

外交軍事セクション論文要旨

本稿の目的は、自律型致死兵器システム（LAWS）によって生じる「責任の空白」を有意な人間による制御（MHC）がどの程度補完できるのかを明らかにすることである。AIを軍事的に活用する動きが世界で進む中で、国際的な議論の焦点が当たるLAWSは、法的責任や倫理的懸念といった課題が指摘されており、管理のない状態は深刻な国際秩序の危機と捉えられる。LAWSの開発・運用が拡大する前にいかにLAWSを規制・管理すべきか提起を行うことは、持続的な国際平和のために重要な意義を持つのではないかと考える。

本稿ではまず、LAWSの定義が曖昧であることを指摘した上で、その軍事的優位性について分析を行った。LAWSに対する認識の齟齬は運用・規制に関する合意形成を困難にしている。しかし本兵器がもたらすコスト削減・軍事力強化・効率性向上といった利点により、実戦配備は近い将来現実のものとなりうると考察した。さらに本稿では、LAWSが自由主義の根幹である責任をいかに空白化させ、国際秩序を脅かすかを倫理・構造・法の観点から検討した。その後、この責任の空白への応答として注目されるMHCについて論じ、MHCは判断主体としての人間の制御を明確にし、責任の空白を補完する枠組みとなることを示した。一方で、LAWSの軍事的優位性や安全保障のジレンマの影響により、MHCの原則を実際に適用するには限界があることも明らかとなった。

以上の検証から本稿では、LAWSが生み出す責任の空白に対しMHCが理念としては有効であると考えられる一方、実効性に限界があることを示した。そして、第三者審査機関の設立をMHC制度化の基盤として提案することで、その有効性と実現可能性を提案した。

責任なき兵器 ～LAWSが自由主義にもたらす脅威と MHCを超えた管理の可能性～

目次

序章	3
第1章 LAWSとは.....	4
第1節 LAWSの定義	4
第2節 実戦投入と急速な一般化.....	5
第3節 LAWSがもたらす軍事的優位性	6
第4節 まとめ	7
第2章 LAWSによる責任の空白化と自由主義.....	7
第1節 自由主義と責任	7
第2節 責任の空白化	9
第1項 テクノロジー倫理的観点	9
第2項 構造的観点	9
第3項 法的観点	9
第3節 まとめ	10
第3章 MHCについて.....	11
第1節 MHCとは	11
第2節 責任の空白を補完しうるMHC	12
第3節 安全保障のジレンマ	13
第4節 まとめ	15
第4章 MHCの実効性確保に向けて.....	15
第1節 IAEAから見る第三者審査機関の意義.....	15
第2節 MHCに関する第三者審査機関の三機能.....	16
第3節 まとめ	17
終章	18

序章

近年、人工知能（Artificial Intelligence: AI）を軍事的優位性のために積極的に活用する動きが世界各国で進んでいる¹。中でも、自律型致死兵器システム（Lethal Autonomous Weapon Systems: LAWS）²は、国際的な議論の中心となっている。2013年には国際NGOであるヒューマン・ライツ・ウォッチ（Human Rights Watch）が殺人ロボット阻止キャンペーンを設立し、完全自律型兵器に対する予防的かつ包括的な禁止を求めた³。その後も特定通常兵器使用禁止制限条約（Convention on Certain Conventional Weapons: CCW）における政府専門家会合（the Group of Governmental Experts: GGE）を通じて、国際的な議論が継続的に行われている⁴。

こうした議論においては、兵器の自律性に関して法的責任の課題や倫理的な懸念が指摘されている⁵。この課題の解決に寄与する規範的概念として注目されているのが「優位な人間の制御（Meaningful Human Control: MHC）」⁶である。MHCは人間の制御を通じて、戦争における人間の尊厳の保護・国際人道法（International Humanitarian Law: IHL）の遵守・責任の空白の回避に貢献するとされている⁷。

フィリッポ・サントニー・デ・シオ（Filippo Santoni de Sio）らは、自律型兵器システムは責任の空白を生みうるとし、倫理的・政治的・法的議論においてMHCは重要な役割を果たしてきていると主張する⁸。しかし、MHCの理念を実際に適用することには多くの困難が伴うと考えられる。特に、LAWSによる軍事的優位性は、人間の制御や責任の重要性を強調するMHCとの間に深刻な対立を生じさせる。

本稿ではこうした状況を踏まえ、LAWSの登場を単なる技術的進化としてではなく、現在の国際秩序を根本から揺るがす「国際秩序の危機」として捉える立場を取る。その上で「MHCはLAWSにおける責任の空白をどの程度埋めることができるのか」という問いを設定し、これに対する検証を行う。本稿の仮説は、「MHCはLAWSにおける責任の空白に対して、倫理的・制度的に責任を支える枠組みである一方、LAWSの軍事的優位性や安全保障のジレンマの影響により、MHCの原則を実際に適用するには限界がある」というものである。

¹King, Anthony. (2024) “Digital Targeting: Artificial Intelligence, Data, and Military Intelligence,” *Journal of Global Security Studies*, 9(2), pp.1-2.

²現状ではLAWSの明確な定義は定まっておらず、この点については第1章で論じる。また、直接LAWSと言及していない文献を引用する場合は、元の文献が使用している語句を使用する。

³Human Rights Watch. (April 23, 2013) “Arms: New Campaign to Stop Killer Robots: Swift Action Needed to Prevent Fully Autonomous Weapons,” <https://www.hrw.org/news/2013/04/23/arms-new-campaign-stop-killer-robots>, (accessed June 7, 2025).

⁴外務省（2024年6月24日）「自律型致死兵器システム（LAWS）について」https://www.mofa.go.jp/mofaj/dns/ca/page24_001191.html、（2025年6月3日閲覧）。

⁵Human Rights Watch. (2016) “*Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban*,” https://www.hrw.org/sites/default/files/report_pdf/arms1216_web.pdf, p.4.

⁶MHCの日本語訳については学術的に統一されたものは存在せず複数併存している。本稿では、「有意な人間による制御」を採用する。

⁷Human Rights Watch. (April 11, 2016) “Killer Robots and the Concept of Meaningful Human Control,” <https://www.hrw.org/news/2016/04/11/killer-robots-and-concept-meaningful-human-control>, (accessed June 6, 2025).

⁸Santoni de Sio, Filippo., & van den Hoven, Jeron. (2018) “Meaningful Human Control over Autonomous Systems: A Philosophical Account,” *Frontiers in Robotics and AI*, 5(15), pp.4-5, 12.

本稿の構成は以下の通りである。第1章では、LAWSの定義や自律性、それがもたらす軍事的優位性について検討する。第2章では、LAWSが引き起こす責任の空白とそれが自由主義に基づく現国際秩序に与える影響について論じる。第3章では、MHCとは何かを示し、責任の空白をどの程度埋めうるのか検証する。さらに、安全保障のジレンマとの関係も分析し、MHC適用の壁について議論する。最後に第4章では、MHCの実効性を確保するための制度的枠組みの可能性について考察を行う。

第1章 LAWSとは

本章では、各国・各機関のLAWSに対する認識の違いを示すことで、LAWSの定義の曖昧さを明らかにする。その後、自律性を伴う兵器の実例を踏まえた上で、LAWSがもたらす軍事的優位性を検討し、各国が開発・運用を進める背景を考察する。

第1節 LAWSの定義

AI兵器の開発・導入が進む中、LAWSは国際的論争の的となっている。一方で、現在この兵器には明確な一律の定義が存在せず、各国・各機関が様々な声明を出している。

国/機関	定義
日本	一度起動すれば、操作者の更なる介入なしに標的を識別し、選択し、殺傷力を持って交戦することができるという特徴を備えている兵器システム
米国	一度起動されればオペレーターの介在なしに目標を選定し交戦することができる兵器システム
中国	致死性・自律性・停止不可能性・無差別性・進化性の5つの特徴を備えた兵器
イギリス	文脈に即した人間の制御がなく、また人間の責任や説明責任の枠外で作動する兵器
ドイツ	運用に関する判断から人間の要素を完全に排除した兵器システム
赤十字国際委員会	人間の介入なしに目標を選定し、攻撃を実行するすべての兵器

