GCSP
Geneva Centre for
Security Policy

MARCH 2024

The Law of Targeting's Mechanisation and Objectivisation through the Use of Artificial Intelligence

Tobias Vestner

Forthcoming in Irene Couzigou and Edouard Fromageau (eds), *Technological Change and International Law: Testing the Adaptability of International Law*, Edward Elgar Publishing

Geneva Centre for Security Policy

The Geneva Centre for Security Policy (GCSP) is an international foundation that aims to advance global cooperation, security and peace. The foundation is supported by the Swiss government and governed by 54 member states. The GCSP provides a unique 360° approach to learn about and solve global challenges. The foundation's mission is to educate leaders, facilitate dialogue, advise through in-house research, inspire new ideas and connect experts to develop sustainable solutions to build a more peaceful future.

About the Author

Dr Tobias Vestner is the Director of the Research and Policy Advice Department and the Head of the Security and Law Programme at the GCSP. He is an Honorary Senior Research Fellow at the University of Exeter, a Fellow at Supreme Headquarters Allied Powers Europe, and a Non-Resident Fellow at the United Nations Institute for Disarmament Research. He serves as reserve Legal Advisor at the Swiss Armed Forces Staff.

Acknowledgement

Special thanks go to Nicolò Borgesano for excellent research assistance and Dr Stuart Casey-Maslen, Federico Mantellassi, and Alex Potter for most valuable comments on earlier versions.

Geneva Centre for Security Policy
Maison de la paix
Chemin Eugène-Rigot 2D
P.O. Box 1295
1211 Geneva 1
Switzerland
Tel: + 41 22 730 96 00
E-mail: info@gcsp.ch
www.gcsp.ch

For citation: Tobias Vestner, 'The Law of Targeting's Mechanisation and Objectivisation through the Use of Artificial Intelligence', in Irene Couzigou and Edouard Fromageau (eds), *Technological Change and international Law: Testing the Adaptability of International Law*, (Edward Elgar Publishing, forthcoming).

Abstract

Armed forces are increasingly introducing artificial intelligence (AI) for targeting purposes. This leads to the question of what implications the use of AI technologies will have for the law of targeting under international humanitarian law. This chapter argues that the use of AI in military operations leads to a 'mechanisation' and 'objectivisation' of the law of targeting. It analyses the relatively indeterminate elements of the law of targeting's principles and rules of distinction, proportionality in attack, and precautions in attack and defence. It also contrasts them with recent technical developments related to AI by using examples of current technologies. Thereby, it identifies and demonstrates which and how elements of the law of targeting are likely to become more objective through the use of AI. It concludes that the law is subject to an evolutionary process driven by this emerging technology.

Key words: artificial intelligence, emerging technology, international humanitarian law, law of armed conflict, law of targeting, evolution of international law

Contents

Introduction

Armed forces are increasingly introducing artificial intelligence (AI) for targeting purposes. This leads to the question of what implications the use of AI technologies will have for the law of targeting, namely the rules under international humanitarian law (IHL) that govern the process of selecting and attacking objects and persons in warfare.[1] While current debates and scholarly analyses tend to focus on how IHL applies to the use of AI systems,[2] this chapter takes the opposite approach by assessing how the use of AI technologies influences the law. It thus examines the adaptability of IHL to military AI applications and finds that IHL is evolving based on this technological development.

At the outset, the law of targeting tends to be relatively indeterminate. Many provisions governing targeting are open-textured and highly context-dependent. This includes the widely accepted notion that targeting decisions must comply with the relatively subjective standard of a 'reasonable commander'. Accordingly, commanders, operators, and soldiers have a significant margin of appreciation regarding the application of the law of targeting.[3] In contrast, the use of AI in targeting requires precise guidance and delimitations. The IHL rules need to be incorporated or represented as technical parameters for guiding the systems to properly analyse data and reach conclusions.[4] This requires coding and converting qualitative judgements into quantitative ones.[5] The use of AI thus demands clearer specifications and parameters than those of the current law of targeting.

This chapter examines how the use of AI in military operations affects the law of targeting based on the premise of this tension between the law of targeting's relative indeterminacy and AI's need for clear parameters. The chapter first reviews armed forces' use of AI for targeting purposes and describes how the introduction of AI generally leads to a 'mechanisation' of the targeting process. It then develops how the military use of AI plausibly leads to an 'objectivisation' of the law of targeting.[6]

The chapter then analyses the relatively indeterminate elements of the law of targeting's principles and rules of distinction, proportionality in attack, and precautions in attack and defence, contrasting them with recent AI-related technical developments by using examples of current technologies. Thereby, it identifies and demonstrates which and how elements of the law of targeting

1 See, eg, Stuart Casey-Maslen and Steven Haines, *Hague Law Interpreted: The Conduct of Hostilities under the Law of Armed Conflict* (Hart Publishing 2018); William H Boothby, *The Law of Targeting* (OUP 2012); Ian Henderson, *The Contemporary Law of Targeting: Military Objectives, Proportionality and Precautions in Attack under Additional Protocol I* (Martinus Nijhoff 2009).

2 See notably Afonso Seixas-Nunes, *The Legality and Accountability of Autonomous Weapon Systems: A Humanitarian Law Perspective* (CUP 2022) ch 3; Elliot Winter, 'The Compatibility of Autonomous Weapons with the Principles of International Humanitarian Law' (2022) 27(1) Journal of Conflict and Security Law 1; 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' (SIPRI 2021); Tim McFarland, *Autonomous Weapon Systems and the Law of Armed Conflict: Compatibility with International Humanitarian Law* (CUP 2020).

3 For a general assessment, see Casey-Maslen and Haines (n 1); Sigrid Redse Johansen, *The Military Commander's Necessity: The Law of Armed Conflict and Its Limits* (CUP 2019); Michael N Schmitt and Michael Schauss, 'Uncertainty in the Law of Targeting: Towards a Cognitive Framework' (2019) 10 Harvard National Security Journal 148.

4 Boulanin, Goussac, and Bruun (n 2) 19.

5 Ashley Deeks, 'Coding the Law of Armed Conflict: First Steps' (2020) Virginia Public Law and Legal Theory Research Paper No. 2020-49 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3612329> accessed 16 October 2023.

6 Sassòli had first made the point that the use of AI systems may lead to a clarification of IHL, yet without specifying which elements of the law of targeting would be affected and how this would materialise. See Marco Sassòli, 'Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified' (2014) 90 International Law Studies 308, 339.

are likely to become more objective by the use of AI. The chapter concludes by discussing broader consequences of this evolution of IHL based on recent technological developments.

