine learning systems can function beyond their initial input to respond to environmental conditions that may be more responsible or informed than hand-coded systems. The armed forces would likely enjoy greater operational advantage in using machine

65. McFarland, *Autonomous Weapons Systems*; see especially Chapter 4.
66. Some studies have shown that the average error rate in the software industry is approximately 15-50 errors per 1,000 lines of code. To compare with a modern weapon system, the F-35 Joint Strike Fighter uses over 20 million lines of code (McFarland, *Autonomous Weapons Systems*). See also Roberto Cordeschi, "Automatic Decision-Making and Reliability in Robotic Systems: Some Implications in the Case of Robot Weapons," *AI & Society* 28 (2013); William Bialek, Ilya Nemenman, and Naftali Tishby, "Predictability, Complexity, and Learning," *Neural Computation* 2409 (2001).
67. McFarland, *Autonomous Weapons Systems*.
68. Defense Science Board "The Role of Autonomy in DoD Systems" Task Force Report, US Department of Defense (July 2012).
69. McFarland, *Autonomous Weapons Systems*.
70. Vestner and Rossi, "Legal Reviews."

learning systems due to this flexibility and capacity for environmental response, but they would then face the predictability challenge of machine learning systems.

As discussed previously, machine learning systems operate within an initial programmed structure that allows the system the flexibility to learn and respond to changing conditions and environments. This feature makes machine learning systems inherently less predictable than hand-coded systems. For a machine learning system to "learn," it is trained on a substantial amount of data. Yet current machine learning systems still lack deterministic behavior, even when confronted with inputs on which the system was trained or similar to input on which the system was trained.[71] As such, while the current state of machine learning systems remains brittle, this does not mean that future systems cannot improve in this regard.

6.3. Algorithmic Bias

The third AI quality challenging IHL compliance is algorithmic bias. Algorithmic bias is when a system "exhibits behavior and biases that result from a variety of decisions and inputs."[72] Here, system performance is conditioned on biases which could stem from multiple sources throughout the system development cycle, such as biases from the original programmer or biases within the data set used for training purposes.

Algorithmic bias would have implications for IHL distinction, proportionality, and precaution requirements. A biased AI system may have misguided preconceptions regarding who is a combatant and civilian, leading to inaccurate distinction calculations. Similarly, a biased system could inadvertently prioritize particular aspects of military necessity, leading to miscalculated proportionality assessments and ultimately en-

71. Vestner ad Rossi, "Legal Reviews," especially pp. 537-41.
72. United Nations Institute for Disarmament Research, "Algorithmic Bias and the Weaponization of Increasingly Autonomous Technologies: A Primer" No. 9 (2018): 1, https://unidir.org/publication/algorithmic-bias-and-weaponization-increasingly-autonomous-technologies; Will Knight, "Forget Killer Robots: Bias is the Real AI Danger," *MIT Technology Review*, October 3, 2017; Nema Milaninia, "Biases in Machine Learning Models and Big Data Analytics: The International Criminal and Humanitarian Law Implications," *International Review of the Red Cross* 102 (913) (2020).

gaging in risky attacks unnecessarily endangering civilians. And, finally, algorithmic bias could potentially affect the constant care standard for precautionary measures or impact the assessment to determine all means to take alterative measures and mitigate the risk to the civilian population.

Bias in training data can occur in two stages: data collection and data preparation.[73] Data collection can infuse bias by either collecting data that is unrepresentative of reality or data that reflects existing prejudices. Bias can also exist in data preparation through determinations of what attributes you want the algorithm to consider; "this is what people often call the 'art' of deep learning: choosing which attributes to consider or ignore can significantly influence your model's prediction accuracy."[74]

Algorithmic bias has significant implications for on-the-ground performance. Studies on algorithmic bias offer valuable insight into bias translating into performance, and the negative consequences. For example, a University of Virginia researcher was training an image recognition machine learning model, but the image data that was used to train the system disproportionately associated images of kitchens with women. This subsequently led the machine to form biased conclusions about gender and household duties.[75] In another image recognition study, algorithmic bias emerged due to the disproportionate number of training images used of men with guns, which led the AI to draw particular conclusions regarding gender and violence. It is not difficult to imagine the implications of a biased LAWS in a battlefield environment. It is critical not only to demonstrate legal compliance, but for operational effectiveness for the system to have accurate and unbiased training. This is also why maintaining as much control over the design, development, and training processes is critical for accurate and reliable warfighting capabilities.

73. Karen Hao, "This Is How AI Bias Really Happens—And Why It's So Hard to Fix," *MIT Technology Review*, February 4, 2019, https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/.
74. Hao, "AI Bias."
75. Milaninia, "Biases;" Tom Simonite, "Machines Taught by Photos Learn a Sexist View of Women," *Wired*, August 21, 2017, https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/.

Bias can also stem from the original designer or programmer. Programmer bias can begin in the earliest stages of development, when objectives are transformed into quantified measures in the learning model. For LAWS, the objectives that must be quantified are extremely difficult and leave room for subjective human judgement to creep in. Essentially, "the ability of AI-enabled weapon systems to comply with IHL will depend in part on whether the tasks in question are susceptible to being described as numerically quantifiable for objective standards."[76] The original determinations for quantifying such tasks will ultimately have an important impact on machine performance and battlefield outcomes.

To be sure, algorithmic bias is a serious concern when introducing LAWS (and AI more broadly) into the military procedure. But it is not necessarily an insurmountable concern; bias can be mitigated through sufficient system training.

6.4. Mitigating the Challenges: Considering TEVV Overhaul

Each of the technological qualities—transparency, predictability, and bias—could pose legal challenges for IHL compliance, and thus acquisition and procurement. There are steps that states can take to overcome these challenges and adopt and integrate AI weapon systems in a responsible and effective way. One way to accomplish this is through rethinking the national defense testing, evaluation, verification, and validation (TEVV) processes, and exploring opportunities for project-based acquisition. TEVV is the combination of two separate processes that are critical parts of acquisition and ensure that new weapons meet the legal criteria and safety measures required for weapon adoption and integration. Testing and Evaluation (T&E) occurs at an earlier stage of weapons development and ensures new technology or systems perform as expected in the intended environment.[77] The second phase is the Verifi-

76. L.C. Alan L. Schuller, "Artificial Intelligence Effecting Human Decisions to Kill: The Challenge of Linking Numerically Quantifiable Goals to IHL Compliance," *I/S: A Journal of Law and Policy 15*, no. 1 (2019).

77. According to the United States DOD, "test and evaluation shall be structured to provide essential information to decision-makers, assess attainment of technical performance parameters, and determine whether systems are operationally effective, suitable, survivable,

cation and Validation (V&V), which ensures that a system reflects the developer's intended description and specification, and that it reflects the real-world environment conditions for its intended use; additionally, validation confirms the system is the correct model for those who will use it.[78]

Taken together, the TEVV processes follow a cycle of technological adoption beginning with development and training and continuing to ensure new weapon systems are consistent with national and international criteria or standards. Currently, TEVV for conventional weapons follows a more "linear process" in which "companies must pass through a series of acquisition phases and milestone decision points—moving from prototyping/technology maturation to manufacturing and development to production and deployment."[79]

A linear process is ill suited for the acquisition of LAWS in part due to IHL requirements. Some defense experts have argued for TEVV restructuring and overhaul to better accommodate emerging technology (e.g., AI, machine learning).[80] For example, former US defense officials argued for major changes to the TEVV procedures at the US DoD in favor of processes that are more collaborative with industry experts and instead institute iterative TEVV processes. Some even argue that current approaches to TEVV function more as barriers to fielding AI systems to be operational in an effective timeframe.[81] Acquisition and procurement processes that are flexible for iterative TEVV processes, meaning TEVV procedures that continue to retrain themselves and pass multiple V&V phases, will contribute to mitigating transparency, prediction, and bias

and safe for intended use. The conduct of test and evaluation, integrated with modeling and simulation, shall facilitate learning, assess technology maturity and interoperability, facilitate integration into fielded forces, and confirm performance against documented capability needs and adversary capabilities as described in the system threat assessment." U.S. Department of Defense, Test and Evaluation Management Guide 220 (6th ed. 2012), para. 2.1 on 23. Quoted in Vestner and Rossi, "Legal Reviews."

78. U.S. Department of Defense, Test and Evaluation Management Guide 220 (6th ed. 2012); Vestner and Rossi, "Legal Reviews."
79. Michèle Flournoy, Gabrielle Chefitz, Avril Haines, "Building Trust through Testing: Adapting DOD's Test & Evaluation, Validation & Verification Enterprise for Machine Learning Systems, including Deep Learning," October 2020, p. 7, https://cset.georgetown.edu/wp-content/uploads/Building-Trust-Through-Testing.pdf.
80. Most directly, this issue was discussed in Flournoy et al., "Building Trust," in the context of DOD-specific TEVV policies that should be restructured.
81. Flournoy et al., "Building Trust," especially p. 3.

issues. As the previous section detailed, these measures are important for demonstrating IHL legal compliance and minimizing the inherent risks of lethal autonomous weapons posing further legal questions or uncertainty for armed forces.