【表1 各国・機関のLAWSの定義】

(出典：外務省「自律型致死兵器システム (LAWS) について」；U.S. Department of Defense “*AUTONOMY IN WEAPON SYSTEMS*”；People’s Republic of China “*Position Paper*”；United Kingdom “*Input to UN Secretary-General’s Report on Lethal Autonomous Weapons Systems (LAWS)*”；Federal Foreign Office “*German commentary on “operationalizing all eleven guiding principles at a national level as requested by the chair of the 2020 Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS) within the Convention on Certain Conventional Weapons (CCW)*”；International Committee of the Red Cross “*What you need to know about autonomous weapons*”より筆者作成。)

表1のようにLAWSに対する姿勢が多様化している背景には、自律という概念の複雑性があげられる。ヒューマン・ライツ・ウォッチ (Human Rights Watch) は、完全自律型兵器 (Fully Autonomous Weapons: FAWS) を人間の制御の度合いによって以下の3種類に分類している。

- (a) Human-in-the-Loop Weapons
- (b) Human-on-the-Loop Weapons
- (c) Human-out-of-the-Loop Weapons⁹

この分類において自律型の兵器とは、遠隔で人が操る半自律システムから、完全に自律行動する兵器までを幅広く指す。人間の意思決定が介在しない (c) が完全自律であることは自明である一方で、ジェフリー・サーンハー (Jeffrey Thurnher) は (b) について「人間の監視が交戦を形式的に追認するものであれば、その監視は有効ではない」¹⁰と述べており、監視の形態によっては (b) も完全自律の類いに含まれると考えられる。また (a) に関しては、人間が意思決定に関与するため、完全自律とは捉えられない。

このように、自律性の解釈次第でLAWSに含まれる兵器の範囲が大きく変動するため、LAWSの運用・規制に関する国際的な合意形成は難航している。実際、国連軍縮局 (United Nations Office for Disarmament Affairs: UNODA) も「現在、一般的に合意されたLAWSの定義は存在しない」¹¹と明言しており、LAWSに関する議論は混迷を極めている。

第2節 実戦投入と急速な一般化

⁹Human Rights Watch. (2012) “*Losing Humanity: The Case against Killer Robots*,” <https://www.hrw.org/reports/arms1112ForUpload.pdf>, p.2.

¹⁰Thurnher, S. Jeffrey. (2012) “No One at the Controls: Legal Implications of Fully Autonomous Targeting,” *Joint Force Quarterly*, 67(4), p.83.

¹¹United Nations Office for Disarmament Affairs. (n.d.) “Lethal Autonomous Weapon Systems(LAWS),” <https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/>, (accessed June 6, 2025).

LAWSに対する議論が混迷を極める中、一定の自律性を備えた兵器はすでに複数の紛争において実戦配備され、その運用が現実のものとなっている。例えばイスラエル軍は、爆撃の際にラベンダーと呼ばれるAI標的システムを用いている¹²。その実態については、AIの選んだ殺害リストが精査されることはなく、人間の制御は「形式的な承認印」にすぎなかったとされている¹³。つまりラベンダーは人間の意思決定が実質的に形骸化している点で、高い自律性を持つ兵器として捉えられる。さらに、徘徊型自律兵器と呼ばれる兵器も、ここ数年で複数の紛争において使用されている。この兵器は従来のミサイル兵器とは違い、発射時に標的の正確な位置情報を必要とせず、事前に設定されたターゲットを見つけるまで空域を徘徊し、標的を発見次第人間の介入なしで攻撃を行う自律性を備えた兵器である¹⁴。2014年のリビア内戦では、本兵器が操作者との接続なしに標的を攻撃するよう設定されていた可能性が報告されている¹⁵。また、ナゴルノ・カラバフ紛争や、2022年以降のロシアによるウクライナ侵攻においても、本兵器の使用が確認されている¹⁶。一定の自律性を備えた兵器を投入することは、もはや戦場において常態化していると言える。

第3節 LAWSがもたらす軍事的優位性

前節で示したように、一定の自律性を備えた兵器は一般化しており、その背景にはLAWSがもたらす軍事的な利点があると考えられる。自律性を巡る懸念や倫理的問題があるものの、各国は戦場における人的リスクの低減や任務の効率化といった具体的なメリットを重視し導入を進めている。

例えば米国防総省は、無人兵器が、単調な任務ではオペレーターの負担を減らし、劣悪な環境での任務では人間のリスクを最小化し、高い危険を伴う任務では失敗時の人的・政治的コストを大幅に削減するという3点において、人間よりも任務に適しているとしている¹⁷。さらに、心身への負担が大きい任務を無人兵器が代替することで、人的コストが抑えられるだけでなく、結果として戦力の底上げにも繋がる。また、ロナルド・アーキン (Ronald Arkin) は自律性を備えた兵器が自己犠牲的に使われることもあるとしている¹⁸。機械が自己を守る行動を優先することではなく、より積極的な交戦が可能になる。すでに戦場では自爆ドローンなどの兵器が自己犠牲的な目的で使用されており、このような兵器に高い自律性が備われば、戦術的・戦略的な優位性が大幅に向上する可能性がある。

LAWSの導入は前述したような総戦力の増加に留まらず、任務の効率を上げる効果も期待できる。ロシアはLAWSについて、人間よりも効率的にタスクをこなせるとしており、特に人間の操縦者の誤りや、精神的・肉体的状態、道徳的・宗教的・倫理的態度に起因する負の影響を大幅に削減すると主張す

¹² +972Magazine. (April 3, 2024) “Lavender’: The AI machine directing Israel’s bombing spree in Gaza,” <https://www.972mag.com/lavender-ai-israeli-army-gaza/>, (accessed June 6, 2025).

¹³ Ibid.

¹⁴ Michel, Arthur. Holland, & Gettinger, Dan. (2017) “Loitering Munitions,” <https://dronecenter.bard.edu/files/2017/02/CSD-Loitering-Munitions.pdf>, p.4.

¹⁵ UN Security Council, Final report of the Panel of Experts on Libya established pursuant to resolution 1973 (2011), UN Document No.S/2021/229, (March 8, 2021), p.17.

¹⁶ Bode, Ingvald., & Watts, F. A. Tom. (2023) “Loitering Munitions and Unpredictability Autonomy in Weapon Systems and Challenges to Human Control,” https://www.researchgate.net/publication/371351342_Loitering_Munitions_and_Unpredictability_Autonomy_in_Weapon_Systems_and_Challenges_to_Human_Control, pp.52-66.

¹⁷ U.S. Department of Defense. (2007) “Unmanned Systems Roadmap 2007-2032,” <https://apps.dtic.mil/sti/pdfs/ADA475002.pdf>, p.19.

¹⁸ Arkin, C. Ronald. (2009) *Governing Lethal Behavior in Autonomous Robots*, Boca Raton: CRC Press, pp.22-23.

る¹⁹。学習データの偏りや意思決定の不透明さといった懸念はあるものの、非感情的に一定の基準に基づく判断を下すことができるであろう。

上記のメリットを享受する上で、人間の制御は大きな壁である。ゲイリー・マーチャント（Gary Marchant）らは、システムの複雑性や思考スピードが向上する中で、「比較的遅い人間の意思決定を介入させることは、性能のレベルにおいてますます制限をかけ、問題を引き起こすことになるだろう。」²⁰と述べている。兵器の自律性がもたらす利益を最大化するには、人間の制御は不要どころか邪魔なのである。

第4節 まとめ

本章では、LAWSの定義の曖昧さを示し、特に自律性の解釈次第で当該兵器の範囲が大きく変動することを明らかにした。国・組織間での認識のばらつきは、LAWSの運用・規制に関する合意形成を困難にしている。しかし実際の紛争においては、一定の自律性を伴う兵器がすでに使用されており、LAWSの実戦配備は近い将来現実のものとなるだろう。兵器導入の背景には、LAWSがコストの削減・軍事力の強化・効率性の向上といった軍事的利点をもたらすことがあげられる。また、人間の制御ありきではLAWSの特徴を最大限に活かすことができず、完全な自律状態が望ましいことも示した。

第2章 LAWSによる責任の空白化と自由主義

前章では、LAWSが軍事的優位性をもたらす点に注目した。一方で、LAWSは責任の空白という深刻な問題を引き起こす。フィリップらはこの問題に関して、「AIにおける『責任の空白』という概念は（中略）『学習オートマトン』²¹が不測の事態に対して人々に道徳的責任を帰属させることを難しく、あるいは不可能にするかもしれない懸念を示すために提起された。」²²と述べており、責任の空白が自律性に起因することを明確にしている。本章では、自由主義の根幹にある責任原理に着目しつつ、それがLAWSによっていかに空白化し、国際秩序を脅かすかを倫理・組織・法の観点から検討する。

第1節 自由主義と責任

¹⁹Russian Federation, Approaches of the Russian Federation to the issue of emerging technologies in the area of lethal autonomous weapons systems, CCW/GGE.1/2024/WP.2, (March 1, 2024), p.4.