1. Using AI in targeting

Targeting is the process of selecting and prioritising targets as well as matching appropriate responses and effects according to operational requirements, capabilities, or limitations.[7] Targeting operations require the assessment of targets, weapons, execution modalities, possible collateral damage and incidental injury, and location evaluations.[8] It encompasses several processes, including the translation of direction at the strategic and operational levels into guidance at the tactical level.[9] Targeting is 'planned' if directed against targets that are known to exist in the operational environment. Targeting is 'dynamic' if directed against defined 'targets of opportunity' not precedingly selected for action but meeting the relevant criteria to achieve operational objectives.[10] US joint doctrine and NATO, for instance, identify six phases in targeting.[11]

Armed forces increasingly use AI - systems that perform tasks which normally require human intelligence, such as recognising patterns, learning from experience, drawing conclusions, making predictions, and taking action[12] - in the context of targeting.[13] Military AI applications can serve reconnaissance and analytical purposes,[14] support target identification and selection,[15] and support commanders' decision-making regarding targeting decisions. Highly autonomous AI systems may even select and engage targets by themselves under minimal human supervision.[16]

---

7 US Joint Chiefs of Staff, *Joint Publication 3-0* (2018) GL–17; Australian Defence Force, *Australian Defence Doctrine Publication 3.14* (2nd edn, 2009) 1–1; see similarly NATO, *Allied Joint Publication 3.9* (edn B, 2021) 1–1; see also UK Chiefs of Staff, *Joint Doctrine Publication 3-00* (3rd edn, 2009) 3–12; Nicholas Tsagourias, 'Targeting in International Humanitarian Law' (Oxford Bibliographies, 28 November 2016) <https://www.oxfordbibliographies.com/display/document/obo-9780199796953/obo-9780199796953-0142.xml> accessed 16 October 2023.

8 Michael N Schmitt and Eric W Widmar, '"On Target": Precision and Balance in the Contemporary Law of Targeting' (2014) 7(3) Journal of National Security Law and Policy 379, 380.

9 NATO (n 7) 1–13.

10 US Joint Chiefs of Staff, *Joint Publication 3-60* (2013) II–1 ff; see similarly ibid 1–11-12.

11 Phase 1 is concerned with the commander's objectives and intent. Phase 2 leads to target development, vetting, validation, and list management and prioritisation. Phase 3 requires the evaluation of available capabilities (such as sensors and weapons systems) against desired effects, and the assessment of the most appropriate options available. Phase 4 follows with the decision and force assignment. Phase 5 governs the planning and force execution. In Phase 6, it is assessed whether the desired effects are created, if objectives are achieved, and what next steps are required. See US Joint Chiefs of Staff (n 10) II–6 ff; NATO (n 7) 1–14-21; see also Ilse Verdiesen, Filippo Santoni de Sio, and Virginia Dignum, 'Accountability and Control Over Autonomous Weapon Systems: A Framework for Comprehensive Human Oversight' (2020) 31(1) Minds and Machines 137, 154.

12 Defense Science Board, 'Summer Study on Autonomy' (2016); see similarly, House of Lords Select Committee on Artificial Intelligence, 'AI in the UK: Ready, Willing and Able?' (2018) 14, para 10 <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf> accessed 16 October 2023; see also the table of definitions found in Stuart J Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (3rd edn, Pearson 2010) s 1.1.

13 See Forrest E Morgan and others, 'Military Applications of Artificial Intelligence Ethical Concerns in an Uncertain World' (RAND 2020) s 'Findings', xii ff; see also the current trends in Anthony King, 'AI at War' (*War on the Rocks*, 27 April 2023) <https://warontherocks.com/2023/04/ai-at-war/> accessed 16 October 2023.

14 See Maggie Gray and Amy Ertan, 'Artificial Intelligence and Autonomy in the Military: An Overview of NATO Member States' Strategies and Deployment' (CCDCOE 2021) s 7.3 'Data Analytics'; Karel van den Bosch and Adelbert Bronkhorst, 'Human-AI Cooperation to Benefit Military Decision Making' (NATO 2018) s 3 -1-4-5.

15 See Morgan and others (n 13) 17; Anastasia Roberts and Adrian Venables, 'The Role of Artificial Intelligence in Kinetic Targeting from the Perspective of International Humanitarian Law' (CCDCOE 2021) 48; ICRC, 'Artificial intelligence and machine learning in armed conflict: A human-centred approach' (2019) 3, and s 3.3; see also generally Brendan Cook, 'The Future of Artificial Intelligence in ISR Operations' (2021) 35(SE) Air & Space Power Journal 41.

16 Autonomy is understood as the ability of machines to perform tasks without human intervention through the interaction of sensors and computer programming with the environment. See 'Key Findings' in Vincent Boulanin and Maaike Verbruggen, 'Mapping

Current examples of military AI applications include Project Maven, which combines AI, deep learning, and computer vision to detect, classify, and track objects within full motion video (FMV) images;[17] the Athena AI Defence target classifier;[18] the Pleora project that enables the detection, identification, and classification of tanks through sensor data and machine learning capabilities;[19] the Advanced Targeting and Lethality Aided System (ATLAS) AI targeting system;[20] and the Targeting Long-range Identification Optronic System (TALIOS) system attached to the French Rafale to analyse sensor and image data for automatic target detection and recognition.[21] Autonomy is also a feature of certain military systems, such as loitering munitions and drones.[22]

The use of AI and autonomy is, however, still confronted with technical and operational challenges. Major issues lie with the unpredictability and understandability of certain systems due to the complexity of the operating environment or the underlying systems' architecture,[23] such as deep learning neural networks,[24] self-learning,[25] continuous or continual learning,[26] or reinforcement learning[27] approaches. Another major challenge is that human operators and commanders may over-rely on information and recommendations provided by AI systems, which is known as automation bias.[28] Moreover, states' definitions of the appropriate levels of human control and judgement over autonomous AI systems remain vague and diverse.[29]

---

the Development of Autonomy in Weapon Systems' (SIPRI 2017) vii; for a broader discussion on definitions of autonomy in weapon systems see Seixas-Nunes (n 2) ch 3; for definitions adopted by states and international organisations, see also generally Mariarosaria Taddeo and Alexander Blanchard, 'A Comparative Analysis of the Definitions of Autonomous Weapons Systems' (2022) 28(5) Science and Engineering Ethics 36.

17 US Defense Technical Information Center, 'Exhibit R-2A, RDT&E Project Justification: PB 2019 Office of the Secretary of Defense' (2018) 3-4.

18 'Athena AI' (*Athena Defence,* 2021) <https://athenadefence.ai> accessed 16 October 2023.

19 'Detect Tanks Using AI' (*Pleora Technologies*) <https://www.c4isr.ai/insights/detect-tanks-using-ai/> accessed 16 October 2023.

20 Rojoef Manuel, 'US Army Showcases AI-Based Target Recognition Aboard M1 Abrams Tank' (*The Defence Post*, 17 February 2023) <https://www.thedefensepost.com/2023/02/17/us-target-recognition-abrams-demonstration/> accessed 16 October 2023; Oliver Parken, 'M1 Abrams Tank Tested With Artificial Intelligence Targeting System' (The Drive, 14 February 2023) <https://www.thedrive.com/the-war-zone/m1-abrams-tank-tested-with-artificial-intelligence-targeting-system> accessed 16 October 2023.

21 Gray and Ertan (n 14) 20.

22 See the relevant examples in Daan Kayser, 'Increasing autonomy in weapons systems: 10 examples that can inform thinking' (Automated Decision Research and PAX 2021); see also Boulanin and Verbruggen (n 16).

23 Arthur Holland Michel, 'The Black Box, Unlocked: Predictability and Understandability in Military AI' (UNIDIR 2020) 6-7, 9; ICRC (n 15) 10-11.

