Security Intelligence in the Age of AI

Navigating Legal and Ethical Frameworks

Edited by

Pushan Kumar Dutta
Bhupinder Singh
Christian Kaunert
Annita Larissa Sciacovelli

Security Intelligence in the Age of AI

This page intentionally left blank

Security Intelligence in the Age of AI: Navigating Legal and Ethical Frameworks

EDITED BY

PUSHAN KUMAR DUTTA
Amity University, India

BHUPINDER SINGH
Sharda University, India

CHRISTIAN KAUNERT
Dublin City University, Ireland
University of South Wales, UK

AND

ANNITA LARISSA SCIACOVELLI
University of Bari Aldo Moro, Italy

United Kingdom – North America – Japan – India – Malaysia – China

Emerald Publishing Limited
Emerald Publishing, Floor 5, Northspring, 21-23 Wellington Street, Leeds LS1 4DL

First edition 2025

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-83608-157-9 (Print)
ISBN: 978-1-83608-156-2 (Online)
ISBN: 978-1-83608-158-6 (Epub)

Contents

This page intentionally left blank

About the Book

Security Intelligence in the Age of AI: Navigating Legal and Ethical Frameworks is an edited volume exploring how artificial intelligence (AI) may be increasingly integrated into security intelligence practices, as well as the new, unforeseen problems this presents. With the ever-evolving AI technologies, hackers are finding new opportunities through this technology to engage in clever ways of breaching networks. But this rapid advancement also highlights substantial legal and ethical considerations that need to be managed cautiously in order to achieve responsible use. In this book, leading academics, practitioners and policy experts provide valuable insight across a range of AI security matters for building fairer and stronger future societies. These include the need for transparency in AI systems, bias reduction strategies and fairness analysis in algorithmic decision-making. Thematic Focus: Methodologies of surveillance and data processing in the framework for longitudinal research with an interdisciplinary perspective experience AI processes that are predictive rather than prescriptive. The broader review roundtable finally discusses AI and its implications on human rights and national securities balancing between security measures and civil liberties.

The volume also examines the policy environment that impacts AI and national security, canvassing existing authorities and where those gaps might work. The contributors offer timely perspectives on various challenges of international governance, with a focus on military AI and dual-use technologies. This text includes number of important teachings for practitioners that are drawn from several legal and ethical conundrums affecting current AI security ventures by utilising case studies. Developed to provide security managers, policymakers and academics with an authoritative but practical manual on AI in the best tradition of informative handbooks. The array of viewpoints provided by these chapters allow the reader to gain a more complete view of how best to approach these inter connections between technology, law and ethics in the field of security intelligence. In the end, *Security Intelligence in the Age of AI* stands as an important asset for anyone looking to utilise AI in a way that respects security frameworks. The necessity of working together on behalf of all stakeholders to develop a strong do no harm for the future age relying more heavily on legal mores to set out an ethics-based guideline book in this ever-more-digital world. Tackling these pressing challenges, this volume will extend our understanding of how society can reap the benefits from technological advances without compromising the basic rights and values.

This page intentionally left blank

Preface

With AI revolutionising the world we know, security intelligence is a core focal point. In doing so, the convergence of AI and security frameworks opens up new possibilities for tuning threat detection, improving line-of-defence fortifications and anticipating risks. But the fast pace of AI innovation comes with a lot of growing pains, especially in sorting through the intricacies of current legal and ethical systems regulating their use. With increasing access to data and decision making by AI systems, privacy, accountability and bias issues are becoming more urgent. Keeping AI security intelligence within reasonable limits is not only a technological but also a legal requirement. It is the goal of this book, *Security Intelligence in the Age of AI: Navigating Legal and Ethical Frameworks*, to provide a holistic discussion of these multifaceted concerns. When these distinct but intertwined lines are unravelled as AI, Law and Ethics: The Law, the Promises and the Perils seeks to provide pragmatic advice about how policymakers, security professionals and technologists should approach them. Through the introduction of AI-driven security systems, henceforth surveillance, predictive policing and cybersecurity also address their implications and argue the importance of comprehensive legal norms governing these novelties. In an age where AI can redefine security dynamics, this book is a well-timed study to protect the legal and ethical issues that need to define a future for safety collaboration. The intent of the paper is to challenge readers on questions they should be asking about the liability and moral implications of AI in security, for an equally innovative and protective framework stringent enough to promote both.