²⁰Marchant, E. Gary. et al. (2011) "International Governance of Autonomous Military Robots," *Science and Technology Law Review*, 12, p.275.

²¹環境の情報を元に、自身の行動を最適化していく機械学習の一種である。

²²Santoni de Sio, Filippo., & Mecacci, Giulio. (2021) "Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them," *Philosophy & Technology*, 34, p.1057.

現在の国際連合や世界貿易機関をはじめとする国際機関・国際規範、また、それらに基づく国際秩序は、第二次世界大戦後、自由主義を標榜する米国ならびにその同盟国主導で構築されてきた²³。現行の国際秩序は自由主義に基づいていると言えよう。人権の尊重・民主主義・法の支配といった原則を中核とするこの秩序は、個人の自由を最大限に尊重することを根本理念としている²⁴。しかし、その機能と正当性を支えるには責任の概念が不可欠である。自由と責任は不可分であり、自由を正当化し持続可能なものとする条件として責任が求められる。

自由主義の思想的源流に立ち返ると、ジョン・ロック (John Locke) は、自然状態における個人の自由を強調しつつも、その自由は理性によって認識される他者の生命・自由・財産を尊重するという自然法の義務に制約されており、個人はこの義務を受け入れることによって道徳的主体として社会秩序を支え、他者との平和的な関係を築くことができると論じた²⁵。すなわち、自由とは無制約な権利ではなく、他者との共存を前提とした責任ある自己決定であると言える。

イマニュエル・カント (Immanuel Kant) もまた、自由主義の倫理的基盤として自律性を重視した。彼は、人間が理性によって普遍的な道徳法則を認識し、その法則に従うことが道徳的行為の根源であると考えた。要約すると、この行動が外部からの強制によるのではなく、自らの理性的な意志に基づいていることである。つまり、真の自由とは自己の理性に基づいて道徳法則を自律的に選び取ることであり、この自律的な意思決定が行為者に責任を帰属させる根拠となると説いたのだ²⁶。カントをとっても、自由と責任は不可分であると言えるだろう。

責任の重要性は、個人の倫理的行動にとどまらず、国家や国際機関などにも及ぶ。セス・レイザー (Seth Lazar) は「正当性及び権威は、本質的に公開性に依存する。公開性の要件を満たすためには、権力が正当かつ適切な権威をもって行使されているかどうかを判断できることが可能でなければならない。」²⁷と主張する。つまり、権力者はその意思決定を説明する義務を負い、この説明責任と透明性の原則が自由主義に基づく国際秩序の正当性を支えている。例えば国際連合や国際裁判所は、国家の行動を国際的価値に照らして説明可能なものにする制度として捉えられる。

結局のところ自由主義に基づく秩序は、単に自由を保障するだけでなく、その自由が無秩序や専断に偏らないように責任を課すものでもある。自由が濫用されず、現秩序が正当性を保ち続けるためには、自由を行使する個人・国家・国際機関などのあらゆる主体が、それぞれの行動に倫理的かつ制度的責任を負うという構造が不可欠なのである。自由の理念を支えるもの、それはまさに責任の原理なのである。

²³細谷雄一訳 (2012) 『リベラルな秩序か帝国か (上) ——アメリカと世界政治の行方』勁草書房、(Ikenberry, G. John. (2006) *Liberal Order and Imperial Ambition: Essays on American Power and International Order*, Cambridge: Polity Press)、pp.26-27。

²⁴UN General Assembly, Universal Declaration of Human Rights, A/RES/217(III) (December 10, 1948), p.71、加藤節訳 (2010) 『完訳統治二論』岩波書店、(Locke, John. (1690) *Two Treatises of Government*, London: Awnsham Churchill)、pp.384-403、小室直樹訳 (1992) 『憲法序説』学陽書房、(Dicey, Venn. Albert. (1885) *Introduction to the Study of the Law of the Constitution*, London: Macmillan and Co.)、pp.183-192、西山千明訳 (1992) 『隷属への道』東京創元社、(von Hayek, August. Friedrich. (1944) *The Road to Serfdom*, London: Routledge)、pp.276-301。

²⁵加藤 (2010) 前掲書、pp.296-309。

²⁶波多野精一ほか訳 (1959) 『実践理性批判』岩波文庫、(Kant, Immanuel. (1788) *Kritik der praktischen Vernunft*, Riga: Johann Friedrich Hartknoch)、pp.90-100。

²⁷Lazar, Seth. (2024) "Legitimacy, Authority, and Democratic Duties of Explanation," in Sobel, David., & Steven Wall. ed s., *Oxford Studies in Political Philosophy Volume 10*, Oxford: Oxford University Press, p.42.

第2節 責任の空白化

本節では、どのようにしてLAWSは自由主義の要である責任を空白化させるのかを、テクノロジー倫理・構造・法の3つの観点からその可能性を明らかにしていく。

第1項 テクノロジー倫理的観点

まずテクノロジー倫理的観点である。自由主義社会は、責任を前提として成立しているが、LAWSの自律性はこの前提を脅かす。ルーシー・サッチマン（Lucy Suchman）の「状況に埋め込まれた行為」²⁸の理論によれば、LAWSの判断は人間の内省や道徳的熟慮を伴わず、単なるプログラムの実行に過ぎない。そのため予期せぬ結果が生じても、特定の人間に行為の意図や責任を帰属させることは困難である。また、ルシアーノ・フロリディ（Luciano Floridi）が指摘するように、AIエージェントは意味づけや価値判断を欠くため、真の責任主体にはなりえない²⁹。そのためLAWSによる攻撃における責任の主体が不明確であり、責任の空白が生じてしまうのだ。

第2項 構造的観点

次にLAWSを構造的観点から見ると、その運用に関わる主体が多岐にわたるため、責任の所在が分散し、希薄化する恐れがある。チャールズ・ペロー（Charles Perrow）は、組織内では単一の大きな目的のもとに複数の異なる目標が設定され、各部署はそれぞれの目標に基づいて仕事を進めることを指摘した³⁰。ゆえに、分業化が進むほど本質的な判断は困難となる。特に即断即決が求められる軍隊のような組織では意思決定システムが簡略化され、各個人は自らの行動を目標達成のためと正当化しやすくなり、組織全体として責任の所在は曖昧になりやすい³¹。

こうした構造の中でLAWSが導入されると、問題はさらに深刻化する。LAWSには開発者・プログラマー・製造者・システム評価者・指揮官など多くの人間が関与してはいるものの、最終的な意思決定は人間ではなくAIによって行われる。分業的な組織構造の中で判断が分散し、人間の関与が事前的・間接的なものに留まるほど、誰が最終的な決定に対する責任を負うのかが曖昧になる。つまり、LAWSの自律性とそれを取り巻く分業的な構造が相まって、責任の空白を生じさせるのである。

第3項 法的観点

²⁸人間の行動は、予測不可能な状況の変化、内省や道徳的常識に常に対応しながら、その場で「即興的」に生成されるという理論。佐伯胖ほか訳（1999年）『プランと状況的行為——人間-機械コミュニケーションの可能性』産業図書（Suchman, A. Lucy. (1987) *Plans and Situated Actions: The Problem of Human Machine Communication*, Cambridge: Cambridge University Press.）、pp.160-200。

²⁹Floridi, Luciano. (2013) *The Ethics of Information*, Oxford: Oxford University Press, pp.134-160.

