

自律型致死兵器システム (LAWS) 規制議論の 動向と今後の展望

—— 2023 年度 ——

岩 本 誠 吾

1. はじめに
2. CCW 枠内での議論
 - (1) GGE 第1会期
 - (a) GGE 第1会期提出の作業文書
 - (b) GGE 第1会期での議論
 - (2) GGE 第2会期
 - (a) 第2会期前の非公式協議
 - (b) GGE 第2会期提出の作業文書
 - (c) GGE 第2会期での議論
 - (3) GGE 報告書と CCW 締約国会議
3. CCW 枠外での議論
 - (1) 米国防総省指令の改訂
 - (2) 国際会議と地域会議
 - (a) オランダ・韓国共催「軍事領域での責任ある AI (REAIM) サミット」
 - (b) コスタリカ主催「CELAC 地域会議」
 - (c) ルクセンブルグ主催国際会議
 - (d) トリニダード・トバゴ主催「CARICOM 地域会議」
 - (e) フィリピン主催「インド太平洋地域会議」
 - (3) 国連事務総長と国連総会
 - (a) 国連事務総長イニシアチブ
 - (b) 国連総会での議論
4. まとめにかえて —— 2026 年に向けて
 - (1) 三つ巴対立から 2 項対立への移行
 - (2) CCW 枠外での議論の活性化

(3) 2026 年が勝負年

* 「添付翻訳資料（全文または部分翻訳）」

資料1：「軍事領域での責任ある AI 行動要請（REAIM Call to Action）」

資料2：米国提案の「AI および自律性の責任ある軍事利用に関する政治宣言」

資料3：「自律兵器の社会的・人道的影響に関するラミ・カリブ会議コミュニケ」（ベレン共同声明）

資料4：「AWS に関する CARICOM 宣言」

1. はじめに

近未来兵器と称される「自律型致死兵器システム（Lethal Autonomous Weapons Systems, LAWS^{ローズ}）」の規制問題は、特定通常兵器条約（Convention on Certain Conventional Weapons, CCW）締約国会議で取り上げられた2013年以降、具体的で実質的な成果なく、現在まで引き続き議論されている。LAWS 規制議論の10年目に当たる2022年に、政府専門家会合（Group of Governmental Experts, GGE）の会合は第1会期（3月7～11日）、会期間非公式オンライン会合（4月26・27日、6月1・3日、6月27・29日）および第2会期（7月25～29日）が開催され、最終日にGGE 報告書（CCW/GGE.1/2022/2, 31 August 2022）が採択された。CCW 締約国会議（2022年11月16～18日）は、そのGGE 報告書を基に議論し、最終日に最終報告書（CCW/MSP/2022/7, 24 November 2022）を採択した。

2022年のCCW 枠内でのLAWS 規制の議論は、結果として、法文書によるLAWS 規制派、グッド・プラクティス編集という非法文書によるLAWS 規制派および追加的な法的措置不要派の三つ巴の対立を解消させ

(1) 岩本誠吾「〈資料〉特定通常兵器条約（CCW）締約国会議における自律型致死兵器システム（LAWS）規制議論の動向——2022年度——」『京都産業大学世界問題研究所紀要』38巻2023年3月99-120頁。

ることなく、そのまま 2023 年の CCW 会合に引き継がれた。CCW 締約国会議が LAWS 規制問題の最適なフォーラムであるとの意見がある一方で、硬直化した論争状況を打開すべく、2022 年 10 月に、結果として国連総会決議までには至らなかったが、LAWS に関する 70 か国の共同声明⁽²⁾が国際連合（国連）総会第 1 委員会に提出された。さらに、オランダが、2022 年 9 月に国連総会において 2023 年 2 月に「軍事領域での責任ある人工知能（Artificial Intelligence, AI）サミット」の開催を告知し⁽³⁾、そして、CCW 締約国会議の中でコスタリカやルクセンブルグが 2023 年 2 月や春に自律兵器システムに関する地域会合の開催を予告する⁽⁴⁾といった新たなフォーラムでの LAWS 規制議論の予兆が見られた。