This page intentionally left blank

Chapter 1

Can the Red Cross Redefine the Battlefield? Examining the ICRC's Influence on Autonomous Weapon Systems in International Humanitarian Law

Akash Bag[a], *Anwesha Ghosh*[b] *and Tejaswini Tripathy*[c]

[a]Adamas University, India
[b]Amity University, India
[c]National Law University Odisha, India

Abstract

Evaluating the ethical recommendations on artificial intelligence (AI) adopted by UNESCO member states during the 2021 general conference reveals significant insights into integrating historically disadvantaged groups within human rights and AI governance. The paper utilises a dual methodological approach: This analysis describes the aims and objectives of the UNESCO recommendations and examines their practical consequences and accessibility. At the heart of this examination lies the belief that the recommendations support an open universalism based on the inherent ethic of human rights. This method reflects postcolonial theories that emphasise enduring differences in representation and authority among global groups. The study examines how AI participants shape a system often favouring Eurocentric standards and values while affecting AI technology direction and growth. The research shows that even if the proposals endorse universal rights frameworks in AI technology, they clearly emphasise aiding the least developed countries with education and technology. Facilitating this educational outreach is the responsibility of the most advanced countries. Analysis showcases major deficiencies within the recommendations that do not address the needs and desires of communities that reject AI implementation. In addition, the previous supremacy of 'whiteness' in AI technology creates a constant barrier to genuine inclusion in AI governance and moral standards.

Security Intelligence in the Age of AI, 1–21

doi:10.1108/978-1-83608-156-220251001

Keywords: Artificial intelligence ethics; UNESCO recommendations; post-colonial theory; human rights; technological equity; critical idea analysis

Introduction

Autonomous weapon systems (AWSs) have become essential in discussions concerning international humanitarian law (IHL) because they can decide in combat without human oversight. Many issues centre on whether these systems follow essential principles of IHL regarding human control and responsibility in decisions leading to human loss (Asaro, 2012). Ethical discussions point out the dangers of violating human dignity and ethical values during conflicts, resulting in some pushing for a total prohibition on AWS (Sparrow, 2007). Regulatory discussions receive great support from the International Committee of the Red Cross (ICRC) as it examines ways to restrict the use of AWS from coinciding with IHL standards (Ivanov et al., 2021). The conversations investigate how AWS could restructure global regulations and challenge its credibility (Bode & Huelss, 2018). The focus is on the principle of distinction as a critical element of IHL that ranks combatants and civilians sharply apart (ICRC, 2023). This principle highlights the necessity to protect ethical norms in designing and implementing AWS (Rosert & Sauer, 2019). Furthermore, the term 'meaningful human control' about AWS has taken centre stage to forestall the occurrence of IHL violations through appropriate governance structures (Bode, 2023). Current laws are being reshaped to meet the unique challenges introduced by AWS, and they are being transformed into valuable suggestions for engineers and policymakers (Letendre, 2016).

The legal dialogue assesses how current IHL applies to conflicts, including advanced artificial intelligence (AI) technologies in AWS (Belikova & Akhmadova, 2021). The quick progress and future deployment of AWS by significant powers create a pressing need for defined international regulations and pacts to manage its implementation (Wyatt, 2020). The complex relationship among technology growth, ethics and existing laws urges the need for firm regulations to enforce the development and use of these systems in line with global humanitarian principles. The use of AWS in military strategies signals an essential change in current combat methods. To uphold humanitarian principles, these systems require adherence to the rules established by IHL (Szpak, 2020). Illustrating ongoing legal discussions and worldwide forums intends to confront likely difficulties with critical ideas such as distinction and proportionality that could lead some AWS applications to conflict with international regulations. It analyses the importance of establishing the latest legal instruments to oversee fully self-governing weapons (Maskun & Ramli, 2018) and the developing connection between troops and these modernly autonomous platforms (McFarland, 2015). This signals an essential challenge for a complete analysis and monitoring to verify that AWS aligns with recognised legal structures and observes the norms of IHL.