24 See 'What are neural networks?' (IBM) <https://www.ibm.com/topics/neural-networks> accessed 16 October 2023; on deep learning, see Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning (The MIT Press 2016) 8.

25 'Capacity for self-learning' in Michel (n 23) 7; see also Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, 'Chairperson's Summary' (2021) CCW/GGE.1/2020/WP.7 <https://documents.unoda.org/wp-content/uploads/2020/07/CCW_GGE1_2020_WP_7-ADVANCE.pdf> accessed 16 October 2023; see, eg, Timothée Lesort and others, 'Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges' (2020) 58 Information Fusion 52, s 6.2.4. 'Learning algorithms stability', 63-64.

26 David Fernández Llorca and others, 'Liability Regimes in the Age of AI: a Use-Case Driven Analysis of the Burden of Proof' (2023) 76 Journal of Artificial Intelligence Research 613, 623; see also continual learning in Lesort and others (n 25) 5.

27 Richard S Sutton and Andrew G Barto, *Reinforcement Learning: An Introduction* (2nd edn, The MIT Press 2018) 1-2.

28 Mary L Cummings, 'Creating Moral Buffers in Weapon Control Interface Design' (2004) 23(3) IEEE Technology and Society Magazine 28, 41; Mary L Cummings, 'Automation and Accountability in Decision Support System Interface Design' (2006) 32(1) Journal of Technology Studies 23; Peter M Asaro, 'Modeling the Moral User' (2009) 28(1) IEEE Technology and Society Magazine 20, 22; see also generally Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt, 'Automation bias: a systematic review of frequency, effect mediators, and mitigators' (2012) 19(1) Journal of the American Medical Informatics Association 121.

29 See, eg, the explanation of 'semi-autonomous' weapon systems found in US Department of Defense, *DoD Directive 3000.9: Autonomy in Weapon Systems* (2023) 23. There are dozens of states that submitted definitions of (meaningful) human control at the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS) in 2023. See, eg, 'Draft Protocol on Autonomous Weapon Systems (Protocol VI)' (*Reaching Critical Will,* 10 May 2023) <https://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/documents/WP6.pdf> accessed 16 October 2023. Note also the absence of a proper definition within the conclusions reached in the final report at the 2023 GGE on LAWS. See GGE on LAWS, 'Report of the 2023 session' (2023) CCW/GGE.1/2023/2 4 <https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_2_Advance_version.pdf> accessed 16 October 2023.

2. AI mechanises the targeting process

The result of the military use of AI is that the battlefield and its engagement by armed forces are increasingly processed by, involve, and depend on data and algorithms. This is in line with the NATO Science & Technology Organization's finding that the fourth industrial revolution has led to more intelligent, interconnected, decentralised, and digital (I2D2) technologies. These may drive specific military capability development trends, notably intelligent autonomous action and precision warfare.[30] The consequence is a certain degree of 'mechanisation' of the targeting process, meaning that processes and procedures become more mechanical in character.

The use of AI in targeting leads to a further digitalisation of the targeting process, involving increased reliance on computer systems and algorithms. AI also heavily relies on data, which can be defined as 'a representation of facts, concepts or instructions in a manner suitable for communication, interpretation, or processing by humans or by automatic means'.[31] This is notably the case for machine learning,[32] where data are the basis for training the systems to perform desired tasks,[33] and which requires that the data be thoroughly collected, cleaned, validated, and integrated.[34]

Making reality understandable and manageable by algorithms thus requires breaking it down and representing it in the form of data and coding. The same applies to military operational contexts, missions, tasks, and instructions.[35] Thus, the introduction and use of military AI rely on technical configurations and requirements. NATO, for instance, is working on certain certification protocols and standards.[36] The US has also developed a strategy for the implementation of responsible AI and is establishing proper data management for strategic aims.[37] The UK is undertaking a similar path by developing national strategies on the ambitious, safe, and responsible use of military AI.[38]

The mechanisation of the targeting process does not mean that the human element (including the human control of autonomous systems) is, *per se*, lessened by AI. Current practices and rationales regarding targeting can be programmed into

30 Dale F Reding and others, 'Science & Technology Trends 2023-2043: Across the Physical, Biological, and Information Domains' (NATO, Science & Technology Organization, 2023) 10-14.

31 IEEE, 'IEEE Standard Glossary of Software Engineering Terminology' (1990).

32 Machine learning is known as the field of study giving computers 'the ability to learn without being explicitly programmed'. Mariette Awad and Rahul Khanna, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers* (Apress 2015) 1.

33 This is opposed to 'rule-based AI systems'. See Paul Scharre, *Four Battlegrounds: Power in the Age of Artificial Intelligence* (W W Norton & Company 2023), Preface and ch 2.

34 See, eg, Steven Euijong Whang and others, 'Data collection and quality challenges in deep learning: a data-centric AI perspective' (2023) 32(4) The VLDB Journal 791, 792; Lina Zhou and others, 'Machine Learning on Big Data: Opportunities and Challenges' (2017) 237 Neurocomputing 350, s 3 'Data preprocessing opportunities and challenges'.

35 See, eg, Tobias Vestner, 'From Strategy to Orders: Preparing and Conducting Military Operations with Artificial Intelligence', in Robin Geiss and Henning Lahmann (eds), *Research Handbook on Warfare and Artificial Intelligence* (Edward Elgar Publishing 2024).

36 NATO starts work on Artificial Intelligence certification standard' (*NATO*, 7 February 2023) <https://www.nato.int/cps/en/natohq/news_211498.htm> accessed 16 October 2023.

37 US Department of Defense, 'Responsible Artificial Intelligence Strategy and Implementation Pathway' (2022); see also US Department of Defense, 'DOD Adopts Ethical Principles for Artificial Intelligence' (*US Department of Defense*, 24 February 2020) <https://www.defense.gov/News/Releases/release/article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/> accessed 16 October 2023; US Department of Defense, 'DoD Data Strategy' (2020).

38 UK Ministry of Defence, 'Defence Artificial Intelligence Strategy' (2022); UK Ministry of Defence, 'Ambitious, safe, responsible: our approach to the delivery of AI-enabled capability in Defence' (2022).

AI-enabled systems when tasks are delegated to them. Yet, even when existing practices and rationales are maintained, the very fact of programming them into the digital sphere makes the resulting AI-enabled processes mechanical. This mechanisation of the targeting process affects how the law of targeting is implemented.

3. AI objectivises the law of targeting

The use of AI systems for targeting purposes needs to be done without undermining states' compliance with international law, in particular IHL. While states have officially recognised the applicability of IHL to the use of AI systems in the framework of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems,[39] significant debate remains regarding how the use of AI systems can be in line with IHL and its law of targeting. In this context, Human Rights Watch and others have argued that the targeting rules under IHL could not be quantified and therefore human operators would always need to make targeting decisions in line with current targeting practices.[40]

The increased mechanisation of the targeting process through the use of AI does generally affect the implementation of the law of targeting, however. This is because for AI systems to operate in this legal framework, its elements need to be specified, quantified, and programmed into the systems' functioning. Such coding of IHL-compliant algorithms requires the coding and conversion of qualitative elements of the law into quantitative ones.[41] To the extent that qualitative judgements are managed by, or delegated to, AI systems, this leads to a general shift in how the targeting rules are applied. Notably the current standard for making targeting decisions, namely that of a 'reasonable military commander'[42] who acts in 'good faith'[43] based upon available information, will

39 Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, 'Report of the 2019 session' (2019) CCW/GGE.1/2019/3 Annex IV, Guiding Principle (a), 13 <https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_2_Advance_version.pdf> accessed 16 October 2023.