そして、AI に関連して、2022 年 11 月 30 日に Open AI 社が生成 AI 「チャット GPT」を世界に公開したことで、AI が我々現代人にとって一層身近な存在となった。その便利さと共に、そのリスク（フェイク・ニュースの作成、詐欺メールおよびフィッシング詐欺の犯罪に悪用される危険性など）も実感として認識されるようになった。我々は、加速度的な AI の研究開発の進展を目の当たりにして、それが AI の軍事利用や兵器利用に直結するとの脅威認識を一段と高めることとなった。

実際に、AI の軍事利用は、既に AI 意思決定支援システム（AI-Decision Support System, AI-DSS）として戦場で行われている。たとえば、2022 年 2 月からのロシア・ウクライナ戦争において、ウクライナは、米国のデータ解析企業パランティア・テクノロジーズの AI プラットフォーム「ゴサム（Gotham）⁽⁵⁾」を活用して、ロシアに善戦している。また、2023

(2) 70 か国共同声明に関して、「岩本前掲資料」119 頁注 53 参照。

(3) Government of the Netherland, “The Netherland to host international summit on artificial intelligence,” 21-09-2022, <https://www.government.nl/ministries/ministry-of-foreign-affairs/news/2022/09/21/the-netherlands-to-host-international-summit-on-artificial-intelligence>, accessed 6 March 2024.

(4) 「岩本前掲資料」112 頁。

(5) 西側諸国の偵察衛星・早期警戒管制機・無人偵察機・電子偵察機などから入手した情報や一般人の SNS などからの敵の位置情報や被害状況などの膨大な情報がスターリンクを

年10月からのイスラエル・ガザ戦争では、イスラエルは、AIに基づく標的設定システム「ハブソラ (Habsora、=The Gospel・福音)⁽⁶⁾」により大量に生成される有力標的 (power targets、テロリストのアセット) を自動的に空爆や砲撃の対象にしている。

本稿は、2022年でのLAWS規制議論の動向を踏まえて、2023年にCCW 枠内およびCCW 枠外においてどのような議論が行われたのかを概観する (【2022~26年LAWS・軍事AI関連年表】参照)。そして、それらの議論の中から、LAWS規制に関する新たな兆しを読み取るとともに、LAWS規制に関する今後の展望を描いてみることを目的とする。

、通じて収集され、AIがそれを瞬時に分析し、敵の正確な位置情報 (1日300か所) や効率的な攻撃法を立案する。軍司令官は、提示された複数の軍事作戦案の中から最適の作戦を選択し、前線の将校に命令するというシステムである。TBS NEWS DIG 「「こっそり歴史を変えた」〜ウクライナ善戦の裏でアメリカ民間企業“魔法の力”【報道1930】」2023年2月8日 <https://newsdig.tbs.co.jp/articles/-/318447?page=2>, 2024年4月2日閲覧; 大前研一「最新AIのChatGPTの衝撃! 天才経営者は何を考えているのか」『プレジデント』2023年3月31日号81頁。日本も、バルンティア・テクノロジーズのAIプラットフォームを念頭に置いているのか、「防衛力整備計画」の中に、「複雑かつ高速に推移する戦闘様相に対して、人工知能 (AI) により行動方針を分析し、指揮官の意思決定を支援する技術を装備品に反映するするための研究を行う」と規定している。「防衛力整備計画について」2022年12月16日 (国家安全保障会議決定・閣議決定) 別紙24項IX 2(6)ウ。