In ongoing debates about AWS legality, worldwide stakeholders, including the United Nations and the EU, express considerable worry. A key topic in these talks involves settling the conflicts surrounding AWS and their consequence for IHL (Shaw, 2003). In these discussions, the ICRC holds a key position influenced by the Geneva Conventions. The ICRC's perspective on AWS, particularly outlined in their May 2021 position, accentuates their critical role as a steward of IHL. The ICRC's historical responsibility to oversee the implementation of the Geneva Conventions emphasises the significance of their input and positions their recommendations as the cornerstone for ongoing discussion. The ICRC's position on AWS is used as the basic framework for analysis in this piece, which acknowledges that although IHL is still the major emphasis, other humanitarian laws and the laws of war initiation (*jus ad bellum*) will not be covered. The discourse will also exclude a detailed examination of the impacts of specific customary rules and treaty laws.

Furthermore, the discussion on AWS will be considered relevant across various forms of conflicts, both international and non-international, as governed by customary law. However, the application of AWS in specific types of conflicts, such as naval engagements, will not be addressed. Throughout the paper, the term 'IHL' will be consistently used to denote what may also be referred to as the Law of Armed Conflict or *jus in Bello*, while the use of the term 'AWS' will be prevalent, with 'Lethal Autonomous Weapon Systems' (LAWSs) used interchangeably where necessary. This approach ensures clarity and maintains the focus on the legal dimensions and implications of AWS deployment in contemporary conflict scenarios (Gunawan et al., 2022). In IHL, there needs to be an agreed methodical framework. Unlike national legal systems that operate within a hierarchical structure, international law functions horizontally, where no single source holds absolute authority, except for *jus cogens* norms, which are peremptory and overriding (Cassese, 2005). Further, this approach reflects the nature of international law, where court decisions and scholarly opinions serve primarily as interpretive aids rather than direct sources of law, contrasting with their often more substantive role in national legal contexts (Zimmermann et al., 2019, Article 38).

Therefore, the interpretation and application of international law rely heavily on the Vienna Convention on the Law of Treaties (1969), which mandates that treaties be interpreted in 'good faith'. In this landscape, the International Court of Justice (ICJ) statutes stand out as a recognised source that introduces a hierarchy within IHL. However, the significance of customary international law must be balanced. Customary law emerges from consistent state practice accompanied by a belief in its legal obligation (*opinio juris*) (Sinclair, 1984, Article 31–32). A key issue in addressing AWS within the framework of IHL arises from the absence of direct regulation by either treaty or established customary law (Henckaerts, 2005). This lack of clear legal guidance poses significant methodological challenges in defining AWS legally and weighing diverse scholarly opinions across multiple academic disciplines. The paucity of publicly available information and restriction to accessible data on military technology further complicates this analysis by imposing an inherent limit on the depth of analysis and the scope of conclusions

that can be drawn. This limitation is made worse by the dearth of scholarly legal research on AWS.

Autonomous Weapon Systems and International Humanitarian Law

The Current State of IHL, the General Laws of Prohibition and the Idea of 'Ruling Out' AWS

The initial recommendation addresses AWS, which lacks sufficient predictability. These systems are not fundamentally new weapons but represent novel approaches to weapon control (McFarland, 2020), potentially integrating both pre-set programming and elements of deep learning (Schwarz, 2021). Consequently, AWS may operate with a degree of autonomy in selecting targets within predefined parameters. The United States has indicated a readiness to deactivate any system that misbehaves (Losey, 2021). Given the inherent unpredictability of warfare, where operational conditions can shift rapidly, achieving complete predictability in weapon behaviour is implausible. Military strategists generally prefer highly predictable weapons, though it is recognised that this is not feasible in all scenarios (Klare, 2019). This recommendation advocates explicitly for prohibiting the deployment of AWS that cannot be comprehensively understood, aligning with the broader objective of ensuring compliance with IHL (International Committee of the Red Cross, 2021).