Iterative acquisition processes with companies that specialize in the development of relevant algorithmic systems will promote greater transparency, because the system can be designed and developed in cooperation with Danish defense input. An iterative model likely enhances the accuracy and predictability of the system through more training on classified defense data that will help the system maintain conclusions consistent with battlefield-similar conditions. And, finally, an iterative acquisition model can mitigate bias through access to defense data and continuous LAWS training programs. Restructuring TEVV processes toward an iterative model may be a useful step for Denmark to comply with its IHL legal requirements. As previously discussed, LAWS are fundamentally a new weapon, and Danish MOD legal experts may want to examine these new weapons systems in light of their transparency, predictability, and bias. Including IHL legal experts in the very beginning of LAWS development and prioritizing cooperation and coordination between the DALO and MOD legal office will ensure acquisition frameworks that are built to suit AI development as well as suited to address unprecedented legal questions.

Prioritizing an iterative TEVV process is also possible and consistent with shifting toward project-based acquisition approaches. Danish procurement stakeholders could collaborate with a range of AI and defense manufacturers on tailored projects for LAWS development that incorporate IHL legal expertise from the beginning of development.[82] Additionally, LAWS systems training would be secured with in-house datasets (rather standard training data produced by manufacturers) to use real-world training data reflecting Danish standards and to mitigate predictability and bias issues.

82. Peter Margulies, "Making Autonomous Weapons Accountable: Command Responsibility for Computer-Guided Lethal Force in Armed Conflicts," in *Research Handbook on Remote Warfare* (ed. Jens David Ohlin) (Northamton MA: Edward Elgar, 2017) makes a similar point that "the demands of IHL will not allow a military unit to use an 'off the shelf' autonomous weapon system," p. 432.

Switching to a project-based acquisition approach obviously incurs costs. This approach requires more time, enhanced costs, and increased engagement with the civilian defense and technology industries. For DALO, with limited resources, this change would require significant Danish investment. Nevertheless, this is the most straightforward option to address and resolve some of the challenges inherent to LAWS and the legal issues associated with their use. As this report has already demonstrated, Danish allies have adjusted to these issues by establishing AI systems to fulfill legal and ethical frameworks at the international and national levels, respectively. This also opens new avenues of joint acquisition and collaboration with partners who have begun to innovate acquisition systems to better accommodate the realities of AI. Working alongside partners can alleviate much of the upfront costs associated with development and ensure greater interoperability with the AI systems of critical partners. As mentioned previously, Denmark can utilize its participation in the AI Partnership for Defense or work through NATO as a forum to promote allied standards for the development and implementation of future autonomous weapon systems. NATO is vocal about the importance of AI to its future strategies and capabilities, but it has not yet outlined a specific strategy for NATO-wide standardization.[83] Future planning for a Danish approach to LAWS that considers these steps will bring Denmark to the forefront of AI policy and toward implementing lawful and responsible LAWS.

6.5. Conclusions

This report presents two legal issues critical to the acquisition of LAWS. This section tackled the first of these issues: the inherent qualities of the technology potentially impeding IHL compliance. Examining LAWS transparency, predictability, and bias, this section has argued

83. In October 2021, NATO released its first AI Strategy outlining the six principles to guide allied development and deployment of AI: lawfulness, responsibility and accountability, explainability and traceability, reliability, governability, and bias mitigation. However, this initial AI strategy "only" calls for allied cooperation and does not go further toward guiding principled action; https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html.

that demonstrating IHL compliance will be particularly complicated by these three qualities; qualities which were not as relevant in the conventional weapons context.

This section has offered three policy recommendations for Danish officials to mitigate these issues, although they are also applicable for other national contexts. First, the Danish MoD acquisition office (DALO) and MoD legal offices can enhance coordination on LAWS acquisition.[84] Incorporating IHL legal expertise into early stages of weapons design and development will streamline the legal process and ensure that Denmark has approached autonomous systems with international law at the core. Second, Denmark could address these legal issues through a project-based acquisition model to give the greatest control in reducing the risks associated with AI transparency, predictability, and bias. Project-based solutions can be designed in cooperation with partners (e.g., partnerships like the AI Partnership for Defense or NATO). Additionally, as the next section demonstrates, this also contributes to minimizing an increased risk of legal responsibility or potentially even criminal liability. Second, Danish TEVV processes would benefit from an iterative model for LAWS to achieve the greatest accuracy and fulfill IHL requirements.

The next section tackles the second point of legal uncertainty for LAWS, which is the complexities of determining responsibility and accountability, and how unlawful conduct by a LAWS can be understood through multiple international legal accountability regimes.

84. This is especially true for the Article 36 Weapons Review; see Yde, "Autonome våbensystemer."

7

Responsibility under International Law & LAWS

One of the most pressing issues for the adoption and integration of LAWS is the uncertainties regarding responsibility and accountability. Since machines cannot be held legally accountable, the perceived "accountability gap" regarding who is responsible for machine violations is a critical component for LAWS acquisition and procurement. There are many stakeholders in the lifecycle from LAWS development to implementation; acquisition officials will be pressed to consider responsibility within four frameworks—state responsibility, individual responsibility, commander responsibility (a subset of individual responsibility), and corporate responsibility. Each vantage point plays a role in machine development and performance. This section outlines key considerations for acquisition stakeholders within each framework of responsibility.

7.1. International Law of State Responsibility

The law of state responsibility is a central institution in international law and defines when a state has breached an international obligation, the consequences of the breach, and the appropriate measures to be taken to implement the consequences of the breach. The rules on state responsibility are articulated in the International Law Commission (ILC) ar-

ticles on "Responsibility of States for Internationally Wrongful Acts."[85] The ILC articles specify the principles governing the responsibility of states in instances where the state has committed an internationally wrongful act.[86] Article 1 of the ILC articles outlines how "every internationally wrongful act of a state entails the international responsibility of that state, and thus gives rise to the new international legal relations additional to those which existed before the act took place."[87] An internationally wrongful act occurs when two conditions are met: (1) conduct consisting of an action or omission is attributable to the state under international law; and (2) when that action or omission constitutes a breach of an international obligation.[88]

The first element of an internationally wrongful act, attribution to the state, only applies to conduct that is "attributed to the State...that is of its organs of government, or of others who have acted under the direction, instigation or control of those organs, i.e. as agents of the State."[89] An "organ of the state" refers to any person or entity that carries that status in accordance with the internal law of that state.[90] A LAWS does not necessarily in itself constitute an organ of the state that is capable of acting on behalf of the state. The ILC has clarified, "the 'act of the State' must involve some action or omission by a human being or group."[91] The more appropriate organ of the state is the commander who makes the decision to deploy LAWS and determines the conditions to use LAWS. This is discussed in more detail below.

85. International Law Commission, "Responsibility of states for internationally wrongful acts," draft articles, text adopted by the Commission at the 53 session, Apr. 23-June 1 and July 2-Aug. 10 2001, adopted by the United Nations General Assembly through Resolution A/RES/56/83 of Dec. 12, 2001; see also Vincent Boulanin, Netta Goussac, and Laura Bruun, "Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction, Stockholm International Peace Research Institute (June 2021).
86. International Law Commission, Draft Articles on Responsibility of State for Internationally Wrongful Acts, with commentaries (2001), para. 2. [Hereinafter ILC articles, with commentaries (2001)].
87. ILC Articles, with commentaries (2001), para. 3.
88. International Law Commission, Responsibility of States for Internationally Wrongful Acts, A/56/49, December 2001, Article 2. [Hereinafter ILC Articles (2001)].
89. ILC Articles, with commentaries (2001), Chapter 2, para. 2.
90. ILC Articles (2001), Article 4.
91. ILC Articles, with commentaries (2001), Article 2(5).

The second element of an internationally wrongful act, a breach of the state's international legal obligations, occurs when a state's actual conduct violates a primary rule of international law that is binding on it.[92] Because the law on state responsibility is intentionally written to be broad and applicable in many circumstances, the origin of the obligation matters not; for the purpose of this report, however, the obligations to consider are the obligations within IHL. State responsibility is a distinct framework from other frameworks discussed below and does not have the distinction between "civil" or "criminal" responsibility found in internal legal systems.

States that have committed internationally wrongful acts have two obligations: They must cease the conduct that is in breach of international legal obligations if it is continuing, and the state must make full reparation for the injury caused by the internationally wrongful act.[93]

Based on these requisite conditions of the law of state responsibility, there are three ways a state will assume responsibility for unlawful behavior by a LAWS.[94] First, when a state agent deploys LAWS and the machine violates IHL rules. This is the most straightforward example of the action being attributable to a state organ and constituting a breach of an international legal obligation. Second, state responsibility could be incurred if the "authorization, acquiescence, complicity or acknowledgement from state agents, a non-state actor deploys LAWS which violates protected rights."[95] In this circumstance, non-state actors could be groups or contractors working under or alongside state organs in which the state acknowledges responsibility for unlawful outcomes of non-state LAWS. Within these first two conditions, states are responsible for the unlawful conduct and have a duty to provide reparations to individuals or other entities for an internationally wrongful act committed by a machine.[96]

The third circumstance requires a deeper assessment. A state can incur responsibility where a private technology firm (or any private entity)

92. ILC Articles, with commentaries, Chapter 3, para. 3.
93. ILC Articles (2001), Articles 30 and 31, respectively.
94. This point draws from Thompson Chengeta, "Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law," *Denver Journal of International Law and Policy 45*, no. 1 (2016), particularly pp. 47-49.
95. ILC Articles (2001), 40-42.
96. ILC Articles (2001), Article 31.

contributes to the manufacturing of LAWS and that system, due to potentially low development standards, violates international legal obligations. In this case, both Articles 5 and 11 of the ILC report designate state responsibility.