³⁰弥永真生ほか（2018）『ロボット・AIと法』有斐閣、pp.300-301。

³¹岡田至雄訳（1973）『企業の行動科学7 組織の社会学』ダイヤモンド社、（Perrow, Charles. (1972) *Complex Organizations: A Critical Essay*, New York: McGraw-Hill）、p.315。

最後に法的観点からLAWSを考察する。第1節で示したように、現在の国際秩序を支える基盤の1つに法の支配がある。これは、すべての権力に対して法の優越を認め、権力行使における恣意性を排除する考えである³²。

法的枠組みにおけるLAWSの議論は、2014年にCCWにおいて始まり、2016年にはGGEの枠組内でLAWSの技術的特徴やIHLとの整合性などが議論されてきた。2019年には、GGEにおいて11の指針（Guiding Principles）が採択され、IHLはLAWSを含む全ての兵器システムに適用されることが明記された³³。しかし、LAWSに対してIHLは実効的に適用可能なのだろうか。

例えば、国際武力紛争に適用されるジュネーヴ諸条約第1追加議定書第48条は、民間人と戦闘員、民物と軍事目標の明確な区別を求め、第51条5項(b)では、軍事的利益に比して過剰な文民被害が予測される攻撃を無差別攻撃と定義している³⁴。これらの規定は、いずれも状況・文脈に応じた人間による判断を前提としていると読み取れ、IHLは本質的に人間の価値判断に基づいていることがわかる。つまり、自律的に判断を行うLAWSにIHLを適用することは困難であり、IHLのみではLAWSに対して法の支配を十分に機能させることはできないのではないかと。

既存法の適用が難しい可能性があることに加え、新しい枠組みを作り出すことも容易ではない。CCWでは全会一致制が採用されているため、少数の反対国の存在により新たな国際規範の形成が困難となっている³⁵。また、クラスター爆弾のように実戦で使用されてから明確なルールが策定されるまでに大幅な時間差がある点にも見られるように、急速に発展する軍事技術に対し、国際法が後追いで対応せざるをえないという法的な課題がLAWS規制にも顕著に表れている³⁶。以上から、法の支配が十分に行き届いているとは考えづらく、これまでの規範にそぐわない形でのLAWSの利用が黙認されてしまう危険性がある。

本節では、テクノロジー倫理的観点・構造的観点・法的観点において、LAWSが責任の空白化や、法の支配の形骸化を生むことを示した。このような状況は、社会が無秩序や恣意的な判断が横行する状態に陥る可能性を孕んでおり、自由主義的な観点から見て健全な状態とは言い難い。

第3節 まとめ

³²外務省（2023年6月30日）「第3章 国益と世界全体の利益を増進する外交」https://www.mofa.go.jp/mofaj/gaiko/bluebook/2023/html/chapter3_01_06.html、（2025年6月7日閲覧）、郭舜（2015）「国際法の課題としての世界正義」『世界法年報』34、p.50。

³³United Nations, Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, CCW/GGE.1/2019/3, (September 25, 2019), p.13.

³⁴United Nations. (June 8, 1977) “Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol 1),” <https://www.ohchr.org/en/instruments-mechanisms/instruments/protocol-additional-geneva-conventions-12-august-1949-and>, (accessed May 17, 2025).

³⁵小金澤鋼一（2023）「国連における自律型兵器規制・禁止に関する論争の動向と現状」『日本の科学者』58(4)、pp.5-7。

³⁶足立研幾（2007）「通常兵器ガヴァナンスの発展と変容——レジーム間の相互作用を中心に」『国際政治』2007(14)8、p.109、軍縮会議日本政府代表部（2023年6月8日）「クラスター弾に関する条約」https://www.disarm.emb-japan.go.jp/itpr_ja/11_000001_00398.html、（2025年6月7日閲覧）。

本章では、自由主義秩序の根幹をなす責任の概念が、LAWSの出現によってどのように侵食されるかを考察し、ロックやカントの思想が示唆するように、自由は責任と不可分であり、個人の自律性及び国家統治の正当性は、この原理によって基礎づけられると考えた。

しかしながら、LAWSはテクノロジー倫理の観点から責任主体たりえず、さらに、LAWSの自律性とそれを包囲する組織の分業構造は、責任の所在に本質的な空白を生じさせる。加えて、LAWSに対する法の支配は不十分であり、従来の兵器使用原則からの逸脱を看過する可能性を内包すると考えられる。

以上の考察は、LAWSが自由主義に基づく秩序を根本的に揺るがす存在であることを示唆しており、ゆえにその適切な運用と規制に関する喫緊の議論、並びに積極的な対応策の模索が不可欠であるのだ。

第3章 MHCについて

前章では、LAWSにおける責任の空白の問題を指摘したが、この問題に対する応答の一つとして国際的に注目を集めているのが、MHCの概念である。本章では、まずMHCとは何かを明らかにし、次にこの概念がいかにしてLAWSの責任の空白を補完しうるかを論じる。さらに、MHCの実効性を安全保障のジレンマを用いて分析する。これにより、MHCの規範的意義と現実的限界の双方を明確にする。

第1節 MHCとは

MHCとは、2013年にArticle36によって提唱された概念であり、その実現に必要な要素として以下の3点をあげている。

(a) 情報：人間オペレーターや攻撃計画を担当するものは、攻撃対象地域に関する十分な情報、なぜ特定の対象が攻撃対象として提案されたのか、任務の目的、そしてその攻撃によって生じる即時的及び長期的な兵器の効果に関する情報を持っている必要がある。(b) 行動：攻撃の開始には、オペレーターによる積極的な行動が必要である。

(c) 責任：情報を評価し攻撃を実行する責任者は、攻撃の結果に対して責任を持つ必要がある³⁷。

つまりMHCは、兵器の使用において人間が十分な情報を持ち、攻撃を許可し、その結果に対して責任を負うことを求める概念として提唱された。

³⁷Article36. (2013) “Killer Robots: UK Government Policy on Fully Autonomous Weapons,” https://article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf, pp.3-4.

このような理念を基とした国際的な議論も進んでいる。例えば、2020年のGGEにおける議長総括では、「人間と機械の相互作用は、将来の運用的・規範的枠組みを構築する上での基礎である」³⁸と強調されており、GGEにおいても人間の意思を反映する重要性が確認されている。

しかし、MHCの定義は依然として国際的に確立されていない。国連軍縮研究所も「遵守に必要な人間の制御の形態と程度は未だ定義されていない」³⁹と指摘している。この背景には、定義の明確化に反対する国家の存在がある。米国・フランス・ロシア・英国といった主要開発国は、MHCという用語の定義に慎重な態度を示している⁴⁰。これらの国々はすでにLAWSの開発を進めており、開発に対する制限を回避するためだと考えられる⁴¹。

しかし、「人間の制御が倫理的に必要である」という認識は広く共有されている。ヘザー・ロフ (Heather Roff) は「最終的にどのような用語が用いられるにせよ、人間の制御なしに作動する兵器は、いかなる主体にとっても望ましくないという点では、国際的な合意が存在する」⁴²と主張している。

要するに、MHCの定義は未だ曖昧であり、定義を明確化すべきかに関しても議論の余地がある。しかし兵器の使用において、人間が有意に制御することを求めるという理念自体は広く国際的に受け入れられている。

第2節 責任の空白を補完しうるMHC

前章第2節において倫理・構造・法の3つの観点からLAWSは責任の空白を引き起こすことを示した。本節では、その空白をMHCがいかにして補完しうるのかを検討する。

倫理的観点において、LAWSのようなAIエージェントは倫理的責任主体にたりえないことを指摘した。しかしMHCは十分な情報が与えられる事を担保する仕組みであるため、人間が自身で意味づけすることや、価値基準に沿った判断を下すことが可能になる。人間らしさが存在しないがための倫理的課題であったが、AIと人間の判断基準の尺度が接続すれば、責任の所在を明確にできる。

次に構造的観点においては、LAWSの自律性と組織の分業的な構造が複合的に作用して責任の所在が曖昧になることについて言及した。しかしながら、Article36は、MHCにおける重要な要素として兵器システムの設計上の予測可能性・信頼性・透明性の3要素を提示している⁴³。つまり、MHCを遵守するためには、兵器システムの開発の段階から、人間がその作動原理や行動を理解・予測できる設計がなされ、意図通りに安定して機能すること、そしてその運用過程や意思決定の履歴が外部から検証可能であ

³⁸United Nations, Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, CCW/GGE.1/2020/WP.7, (April 21, 2021), p.8.