40 'Losing Humanity' (Human Rights Watch 2012) 33; Bill Boothby, 'How Far Will the Law Allow Unmanned Targeting to Go?' in Dan Saxon (ed), *International Humanitarian Law and the Changing Technology of War* (Martinus Nijhoff Publishers 2013) 57.

41 Deeks (n 5).

42 See Ian Henderson and Kate Reece, 'Proportionality under International Humanitarian Law: The "Reasonable Military Commander" Standard and Reverberating Effects' (2018) 51(3) Vanderbilt Journal of Transnational Law 835, 841 and 845; Yves Sandoz, Christophe Swinarski, and Bruno Zimmerman (eds), *Commentary to the Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, 8 June 1977* (Martinus Nijhoff Publishers 1987) [AP I Commentary] para 1931, cited by Geoffrey S Corn, 'Humanitarian Regulation of Hostiles: The Decisive Element of Context' (2018) 51(3) Vanderbilt Journal of Transnational Law 763, 769; Laurie R Blank, 'New Technologies and the Interplay between Certainty and Reasonableness' (2017) Emory Legal Studies Research Paper No. 19, 4 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3065745> accessed 16 October 2023; see, eg, US Department of Defense, *Law of War Manual* (2023) ss 5.4.3.2, 5.4.3 (assessment), and 5.3.2 (decision); see Canada National Defence, *Law of Armed Conflict at the Operational and Tactical Levels* (2001) 4–4; Jeroen van den Boogaard, *Proportionality in International Humanitarian Law: Refocusing the Balance in Practice* (CUP 2023) ch 8; Jason D Wright, '"Excessive" ambiguity: analysing and refining the proportionality standard' (2012) 94(886) International Review of the Red Cross 819, 846; and Henderson (n 1) 222-223; Amichai Cohen and David Zlotogorski, *Proportionality in International Humanitarian Law: Consequences, Precautions, and Procedures* (OUP 2021) 102-103.

43 Johansen (n 3) 77-78; see, eg, US Department of Defense (n 42) s 1.10.1.1; for a general analysis of 'good faith', see Robert Kolb, *Good Faith in International Law* (Hart Publishing 2017); see, eg, with regard to the principle of distinction and feasibility of precautionary measures, John Merriam, 'Affirmative Target Identification: Operationalizing the Principle of Distinction for U.S. Warfighters' (2016) 56(1) Virginia Journal of International Law 83, 123, citing the LOAC manuals on subjective conditions, and ibid 123-125; US Department of Defense (n 42) 202; Cohen and Zlotogorski (n 42) 101, also citing Yoram Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict* (4th edn, CUP 2022) 132-133; Michael Bothe, Karl Josef Partsch, and Waldemar A Solf, *New Rules for Victims of Armed Conflicts Commentary on the Two 1977 Protocols Additional to the Geneva Conventions of 1949* (Martinus Nijhoff 2013) 253; see also US Department of Defense (n) s 2.2.3.3; Corn (n 42) 770, citing Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (8 June 1977) 1125 UNTS 3 [AP I] art 57; see also Robert D Sloane, 'Puzzles of Proportion and the "Reasonable Military Commander": Reflections on the Law, Ethics, and Geopolitics of Proportionality' (2015) 6 Harvard National Security Journal 299; Boothby (n 1) 172.

likely shift to a standard of statistical 'certainty' according to pre-defined levels of probabilities.[44]

Such a shift towards a more mechanised application of the law of targeting leads to its 'objectivisation'. As the law of targeting tends to be relatively indeterminate, which gives states a certain latitude and discretion, its specification and quantification for the proper use of AI systems forces states to clarify meticulously how they apply the law of targeting to specific situations. It is possible that this clarification would only apply to AI systems, which could lead to a new *sui generis* regime in the law of targeting based on algorithmic parameters that is applicable to AI systems only, which has been termed the 'special computer law on targeting'.[45] More plausible, however, is that the further mechanisation of the targeting process through AI influences the overall manner by, and the basis on which, even humans take decisions. The more AI applications and tools penetrate the targeting process, the more commanders and operators will rely on, and become influenced by, the mechanisation of the targeting process and the objectivisation of the law of targeting.

4. Objectivising the principle of distinction

The mechanisation and objectivisation of the law of targeting through the use of AI directly concerns the principle of distinction. Applicable at all echelons[46] and consisting of informational, decisional, and executional components,[47] the principle requires that parties to a conflict distinguish between combatants and civilians as well as between military objectives and civilian objects. It is legal to directly target military objectives[48] and combatants[49] as long as they are not wounded, sick, shipwrecked,[50] or *hors de combat*.[51] Civilians, unless they directly participate in hostilities or are members of an armed group,[52] and medical personnel and chaplains cannot be legally targeted.[53]

44 Blank (n 42).

45 Masahiro Kurosaki, 'Toward the Special Computer Law of Targeting: "Fully Autonomous" Weapons Systems and the Proportionality Test' in Claus Kreß and Robert Lawless (eds), *Necessity and Proportionality in International Peace and Security Law* (OUP 2020) 429.

46 See also Dinstein (n 43) 119.

47 Merriam (n 43) 102, 106.

48 AP I (n 43) arts 43, 48, 51; 'Customary IHL' (*ICRC*) Rules 7, 8, 9 <https://ihl-databases.icrc.org/en/customary-ihl> accessed 17 October 2023; see also International Law Association Study Group on the Conduct of Hostilities in the 21st Century, 'The Conduct of Hostilities and International Humanitarian Law: Challenges of 21st Century Warfare' (2017) 93 International Law Studies 322, 333-334.

49 AP I (n 43) art 52; see also the relevant rules as reflected in customary international humanitarian law in 'Customary IHL' (n 48) ch 1.

50 Marco Sassòli, *International Humanitarian Law* (Edward Elgar Publishing 2019) para 8.07; Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field (12 August 1949) 75 UNTS 31 [GC I] art 12; Convention (II) for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea (12 August 1949) 75 UNTS 85 [GC II] art 12; AP I (n 43) art 10; see also the practice related to Rule 110 and 111 in 'Customary IHL' (n 48).

51 AP I (n 43) art 41; 'Customary IHL' (n 48) Rule 47.

52 AP I (n 43) art 48; 'Customary IHL' (n 48) Rule 1; 'The Rule of Distinction in Attack: Persons' in Casey-Maslen and Haines (n 1); AP I (n 43) art 43(1) and 50(1); see also 'Customary IHL' (n 48) Rule 5; see also the concept of civilian IACs and NIACs as considered in Nils Melzer, 'Interpretative Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law' (ICRC 2009) Part 1.