(6) Yuval Abraham, “A mass assassination factory: Inside Israel’s calculated bombing of Gaza,” +972 Magazine, November 2023, https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/?utm_source=972+Magazine+Newsletter&utm_campaign=9e50be135f-EMAIL_CAMPAIGN_9_12_2022_11_20_COPY_01&utm_medium=email&utm_term=0_f1fe821d25-9e50be135f-318845989, accessed 2 April 2024. 川上泰徳「イスラエル軍のガザ攻撃で市民死者の激増の背景にあるAI標的生成システム: イスラエルメディア調査報道」2023年12月2日。
<https://news.yahoo.co.jp/expert/articles/12fde146311fa7dc0e56c37a607d34096a8badbe>, 2024年4月2日閲覧; 「新世代AI 変わる戦場まるで『ゲーム』薄れる抵抗感広がる犠牲」朝日新聞2024年3月25日1・2面。「ハブソラ (ゴスペル)」の国際人道法上の問題点を分析した論考として、黒崎将広「国際法における軍事AI問題の本質——攻撃目標選定支援プラットフォームを手掛かりに」『法学教室』2024年4月号 (523号) 32-36頁。

自律型致死兵器システム（LAWS）規制議論の動向と今後の展望

【2022～26年 LAWS・軍事 AI 関連年表】

| 年 | 月日 | LAWS・軍事 AI 関連事項（CCW、国連内、国連外、国内など） |
|------|----------|---|
| 2022 | 3/7～11 | CCW・GGE 第1会期 |
| | 7/25～29 | CCW・GGE 第2会期 |
| | 10/7 | 人権理事会決議「軍事領域での新興技術の人権上の含意」 |
| | 10/21 | 国連総会第1委員会での70か国共同声明 |
| | 11/16～18 | CCW 締約国会議 |
| | 11/30 | Open AI 社が生成 AI 「チャット GPT」 を公開 |
| 2023 | 1/25 | 米国 DoD 指令 3000.09 改訂版 |
| | 2/15、16 | オランダ・韓国共催 REAIM サミット（ハーグ） 米国提唱の「AI・自律性の責任ある軍事利用政治宣言」構想公表 「軍事領域での責任ある AI 行動要請」（60か国賛同） |
| | 2/20 | CCW・GGE オンライン非公式協議 |
| | 2/23、24 | コスタリカ主催の CELAC 地域会議：33か国「ベレン共同声明」 |
| | 3/6～10 | CCW・GGE 第1会期（5つの作業文書提出） |
| | 4/20 | CCW/GGE 会期間第1回オンライン非公式協議 |
| | 4/25、26 | ルクセンブルグ主催国際会議 |
| | 5/2 | CCW/GGE 会期間第2回オンライン非公式協議 |
| | 5/15～19 | CCW・GGE 第2会期（GGE 報告書 CCW/GGE. 1/2023/2） |
| | 7/22 | 事務総長「我々の共通の課題：政策概要 9 平和のための新課題」 |
| | 8/28 | ヒックス米国防副長官による「レプリケーター」構想発表 |
| | 9/5、6 | トリニダード・トバゴ主催地域会議（CARICOM 宣言）15か国・地域 |
| | 10/5 | 国連事務総長と ICRC 総裁の共同要請（2026年までに法文書締結） |
| | 11/1 | 国連総会第1委員会で LAWS 決議採択（164:5:8） |
| | 11/9 | 米国作成「AI・自律性の責任ある軍事利用政治宣言」（54か国賛同） |
| | 11/13 | 米国提唱の AI・自律性軍事利用政治宣言の初会合（ニューヨーク） |
| | | 11/15～17 |
| | 12/4、5 | 国連・韓国共催「第22回軍縮・不拡散合同会議」 |
| | 12/13～14 | フィリピン主催「インド太平洋地域会議」24か国オブザーバー5か国 |
| | 12/22 | 国連総会で LAWS 決議採択（152:4:11） |
| 2024 | 2/26、27 | REAIM 地域（アジア）協議（シンガポール） |
| | 3/4～8 | CCW・GGE 第1会期 |
| | 3/19、20 | AI・自律性の責任ある軍事利用政治宣言第1回全体会議 |
| | 8/26～30 | CCW・GGE 第2会期 |
| | 9/9、10 | 第2回 REAIM サミット（ソウル） |
| | 9? | 事務総長の LAWS 報告書を国連総会に提出 |
| | ? | 国連総会第1委で LAWS 議論 |
| | | 11/13～15 |
| | 12/? | 国連総会本会議で LAWS 議論 |
| 2025 | ? | CCW・GGE の10日間開催 |
| | 8 or 9? | 国連人権理事会第60会期に諮問委員会の研究報告書提出 |
| | 11? | CCW 締約国会議 |
| 2026 | ? | CCW・GGE の10日間開催 |
| | ? | CCW 第7回再検討会議 |