³⁹Spaziani, Alice et al. (2021) "Unidir on Lethal Autonomous Weapons: Mapping our Research to the Discussions of the GGE on LAWS," p.4.

⁴⁰Acheson, Ray., & Pytlak, Allison. (2019) "CCW Report, Vol.7, No.7," <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2019/gge/reports/CCWR7.7.pdf>, p.2.

⁴¹Ibid., pp.2-3.

⁴²Roff, M. Heather. (2016) "Meaningful Human Control or Appropriate Human Judgment? The Necessary Limits on Autonomous Weapons," https://article36.org/wp-content/uploads/2016/12/Control-or-Judgment_-Understanding-the-Scope.pdf, p.2.

⁴³Article36. (2016) "Key elements of meaningful human control," <https://www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf>, p.4.

ることが不可欠である。したがって、MHCはどの段階で誰がどのように関与したのかを説明可能にする役割を果たすものであり、この観点において責任の所在を明確にする枠組みを提供する。

最後に法的観点について述べる。ヒューマン・ライツ・ウォッチは、MHCがIHLの原則、特に区別と均衡性の原則と合致し、法の遵守を促進するものであると主張している⁴⁴。ゆえに、MHCはIHLの遵守を技術的に担保する手段としても位置づけられる。最終的な意思決定に人間が関与することで、非戦闘員との区別原則や攻撃の均衡性といったIHLの基本原則をAIシステムが逸脱するリスクを抑制できる。加えて、事後的な説明責任や法的追及の場面においても、誰がどのように意思決定に関与したのかを特定可能にし、国際法に基づく責任追及の基盤となりうる。

このように、倫理的・構造的・法的観点のいずれにおいても、MHCはLAWSが内包する責任の空白を補完する規範的枠組みとして機能するものと考えられる。

第3節 安全保障のジレンマ

前節において、MHCの概念はLAWSの導入によって生じる責任の空白を補完しうる有効な枠組みであることを示した。しかし、MHCは理念としては魅力的であっても、実際の国際政治の文脈においてその適用を実現するのは極めて困難である。このことを安全保障のジレンマを用いて説明する。

安全保障のジレンマは、しばしば囚人のジレンマと類似の非協力的状況に陥るとされるが、両者は本質的に異なる性質を持つ概念である⁴⁵。まず囚人のジレンマは、当事者間での会話ができない状況を前提とする⁴⁶。したがって、集団にとって合理的な協調選択が存在していても、個人が合理的な行動をとった結果として非協力の均衡に陥るという構造を持つ⁴⁷。一方、安全保障のジレンマとは、自らの安全を追求する行為が、他者には脅威と認識され、それがさらなる不信を招くことで、結果的に対立と不安定化を生む構造を指す⁴⁸。ここでは、当該者間のコミュニケーションが可能であるため協調の余地が存在する⁴⁹。ゆえに、各国の選択は固定的なものではなく、相互作用の中で動的に変化する。すなわち、安全保障のジレンマは固定されたものではなく、利得・コストの変動によって構造的に変化するため、囚人のジレンマのような静的ゲームとは異なり、協調への移行が可能である。

【表2 安全保障のジレンマ】

	B：協力 (C)	B：無視 (D)
A：協力 (C)	r, r	0, $t + \beta$

⁴⁴Human Rights Watch. (April 11, 2016) op. cit.

⁴⁵鴨武彦訳 (1984) 『安全保障のジレンマ——核抑止・軍拡競争・軍備管理をめぐって』有斐閣、(Russett, M. Bruce. (1984) *The Prisoners of Insecurity: Nuclear Deterrence, The Arms Race, and Arms Control*, San Francisco: Chandler Publishing)、pp.142-143。

⁴⁶岡田章 (2014) 『ゲーム理論・入門——人間社会の理解のために (新版)』有斐閣、pp.91-92。

⁴⁷同上書、pp.95-96。

⁴⁸Herz, H. John. (1950) "Idealist Internationalism and the Security Dilemma," *World Politics*, 2(2), p.157.

⁴⁹鴨 (1984) 前掲書、p.142。

A : 無視 (D)	$t + \alpha, 0$	$w + \alpha, w + \beta$
------------	-----------------	-------------------------

※r: AB両国がMHCを適用した際の利得、t: 一方がLAWSを使用し、他方がMHCを適用し続けた際に、使用した側が得る利得、w: AB両国がLAWSを使用した際の利得、 α : A国がMHCを適用しないことのコスト、 β : B国がMHCを適用しない際にA国が被るコスト

(出典：栗崎周平「集団的自衛権と安全保障のジレンマ」；Kydd, H. Andrew. “*Trust and Mistrust in International Politics*”；Schultz, A. Kenneth. “The Politics of Risking Peace: *Do Hawks or Doves Deliver the Olive Branch?*,”より筆者作成。)

表2は、安全保障のジレンマを示すゲームの利得表である。ここで、両国がMHCを適用する場合を(CC)、一方がLAWSを導入し他方がMHCを適用する場合を(DC)または(CD)、両国がLAWS導入に傾く場合を(DD)と定義する。以下では、A国の視点から考察する。

まず利得の関係として、 $t + \alpha > r > w + \alpha > 0$ が前提とされる。この関係は以下のような論理に基づいている。

第一に、 $w + \alpha > 0$ 。(DD)においては両国がLAWSを導入するため戦略的優位は得られないが、相手に一方的な利得を許す状況を回避できるため、 $w + \alpha$ は負にはならない。

第二に、 $r > w + \alpha$ 。(CC)・(DD)はどちらも軍事的にはバランスが取れている。しかし、(DD)においては責任の空白・誤作動のリスク・交戦のハードルの低下といったコストが生じるため、(C)に比して利得が低い状態にある。

第三に、 $t + \alpha > r$ 。(DC)は、短期的には戦略的優位を確保し、協調よりも高い利得をもたらす。第1章第3項で述べたとおり、LAWS導入には多くの軍事的利点がある上、現在は法整備が不十分であるため、導入に伴うコスト(α)は限定的である。さらに、LAWSが実戦投入されていない現段階では、先行導入による利得(t)は大きくなり、国家にとって先制的に優位を獲得するインセンティブは極めて高くなる。

このような状況では、両国はともに(D)戦略を選ぶインセンティブを持つ。まず、相手国が先にLAWSを導入する(CD)または(DC)という最悪のシナリオを回避するため、少なくとも(DD)の状態に持ち込みたいという安全保障上の欲求が働くため。さらに、自国のみがLAWSを導入すれば、協調状態(CC)よりも高い利得($t + \alpha$)を得られる可能性が存在するためである。これらの動機が重なることで、本来は(CC)がパレート最適⁵⁰であるにも関わらず、結果としては(DD)という非協力的かつパレート非効率な状態がナッシュ均衡⁵¹となる。ゆえに、両者は囚人のジレンマと類似の非協力状態に陥っていると言える。

では、このようなジレンマ状態をいかにして乗り越えるか。先に述べたように安全保障のジレンマは利得・コストの調整によって、協調状態に近づけることができる。この点に関して、ロバート・ジャービス(Robert Jervis)は、以下の三点が協調状態の形成に有効としている。

⁵⁰パレート最適とは、誰かの効用を犠牲にしなければ、他の誰かの効用を高めることができない状態。

⁵¹ナッシュ均衡とは、相手が戦略を変更しない限り、どのプレイヤーも自分だけが戦略を変更しても利得を増やせないような状態。

- (a) 相互協力の利得を高める、または裏切り（非協力）による追加的利益を減少させる
- (b) 非協力戦略の利得を減少させる、またはそのコストを高める
- (c) 協力行動を促すような制度的・認知的枠組みを整備する

この理論をMHCの文脈に応用するならば、MHCを適用した際の利得（ r ）を高めるとともに、それ以外のコスト（ α ）を引き上げることで、 $r > t + \alpha$ 、 $r > w + \alpha$ とすることができる。MHC適用による協調（CC）がナッシュ均衡へと近づき、各国は戦略的合理性から協力的選択をとるようになる。