The second point under consideration advocates for the ICRC to restrict the use of AWS or LAWS against human targets. Implicit in this recommendation is the differentiation between targeting individuals exclusively and engaging military objectives where civilians may be present. This aligns with the ICRC's broader agenda, mirroring initiatives like the 'Stop Killer Robots' campaign, which seeks to kerb or completely prohibit AWS development and deployment (Sharkey, 2018). The underlying concerns driving these campaigns primarily revolve around the potential for AWS to target humans, highlighting a significant ethical challenge (Wareham, 2023). The need to uphold the principles of IHL, which prioritises protecting soldiers who are *hors de combat* – out of the fight – and civilians, exacerbates these worries. Ensuring that AWS complies with these IHL principles is a paramount concern for military leaders and governments considering deploying such technologies (Seixas-Nunes, 2022).

However, there is a counterargument emphasising AWS's technical superiority, particularly in its ability to process information rapidly (Kallenborn, 2021). In scenarios where combatants may quickly become *hors de combat*, human soldiers might fail to recognise this status change due to combat stress or delays in processing what has occurred. Conversely, AWS could potentially identify and react to such changes almost instantaneously, thus theoretically increasing the survival likelihood of those no longer active in combat (McFarland, 2020). This suggests that, under certain conditions, AWS might offer a tactical advantage in complying with the principles of IHL, specifically in terms of rapidly updating

engagement decisions to reflect changing conditions on the battlefield (Christie et al., 2023).

Regulate, Do Not Prohibit

The third recommendation (International Committee of the Red Cross, 2021) delves into the realm of AWS design and use, highlighting a clear desire to establish international regulations suiting the principles of IHL. Despite discussions among the High Contracting Parties of the Convention on Certain Conventional Weapons (CCW), and the release of 11 guiding principles for LAWS in 2019 (Wareham, 2023), there have been scanty governmental initiatives towards crafting specific regulations. The initial subsection of this recommendation proposes that AWS targeting be strictly confined to military objectives, which would enhance the existing restrictions under IHL. Presently, IHL mandates targeting only military objectives (Solis, 2010), though under certain conditions, civilian objects may also be targeted if they are temporarily being used for military purposes (Henckaerts, 2005, Rule 9). The proposal suggests a stringent application, presumably to prevent any targeting of civilian structures like schools, even if used by combatants, thereby aiming to minimise breaches of IHL (Bothe, 2013, Article 52 (2)). However, the likelihood of states agreeing to restrictive use of sophisticated military technology remains low. Further, the second point suggests setting limitations on the situations in which AWS can be used. Typically, all weapons are subject to specific use limitations due to their inherent illegality or specific conditions under which they can be legally employed (Seixas-Nunes, 2022). The proposal supports the idea of putting temporal constraints on AWS deployment for a set amount of time or the duration of a particular mission. This idea extends to debating whether such constraints are already encapsulated within existing IHL norms. However, even if such limitations exist, this does not preclude the establishment of new, AWS-specific agreements (International Committee of the Red Cross, 2021).

The third bullet point within the recommendation delineates parameters for the lawful deployment of AWS. It suggests restricting AWS use to scenarios devoid of civilians or civilian objects, echoing existing IHL, which mandates a high certainty that only lawful targets are engaged (Boothby, 2016). This recommendation potentially establishes stricter criteria for AWS than current norms for other weapons, although interpretations may vary. This topic ties into broader distinction issues within IHL, where the dynamic definition of what constitutes a military object complicates strict adherence to such limitations (Gunawan et al., 2022). The fourth and final recommendation addresses human involvement in AWS operations. There is no universally accepted definition of what constitutes sufficient human interaction with AWS, nor clear rules governing their use. This recommendation aligns with the general stance of most states under IHL, which advocates for some level of human control over AWS (Taddeo & Blanchard, 2022). While there is broad agreement that AWS should be under 'adequate human control', the interpretations of what this entails can vary

significantly between nations (Taddeo & Blanchard, 2022). This lack of uniform understanding accentuates the need for more straightforward guidelines, suggesting that the ICRC's call for more stringent control measures is relevant and justified.