著者作成

2. CCW 枠内での議論

CCW の GGE は、第 1 会期前の 2 月 20 日にオンラインによる非公式協議を開催した。⁽⁷⁾ そこでは、ロシアも米国も、異なる立場の共通性を見出す必要性を強調し、実質的な成果を期待すると発言した。また、禁止すべき自律兵器システム (Autonomous Weapons Systems, AWS) と規制すべき他の AWS についての議論を分離する米国の様式は、他国の提案する 2 層アプローチを考慮に入れたものになるという。NGO の殺人ロボット禁止キャンペーン (Stop killer Robots, SKR) によれば、有意な人間による制御 (meaningful human control, MHC) の幾つかの要素 (たとえば、予測可能性、理解可能性、説明可能性、信頼可能性、追跡可能性) はまだ改善する余地があるとしても、特定の兵器は禁止されなければならないことにほとんどの代表が同意していることに注目した。加えて、人間の標的化の禁止が諸提案に十分に組み込まれていないと強調し、最後に、単なる原則やプラクティスに留まらず、法的拘束力のある規則が AWS の脅威に対する唯一可能な対策であると繰り返し述べた。

(1) GGE 第 1 会期

(a) GGE 第 1 会期提出の作業文書

GGE 第 1 会期が、3 月 6 日から 10 日まで 5 日間開催され、それに 5 つの作業文書 (パレスチナ、オーストリア、パキスタン、豪・加・日・韓・英・米 6 か国共同、露) が提出された。以下、それぞれの作業文書を概観する。

(i) パレスチナ作業文書⁽⁸⁾

同文書によれば、AWS とは、「人間の使用者による起動後に、人間の

(7) Laura Varela, "Report on the GGE Virtual Informal Consultation," *CCW Report*, 27 February 2023, Vol. 11, No. 1, pp. 1-2. <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/reports/CCWR11.1.pdf>, accessed 16 February 2024.

(8) 「パレスチナ国による自律兵器システムの規範的・運用上の枠組み提案」CCW/GGE.ノ

介入なく標的を選定しそれに武力攻撃するためにセンサーのデータ処理を利用するシステム」を指す（3項）。ただし、当該システムの起動後の名ばかりの人間の入力（nominal human input）は、人間の介入に該当しない（4項）という。そのシステムの法的、倫理的、人道のおよび安全保障上のリスクに対処するために、禁止と規制の両方が必要とされる（5項）。（a）人間を直接に標的化するように設計されもしくは使用される自律兵器、または（b）有意な人間による制御（MHC）で使用できない自律兵器の開発および使用は禁止される（6項）。言い換えれば、起動後に、自律プロセスで標的を選択し、武力でそれを攻撃しようとするAWSは禁止される（36項）。有意な人間による制御には、AWSが予測可能性（Predictable）、信頼可能性（Reliable）、理解可能性・説明可能性（Understandable and Explainable）および追跡可能性（Traceable）といった要件すべてが満たされなければならない（7項）。AWSが有意な人間による制御によって使用できるよう確保するために、積極的な義務や制限を含む規制が必要となる（8項）。そして、国際法的に拘束力のある文書形式で禁止と規制が組み合わされるべきである（9項）という（【図1：パレスチナ作業文書のAWS規制の2層（禁止・規制）アプローチ】参照）。