本節では、現在のMHCの適用・非適用の議論が安全保障のジレンマにおいて非協力状態に陥っていることを示した。LAWS導入による軍事的優位性がMHC適用の利得を上回りジレンマが発生している一方で、利得・コストの調整によって協力が合理的に選択される道筋も存在する。技術の急速な進歩により、LAWSの開発を完全に封じることは不可能であるが、その使用や制御の意図に影響を与えることは可能である。

第4節 まとめ

本章では、LAWSにおける責任の空白への応答として注目されるMHCについて論じた。まず、MHCの定義は依然として曖昧であるものの、その理念は国際的に受け入れられていることを示した上で、MHCがいかに責任の空白を補完しうるのかを検討した。MHCは判断主体としての人間の制御を明確にし、説明責任の構造を担保することで、倫理・構造・法的な責任を支える枠組みとなる。しかし、各国が自己利益を優先するがゆえに、MHCを実際に適用することが困難であることを、安全保障のジレンマの観点から検証した。

以上より、序章で提示した問いである「MHCはLAWSにおける責任の空白をどの程度埋めることができるのか」に対して、本章は、MHCはLAWSにおける責任の空白に対して、倫理的・制度的に責任を支える枠組みである一方、LAWSの軍事的優位性や安全保障のジレンマの影響により、MHCの原則を実際に適用するには限界があるという仮説を裏付ける内容となった。

第4章 MHCの実効性確保に向けて

本章では、前章で論じた安全保障のジレンマを緩和し、各国が実効的にMHCをLAWSに適用するための制度的対応として、第三者審査機関の設立可能性を検討する。

第1節 IAEAから見る第三者審査機関の意義

前章で示した安全保障のジレンマ状態を緩和する方策として有望なのは、第三者審査機関の設立であると考えられる。これにより、各国のLAWS使用及びMHC適用の透明性が確保され、信頼関係が醸成されることで、MHC遵守の動機づけが強化されることが期待できる。第三者審査機関の設立において参考と

なるのが、国際原子力機関（International Atomic Energy Agency: IAEA）⁵²の制度的枠組みである。核兵器に用いられる原子力とLAWSに用いられるAI技術は、その特性や応用分野において多くの違いを持つものの、いずれも平和的利用と軍事的応用の双方が可能であるという点で共通している。したがって、IAEAが原子力の平和的利用を推進しつつその軍事転用を防止してきたことは、LAWSに関する第三者審査機関の構想においても示唆に富む⁵³。

IAEAの設立の目的は、核不拡散及び原子力の平和的利用の促進であり、この二重の目標を達成するため、保障措置・技術協力・安全確保という三本柱に基づいた活動を展開している⁵⁴。保障措置では、IAEA事務局が検証機能を担い、核物質などが軍事転用されないように、査察や監視といった検証を行っている⁵⁵。技術協力では、全世界における原子力の平和的利用のための研究・開発・実用化を奨励・援助し、加盟国間の役務や物質・施設の供給を仲介する⁵⁶。安全確保では、国連機関などと協力して、健康の保護や人命・財産の安全を確保するための安全基準を定めている⁵⁷。

このようにIAEAは三つの柱を通じて、国家間の不信や対立を和らげ、協調を促す環境を整えている。こうしたIAEAの多層的な活動は、原子力分野におけるリスクと利益を適切に調整し、相互の信頼構築を支える有効な枠組みとして機能している。

この例からは適切に設計された第三者審査機関の介入が協調を促進しうることがわかり、同様の機関の設立がMHCの実効性を担保する重要な一歩となることを示唆している。

第2節 MHCに関する第三者審査機関の三機能

想定されるMHCにおける第三者審査機関は、以下三つの機能を軸に構築されることが望ましいと考える。第一に、MHCの履行監視機能である。各国におけるLAWSの開発・運用プロセスに対して、外部の専門家チームが査察・検証を行うことがMHCの遵守状況を明らかにするとともに、不正な運用が発覚する可能性を高め、各国の非協調的行動に対する抑止力として機能することが期待できる。第二に、AI技術の安全な提供・普及機能である。IAEAが平和利用を前提とした技術援助をしてきたように、MHC準拠を前提としたAI技術の支援を行うことで、開発支援・訓練・制度構築を促進し、MHCを適用することの利益を明確にする。特に途上国においては、信頼性のある経路から技術を享受し、不透明な第三勢力から軍用AIを導入するリスクを軽減することができるだろう。また、無秩序な技術拡散を抑制できるため、すでに高度な自律性を伴う兵器を有する先進国にとっても有益だと考えられる。第三に、事故リスクの軽減と評価体制の構築機能である。誤爆や機能障害といったリスクを国際的に共有し、事故の

⁵²Chesterman, Simon. (2021) "Weapons of mass disruption: artificial intelligence and international law," *Cambridge International Law Journal*, 10(2), p181.

⁵³外務省（2023年4月27日）「国際原子力機関（IAEA）の概要」https://www.mofa.go.jp/mofaj/gaiko/atom/iaea/iaea_g.html、（2025年6月5日閲覧）。

⁵⁴外務省（2024年12月10日）「技術協力・原子力科学技術応用」https://www.mofa.go.jp/mofaj/dns/inec/page22_003032.html、（2025年6月3日閲覧）、渡邊直行（2016）「国際保健医療における国際原子力機関（IAEA）の取り組みについて」『保健医療科学』65(4)、p.427。

⁵⁵外務省（2024年5月21日）「国際原子力機関（IAEA）保障措置」<https://www.mofa.go.jp/mofaj/gaiko/atom/iaea/kyoutei.html>、（2025年6月5日閲覧）。

⁵⁶外務省（2023年4月27日）前掲webサイト。

⁵⁷外務省（2023年4月27日）前掲webサイト。

未然防止に向けた規約を策定・整備することで、AI兵器運用における安全性と信頼性を高めることも期待される。

このように、IAEAの制度設計を参照した第三者審査機関は、MHCの文脈においても、国家間の不信と対立を和らげ、協調を促す枠組みとして機能しうる。協調による利得の増加や裏切りに伴うコストの上昇、そして裏切られた場合の損失の低減という三方向からの調整により、各国が戦略的合理性に基づいて行動する場合でも、MHCを採用・維持する動機を構造的に形成できるのである。

では、MHCの議論の中心的な場となっているGGEは第三者審査機関的な役割を果たせるのだろうか。現在のGGEの機能は主にLAWSに関する規範形成に留まっている⁵⁸。ゆえに、規範が実際に守られているかを評価することができず、各国が自国の安全保障上の都合に応じてMHCを解釈し、運用にばらつきや不透明性が生じている。この問題の危険性は、生物兵器禁止条約（Biological Weapons Convention: BWC）の事例からも明確である。BWCは生物兵器の開発・生産・貯蔵・保有を戦時・平時を問わず包括的に禁止する唯一の多国間の法的枠組みであるが、第三者審査機関が存在せず、加盟国の遵守状況を客観的に監査できない⁵⁹。条約違反が疑われても証明が困難であり、実際にBWC締約国であるイラクが生物兵器を秘密裏に研究していたことが発覚するなどの問題も生じている⁶⁰。このことは、兵器規制において規範の策定だけでは不十分であり、検証・監視機能を伴う制度が不可欠であることを示している。

以上の背景から、MHCに関して、規範形成と並行して遵守状況を検証できる第三者審査機関の設立が求められるのではないか。各国が恣意的な解釈によって規範を軽視することを防ぎ、MHCの実効性を高めることに繋がる。

さらに、なぜLAWS自体の規制を行う機関ではなく、あえてMHCの遵守を検証する機関を提案するのかについても説明が必要である。第1章で論じた通り、LAWSは軍事的な戦略価値や効率性が高く、多くの国が開発・運用を続ける強い動機を持っている。そのため、LAWS全体の開発や保有の規制が進むと、各国の安全保障上の利得が減少し、協調への参加動機が低下すると考えられる。したがって、LAWSの全面的な規制ではなく、MHCという実践的かつ倫理的指標の遵守を求める形で協調を促す方が現実的であろう。MHCを基軸とすることで、最低限の軍事的利益は維持しつつ、その運用に一定の制約と透明性を設けることができる。これにより、多くの国が参加しやすい協調体制が構築される。