53 AP I (n 43) art 43(2); 'Customary IHL' (n 48) Rules 25 and 27; see Laurent Gisel, 'The protection of medical personnel under the Additional Protocols: the notion of "acts harmful to the enemy" and debates on incidental harm to military medical personnel' (International Institute of Humanitarian Law 2017) 2-4; R Scott Adams, 'Lancelot in the Sky: Protecting Wounded Combatants from Incidental Harm', Harvard National Security Journal, 8 August 2017.

4.1 Contribution to military action and military advantage

The principle of distinction contains several relatively indeterminate notions that states need to specify when using AI in targeting. The determination of military objectives is largely based on objective criteria, namely the nature, location, purpose, or use of the given object, but also relies on relative vague elements. Article 52(2) of the AP I notably states that military objectives need to 'make an *effective contribution to military action* and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offer a *definite military advantage* [emphasis added]'. This makes the definition of a military objective relatively abstract.[54] Moreover, the criteria of nature, location, purpose, or use are context-dependent but can also be redundant if an object is evaluated as making an effective military contribution.[55]

The notion of 'definite military advantage' is also not defined by treaty law, although it is considered as excluding purely political advantages and cannot be potential or unspecified.[56] In addition, there should be a high likelihood that the military advantage being sought will be attained.[57] It has been argued that 'definite' roughly corresponds to 'concrete', as in 'perceptible to the senses', or 'specific' – and therefore neither general, nor hypothetical nor speculative.[58] While the concept encompasses a wide range of tactical gains and military considerations,[59] an expert manual on international law applicable to air and missile warfare suggests a relatively broad notion of 'any consequence of an attack which directly enhances friendly military operations or hinders those of the enemy'.[60]

With the use of AI programmes in the context of targeting, military objectives' contributions to military action and the military advantages of planned attacks are likely to be given quantitative values. AI may notably be used to calculate, simulate, and predict the military consequences of attacks, including the anticipation of the reduction of an adversary's military materiel and capacity, or by mapping future proper forces' advancement in a given territory based on the attack. AI systems that assess and attribute the military significance of enemy threats and targets for military objects in the air and on the ground already exist.[61]

Other existing AI applications, such as computer vision,[62] also lead to an objectivisation of information related to lawful targets. Computer vision supports

54 Dinstein (n 43) 118.

55 Henderson (n 1) 53-54; see also Seixas-Nunes (n 2) 175.

56 Dinstein (n 43) 5; AP I Commentary (n 42) para 2024.

57 International Law Association Study Group on the Conduct of Hostilities in the 21st Century (n 48) 365, citing the AP I Commentary (n 42) para 2019.

58 Bothe, Partsch, and Solf (n 43) 407.

59 Van den Boogaard (n 42) 18.

60 Ibid 17, fn 59.

61 See, eg, Ruojing Zhao and others, 'Dynamic Air Target Threat Assessment Based on Interval-Valued Intuitionistic Fuzzy Sets, Game Theory, and Evidential Reasoning Methodology' (2021) 2021 Mathematical Problems in Engineering 1; Hoyeop Lee and others, 'Threat Evaluation of Enemy Air Fighters Via Neural Network-based Markov Chain Modeling' (2017) 116 Knowledge-Based Systems 49; see also the Aegis combat system, which is able to independently identify, target, prioritise, and engage a large number of surface and airborne threats. Michael Mayer, 'The New Killer Drones: Understanding the Strategic Implications of Next-Generation Unmanned Aerial Combat Vehicles' (2015) 91(4) International Affairs 765, 772.

62 See, eg, Congressional Research Service, 'Artificial Intelligence and National Security' (2018); Larry Lewis and Andrew Ilachinski, 'Leveraging AI to Mitigate Civilian Harm' (CNA 2022) 29-30; James Johnson, *Artificial intelligence and the future of warfare: The USA, China, and strategic stability* (Manchester University Press 2021) 20-21.

sensors to identify targets by deriving information from digital images, videos, and other visual inputs.[63] A system may determine and locate objects in a picture or a video frame (a so-called 'object detection' task) and detect the type of object (a so-called 'image segmentation' task).[64] Computer vision can be combined with other AI techniques (such as deep learning and convolutional neural networks).[65] This contributes to the mechanisation and objectivisation of the principle of distinction.

4.2 Direct participation in hostilities

Distinguishing persons into those that may be targeted and those that may not is often relatively straightforward under IHL because it is based on generally objective criteria, such as uniforms and distinctive signs.[66] The situation is different, however, when civilians directly participate in hostilities and thereby can be lawfully targeted. According to the ICRC Interpretative Guidance, which remains controversial, three conditions need to be fulfilled for civilian direct participation in hostilities (DPH) to apply, namely a 'threshold of harm', 'direct causation', and a 'belligerent nexus'.[67] The first two elements can be observed objectively. The 'belligerent nexus', however, is difficult to properly identify[68] because it is strongly related to the perceived intent of the person.[69]

With the increased use of algorithms for targeting purposes, it is plausible that civilian DPH will be deduced from a person's observable behaviour, similar to current 'pattern-of-life analyses' that are based on a person's movements and social interactions,[70] or directly inferred from the effects of the person's actions. This can be done by pre-defining scenarios or typologies for when a certain action amounts to civilian DPH, for instance. Such an approach to dealing with civilian DPH would resemble the current practice of 'signature strikes', where people are targeted because their activities are believed to fit a particular behavioural profile, yet without knowing their individual identities.[71]

'Target profiles', which are patterns of sensor data that represent targets,[72] are AI tools that are relevant in this context. When sensor inputs match or fall within a certain target profile, a system recommends applying force (or may apply force

63 'What is Computer Vision?' (*IBM*) <https://www.ibm.com/topics/computer-vision> accessed 18 October 2023; see also S R Vijayalakshmi and S Muruganand, *Embedded Vision: An Introduction* (Mercury Learning and Information 2019) 2.

64 Xin Feng and others, 'Computer vision algorithms and hardware implementations: A survey' (2019) 69(C) Integration 309, 311-314; see also 'image classification', ibid 310-311.

65 See generally ibid 309; Aryan Karn, 'Artificial Intelligence in Computer Vision' (2021) 6(1) International Journal of Engineering Applied Sciences and Technology 249; Alexander Egiazarov and others, 'Firearm Detection and Segmentation Using an Ensemble of Semantic Neural Networks' (*ArXiv*, 11 February 2020) <https://arxiv.org/pdf/2003.00805.pdf> accessed 18 October 2023.

66 AP I (n 43) art 44(3); see also Seixas-Nunes (n 2) 172-173 on 'grey areas'.

67 Melzer (n 52).

68 See, eg, the case of individual self-defence in relation to the element of belligerent nexus as found in the ICRC Interpretative Guidance. Melzer (n 52) 61. See also ibid 70-71.

69 Blank (n 42) 14; see also Marcello Guarini and Paul Bello, 'Robotic Warfare: Some Challenges in Moving from Noncivilian to Civilian Theaters' in Patrick Lin, Keith Abney, and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (The MIT Press 2011) 131.