> *Article 5: The conduct of a person or entity which is not an organ of the State...but which is empowered by the law of that State to exercise elements of that governmental authority shall be considered an act of the State under international law, provided the person or entity is acting in that capacity in the particular instance.*[97]

For the context of Article 5, it is important to reiterate that most AI innovation occurs in the civilian technology sector. The American development strategy to be "fast followers" of civilian innovation confirms the futility of trying to "stay ahead" of AI development from tech firms. At issue in Article 5 is whether the public-private partnership for LAWS development qualifies as "governmental authority."[98] Due to the integrated nature of the designers and developers of the parameters and ultimate behavior of the weapon, there is a strong case for exercising governmental authority in this case. In non-autonomous systems, the ultimate outcome or behavior resulting from the use of a weapon was allocated to a human soldier—or an agent of the state—meaning that the question of weapon manufacturers exercising governmental authority was less relevant. But autonomous systems, particularly machine learning systems, behave based on parameters and input implemented in early stages of development. This process is certainly informed, or even instructed, by state legal requirements, rules of engagement, or operational necessities that come from state organs. Nonetheless, non-state organs may possibly have a heavy hand in coding weapon parameters or a large role in weapon behavior. This would likely fall under a category

97. ILC Articles (2001), Article 5.
98. Article 9 of the ILC Report also deals with circumstances of non-state entities and "governmental authority," but ILC commentaries clarifies that Article 9 is for circumstances in which unlawful behavior by non-state entities occurs in times "such as revolution, armed conflict or occupation, where the regular authorities dissolve, are disintegrating, have been suppressed." These are circumstances in the absence of government. These circumstances are beyond the scope of this analysis, and Article 9 is therefore not included.

of government authority that would have been allotted to state organs with conventional weapons.

> *Article 11: Conduct which is not attributable to a State under the preceding articles shall nevertheless be considered an act of that State under international law if and to the extent that the State acknowledges and adopts the conduct in question as its own.*[99]

Article 11 provides state attribution to conduct that was not necessarily attributable to the state at the moment of commission. Essentially, if a private corporation developed LAWS that led to significant IHL violations, the state, through acquisition and procurement of that weapon system, would acknowledge and adopt the conduct as its own, effectively giving rise to state responsibility. As will be discussed in the section on corporate responsibility, private corporations must detail the risks inherent to a weapon system, and the state thus assumes responsibility for the use of the systems.

The preceding analysis suggests that any violation of IHL committed by LAWS will give rise to state responsibility. This is not controversial in the context of LAWS that are employed in armed conflicts (unlike other frameworks of responsibility discussed below), but the degree of the responsibility that the state must assume comes down to the quality of the LAWS that officials decide to procure.

There are two options for acquisition and procurement officials: whether to pursue OTS LAWS or to invest in tailored systems to meet a higher threshold of LAWS standards. On the one hand, procuring OTS LAWS may meet the general international legal standards for baseline interoperability with military partners and allies. However, OTS systems can also meet the minimum requirements to be IHL-compliant but nonetheless contain, *inter alia*, a high risk of unpredictability or a lack of transparency. Depending on the conditions of the armed conflict and circumstances of the location of hostilities, this will certainly increase the risk of state responsibility and incur significant legal and reputational costs. An acquisition strategy based on developing tailor-made systems will give states like Denmark greater control over the

99. ILC Articles (2001), Article 11.

TEVV process and assure a high threshold of machine behavior to mitigate these increased risks of state responsibility. While development projects are certainly costlier in the development stage, they produce more reliable and effective systems for operations.[100]

7.2. Individual Criminal Responsibility

Acquisition officials must additionally consider the risks for individual criminal responsibility for violations committed by LAWS as a framework with the potential to establish some kind of legal responsibility for their armed forces. Establishing individual criminal responsibility for war crimes requires a high degree of *mens rea* (discussed below). For acquisition officials, this framework is difficult to establish in the LAWS context, and the risk for individual criminal responsibility is much less than state responsibility.[101]

Individual criminal responsibility is a basic tenet of international criminal law (ICL) and overall requires the presence of *mens rea*[102] before an individual incurs criminal responsibility, as discussed below. A variety of relevant individuals can fall within this legal framework, and analyses of individual criminal responsibility usually refer to individual military operators and commanders. While command responsibility will be addressed separately below, this section specifically addresses the risks for individual military operators being held accountable for war crimes resulting from LAWS violating IHL.[103]

100. For more on state responsibility and the delineation of specific and general rules under international law, see Astrid Kjeldgaard-Petersen and Cornelius Wiesener, *State Responsibility for the Misconduct of Partners in International Military Operations: General and Specific Rules of International Law* (Copenhagen: Djøf Publishing, 2021).
101. See expanded analysis of individual criminal responsibility at Carrie McDougall, "Autonomous Weapon Systems and Accountability: Putting the Cart before the Horse," *Melbourne Journal of International Law 20* (2019); McFarland, *Autonomous Weapons Systems*, chapter 7; Swati Malik, "Autonomous Weapon Systems: The Possibility and Probability of Accountability," *Wisconsin International Law Journal* (2018).
102. *Mens Rea* refers to the mental element of a person's intention to commit a crime, rather than the conduct of the accused.
103. UN General Assembly, *Rome Statute of the International Criminal Court (last amended 2010)*, July 17, 1998, ISBN No. 92-9227-227-6 [*hereafter Rome Statute*]. It is possible to imagine LAWS committing the other international crimes included in the Rome Statute

The risk of LAWS behavior resulting in unlawful outcomes is of particular concern, as these unlawful outcomes can come in numerous and unpredictable ways. By way of example, one scenario that garners significant concern and attention is the risks in intentionally targeting the civilian population, a crime in both international armed conflicts (IAC) and non-international armed conflicts (NIAC). Under Article 8 of the Rome Statute of the International Criminal Court (the Rome Statute), deliberately attacking civilians is a war crime. To establish this crime, five elements must be satisfied: (1) the perpetrator directed an attack, (2) the object of the attack was a civilian population or individual civilians not taking direct part in hostilities, (3) the perpetrator intended the civilian population or individuals not taking direct part in hostilities to be the object of the attack, (4) the conduct took place in the context of and was associated with an international armed conflict/armed conflict not of an international character, (5) the perpetrator was aware of factual circumstances that established the existence of an armed conflict. Article 30 of the Rome Statute requires that, unless otherwise stated, the elements of war crimes must be "committed with intent and knowledge."[104] Article 30 (2)(b) further provides that if a person has intentions relating to a consequence, then "that person means to cause that consequence or is aware that it will occur in the ordinary course of events."[105]

It should be noted that ICL has not generally recognized recklessness or negligence as sufficient to establish criminal responsibility for most international crimes, aside from a handful of exceptions.[106] There are a small number of crimes that represent exceptions to Article 30 and

(genocide, aggression, and crimes against humanity). But the literature focuses heavily on the potential for LAWS to commit war crimes, so this section addresses these core concerns.

104. *Rome Statute art*. 30; see also McDougall, "Autonomous Weapon Systems," for expanded analysis on this subject.

105. *Rome Statute art*. 30(2)(b). McDougall, "Autonomous Weapon Systems," specifies the term "means to" in paragraph 2 is generally considered equal to direct intent, or *dolus directus*, in the first degree. Additionally, the "reference to an awareness that a consequence 'will occur in the ordinary course of events' is generally equated to oblique intent…in the second degree. Knowledge is defined in art 30(3) as meaning 'awareness that a circumstance exists or a consequence will occur in the ordinary course of events.' In *Prosecutor v Germain Katanga*, the International Criminal Court Trial Chamber held that the latter required 'virtual certainty.'" McDougall, "Autonomous Weapon Systems," 9.

106. McDougall, "Autonomous Weapon Systems;" Rebecca Crootof, "War Torts: Accountability for Autonomous Weapons," *University of Pennsylvania Law Review 164*, no. 6 (2016).

provide a lower mental elements threshold, one of which is individual criminal responsibility within command responsibility, which will be detailed below.[107]

Some scholars have identified certain scenarios in which individuals could still (relatively) clearly be held responsible for war crimes committed by LAWS, two of which are worth highlighting here.[108] First, a clear case where an individual could incur criminal responsibility is intentionally programming targeting parameters to violate international targeting obligations.[109] This scenario would likely, and most clearly, satisfy the mental elements and necessary intention to be considered a war crime. Second, a case in which a commander or senior official was to authorize deployment of a LAWS they knew, or owing to the circumstances should have known, to be unpredictable in complex conditions and could be virtually certain of unlawful conduct.[110] This scenario would necessarily require the commander or individual giving authorization to be aware of the circumstances and likelihood of legal violations by the LAWS (more on command responsibility below).