第3節 まとめ

本章では、MHCの実効性を確保する制度的対応として、IAEAを参考にした第三者審査機関の設立可能性と有効性を検討した。履行監視・AI技術の支援・事故防止と評価体制の三つの機能を備えた第三者審査機関は、国家間の協調的行動を促進する動機を提供しうるだろう。また、GGEやBWCの事例からも、実効的な検証制度の必要性を示した。

⁵⁸外務省（2024年3月15日）「特定通常兵器使用禁止制限条約——自律型致死兵器システムに関する政府専門家会合の開催」https://www.mofa.go.jp/mofaj/dns/ca/pagew_000001_00431.html、（2025年6月6日閲覧）。

⁵⁹外務省（2024年2月16日）「生物兵器禁止条約（BWC）概要」<https://www.mofa.go.jp/mofaj/gaiko/bwc/bwc/gaiyo.html>、（2025年6月3日閲覧）。

⁶⁰外務省（2003年10月）「イラクにおける大量破壊兵器問題（参考）」https://www.mofa.go.jp/mofaj/area/iraq/th_heiki.html、（2025年6月5日閲覧）。

終章

本稿では、「MHCはLAWSにおける責任の空白をどの程度埋めることができるのか」という問いに対して、「MHCは倫理的・制度的に責任を支える枠組みである一方、MHCの原則を適用するには限界がある」という仮説に基づき検証を行った。

第1章ではLAWSの定義が曖昧であることを指摘しつつ、LAWSの導入がもたらす軍事的優位性に注目した。次に第2章では倫理・構造・法の観点から、LAWSがいかに自由主義に基づく秩序において責任の空白を生み出すかを考察した。第3章ではMHCの意義を明らかにすると同時に、安全保障のジレンマを用いてMHCの適用が容易ではないことを検証した。

これらの検討を踏まえ本稿は、MHCは倫理的・制度的な説明責任を支える重要な枠組みではあるものの、安全保障のジレンマに直面する国家の戦略行動により、その適用が困難になるという結論に至った。つまり、各国が自国の利益を追求する限り、MHCの理念と実際の運用との間に乖離が生じることは避けられない。

以上を踏まえ、第4章ではLAWSの管理を実現する仕組みが必要であるとし、MHCを適用させるための制度的対応として、第三者審査機関の設立を提案した。MHCに基づく第三者審査機関の設立は、規制が困難な現状においても最低限の人道的・法的基準を守るための現実的な対応策となりうる。LAWSの全面的禁止が困難である状況下において、管理に向けた第一歩として国際社会はMHCを軸とした制度的枠組みの構築に取り組むべきである。

しかしながら、本稿の提案にはいくつかの課題が残っている。まず、MHCという概念自体の明確な定義は未だ存在せず、その内容や運用基準について国際的合意を形成することも、今後の制度設計における前提条件である。この定義の曖昧さが制度的管理の実効性を阻むリスクもあるため、MHCへの理解をどのように統一していくかが今後の重要な研究課題である。

また、本稿で提案した第三者審査機関については、現実的に設立可能かどうか、また設立にあたって何が必要かについても具体的に検討する必要がある。本稿がモデルとしたIAEAにおいても、1990年代初頭のイラクや北朝鮮の核開発疑惑を受け、包括的保障措置だけでは未申告の核活動を検知・抑止できないという限界が露呈している⁶¹。第三者審査機関の判断が実際に国際的拘束力を持ちうるか再考する必要がある。

以上を踏まえれば、本稿の提案はMHCを軸としたLAWSの管理に向けた一つの可能性を示すものであるが、その実効性を担保するためには、過去の成功事例に基づく仕組みを効果的に適用しつつ、現代特有の課題についても十分に考慮することが求められる。現行の議論に盲目的になるのではなく、LAWSやMHCの長短を把握した上で、より実践的なアプローチを能動的に創出する努力が必要だ。また、安全保障のジレンマをはじめとする国際政治の課題も十分に踏まえ、妥協点を模索していくことも求められる。本稿で扱ったLAWSは、今後国際社会により大きな影響を及ぼすものであり、継続的な議論が不可欠である。

⁶¹外務省（2016）「軍縮・不拡散外交（第七版）」<https://www.mofa.go.jp/mofaj/files/000145531.pdf>, p.86。

参考文献

<邦文著書>

岡田章（2014）『ゲーム理論・入門——人間社会の理解のために（新版）』有斐閣。

弥永真生・穴戸常寿・工藤郁子・大屋雄裕・山本龍彦・横田明美...岩本誠吾（2018）『ロボット・AIと法』有斐閣。

<邦文雑誌論文>

足立研幾（2007）「通常兵器ガバナンスの発展と変容——レジーム間の相互作用を中心に」『国際政治』2007(148)、104-117。

郭舜（2015）「国際法の課題としての世界正義」『世界法年報』34、34-57。

栗崎周平（2017）「集団的自衛権と安全保障のジレンマ」『年報政治学』68(2)、36-64。

小金澤鋼一（2023）「国連における自律型兵器規制・禁止に関する論争の動向と現状」『日本の科学者』58(4)、4-9。

渡邊直行（2016）「国際保健医療における国際原子力機関（IAEA）の取り組みについて」『保健医療科学』65(4)、424-441。

<邦文報告書>

外務省（2016）「軍縮・不拡散外交（第七版）」<https://www.mofa.go.jp/mofaj/files/000145531.pdf>。

<邦文webサイト・動画>

外務省（2024年12月10日）「技術協力・原子力科学技術応用」https://www.mofa.go.jp/mofaj/dns/inec/page22_003032.html、（2025年6月3日閲覧）。

———（2024年6月24日）「自律型致死兵器システム（LAWS）について」
https://www.mofa.go.jp/mofaj/dns/ca/page24_001191.html、（2025年6月3日閲覧）。

———（2024年5月21日）「国際原子力機関（IAEA）保障措置」
<https://www.mofa.go.jp/mofaj/gaiko/atom/iaea/kyoutei.html>、（2025年6月5日閲覧）。

———（2024年3月15日）「特定通常兵器使用禁止制限条約——自律型致死兵器システムに関する政府専門家会合の開催」
https://www.mofa.go.jp/mofaj/dns/ca/pagew_000001_00431.html、（2025年6月6日閲覧）。

———（2024年2月16日）「生物兵器禁止条約（BWC）概要」
<https://www.mofa.go.jp/mofaj/gaiko/bwc/bwc/gaiyo.html>、（2025年6月3日閲覧）。

———（2023年6月30日）「第3章 国益と世界全体の利益を増進する外交」
https://www.mofa.go.jp/mofaj/gaiko/bluebook/2023/html/chapter3_01_06.html、（2025年6月7日閲覧）。

———（2023年4月27日）「国際原子力機関（IAEA）の概要」
https://www.mofa.go.jp/mofaj/gaiko/atom/iaea/iaea_g.html、（2025年6月5日閲覧）。

——— (2003年10月) 「イラクにおける大量破壊兵器問題 (参考)」

https://www.mofa.go.jp/mofaj/area/iraq/th_heiki.html、(2025年6月5日閲覧)。

軍縮会議日本政府代表部 (2023年6月8日) 「クラスター弾に関する条約」

https://www.disarm.emb-japan.go.jp/itpr_ja/11_000001_00398.html、(2025年6月7日閲覧)。

<邦訳書>

岡田至雄訳 (1973) 『企業の行動科学7 組織の社会学』ダイヤモンド社、(Perrow, Charles. (1972) *Complex Organizations: A Critical Essay*, New York: McGraw-Hill)。

加藤節訳 (2010) 『完訳統治二論』岩波書店、(Locke, John. (1690) *Two Treatises of Government*, London: Awnsham Churchill)。

鴨武彦訳 (1984) 『安全保障のジレンマ——核抑止・軍拡競争・軍備管理をめぐって』有斐閣、(Russell, M. Bruce. (1984) *The Prisoners of Insecurity: Nuclear Deterrence, The Arms Race, and Arms Control*, San Francisco: Chandler Publishing)。

小室直樹訳 (1992) 『憲法序説』学陽書房、(Dicey, Venn. Albert. (1885) *Introduction to the Study of the Law of the Constitution*, London: Macmillan and Co)。

佐伯胖・水川喜文・上野直樹・鈴木永幸訳 (1999) 『プランと状況的行為——人間-機械コミュニケーションの可能性』産業図書、(Suchman, A. Lucy. (1987) *Plans and Situated Actions: The Problem of Human-Machine Communication*, Cambridge: Cambridge University Press)。

西山千明訳 (1992) 『隷属への道』東京創元社、(von Hayek, August. Friedrich. (1944) *The Road to Serfdom*, London: Routledge)。

波多野精一、宮本和吉、篠田英雄訳 (1959) 『実践理性批判』岩波文庫、(Kant, Immanuel. (1788) *Kritik der praktischen Vernunft*, Riga: Johann Friedrich Hartknoch)。

細谷雄一訳 (2012) 『リベラルな秩序か帝国か (上) ——アメリカと世界政治の行方』勁草書房、(Ikkenberry, G. John. (2006) *Liberal Order and Imperial Ambition: Essays on American Power and International Order*, Cambridge: Polity Press)。

<欧文著書>

Arkin, C. Ronald. (2009) *Governing Lethal Behavior in Autonomous Robots*, Boca Raton: CRC Press.