70 Vasja Badalič, 'The metadata-driven killing apparatus: big data analytics, the target selection process, and the threat to international humanitarian law' (2023) Critical Military Studies 1, 5-6; Michael N Schmitt and Jeffrey S Thurnher, '"Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict' (2013) 4(2) *Harvard National Security Journal* 231, 268; see US Army, *The Targeting Process FM 3-60* (2010) B–3-4.

71 See Kevin Jon Heller, '"One Hell of a Killing Machine': Signature Strikes and International Law' (2013) 11(1) Journal of International Criminal Justice 1.

72 Richard Moyes, 'Target profiles' (Article 36 2019) 4.

in case of high levels of autonomy) relying on proper observation, recognition, and judgement that are based on objective criteria.[73] Thus, the technical grounds for objectivising civilian DPH already exist.

4.3 *Hors de combat*

A further concept of the principle of distinction that is not clearly defined in IHL treaties is that of a combatant being *hors de combat* due to defencelessness because of unconsciousness, shipwreck, wounds, or sickness. Whether a person is sufficiently wounded or sick to benefit from the protection against an attack based on his or her *hors de combat* status is generally understood as being a matter of common sense and good faith.[74] This relies on the attacker's honesty and reasonableness. An objective element for such determination is whether a wounded or sick combatant stops fighting as a result of what he or she thinks about his or her health.[75] Alternatively, the determination can be based on two cumulative, more objective conditions, namely that the person needs medical assistance and that the person does not commit any act of hostility.[76]

Although it has been argued that the status of *hors de combat* is highly context-dependent and that objective criteria may not suffice for making appropriate determinations,[77] it is plausible that states will base their assessments on observable behavioural patterns when using AI. The interpretation of humans' intentions during combat, such as whether a wounded or sick combatant intends to not engage in further fighting, is already difficult. With the increased use of algorithms, the intent element will most likely either be disregarded or deduced from more objective elements.

With AI applications for precise identification of persons (including procedures for the 'positive identification' of targets),[78] such as for facial recognition, becoming more performant and targeting decisions eventually becoming increasingly individualised,[79] the use of AI is likely to allow for a better identification of a combatant's need for medical assistance. It has been argued that related technical challenges would not be impossible to address. This has been demonstrated in trials regarding the identification of persons who surrender – a form of *hors de combat* that can be objectively assessed through common manifestations of an intention to surrender, such as by laying down one's weapons and raising one's hands or displaying a white flag.[80] The result of such developments is a more objective understanding and application of the *hors de combat* status.

73 Boulanin and Verbruggen (n 16) 7-11; Elliot Winter, 'The Compatibility of Autonomous Weapons with the Principle of Distinction in the Law of Armed Conflict' (2020) 69(4) International & Comparative Law Quarterly 845, 846.

74 Jean S Pictet (ed), *Commentary on the Geneva Conventions of 12 August 1949* (ICRC 1952) vol 1, 136.

75 Ibid.

76 Sassòli (n 50) para 8.04; see also ICRC, *Commentary on the First Geneva Convention* (Cambridge University Press 2016) para 1350.

77 See Steven Umbrello and Nathan Gabriel Wood, 'Autonomous Weapons Systems and the Contextual Nature of *Hors de Combat* Status' (2021) 12(5) Information 216.

78 US Joint Chiefs of Staff (n 10) II–21.

79 Special Competitive Studies Project, 'Interim Panel Report' (2022) 4 ff; Winter (n 73) 862-865; see also Sascha Brodsky, 'The Air Force's Drones Can Now Recognize Faces. Uh-Oh.' (*Popular Mechanics,* 24 February 2023) <https://www.popularmechanics.com/military/a43064899/air-force-drones-facial-recognition/> accessed 19 October 2023.

80 S Kate Devitt and others, 'Developing a trusted human-AI network for humanitarian benefit' (2023) 4 Digital War 1; for an analysis on AI and surrender, see Robert Sparrow, 'Twenty Seconds to Comply: Autonomous Weapon Systems and the Recognition of Surrender' (2015) 91 International Law Studies 699.

4.4 Doubt

A last indeterminate element underlying the principle of distinction concerns doubt. As per Article 52(3) of the AP I, in case of 'a doubt whether an object which is normally dedicated to civilian purposes, such as a place of worship, a house or other dwelling or a school, is being used to make an effective contribution to military action, it shall be presumed not to be so used'. This is the corollary of Article 50(1) of the AP I regarding persons, which states that 'In case of doubt whether a person is a civilian, that person shall be considered to be a civilian'. AP I does not provide clear guidance on the degree of certainty required,[81] and there is no clearly established customary international rule.[82] As such, the concept relies on the reasonableness of the commander.

Doubt is closely linked to uncertainty.[83] Therefore, it can easily be quantified as probabilistic levels of confidence and certainty. AI applications can then attribute values to sensor data and let the system compute the certainty levels, meaning the statistical likelihood that the target is legitimate.[84] To this end, armed forces need to define the exact threshold of certainty that corresponds to an absence of doubt and for which lower values represent 'doubt'. This can be an 80% certainty that an object makes an effective contribution to military action and that a presumed civilian is not a civilian, for instance.

Such thresholds may be different for different situations and/or targets and/or potential collateral damage. Yet in any case, the threshold needs to be at a minimum level of 50%, because otherwise the uncertainty level would be larger than the certainty level, which would represent uncertainty that corresponds to doubt. Such quantification means an objectivisation of the related elements of the principle of distinction.

5. Objectivising the principle of proportionality

The principle of proportionality is a second principle of the law of targeting which is subject to mechanisation and objectivisation through the use of AI. Article 51(5)(b) and Article 57(2)(a)(iii) and (2)(b) of the AP I codify the principle, which also reflects customary law.[85] The principle prohibits attacks that 'may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated'.[86] If the incidental harm (also referred to as 'collateral damage') is expected to be excessive, the attack must be cancelled or suspended. Evaluations must be made using foresight and the effects of an attack must be substantiated.[87]

81 Schmitt and Schauss (n 3) 155-156; see also Charles P Trumbull, 'Autonomous Weapons: How Existing Law Can Regulate Future Weapons' (2020) 34(2) Emory International Law Review 533.

82 See 'Customary IHL' (n 48) Rules 6 and 10; Dinstein (n 43) 130-131; see also Cohen and Zlotogorski (n 42) 187-188.

83 See 'doubt' (*Merriam-Webster*) <https://www.merriam-webster.com/dictionary/doubt> accessed 6 March 2024.

84 Schmitt and Thurnher (n 70) 263.

85 'Customary IHL' (n 48) Rule 14.