Clearly, the requisite mental elements for individual criminal responsibility would be difficult to establish for LAWS. The machine learning processes driving LAWS decision output provide neither obvious signals of intent nor even which intention should be prioritized. For example, the programmers and operators responsible for data input could both be factored into an arguable "intent" of a machine's decision-making output, but there is no clear or direct link from human intention to machine output, particularly with machine learning systems. Even for hand-coded systems operating under more restricted programming parameters, there could be instances where the LAWS makes decisions based on its own interpretations of the battlefield conditions that are not in line with

107. McDougall, "Autonomous Weapon Systems."
108. McDougall, "Autonomous Weapon Systems." This list is not meant to be exhaustive; rather, to identify that there are scenarios in which a human can clearly be found to have the necessary mental elements for committing a war crime using LAWS. These three are the most convincing.
109. McDougall, "Autonomous Weapon Systems."
110. McDougall, "Autonomous Weapon Systems;" Dan Saxon, *Drones and Responsibility: Legal, Philosophical and Socio-Technical Perspectives on Remotely Controlled Weapons*, eds. Ezio Di Nucci and Filippo Santoni De Sio (2016).

commander, or human, interpretation or intention. Furthermore, as discussed previously, even distinguishing this link of intention through the machine's system requires a high degree of AI transparency and explainability. A high degree of transparency will be vital to assessing and assigning accountability.

Critically, the potential for criminal liability for acquisition and procurement officials has received minimal attention. Procuring a LAWS that meets legal requirements and safety and security standards requires a different assessment than conventional weapons. The nature of autonomous systems will require procurement and acquisition officials to have a deep understanding of the technology and the multitude of risks inherent to LAWS deployment. As discussed, the TEVV processes and iterative training protocols can help mitigate such risks and minimize the likelihood of responsibility.

Nevertheless, it is useful to consider the implications of procuring a LAWS that results in unlawful behavior and the potential for acquisition or procurement officials to be responsible under aiding and abetting. The standards and jurisprudence for aiding and abetting have a long and varied history in international law but ultimately revolve around the responsibility of the actors who contribute or aid the commission of a crime without directly participating in it.[111] The Rome Statute establishes liability for aiding and abetting if the accused, "[f]or the purpose of the commission of such a crime, aids, abets or otherwise assists in its commission or its attempted commission, including providing the means for commission."[112] The ILC, with a more general interpretation, defined aiding and abetting in the 1996 Draft Code of Crimes against the Peace and Security of Mankind as, "[a]n individual shall be responsible for a crime...if that individual...knowingly aids, abets or otherwise assists, directly and substantially, in the commission of such a crime, including providing the means for its commission."[113]

111. See Oona Hathaway et al., "Aiding and Abetting in International Criminal Law," *Cornell Law Review 6*, no. 104 (2019).
112. Article 25(3), *Rome Statute*.
113. International Law Commission, Draft Code of Crimes against the Peace and Security of Mankind, art. 2(3)(d), Rep. of the International Law Commission on the Work of its Forty-Eighth Session, U.N. GAOR, 51st Session, Supp. No. 10, U.N. Doc. A/51/10 (1996). This is the most recent iteration of the Draft Code of Crimes against the Peace and Security of Mankind.

There are two elements required to establish criminal liability of aiding and abetting. The first is actus reus, or the act (or omission) that aids or abets the commission of the crime. This is the conduct itself of aiding or abetting. The second element, *mens rea*, is the specified state of mind in aiding or abetting the principle crime.

International courts and tribunals disagree about the *actus reus* standard, although the tribunals generally agree that the act of aiding is distinct from the act of abetting. The ICC delineates: "aiding implies the provision of practical or material assistance" to the perpetrator, whereas "the notion to abet describes the moral or psychological assistance...to the principal perpetrator."[114] To be liable of aiding or abetting, some of the international criminal tribunals require that the act of aiding or abetting has a "substantial effect" on the commission of the crime. The ICC, by contrast, sets a lower bar in that the aiding or abetting must simply have "an effect" on the principle crime.[115] The latter standard applies in Denmark.

Procuring a LAWS that violates IHL could result in liability for aiding for acquisition officials; although not likely abetting for this circumstance. Within the act of aiding is providing "material assistance," which could include providing a faulty or sub-standard autonomous system. Under the *actus reus* standard, the ICC threshold of having "an effect" on the commission of the crime could be satisfied by providing the weapon itself. As such, within this first element of aiding, procurement or acquisition officials have reason to carefully consider developing national standards and international standards for acquiring LAWS.

The second element, *mens rea,* is the mental state required for liability. Similar to the *actus reus* standard, there is fragmentation among the international courts and tribunals. The ad hoc tribunals and some hybrid tribunals have a relatively low threshold for m*ens rea*, simply that "knowledge that one's conduct assists the commission of the principal crime is sufficient to fulfill the requisite *mens rea* for aiding and abetting."[116] The ICC, by contrast, has a heightened standard that requires "purpose" to establish criminal liability.

114. The Prosecutor v. Jean-Pierre Bemba Gombo, Case No. ICC-01/05-01/13, paras. 88, 89 respectively.
115. Hathaway et al., "Aiding and Abetting."
116. Hathaway et al., "Aiding and Abetting," 1614.

Article 25(3)(c) states,

> *In accordance with this Statute, a person shall be criminally responsible and liable for punishment...if that person...for the purpose of facilitating the commission of such a crime, aids, abets, or otherwise assists in its commission or its attempted commission, including providing the means for its commission.*

Although the Rome Statute does not define "purpose" within this context, the ICC has clarified that it is a higher standard than "knowledge" as was used in the ad hoc tribunals. In *Prosecutor v. Bemba*, the ICC Trial Chamber determined that "'purpose' introduces a higher subjective mental element and means that the accessory must have lent his or her assistance with the aim of facilitating the offense. It is not sufficient that the accessory merely knows that his or her conduct will assist."[117]

As such, under the ICC threshold of a required "purpose," acquisition and procurement officials would need to purposefully incorporate or purchase a LAWS that does not satisfy legal and safety standards and perform unlawfully. Certain scenarios could satisfy the "knowledge" standard, such as procuring a system "virtually certain" to be flawed for the sake of speedy acquisition.[118] Technology firms interested in major defense contracts may attempt to incentivize procurement officials to move forward on purchasing systems, even though they still contain flaws.[119] Obviously there must be more diligence when it comes to weapon systems than more general AI systems, but there is still a risk of hasty acquisitions leading to performance problems and heightening the risks of criminal responsibility for procurement officials as aiding in the commission of crimes committed by a LAWS; although this would still not satisfy the "purpose" requirement, but would likely satisfy the "knowl-

117. Prosecutor v. Bemba, Case No. ICC-01/05-01/13, Trial Judgement Pursuant to Article 74 of the Statute, para. 97, Oct. 19, 2016.
118. Prosecutor v. Thomas Lubanga Dyilo, ICC-01/04-01/06 A 5, Appeals Chamber, 1 December 2014, para 6, "[T]he phrase 'a consequence will occur'...refers to future events in respect of which there is virtual certainty that they will occur."
119. There has been some evidence of bad business practices by technology firms for the sake of speedy contracts; see Pax, "Don't Be Evil? A Survey of the Tech Sector's Stance on Lethal Autonomous Weapons," https://paxforpeace.nl/media/download/pax-report-killer-robots-dont-be-evil.pdf.

edge" standard from the ad hoc tribunal jurisprudence. In order to establish "purpose" for criminal liability, procurement decision-makers would need to demonstrate a subjective desire for the commission of LAWS crimes. While this scenario is possible, it is unlikely.

7.2.1. Command Responsibility

A subset of individual criminal responsibility is the responsibility of the commander. It is vital to consider the framework of command responsibility, as some states, notably including the United States, have expressed their commitment to placing emphasis on the relationship between commanders and LAWS. The US has argued that, rather than simply prioritizing human control over machines, the key issue is instead "ensuring machines help effectuate the intention of commanders and the operators of the system."[120] From this perspective, LAWS would not require human supervision, but rather reflect commander intentions and broader objectives; as with subordinate human troops. This makes the commander-machine relationship all the more important to consider. While Denmark has not expressed a position in line with the US, it is useful to consider the implications for Danish commanders and the potential for rethinking the commander responsibility doctrine.

Command responsibility, or superior responsibility for high-ranking civilians, is not meant to punish commanders for directly participating in criminal behavior that can be shown to have planned, ordered, committed, or aided and abetted crimes undertaken by others.[121] Rather, the doctrine under international law refers to a form of liability for the omission of crimes taken by subordinates; or the failure to prevent or punish crimes.

There are three elements necessary to establish command responsibility.

1. The existence of a superior-subordinate relationship between the defendant-superior and the perpetrators of the underlying offense.

120. Karl Chang, U.S. Mission to International Organization in Geneva, Consideration of the Human Element in the Use of Lethal Force, Address Before the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the Area of LAWS (March 26, 2019).
121. Guénaël Mettraux, *The Law of Command Responsibility* (Oxford: Oxford University Press, 2009).