Floridi, Luciano. (2013) *The Ethics of Information*, Oxford: Oxford University Press.

Kydd, H. Andrew. (2005) *Trust and Mistrust in International Relations*, Princeton: Princeton University Press.

<欧文編集論文>

Lazar, Seth. (2024) "Legitimacy, Authority, and Democratic Duties of Explanation," in Sobel, David., & Wall, Steven. eds., *Oxford Studies in Political Philosophy Volume 10*, Oxford: Oxford University Press, 28-56.

<欧文雑誌論文>

- Chesterman, Simon. (2021) "Weapons of mass disruption: artificial intelligence and international law," *Cambridge International Law Journal*, 10(2), 181-203.
- Herz, H. John. (1950) "Idealist Internationalism and the Security Dilemma," *World Politics*, 2(2), 157-180.
- King, Anthony. (2024) "Digital Targeting: Artificial Intelligence, Data, and Military Intelligence," *Journal of Global Security Studies*, 9(2), 1-16.
- Marchant, E. Gary., Allenby, Braden., Arkin, C. Ronald., Borenstein, Jason., Gaudet, M. Lyn., Kittrie, Orde., ... Silberman, Jared. (2011) "International Governance of Autonomous Military Robots," *Science and Technology Law Review*, 12, 272-315.
- Santoni de Sio, Filippo., & Mecacci, Giulio. (2021) "Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them," *Philosophy & Technology*, 34, 1057-1084.
- Santoni de Sio, Filippo. & van den Hoven, Jeroen. (2018) "Meaningful Human Control over Autonomous Systems: A Philosophical Account," *Frontiers in Robotics and AI*, 5(15), 1-14.
- Schultz, A. Kenneth. (2005) "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?," *International Organization*, 59(1), 6-8.
- Thurnher, S. Jeffrey. (2012) "No One at the Controls: Legal Implications of Fully Autonomous Targeting," *Joint Force Quarterly*, 67(4), 77-84.

<欧文報告書>

- Article36. (2016) "*Key Elements of Meaningful Human Control*,"
<https://www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf>.
- (2013) "*Killer Robots: UK Government Policy on Fully Autonomous Weapons*," https://article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf.
- Acheson, Ray., & Phtlak, Allison. (2019) "*CCW Report, Vol.7, No.7*,"
<https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2019/gge/reports/CCWR7.7.pdf>.
- Bode, Ingvid., & Watts, F. A. Tom. (2023) "*Loitering Munitions and Unpredictability Autonomy in Weapon Systems and Challenges to Human Control*,"
https://www.researchgate.net/publication/371351342_Loitering_Munitions_and_Unpredictability_Autonomy_in_Weapon_Systems_and_Challenges_to_Human_Control.
- Federal Foreign Office. (2020) "*German commentary on "operationalizing all eleven guiding principles at a national level as requested by the chair of the 2020 Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS) within the Convention on Certain Conventional Weapons (CCW)*,"
<https://documents.unoda.org/wp-content/uploads/2020/07/20200626-Germany.pdf>.
- Human Rights Watch. (2012) "*Losing Humanity: The Case against Killer Robots*," https://www.hrw.org/sites/default/files/reports/arms1112_ForUpload.pdf.
- (2016) "*Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban*," https://www.hrw.org/sites/default/files/report_pdf/arms1216_web.pdf.

- Michel, Arthur. Holland. & Gettinger, Dan. (2017) “*Loitering Munitions*,” <https://dronecenter.bard.edu/files/2017/02/CSD-Loitering-Munitions.pdf>.
- People’s Republic of China. (2018) “*Position Paper*,” https://unoda-documents-library.s3.amazonaws.com/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_%282018%29/CCW_GGE.1_2018_WP.7.pdf.
- Roff, M. Heather. (2016) “*Meaningful Human Control or Appropriate Human Judgment? The Necessary Limits on Autonomous Weapons*,” https://article36.org/wp-content/uploads/2016/12/Control-or-Judgment_-Understanding-the-Scope.pdf.
- Spaziani, Alice., Michel, Arthur. Holland. & Anand, Alisha. (2021) “*Unidir on Lethal Autonomous Weapons: Mapping our Research to the Discussions of the GGE on LAWS*,” <https://unidir.org/wp-content/uploads/2023/05/UNIDIR-on-Lethal-Autonomous-Weapons-Final.pdf>.
- U.S. Department of Defense. (2023) “AUTONOMY IN WEAPON SYSTEMS,” <https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf>.
- (2007) “*Unmanned Systems Roadmap 2007-2032*,” <https://apps.dtic.mil/sti/pdfs/ADA475002.pdf>.
- United Kingdom. (2024) “*Input to UN Secretary-General’s Report on Lethal Autonomous Weapons Systems (LAWS)*,” [https://docs-library.unoda.org/General_Assembly_First_Committee_-Seventy-Ninth_session_\(2024\)/78-241-UK-EN.pdf](https://docs-library.unoda.org/General_Assembly_First_Committee_-Seventy-Ninth_session_(2024)/78-241-UK-EN.pdf).

< 欧文webサイト・動画 >

- Human Rights Watch. (April 11, 2016) “Killer Robots and the Concept of Meaningful Human Control,” <https://www.hrw.org/news/2016/04/11/killer-robots-and-concept-meaningful-human-control>, (accessed June 6, 2025).
- (April 23, 2013) “Arms: New Campaign to Stop Killer Robots: Swift Action Needed to Prevent Fully Autonomous Weapons,” <https://www.hrw.org/news/2013/04/23/arms-new-campaign-stop-killer-robots>, (accessed June 7, 2025).
- International Committee of the Red Cross. (July 26, 2022) “What you need to know about autonomous weapons,” <https://www.icrc.org/en/document/what-you-need-know-about-autonomous-weapons>, (accessed June 6, 2025).
- United Nations. (June 8, 1977) “Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol 1),” <https://www.ohchr.org/en/instruments-mechanisms/instruments/protocol-additional-geneva-conventions-12-august-1949-and>, (accessed May 17, 2025).
- United Nations Office for Disarmament Affairs. (n.d.) “Lethal Autonomous Weapon Systems (LAWS),”

<https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/>, (accessed June 6, 2025).

+972Magazine. (April 3, 2024) “‘Lavender’: The AI machine directing Israel’s bombing spree in Gaza,” <https://www.972mag.com/lavender-ai-israeli-army-gaza/>, (accessed June 06, 2025).

<国際機関の決議や文書>

UN General Assembly, Universal Declaration of Human Rights, A/RES/217(III) (December 10, 1948).

United Nations, Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, CCW/GGE.1/2020/WP.7, (April 21, 2021), <https://docs.un.org/en/CCW/GGE.1/2020/WP.7>.

——— Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, CCW/GGE.1/2019/3, (September 25, 2019), https://documents.unoda.org/wp-content/uploads/2020/09/CCW_GGE.1_2019_3_E.pdf.

UN Security Council, Final report of the Panel of Experts on Libya established pursuant to resolution 1973 (2011), UN Document No.S/2021/229, (March 8, 2021).

Russian Federation, Approaches of the Russian Federation to the issue of emerging technologies in the area of lethal autonomous weapons systems, CCW/GGE.1/2024/WP.2, (March 1, 2024), [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_\(2024\)/CCW-GGE.1-2024-WP.2.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2024)/CCW-GGE.1-2024-WP.2.pdf).