Current State of IHL

General Laws of Prohibition and Targeting Law

Within IHL, certain rules explicitly prohibit or regulate the use of specific weapons (Turns, 2006). While it's clear that some types of AWS and LAWS might fall under these regulated categories, applying these rules is not uniformly applicable to all AWS. This ambiguity underscores the necessity for a deeper understanding and clarification of existing IHL norms, particularly because no specific prohibitions are tailored to AWS and LAWS. The reliance on general provisions, such as those banning weapons that cannot distinguish between combatants and civilians, becomes essential (International Committee of the Red Cross, 1949; Bassiouni, 2001, Article 51 (4)). Such indiscriminate weapons are prohibited, alongside those causing superfluous injury or unnecessary suffering (Horvitz & Nehs, 2011). The applicability of AWS varies by context; for example, in naval warfare, the requirement for distinction may pose less of a challenge. Naval AWS needs to identify valid military targets from a distance, a task complicated by the potential for misidentifying the signatures of ships, which could result in targeting protected vessels (McFarland, 2020). This issue of misidentification is not exclusive to AWS; human operators are equally prone to such errors. However, the legal framework for addressing these mistakes by AWS remains unsettled, complicating their use and likely prompting future legal clarifications as these technologies are increasingly implemented (Doswald-Beck, 1995).

The laws of targeting, a core section of IHL, govern the conduct of attacks, specifying who or what may be legally targeted. This section is founded on four basic principles: military necessity, distinction, proportionality and humanity, all of which have a well-established history within IHL and apply to all conflict parties, regardless of their role as aggressors or defenders (Boothby, 2016). While these rules are universally binding, individual states may be subject to additional treaty obligations that further restrict the types of weapons and methods of warfare they can lawfully employ (Army, 1956; International Committee of the Red Cross, 1949). Historically, the concept of military necessity was defined as early as 1861 in the Lieber Code, which allowed only those measures 'lawful according to the modern law and usages of war (Lieber, 1863)' (Yale Law School, 1863, Article 15).

Distinction and Proportionality in Armed Conflict-Distinction

As military technologies have evolved and battlefields have expanded beyond well-defined zones to broader, more indeterminate areas, clear rules of distinction

have become paramount (Boothby, 2016). The most definitive codification of these rules is encapsulated in Additional Protocol I (API) to the Geneva Conventions, mainly Article 48 can be found under the section for 'Basic Rules', which underlines the obligation of parties to a conflict to always differentiate between civilians and combatants, and between civilian objects and military objectives, directing operations solely against the latter (Kalshoven, 1978). Despite not being universally ratified, significant portions of API are recognised as customary law, thus binding all conflict participants (Boothby, 2016). The ICJ has reinforced that the principles of distinction are 'fundamental and intransgressible principles of law (ICJ, 1996)'. The ICJ clarifies that lacking the immediate means to comply with these principles does not justify indiscriminate attacks; rather, it obliges states to abstain from such actions. The application of this principle will necessarily vary depending on the available weaponry, intelligence and other pertinent information that could influence the decision-making process regarding an attack.