2. The superior (commander) knew or had reason to know that a subordinate was about to commit such acts or had done so.
3. The superior failed to take necessary and reasonable measures to prevent such acts or to punish perpetrators thereof.[122]

The first element requires a superior-subordinate relationship, whether it be *de jure*, meaning that the commander's authority comes as a position with the purpose of commanding or leading subordinates (e.g., being appointed, elected, or otherwise assigned to an authoritative position); or a de facto relationship, in which a commander exercises authority based on inter-personal relationships or other factual or personal factors.[123] There are two requirements within this relationship. The first is a chain of command or hierarchical relationship, whether direct or indirect. This requirement does not demand a direct command over subordinates, but rather the commander "by virtue of his position, [must be] senior in some sort of formal or informal hierarchy to the perpetrator."[124] The second requirement is effective control, meaning that the commander must have "the material ability to prevent offences or punish the principal offenders."[125] Effective control is the minimal threshold necessary to establish commander liability for subordinate offences.[126]

In the context of LAWS, we can understand LAWS to be a "subordinate" of the commander, as the machine will be programmed with the commander's intentions and objectives. While a machine does not undergo a training program and process with the commander as with human troops, the LAWS must reflect the strategic objectives from the commander and conduct the mission in compliance with international law and rules of engagement—just as with humans. The effective control requirement is more difficult to establish for three reasons. Recall that effective control doctrine requires the power to take necessary steps to prevent and punish crimes committed by subordinates. Firstly, it is unlikely that a commander, even observing in real-time, would have the

122. Mettraux, *Command Responsibility*, 129.
123. For a deeper discussion of these types of superior-subordinate relationships, see Mettraux, *Command Responsibility*, 138-44.
124. *Halilovic*, ICTY Appeal Judgement, par. 59. 16.10.2007.
125. Mettraux, *Command Responsibility*.
126. See Mettraux, *Command Responsibility*, 156-90, for a deeper examination.

speed necessary to direct machine behavior. Secondly, the commander would likely require a certain degree of technical expertise. It would be necessary to read the computer code and make adjustments as necessary for safety. This could also be done with contracted software programmers or engineers, but doing so can slow the process. Thirdly, there must be a high degree of machine predictability for effective commander control. This is less of an issue for hand-coded systems, but machine learning systems can formulate their own decisions in response to changing environmental conditions. It is unlikely to establish effective control over a machine learning system, which may perpetuate a liability loophole.[127]

The second element is a mental requirement, which the Rome Statute Article 28(a)(i) defines as, "a military commander or a person either knew or, owing to the circumstances at the time, should have known that the forces were committing or about to commit such crimes."[128] This "should have known" standard places a requirement for commanders or superiors to be informed of the risks associated with an operation. The issues for LAWS are related to the first element in that commanders may not necessarily have insight into machine learning systems and how LAWS will choose to adapt to the environment; especially as commanders are unlikely to program the machine themselves.[129] Similar to the ICC "should have known" standard, the International Criminal Tribunal for the Former Yugoslavia (ICTY) held that commanders can be liable if they had information to put them "on notice of the risk" that

127. Some scholars have argued this point. If commanders cannot exercise effective control over machine learning systems but nevertheless authorize the use of force, it may create a legal loophole regarding liability for autonomous weapons. Some scholars highlight this issue and instead propose a standard of "meaningful control" to ensure liability is possible—but this is not currently a legal standard. See Human Rights Watch, "Mind the Gap: The Lack of Accountability for Killer Robots," (2015); Rebecca Crootof, "War Torts: Accountability for Autonomous Weapons," *University of Pennsylvania Law Review* (2016); Peter Asaro, "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-making," *International Review of the Red Cross*, 687 (2012).
128. This standard has evolved over time. The culpability of state of mind under international customary law needed to demonstrate that commanders or superiors had actual knowledge of subordinate crimes. However the ICC statute expands this requirement to commanders "should have known;" see Mettraux, *Command Responsibility*, 193-226.
129. Daniel Hammond, "Autonomous Weapons and the Problem of State Accountability," *Chicago Journal of International Law* 15, no. 2 (2015).

is "sufficiently alarming to justify further inquiry."[130] It is unclear what will constitute a notice of risk for a LAWS, but the language could be a useful framework for incorporating a standard of risk for commanders to evaluate appropriateness of use for a particular system.

The third element requires commanders to take necessary and reasonable measures to prevent or punish subordinates committing offenses, such as through pre-deployment testing or other safety measures. But if a commander, in addition to other elements, covers up unlawful outcomes and keeps the LAWS in circulation despite risks of errors or compromised safety and security standards, they could violate this requirement.

Despite the challenges facing command responsibility, some scholars consider command responsibility as the most appropriate framework to remedy the LAWS accountability gap, even so far that "solving the [L]AWS accountability problem hinges on the doctrine of command responsibility."[131] With human soldiers, commanders have multiple measures to control behavior in operations, such as issuing rules of engagement, applying temporal or geographic limitations for operations, designating protected areas from attack, or raising the level of authority required to authorize attacks with high collateral damage concerns.[132]

While these measures still apply as options when programming a LAWS, the threshold for command responsibility in this context is difficult to establish because the doctrinal elements do not easily translate to autonomous weapons. For example, the doctrine of effective control requires reconsideration. Some scholars and military lawyers have suggested that a standard of "meaningful human control" is necessary to adapt the command responsibility doctrine for LAWS and to fill the broader accountability gap. One approach proposes that meaningful human control would "require LAWS to be designed to allow commanders to apply controls to the overall use of the weapon that are necessary and reasonable to prevent IHL violations."[133] The necessary and reasonable controls

130. Prosecutor v. Strugar, Case No. IT-01-42-A, Appeals Chamber Judgement paras. 297-89; quoted in Crootof, "War Torts."
131. Margulies, "Autonomous Weapons," 406.
132. Matthew T. Miller, "Command Responsibility: A Model for Defining Meaningful Human Control," *Journal of National Security Law and Policy 533* (2021).
133. Miller, "Command Responsibility."

would include the following: (1) a certain technical understanding of what the LAWS platform was designed to do and what the testing record shows as consistent and reliable behavior; (2) determining where and when LAWS will operate and relying on human-machine teaming, where battlefield conditions are too risky for LAWS judgement alone (i.e. high risk of violating distinction and proportionality principles)[134] or (3) adequate supervision of subordinates tasked with maintaining and programming LAWS. However, these control measures contain certain assumptions about commander training and technical expertise as well as immediate access to information to effectively determine temporal or geographic deployment of LAWS.

Another approach to command responsibility and accountability for LAWS with a higher threshold of human-machine engagement is the approach experts call commander "dynamic diligence."[135] This framework contains three requirements. First, it includes continual adjustments to the human-machine interface, which is performed within the command structure and includes experts familiar with the risks and benefits of LAWS. Second, it requires frequent assessments of LAWS performance and compliance with IHL; essentially, consistent validation processes. Third, there must be flexibility in the parameters ordering LAWS output. The parameters could include limits to time, distance, or maximum expected collateral damage.[136]

Procurement and acquisition managers are in a unique position regarding the risks for command responsibility. Ensuring that any procured LAWS satisfy strict standards through vigorous testing will place the commanders employing these systems in a more secure and confident position; not to mention the state more broadly. Some have urged that "military leaders and those responsible for procuring and fielding weapons must also recognize the inherent risk associated with pursuing weapon systems."[137] It is imperative for acquisition and procurement officials to reconceive IHL compliance and the risk of failure in the devel-

134. Miller, "Command Responsibility."
135. Margulies, "Autonomous Weapons."
136. Margulies, "Autonomous Weapons," 437.
137. Geoffrey S. Corn, "Autonomous Weapons Systems: Managing the Inevitability of 'Taking the Man out of the Loop'" in Nehal Bhuta et al. (eds) *Autonomous Weapon Systems: Law, Ethics, Policy*, Cambridge University Press (2016), 219.

opment phase, and "the inputs of military procurement managers, weapons developers and legal advisors must be fully engaged in the weapons development process to ensure commander employing such a weapon system may do so with genuine confidence."[138]

The recommendations outlined in this report aim to offer steps to the Danish MoD toward mitigating risks inherent to LAWS. As is clear at this point, the inherent capability of the technology warrants a reconsideration of the development and TEVV processes, which includes a deep coordination with legal expertise involved in weapons development. These steps will have the subsequent effect toward ensuring that the safest (and legally compliant) systems are available for the Danish forces and that responsibility will not erroneously fall on commanders.

7.3. Hidden Costs for Tech Firms—Strategic Litigation as Barrier to Acquisition

Collaboration and cooperation with civilian defense contractors, and especially the technology industry, is crucial for acquiring and maintaining a technological edge for LAWS.[139] The AI talent and resources in the private sector are responsible for driving design and innovation for solving the important challenges facing the development of military AI. The Danish MoD has recognized the necessity of collaboration and adopted civilian collaboration for technology innovation as an initiative in the August 2021 Danish Government Strategy for Defense Industry.[140] Particularly useful for the future acquisition of LAWS is the Danish approach to drone technology acquisition. The Danish "triple helix" approach combines the research, industry, and state sectors to streamline drone technology acquisition and innovate on current Danish acquisition processes. This model will be a useful starting point for future AI acquisition. As the 2021 Defense Strategy for Industry illustrates,

138. Corn, "Autonomous Weapons Systems," 224.
139. For a particularly insightful analysis of civilian-military cooperation and innovation, see Verbruggen, "Civilian Innovation."
140. Regeringens Strategi for Dansk Forsvarsindustri: Styrket Samarbejde for Dansk Sikkerhed, August 2021, https://fmn.dk/globalassets/fmn/dokumenter/nyheder/2021/-regeringens-strategi-for-dansk-forsvarsindustri-dk-.pdf.

these innovative approaches to acquisition are useful for maintaining operational ability and improving processes to be speedy and flexible for changing technology. But collaboration with civilian industries on LAWS will need to consider the risks of strategic litigation for industry partners and the associated risks for defense collaboration. As this section will demonstrate, the Danish defense industry may not have to worry about criminal liability if a system engages in an unlawful performance, but the risks of strategic litigation may be a barrier, especially for smaller firms, to collaborate with the Danish MoD or other military partners.