86 AP I (n 43) art 51(5)(b), 57(2)(a)(iii), and 57(2)(b).

87 See Bothe, Partsch, and Solf (n 43) 352; Dinstein (n 43) 182.

5.1 Incidental harm

The principle of proportionality remains indeterminate in terms of its mechanism and constituent elements. Proportionality tends to be considered as 'one of the most complex and misunderstood norms in IHL with respect to both interpretation and application'.[88] The rule consists of a balance between two elements that are indeterminate in scope and different in nature, namely 'the foreseeable extent of incidental or collateral civilian casualties or damage' and 'the relative importance of the military objective as a target'.[89] In this vein, it has been said that 'military advantage and civilian casualties are like [...] apples and oranges: a comparison between them is an art, not science'.[90] While the element of 'military advantage' has been discussed above under the principle of distinction, it is noteworthy that the proportionality rule demands 'concrete and direct' military advantage. This increases the necessary degree of certainty and reinforces the chain of causation.[91]

The notion of incidental harm remains indeterminate notably as this concerns the valuing of civilian life, which relates to the fundamental notion of human dignity and the resulting understanding that the value of a person's life and suffering cannot and should not be mathematically defined (nor, to a less problematic extent, should civilian and other protected infrastructure).[92] Nonetheless, there are widely used computer-based tools that have already introduced a degree of digitalisation and certain forms of calculation related to incidental harm in targeting processes.

A common technique used to assess the risk of incidental harm is the 'collateral damage estimation methodology' (CDEM). The methodology encompasses 'the joint standards, methods, techniques, and processes for a commander to conduct CDE [collateral damage estimations] and mitigate unintended or incidental damage or injury to civilian or noncombatant persons or property or the environment'.[93] It generally works with five CDE levels to assist commanders 'in weighing risk against military necessity, and in assessing proportionality',[94] thereby helping commanders to adhere to the law of targeting. As such, it is based on empirical data, probability calculations, historical observations, and complex modelling.

The ultimate decision of whether the estimated collateral damage is acceptable remains with the commander in line with the standard of reasonableness. Yet the fact that the 'CDEM provides a numeric estimate of the number of civilians who may be injured or killed if the attack goes forward'[95] shows that the

88 Schmitt and Thurnher (n 70) 253.

89 Bothe, Partsch, and Solf (n 43) 351.

90 Yoram Dinstein, 'Distinction and Loss of Civilian Protection in International Armed Conflicts' (2008) 38 Israel Yearbook on Human Rights 1, 5.

91 See also Bothe, Partsch, and Solf (n 43) 403; see also the discussion in Cohen and Zlotogorski (n 42) 63 ff, in particular the temporal scope of the anticipated military advantage; Schmitt and Thurnher (n 70) 169-170; International Law Association Study Group on the Conduct of Hostilities in the 21st Century (n 48) 364.

92 See, eg, Jens David Ohlin, Larry May, and Claire Finkelstein (eds), *Weighing Lives in War* (OUP 2017); see also Thomas Hurka, *Drawing Morals: Essays in Ethical Theory* (OUP 2011) ch 15; Ben Clarke, 'Proportionality in Armed Conflicts: A Principle in Need of Clarification?' (2012) 3(1) Journal of International Humanitarian Legal Studies 73.

93 Chairman of the Joint Chiefs of Staff, 'No-strike and the Collateral Damage Estimation Methodology' (2012) 35; see also European External Action Service, 'Avoiding and Minimizing Collateral Damage in EU-led Military Operations Concept' (2015).

94 Chairman of the Joint Chiefs of Staff (n 93) 35.

95 John Cherry, 'Avoiding Collateral Damage on the Battlefield' (*Just Security,* 11 February 2021) <https://www.justsecurity.org/74619/avoiding-collateral-damage-on-the-battlefield/> accessed 6 March 2024.

computer-based assessment of incidental harm has already led to a degree of quantification and objectivisation of incidental harm. This tendency will continue with increasing reliance on AI.

5.2 Excessiveness

The fundamental indeterminacy of the proportionality rule results from the threshold of excessiveness regarding civilian harm. The notion of excessiveness is not defined in treaty law. A very common view is that proportionality 'cannot be measured simply by crunching numbers' and that the values involved 'cannot be compared through the simple use of a formula, as there is no common denominator between them'.[96] In 2016, the ICRC convened an international experts meeting seeking to clarify the interpretation of the proportionality rule, including the assessment of excessiveness, but it did not achieve much clarity.[97] The ICRC has also suggested that there would be an absolute limit for civilian losses, which equates the term 'excessive' with 'extensive'.[98] Yet this view has met heavy criticism,[99] including the assertion that this would represent a misreading of the text.[100]

Although some scholars have argued in favour of an objective standard and there are some examples of clearly disproportionate attacks,[101] states and others tend to remain committed to a subjective or semi-subjective standard. The ICRC Commentary, for instance, argues that the evaluation must not be simply subjective, but must be made in 'good faith' and based on 'common sense'.[102] Others have stressed that combatants would need to 'act honestly' and 'competently'.[103] The US also seems to uphold a subjective understanding of proportionality that is based on good faith in view of the available information at the time, as is shown by the wording of the US Law of War Manual of July 2023.[104]

To the extent that AI systems are used to manage or conduct proportionality assessments, the notion of excessiveness is most likely to be quantified and thereby objectivised in one way or another. A computational model of the proportionality rule's 'weighting exercise' has already been developed.[105] The model contains the two values of expected 'incidental harm' and anticipated 'military advantage'. Incidental harm encompasses the loss of civilian life; civilian injury (physical and mental); and damage to civilian objects. Assessing military advantage is influenced by the ground gained by an attack; the disruption of

96 Cohen and Zlotogorski (n 42) 59 and cited sources; but see Schmitt and Thurnher (n 70) 166 ff.

97 Gisel (n 53) Part III.

98 See AP I Commentary (n 42) paras 1979-1980.

99 Cohen and Zlotogorski (n 42) 100; Johansen (n 3) 320.

100 Dinstein (n 43) 181.

101 See Cohen and Zlotogorski (n 42) 104 and 5; see also Dinstein (n 43) 180.

102 AP I Commentary (n 42) para 2208; see also the discussion on the 'reasonably foreseeable reverberating effects' in Isabel Robinson and Ellen Nohle, 'Proportionality and precautions in attack: The reverberating effects of using explosive weapons in populated areas' (2016) 98(1) International Review of the Red Cross 107.

103 Bothe and others (n 43) 351.

104 'Determining whether the expected incidental harm is excessive does not necessarily lend itself to quantitative analysis because the comparison is often between unlike quantities and values.' US Department of Defense (n 42) s 5.12.3, as explained in Casey-Maslen and Haines (n 1) 181.

105 Tomasz Zurek and others, 'Computational modelling of the proportionality analysis under International Humanitarian Law for military decision-support systems' (2022) Asser Institute Working Paper, 6 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4008946> accessed 11 March 2024.

enemy activities; the diversion of the enemy force's resources and attention; the denial of the enemy's ability to benefit from the military objective's effective contribution to its military action; the own force's preservation by the attack; the lowering of enemy forces' morale; and the protection of civilians.

In the model, weights that 'are represented by a number' are attributed to the values of incidental harm and military advantage, thereby 'representing the relative importance of a value'. Probabilities are also added. Overall, incidential harm is considered excessive in comparison to the military advantage *'if the level of promotion of values [of] life of civilians and civilian objects will be lower than military advantage'.*[106] Such modelling of the weighting exercise under the proportionality rule is a clear case of the objectivisation of 'excessiveness' and the proportionality rule more generally.