To give effect to the principles of IHL, particularly those concerning civilian protection, a clear definition of 'civilian' is essential. The API provides this definition in a way that broadly includes any person not classified under specific categories of combatants detailed in Articles 4A and 43 of the Geneva Conventions and the API itself (International Committee of the Red Cross, 1949; Bassiouni, 2001, Article 50(1)). This negative definition, which essentially categorises individuals as civilians unless proven otherwise, aims to minimise civilian casualties in conflict – a fundamental goal of IHL (Boothby, 2016). The distinction between combatants and civilians is further elaborated in Articles 43 and 44 of the API. Article 43 states that 'The armed forces of a party to a conflict consist of all organised armed forces, groups, and units which are under a command responsible to that party for the conduct of its subordinates, even if that party is represented by a government or an authority not recognised by an adverse party. Such armed forces shall be subject to an internal disciplinary system which, among other things, shall enforce compliance with the rules of international law applicable in armed conflict (International Committee of the Red Cross, 1949)' (Bassiouni, 2001, Article 43). The protocol modifies previous narrower definitions, like those in the Hague Regulations, by emphasising the need for armed forces to display a 'distinctive sign and to carry arms openly', adding a layer of transparency to combat operations (Boothby, 2016).

Moreover, the API commentary notes the flexibility in the term 'organised', suggesting that it encompasses groups collaborating under established command and rules rather than mere ad hoc cooperation (Boothby, 2016). Article 44 (3) addresses the obligations of combatants to distinguish themselves from civilians during attacks or preparations for attacks. This provision reflects a practical recognition of the complexities of modern armed conflicts, where traditional norms of combat visibility are not always possible but where the necessity for some form of distinction remains critical. The distinction between civilians and combatants, as elaborated in Articles 43 and 44 of the API, is crucial for enhancing civilian protection in conflict zones. These articles require combatants to carry arms openly to be recognised as such, a measure that significantly

contributes to safeguarding civilians residing in or near war zones. Despite their value as interpretive tools for understanding the rules of distinction, these parts of the API are contentious and subject to reservations by some ratifying states, affecting the application of these rules and their recognition as customary law. Notably, the UK has reservations about Article 44(3), and the US has not ratified the API, explicitly rejecting Articles 43 and 44, citing concerns over the expanded definitions of prisoners of war and combatants as major obstacles (Boothby, 2016).

These objections, however, do not detract from the utility of these articles as guidelines for this analysis. Under API Article 50, the protocol clearly states that individuals should be considered civilians in cases of doubt, reinforcing the principle of protection (Bassiouni, 2001, Article 50; Bothe, 2013; International Committee of the Red Cross, 1949). This approach contrasts with earlier IHL frameworks, which, due to the more straightforward nature of traditional warfare, found it easier to distinguish between combatants and civilians (Boothby, 2016). The complexities of modern conflict, where the lines between combatant and civilian statuses are increasingly blurred, stress the importance of adhering to strict rules against indiscriminate attacks (Boothby, 2016). These challenges highlight the necessity for rigorous adherence to IHL in contemporary armed conflicts, ensuring that the principles of distinction and protection of civilians remain a priority.

Proportionality

The principle of IHL insists that a fine line between military strength and civilian suffering must stay unchanged. The idea that underlies IHL influences how force is applied and insists that attacks take premeditated measures rather than respond afterwards to their effects (Saul & Akande, 2020). Understanding this principle divides the discussion into the need for a prior evaluation of anticipated outcomes and the frequent confusion that regards proportionality as a simple ratio of harm to advantage (Dinstein, 2004; Saul & Akande, 2020). Explaining proportionality in military applications works better when the idea is organised into two distinct parts. Proportionality must be assessed through the expected outcomes of an attack, consisting of military advancement and the risk to civilians (Parks, 1990). Second, the actual outcome of the attack, whether it results in more or less damage than anticipated, should not influence this initial judgement. Essentially, the legality of an attack is evaluated based on the information available before it takes place, not the results that follow.

Critically, proportionality must be evaluated with the information available during decision-making, prioritising the attacker's judgement on potential military gains against expected civilian losses, the necessity of the military action and other related factors (Boothby, 2016). When the outcome diverges significantly from expectations, causing unforeseen civilian harm with minimal military benefit, subsequent evaluations must revert to the original anticipatory understanding (Saul & Akande, 2020).