Strategic litigation is a tool that organizations can use to bolster awareness of a particular cause, often globally, and motivate other organizations or individuals to influence governmental change. It would involve a campaign of targeted litigation within the defense industry with the prospect of high media attention and the goal of bolstering public awareness. Strategic litigation aims to raise awareness of corporate participation in either weapon system development or arms transfers, which can often result in reputational damage as a result of the campaign.[141]

Such high-profile cases can spread awareness of the risks resulting from deploying LAWS to foster a skeptical or critical public opinion. A public that disapproves of deploying such a weapon system can incentivize politicians to act accordingly and impose limitations or restrictions to quell public concern. This is the goal of strategic litigation; and in this process, the reputations of companies involved in weapons development can be significantly affected by financial repercussions. Importantly, a company does not need to produce low-performance systems to be a target of strategic litigation—it is not about LAWS performance—simply participating in a contract to produce LAWS for the Danish Armed Forces (or any other military) may be enough to be involved in a strategic litigation campaign.

Another reason why strategic litigation is a likely outcome is that criminal responsibility for the design or development of LAWS is unlikely. Civil liability will be more likely than previous weapons manufacturers as part of the strategic litigation campaign. But criminal respon-

141. Weapons development or arms transfers are just examples, but most frequent for strategic litigation campaigns.

sibility is unlikely to occur based on previous attempts to hold weapons manufacturers criminally liable.

There is a history of exploring criminal liability for weapons manufactures, albeit unsuccessfully. Nevertheless, some international legal instruments recognize the criminal responsibility of corporations; for example, the European Convention against Terrorism and the UN Convention against Transnational Organized Crime both recognize the criminal, civil, and administrative liability of corporations.[142] And certainly, there is a significant legal framework for corporate human rights responsibility; however, this is outside the scope of this report.

Defense contractors, specifically weapons manufacturers, are in a unique category due to the legal nature of armed conflicts. Outside of armed conflicts, a number of weapons companies, particularly gun manufacturers, have been sued for the production and release of weapons to the greater population.[143] But defense contractors and manufacturers are largely exempt from civil and criminal liability when it comes to weapons used in an armed conflict. In the US case *Boyle v. United Technologies*, for example, the plaintiff alleged a wrongful death as a result of a defective design in a military-supplied aircraft emergency escape system. But the US Supreme Court determined procurement for military equipment is a "uniquely federal interest" wherein liability under state tort law is displaced if certain conditions are met: (1) the US approved reasonably precise specifications of the procured equipment, (2) the equipment complied with the specifications, (3) the manufacturer communicated any dangers to the United States that was known to the manufacturer but not the United States.[144] Similarly, in another US case involving a suit against a defense contractor, the US court held that defense manufacturers do not have a duty of care once the system is acquired by the US

142. Council of Europe Convention on the Prevention of Terrorism, art. 10, May 16, 2005, 16.v.2005 No. 196; General Assembly Resolution 55/25, art. 10, United Nations Convention against Transnational Organized Crime (Nov. 15, 2000).
143. Most recently, Mexico filed a lawsuit on August 4, 2021, against ten gun companies for flooding the market with military-style weapons that are particularly favored by drug cartels, https://www.nytimes.com/2021/08/04/world/americas/mexico-lawsuit-gun-companies.html.
144. *Boyle v. United Technologies Corporation*, 487 U.S. 500 (1988) para 501.

Armed Forces and cannot be held accountable for the use of weapons against enemy forces.[145]

But it is worth acknowledging the fundamental role that the manufacturers and designers of LAWS play in the machine's outcome, which merits the consideration of manufacturer risk of liability. LAWS developers play a significant role in determining the behavior parameters, or the range of action, that the LAWS can perform as well as determining the behavior of the LAWS after deployment.[146]

While the same legal questions have not been raised in Danish courts, there are some important considerations beyond criminal liability relevant for the Danish MoD in how this responsibility may affect the Danish defense industry. For example, Denmark's largest defense contractor, Terma (among other defense corporations), has been targeted by human rights groups for contributing equipment to a coalition in Yemen accused of committing possible war crimes.[147] The legal issues posed to Terma lie beyond the scope of this report, but it is nevertheless useful to highlight the potential litigation challenges that the Danish defense industry can face in contributing and collaborating with the Danish MoD or partners for LAWS design and development.

LAWS are extremely controversial, particularly within the human rights community that opposes machine capacity for lethal decision making. Some organizations have already started working to raise awareness regarding civilian defense contractors and technology firms involved in autonomous weapons development.[148] Currently, only a hand-

145. *Koohi v. United States* 841 F.2d 1328 (1992) para 1337. The court ruled that "during wartime encounters no duty of reasonable care is owed to those against whom force is directed as a result of authorized military force...neither the United States nor its defense contractors owed any duty to such individuals [enemy combatants]."

146. Tim McFarland and Tim McCormack, "Mind the Gap: Can Developers of Autonomous Weapon Systems Be Liable for War Crimes?" 90 *International Law Studies* 361 (2014).

147. https://danwatch.dk/en/undersoegelse/denmarks-largest-defense-company-contributes-to-possible-war-crimes-in-yemen/; see Amnesty International's blog post for larger analysis accusing the UK of violating export laws, https://www.amnesty.org/en/latest/press-release/2015/12/uk-government-breaking-the-law-supplying-arms-to-saudi-arabia/.

148. For more on this, see "Don't Be Evil? A Survey of Tech Sector's Stance on Lethal Autonomous Weapons," https://paxforpeace.nl/media/download/pax-report-killer-robots-dont-be-evil.pdf.

ful of global technology firms have expressly prohibited their technical products from being used for LAWS.[149]

The degree of reputational risk that strategic litigation will have depends on the size and nature of the firm in question. If Denmark decides to procure OTS autonomous weapon systems, they will likely be from well-established firms that have the financial and reputational capital to withstand strategic litigation campaigns. However, developing systems that work with the technology sector to acquire systems that are at the cutting edge of AI innovation, the risk of strategic litigation may be a barrier to this collaboration. Much of the technology sector develops dual-use technology—that is, AI that has utility in civilian and military sectors—and financial gains are much higher on the civilian side of that development.[150] The gains of military collaboration may not necessarily outweigh the reputational cost that technology firms will want to avoid. Smaller tech companies, or start-ups, that are at the forefront of innovation with high AI talent may be discouraged from accepting defense contractors for these reasons.

This applies to the Danish defense industry. With some exceptions, the Danish defense industry and technology sector may not possess the resources necessary to withstand a strategic litigation campaign and instead prioritize civilian applications of Danish AI development. Denmark has an impressive technology sector with important AI and machine learning innovation; certainly an industry that can be utilized to create responsible and reliable LAWS. However, acquisition officials will need to be aware of the litigation risks for participating companies in order to address and preempt those risks.

In short, corporate criminal liability may not be a major concern when procuring LAWS, but acquisition stakeholders will need to consider the reputational risks in the Danish approach to developing or procuring LAWS.

149. Michael T. Klare, "Few Tech Firms Limit Autonomous Weapons," *Arms Control* (September 2019), https://www.armscontrol.org/act/2019-09/news/few-tech-firms-limit-autonomous-weapons.

150. See Catherine Aiken et al., "'Cool Projects' or 'Expanding the Efficiency of the Murderous American War Machine: AI Professionals' Views on Working with the Department of Defense," CSET *Issue Brief* (November 2020). This report includes a comprehensive survey of AI professionals and the concerns from the technology industry of partnering with the military to develop military AI.

7.4. Conclusions

This section has explored four responsibility frameworks from multiple frameworks to account for the actors involved at multiple stages of the weapon system lifecycle. The first consideration for Danish acquisition officials should be state responsibility, as it is a very comprehensive framework and the least controversial in the deployment of LAWS. In order to reduce the high likelihood of state responsibility from machine performance, Denmark has incentives to strongly consider testing and training procedures.

The second framework is individual criminal responsibility, and this section detailed the difficulties in establishing the mental elements necessary to satisfy war crime requirements. It is not impossible for war crimes to be committed in the employ of LAWS, and this section outlines possible scenarios where the mental elements are satisfied. Nonetheless, the circumstances that satisfy the elements of a war crime are limited in the context of LAWS deployment.

Within individual criminal responsibility is the subset of command responsibility, which is particularly important for Danish acquisition officials to consider because some experts contend this framework is the most appropriate for maintaining criminal liability for unlawful behavior of LAWS. This section discusses potential frameworks to appropriately apply command responsibility in the deployment of LAWS in theater.

Finally, this section demonstrates that corporate criminal liability is unlikely. Due to the core role played by AI designers and programmers, it is necessary to explore the risks of criminal liability, but there is little basis for this to occur. Instead, the civilian defense industry is at risk of strategic litigation, which could function as a barrier to Danish procurement. The 2021 National Defence Industrial Strategy of the Danish Government offered many useful starting points for deeper collaboration with civilian industry, and the Danish "triple helix" approach to drone technology may be a useful starting point for future AI acquisition. Nevertheless, the controversy over autonomous systems makes strategic litigation a concern for dual-use companies and may deter companies from collaborating on military development or manufacturing contracts. Danish acquisition officials will need to consider the reputational stakes involved for collaboration partners, especially for smaller firms, which may not have the interest or resources to withstand a strategic litigation campaign.