6. Objectivising the principle of precautions

The third principle of the law of targeting which is subject to mechanisation and objectivisation through the use of AI is the principle of precautions. Reflecting customary international law and codified in Article 57 of the AP I, the principle demands constant care in military operations to avoid and minimise incidental harm. All feasible precautions must be taken to this end, such as selecting weapons and targets accordingly and warning the civilian population of an imminent attack. As such, the principle is closely linked to the principles of distinction and proportionality.[107]

6.1 Feasibility

The core issue regarding the implementation of the principle of precautions is the question of what measures are 'feasible' in given circumstances. Similar to the issue of 'doubt' discussed above, the level of feasibility of a certain precautionary measure can be quantified and calculated with statistical levels of probability and certainty. Such calculation can also involve expected success rates. It can also factor in the expected cost of a precautionary measure, such as losing the attack's surprise effect or a higher risk to proper troops. If a certain measure has a lower success rate for preventing civilian harm than a pre-defined threshold of 20%, for instance, but stands in stark contrast to military constraints, then this measure may be considered as not feasible.

In addition, AI systems can be used to assist the identification, simulation, selection, and execution of courses of action regarding precautionary measures.[108] With regard to the obligation under Article 57(3) of the AP I that obliges states to select the targets which incur the least danger to civilians and civilian objects

106 Ibid 8; original italics.

107 See Elliot Winter, 'The Compatibility of the Use of Autonomous Weapons with the Principle of Precaution in the Law of Armed Conflict' (2020) 58(2) The Military Law and the Law of War Review 240; Kenneth Anderson, Daniel Reisner, and Matthew Waxman, 'Adapting the Law of Armed Conflict to Autonomous Weapon Systems' (2014) 90 International Law Studies 386; Blank (n 42) 7.

108 See the examples of Zurek and others (n 105) 2; Jeffrey S Thurnher, 'Feasible Precautions in Attack and Autonomous Weapons' in Wolff Heintschel von Heinegg, Robert Frau, and Tassilo Singer (eds), *Dehumanization of Warfare: Legal Implications of New Weapon Technologies* (Springer 2018) 112; Winter (n 107) 272; Peter Margulies, 'The Other Side of Autonomous Weapons: Using Artificial Intelligence to Enhance IHL Compliance' in Ronald T P Alcala and Eric Talbot Jensen (eds), *The Impact of Emerging Technologies on the Law of Armed Conflict* (OUP 2019).

among several options with the same military advantage, current systems for the evaluations of enemy threats may be used to assist such decisions.[109] AI applications can also increase situational awareness, namely the 'perception of elements in the environment, comprehension of their meaning, and projection of their status in near future … within a volume of space',[110] which is essential for the exercise of constant care to spare civilians.

Existing computer tools and techniques related to the CDE are a means of precaution and indicate the related process's mechanisation. The algorithms of the US Department of Defense's (DoD) Joint Technical Coordinating Group for Munitions Effectiveness (JTCG/ME) are already capable of creating scenarios regarding weapons' effects, conducting simulations, and iterating the process multiple times.[111] Similarly, the Digital Precision Strike Suite Collateral Damage Estimation (DCiDE) programme is an algorithm that estimates the number of non-combatants and can predict areas' demographic characteristics based on an analysis of the types of vehicles present in an area.[112]

The US National Geospatial Intelligence Agency (NGA) and the Intelligence Advanced Research Project Activity (IARPA) are also developing algorithms that can 'analyze geospatial information and create accurate 3D object models with real physical properties' to ensure that weaponeering solutions will achieve the intended effect and thereby lower risks of incidental harm.[113] Hence, the trend has already started towards the mechanisation and objectivisation of the employment of precautionary measures and the related concepts of feasibility and constant care.

Conclusion

The law of targeting under IHL is relatively indeterminate, leaving armed forces considerable flexibility regarding its implementation. This is due to the difficulty of finding the balance between military necessity and humanity in complex and dynamic situations with high stakes that characterise armed conflicts. Commanders and soldiers need a regulatory framework that is easily understandable, intuitive, and general enough to allow them to take proper decisions and action on the battlefield. The law of targeting's indeterminacy is also a result of states not wanting to overly restrict their *marge de manoeuvre* through the application of clear and rigid international law.

With the introduction of AI systems in targeting processes, IHL is evolving based on these technological developments. More specifically, the implementation of the law of targeting becomes mechanised and objectivised. This notably concerns the elements of 'military advantage', 'contribution to military action', 'direct participation in hostilities', *'hors de combat'*, 'doubt', 'incidental harm', 'excessiveness',

109 See n 61.

110 Arslan Munir, Alexander Aved, and Erik Blasch, 'Situational Awareness: Techniques, Challenges, and Prospects' (2022) 3(1) AI 55; see also Margulies (n 108) 171.

111 John G Thorne, 'Warriors and War Algorithms: Leveraging Artificial Intelligence to Enable Ethical Targeting (Naval War College Newport 2020) 41-42.

112 Ibid 47.

113 Ibid 43-44.

and 'feasibility'. These relatively indeterminate elements will plausibly become more objective by the use of AI systems because of these systems' requirement for precise and quantified specifications and parameters. Accordingly, even if the wording of these concepts does not change, their implementation will follow another logic and take different considerations into account when using AI, ultimately transforming the concepts' nature and meaning towards a more calculable understanding of the law of targeting.

Such developments may even lead to a more determinate law of targeting. Since many states operate with allies and partners, it is likely that they will share their views and procedures with others. This could result in common practices or even standards. To the extent that states agree on common delimitations with specific numerical values, the law of targeting will become more determinate. States may wish not to agree on common delimitations, notably to keep the law indeterminate and benefit from the resulting *marge de manoeuvre* for taking targeting decisions. Nonetheless, by sharing and jointly applying their views and procedures on how to use AI in targeting, they will contribute to a clearer and more objective understanding of the international rules on targeting.

The mechanisation and objectivisation of the law of targeting by the use of AI will arguably also affect the implementation of the law of targeting when not using AI systems, because the necessary definitions and specifications can also be applied to armed forces' actions in general. It is indeed simpler to apply the same legal concepts, definitions, and specifications across entire organisations, independently of which tools, systems, and weapons are used. Differences between when using AI systems and when not, such as higher levels of required certainty for AI systems, are better made at the level of military doctrine, tactical directives, and rules of engagement.

This leads to the question of whether the mechanisation and objectivisation of the law of targeting will make targeting less humane. As the implementation of the law of targeting becomes more rational, numerical, and specified, there may be less room for human emotions, instincts, and spontaneity. This does not mean, however, that the process becomes less humane *per se*, because the rationales, objectives, and measures to ensure the human dignity of fighters and civilians will continue to be taken into account and applied.[114] Yet it does show that the responsibility to protect human dignity in warfare will accompany the evolution of IHL that results from the military use of AI.

114 For a broader discussion, see Neil Renic and Elke Schwarz, 'Crimes of Dispassion: Autonomous Weapons and the Moral Challenge of Systematic Killing' (2023) 37(3) Ethics & International Affairs 321.

Building Peace Together

Geneva Centre for Security Policy
Maison de la paix
Chemin Eugène-Rigot 2D
P.O. Box 1295
1211 Geneva 1
Switzerland
Tel: + 41 22 730 96 00
E-mail: info@gcsp.ch
www.gcsp.ch

GCSP
Geneva Centre for Security Policy