8

Recommendations

This report presents two novel IHL legal issues relating to the procurement and acquisition of lethal autonomous weapon systems (LAWS). First, the nature of LAWS technology impedes IHL compliance; specifically, AI transparency, predictability, and bias. Second, this report addresses the well-known problem of the "accountability gap" in LAWS performance. This report assesses the responsibility and criminal liability for the range of actors involved in the weapon system lifecycle.

Because the landscape of artificial intelligence and the legal parameters for military applications are changing rapidly, these recommendations can guide Danish decision-makers in future deliberations and planning.

1. **Formulate a policy.** Formulating a national policy regarding a Danish interpretation of LAWS can guide military decision-makers and legal advisers, as AI continues to be important for discussions of future warfighting capabilities. Clarity in policy is all the more relevant as Denmark participates in international coalitions (e.g., the AI Partnership for Defense), but has largely remained silent on the LAWS issue. Danish allies (e.g., US, Australia, and France) have already announced R&D programs, and in some cases an accompanying ethical framework, toward responsible AI weapon systems development. Other Danish allies (e.g., Germany) have instead publicly opposed such weapons development. Without a stated policy, Denmark risks falling behind critical security partners on future dialogues regarding LAWS.
2. **Encourage inter-agency coordination**. Strengthened cooperation and coordination between the Danish Acquisition and Logistics

Office (DALO) and the legal office of the Ministry of Defense can ensure greater IHL compliance. The legal challenges of incorporating LAWS into the Armed Forces will become critical to the Danish warfighting capacity, and early stages of AI design should incorporate IHL. This will ensure that Denmark remains a competitive military ally with responsible, lawful autonomous systems.

3. **Restructure TEVV procurement**. Restructuring the Danish Testing, Evaluation, Validation, and Verification (TEVV) can help Denmark mitigate the inherent challenges to autonomous weapon systems—namely algorithmic transparency, predictability, and bias. One option toward this end is to establish an iterative TEVV process offering accurate AI systems to Danish defense and legal stakeholders and ensuring that IHL standards are front-and-center of autonomous weapon development.
4. **Encourage joint acquisition/collaboration of LAWS through project development**. Project-based acquisition offers the greatest control and flexibility over LAWS design and development relative to purchasing LAWS off the shelf. In order to ensure the maximum IHL compliance and sufficiently trained algorithms, Denmark may want to consider avenues for acquiring LAWS through project-based development. This report offers two starting points. First, Denmark can pursue joint acquisition participation in multi-national defense partnerships (e.g., NATO or the AI Partnership for Defense). Second, the Danish MoD can create formal partnerships with Danish academic and industry experts at the cutting edge of AI design and development.
5. **Mitigate strategic litigation risks**. Global opposition to LAWS development increases the risks of defense and technology partners experiencing strategic litigation and reputational backlash for defense collaboration. The MoD can temper this issue through an actionable collaboration plan. Future LAWS/AI acquisition may want to mirror the Danish "triple helix" collaboration strategy for drone acquisition—which combines research, industry, and the state to explore innovative acquisition potential—from the 2021 Danish National Defence Industrial Strategy. However, civilian collaboration will need to consider the risks of strategic litigation as a barrier for industry, especially smaller technology firms, with the potential to hinder an innovative Danish acquisition approach.

Bibliography

Aiken, Catherine, Rebecca Kagan, and Michael Page. "'Cool Projects' or 'Expanding the Efficiency of the Murderous American War Machine: AI Professionals' Views on Working with the Department of Defense." *Center for Security and Emerging Technology* Issue Brief, (November 2020.)

Amnesty International. "UK Government Breaking the Law Supplying Arms to Saudi Arabia, Say Leading Lawyers." 2015. https://www.amnesty.org/en/latest/press-release/2015/12/uk-government-breaking-the-law-supplying-arms-to-saudi-arabia/.

Anton, Philip S., Brynn Tannehill, Jake McKeon, Benjamin Goirigolzarri, Maynard A. Holliday, Mark A. Lorell, and Obaid Younossi, Strategies for Acquisition Agility: Approaches for Speeding Delivery of Defense Capabilities. Santa Monica, CA: RAND Corporation, 2020. https://www.rand.org/pubs/research_reports/RR4193.html.

Asaro, Peter. "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making." *International Review of the Red Cross* 687 (2012).

Ballais, Renaud, and Renelle Guichard. "Defense Innovation, Technology Transfers and Public Policy." *Defence and Peace Economics* 17, no. 3 (2006).

Bendett, Samuel. "Should the U.S. Army Fear Russia's Killer Robots?" *The National Interest*, (November 2017.)

Bialek, William, Ilya Nemenman, and Naftali Tishby. "Predictability, Complexity, and Learning." *Neural Computation* 2409 (2001).

Boulanin, Vincent, Netta Goussac, and Laura Bruun. "Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction." *Stockholm International Peace Research Institute*, June 2021.

Brooks, Risa. "Technology and Future War Will Test US Civil-Military Relations." *War on the Rocks* (November 26, 2018). https://warontherocks.com/2018/11/technology-and-future-war-will-test-u-s-civil-military-relations/.

Chang, Karl. "U.S. Mission to International Organization in Geneva, Consideration of the Human Element in the Use of Lethal Force." Address Before the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the Area of LAWS, March 26, 2019.

Chengeta, Thompson. "Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law." *Denver Journal of International Law and Policy* 45, no. 1 (2016).

Cordeschi, Roberto. "Automatic Decision-Making and Reliability in Robotic Systems: Some Implications in the Case of Robot Weapons." *AI & Society* 28 (2013).

Corn, Geoffrey S. "Autonomous Weapons Systems: Managing the Inevitability of 'Taking the Man out of the Loop.'" In *Autonomous Weapon Systems: Law, Ethics, Policy*, edited by Nehal Bhuta et al. Cambridge: Cambridge University Press, (2016).

Crootof, Rebecca. "Killer Robots Are Here: Legal and Policy Implications." *Cardozo Law Review* 36 (2014): 1837.

Crootof, Rebecca. "War Torts: Accountability for Autonomous Weapons." *University of Pennsylvania Law Review* 164, no. 6 (2016).

Danwatch. "Denmark's Largest Defense Company Contributes to Possible War Crimes in Yemen." 2020. accessed 2022 https://danwatch.dk/en/undersoegelse/denmarks-largest-defense-company-contributes-to-possible-war-crimes-in-yemen/.

Defense Science Board. "The Role of Autonomy in DoD Systems." *Task Force Report*, US Department of Defense, July 2012.

Deloitte. "Transparency and Responsibility in Artificial Intelligence: A Call for Explainable AI." 2019. https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/innovatie/deloitte-nl-innovation-bringing-transparency-and-ethics-into-ai.pdf.

Department of Defense Memorandum 5000.02. "Getting Defense Acquisition Right." January 2015. https://dod.defense.gov/Portals/1/Documents/pubs/Getting-Acquisition-Right-Jan2017.pdf.

Department of Defense. "Better Buying Power: Acquisition, Technology, Logistics." https://www.ustranscom.mil/dbw/docs/BBP_Fact_Sheet.pdf.

Department of Defense. "Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity." 2018. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

Department of the Air Force and Massachusetts Institute of Technology. "Artificial Intelligence Acquisition Guidebook." February 2022. https://aia.mit.edu/wp-content/uploads/2022/02/AI-Acquisition-Guidebook_CAO-14-Feb-2022.pdf.

Flournoy, Michèle, Gabrielle Chefitz, and Avril Haines. "Building Trust through Testing: Adapting DOD's Test & Evaluation, Validation & Verification Enterprise for Machine Learning Systems, including Deep Learning." October 2020. https://cset.georgetown.edu/wp-content/uploads/Building-Trust-Through-Testing.pdf.

Freedberg, Sydney J. Jr. "Google Helps Chinese Military, Why Not US?" *Breaking Defense*. 2018. accessed 2022 https://breakingdefense.com/2018/06/google-helps-chinese-military-why-not-us-bob-work/.

Gray, Bernard. "Review of Acquisition for the Secretary of State for Defence." October 2009. https://delta.bipsolutions.com/docstore/ReviewAcquisitionGrayreport.pdf.

Hammond, Daniel. "Autonomous Weapons and the Problem of State Accountability." *Chicago Journal of International Law* 15, No. 2 (2015).

Hao, Karen. "This Is How AI Bias Really Happens: And Why It's So Hard to Fix." MIT *Technology Review*, February 4, 2019. https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/.

Hathaway, Oona et al. "Aiding and Abetting in International Criminal Law." *Cornell Law Review* 6, No. 104 (2019).

Horowitz, Michael C. "Artificial Intelligence, International Competition, and the Balance of Power." *Texas National Security Review* 1, no. 3 (2018).

Horowitz, Michael C. "When Speed Kills: Lethal Autonomous Weapon Systems, Deterrence and Stability." *Journal of Strategic Studies* 42, no. 6 (2019).

Horowitz, Michael, Gregory C. Allen, Elsa B. Kania, and Paul Scharre. "Strategic Competition in an Era of Artificial Intelligence." *Center for New American Security*, (2018).

Human Rights Watch & International Human Rights Clinic, Harvard Law School "Advancing the Debate on Killer Robots: 12 Key Arguments for a Preemptive Ban on Fully Autonomous Weapons," May 2014. https://www.hrw.org/sites/default/files/related_material/Advancing%20the%20Debate_8May2014_Final.pdf.

Human Rights Watch. "Losing Humanity: the Case against Killer Robots," 2012. https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots.

Human Rights Watch. "Mind the Gap: The Lack of Accountability for Killer Robots," 2015. accessed 2022, https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots.

Jensen, Eric Talbot. "Autonomy and Precautions in the Law of Armed Conflict." *International Law Studies*, 96 (2020).

Kania, Elsa. "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power." *Center for New American Security*, (November 2017).

Kjeldgaard-Petersen, Astrid, and Cornelius Wiesener. "State Responsibility for the Misconduct of Partners in International Military Operations: General and Specific Rules of International Law." Copenhagen: Djøf Publishing, (2021).

Klare, Michael T. "Few Tech Firms Limit Autonomous Weapons." *Arms Control*, (September 2019). https://www.armscontrol.org/act/2019-09/news/few-tech-firms-limit-autonomous-weapons.

Knight, Will. "Forget Killer Robots: Bias Is the Real AI Danger." *MIT Technology Review*, (October 3, 2017.)

Knight, Will. "Military Artificial Intelligence Can Be Easily and Dangerously Fooled," *MIT Technology Review*, (2019).

Konaev, Margarita, Tina Huang, and Husanjot Chahal. "Trusted Partners: Human-Machine Teaming and the Future of Military AI." *Center for Security and Emerging Technology*, Issue Brief, (February 2021.)

Schuller, Alan L. "Artificial Intelligence Effecting Human Decisions to Kill: The Challenge of Linking Numerically Quantifiable Goals to IHL Compliance." *Journal of Law and Policy for the Information Society* 15, nos. 1-2 (2019): 105.

Lee, Kai-Fu. *AI Superpowers: China, Silicon Valley, and the New World Order*. Boston: Houghton Mifflin Harcourt, 2018.

Lehr, David, and Paul Ohm. "Playing with the Data: What Legal Scholars Should Learn about Machine Learning." *UC Davis Law Review* 51 (2017): 671.

Malik, Swati. "Autonomous Weapon Systems: The Possibility and Probability of Accountability." *Wisconsin International Law Journal*, (2018).

Margulies, Peter. "Making Autonomous Weapons Accountable: Command Responsibility for Computer-Guided Lethal Force in Armed Conflicts." In *Research Handbook on Remote Warfare*, edited by Jens David Ohlin. Northampton: Edward Elgar, (2017).

Maurer, Peter. "ICRC Statement." *International Committee of the Red Cross*, May 12, 2021. accessed 2022 https://www.icrc.org/en/document/peter-maurer-role-autonomous-weapons-armed-conflict.

McDougall, Carrie. "Autonomous Weapon Systems and Accountability: Putting the Cart before the Horse." *Melbourne Journal of International Law* 20 (2019).

McFarland, Tim. *Autonomous Weapons Systems and the Law of Armed Conflict: Compatibility with International Humanitarian Law*. Cambridge: Cambridge University Press, 2020.

McFarland, Tim, and Tim McCormack. "Mind the Gap: Can Developers of Autonomous Weapon Systems be Liable for War Crimes?" *International Law Studies* 90, no. 1 (2014).

Mettraux, Guénaël. *The Law of Command Responsibility*. Oxford: Oxford University Press, 2009.

Milaninia, Nema. "Biases in Machine Learning Models and Big Data Analytics: The International Criminal and Humanitarian Law Implications." *International Review of the Red Cross* 102 (913) (2020).

Miller, Matthew T. "Command Responsibility: A Model for Defining Meaningful Human Control." *Journal of National Security Law and Policy*, 533 (2021).

National Security Commission on Artificial Intelligence (NSCAI), February 2021.

Open AI Blog. "Multimodal Neurons in Artificial Neural Networks." 2021. accessed: 2022 https://openai.com/blog/multimodal-neurons/.

Pax. "Don't Be Evil? A Survey of the Tech Sector's Stance on Lethal Autonomous Weapons." (August 2019). https://paxforpeace.nl/media/download/pax-report-killer-robots-dont-be-evil.pdf.

Regeringens Strategi for Dansk Forsvarsindustri: Styrket Samarbejde for Dansk Sikkerhed. August 2021. https://fmn.dk/globalassets/fmn/dokumenter/nyheder/2021/-regeringens-strategi-for-dansk-forsvarsindustri-dk-.pdf.

Rudin, Cynthia, and Joanna Radin. "Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson from an Explainable AI Competition." *Harvard Data Science Review* 1, no. 2 (2019).

Ryseff, James. "How to (Actually) Recruit Talent for the AI Challenge." *War on the Rocks* (February 2020).

Saxon, Dan. "Autonomous Drones and Individual Criminal Responsibility." In *Drones and Responsibility: Legal, Philosophical and Socio-Technical Perspectives on Remotely Controlled Weapons*, edited by Ezio Di Nucci and Filippo Santoni De Sio. Milton Park: Routlege, 2016.

Schaub, Gary Jr., and Jens Wenzel Kristoffersen. "In, on, or out of the Loop? Denmark and Autonomous Weapon Systems." CMS Report, Centre for Military Studies, February 2017.

Schmitt, Michael N. "Autonomous Weapon Systems and International Humanitarian Law: A Reply to Critics." *Harvard National Security Journal* 4 (2013): 1-37.

Sharkey, Noel E. "The Inevitability of Autonomous Robot Warfare." *International Red Cross Review* 94, no. 886 (2012): 787-99.

Simonite, Tom. "Machines Taught by Photos Learn a Sexist View of Women." *Wired*, August 21, 2017. https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/.

Sprenger, Sebastian. "NATO Tees up Negotiations on Artificial Intelligence in Weapons." *C4ISRNET*. Accessed March 28, 2022. https://www.c4isrnet.com/artificial-intelligence/2021/04/27/nato-tees-up-negotiations-on-artificial-intelligence-in-weapons/.

Stanley-Lockman, Zoe, and Lena Trabucco. "NATO's Role in Responsible AI Governance in Military Affairs." Forthcoming in Oxford Handbook on AI Governance. Available at SSRN: https://ssrn.com/abstract=3939769 (August 2021).

Taylor, Trevor. "Artificial Intelligence in Defence: When AI Meets Defence Acquisition Processes and Behaviours." *RUSI Journal* 164, Nos 5/6 (2019).

Thompson, Loren. "Five Reasons Why Silicon Valley Won't Partner with the Pentagon." *Forbes*, April 2015. https://www.forbes.com/sites/lorenthompson/2015/04/27/five-reasons-why-silicon-valley-wont-partner-with-the-pentagon/?sh=37e3ca8f4de9.

Thurnher, Jeffrey S. "Feasible Precautions in Attack and Autonomous Weapons." In *Dehumanization of Warfare: Legal Implications of New Weapon Technologies*, edited by Wolff Heintschel von Heinegg, Robert Frau, and Tassilo Singer, 99-117. Cham: Springer, 2018.

Trapp, Kimberley. "A Framework of Analysis for Assessing Compliance of LAWS with IHL (API) Precautionary Measures." *Convention on Certain Conventional Weapons (CCW) Informal Meeting of Experts*, 2016. https://discovery.ucl.ac.uk/id/eprint/1513936/1/Trapp_CCW%20Informal%20Meeting%20of%20Experts%20(2016).pdf.

U.S. Department of Defense. "Test and Evaluation Management Guide 220," 6th ed., 2012.

United Nations Institute for Disarmament Research. "Algorithmic Bias and the Weaponization of Increasingly Autonomous Technologies: A Primer," no. 9 (2018). https://unidir.org/publication/algorithmic-bias-and-weaponization-increasingly-autonomous-technologies.

United States Department of DefenseDirective 3000.09, November 21, 2012. Incorporating change May 8, 2017. https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

Van Den Boogaard, Jeroen. "Proportionality and Autonomous Weapons Systems." *Journal of International Humanitarian Legal Studies* 6, no. 2 (2015): 247-83.

Verbruggen, Maaike. "The Role of Civilian Innovation in the Development of Lethal Autonomous Weapon Systems." *Global Policy* 10 No. 2 (2019).

Vestner, Tobias, and Altea Rossi. "Legal Reviews of War Algorithms." *International Legal Studies* 97, no. 1 (2021): 509.

Wong, Jonathan P. "Bad Idea: Overly Focusing on Development and Acquisition Speed." RAND Commentary, December 2020. https://www.rand.org/blog/2020/12/bad-idea-overly-focusing-on-development-and-acquisition.html.

Yde, Iben. "Autonome våbensystemer i danske våbenscreeninger—Nye udfordringer og krav til implementeringen af den folkeretlige våbenscreeningsforpligtelse." Copenhagen: Djøf Publishing, (2021).

ABOUT THE AUTHORS

Lena Trabucco, PhD, is a Postdoc at the Centre for Military Studies and Visiting Researcher at the U.S. Naval War College and Research Affiliate at Cambridge University. Lena researches the intersection of international law and security of emerging technologies.

9 788757 455267