特に、人間を直接に標的化するAWS禁止の下では、センサーを使って外部環境から取得した処理済みデータと照合される標的プロフィールを示すデータ・パターン⁽⁹⁾は、人間を表わすように設計されてはならない（21項）。他方、AWS規制に関して（49項）、予測可能性についてのAWSの「開発」規制では、標的タイプ（軍事目標としての法的特徴が比較的安定的で、AWSの起動後とそれに続く攻撃の間で変化しそうにない標的、たとえば、弾薬、軍用レーダー、軍用艦船）、運用期間（短期間のみ）、運用領域（限定的区域のみ）、攻撃の規模と回数、データ処理方法に制限が設けられ、そして、人間が監視し、システムの起動後にその中断や不活性化

↘ 1/2023/WP.2, 3 March 2023.

(9) たとえば、人間の特徴を表わすデータ・パターンとして、人間の顔、体型、体温、肌の色、歩行・移動の速度が列挙されている（23項）。

【図1：パレスチナ作業文書でのAWS規制の2層（禁止・制限）アプローチ】

| | | 開発（要件） | 使用（要件） |
|-------------|-------------------------|-------------|---|
| 禁 止 | 人的標的 | 開発禁止 | 使用禁止 |
| | 人的・物的標的にMHCなし | 開発禁止 | 使用禁止 |
| A W S | 規制・合法兵器要件 有意な人的制御4要件 | 予見可能性 | 標的タイプ、運用期間、運用領域、規模、回数、人間の監視、起動後の中断・不活性化機能 文民・民用物の存在しない場所 |
| | | 信頼可能性 | 安全装置の装備 使用前の機能検証、起動前の文脈的要因の影響考慮 |
| | | 理解可能性・説明可能性 | 機能説明マニュアル作成 使用者の教育訓練 |
| | | 追跡可能性 | デジタル・ログの組み込み、責任帰属可能な設計 使用責任の帰属設定 |

著者作成

が必要となった場合に介入することができる機能が設定されていることが規定された。さらに、AWSの「使用」規制では、都市のような文民が混在する環境では使用すべきでなく、他方、文民や民用物が存在しない場所、たとえば、航路や漁場から遥か離れた公海上でのみ使用すべきであるという。

信頼可能性に関して、AWSの「開発」規制では、AWSは期待通りの適切な内部機能が保証され、意図せざる危害を引き起こす可能性を最小限に抑え、ハッキングやデータ・スプーフィングに対する安全装置を組み込むなど外的脅威のリスクを最小限にするように設計されることが明記された。AWSの「使用」規制では、使用者はAWSの使用前にそれが正確に機能するかを検証すること、兵器システムの起動前に、自律プロセスの機能に関する文脈的な要因の潜在的な影響（たとえば、晴れ/強風/雨天などの天候によるAWSの特定センサーの機能変動）を考慮することが規定された。

理解可能性・説明可能性に関して、AWSの「開発」規制では、データ処理方法の複雑さに制限を設けてAI・機械学習に基づくデータ処理を回避すること、AWSは十分な情報に基づき有意で注意深い人間による検証なくミッションのパラメーターを変更できないように設計すること、シス

テムがどのように機能するかを説明する詳細なマニュアルと共に開発されること、が規定された。AWSの「使用」規制では、システムが標的として識別する対象物、システムで使用されるセンサーの限界、作動中のデータ処理方法、それらの機能に関する既知の限界と不確実性、武力行使を引き起こす状況を含めて、人間である使用者は、システムの機能に関する適切な訓練が受けられるように明確な手続きが整備されることが規定された。あくまでセンサーがどれほど高性能であっても、センサーの読み取り自体は、標的は何なのか、それがどこにあるのか、それが何のために使われているのか、近くに誰がいるのかといった人間の評価と同等ではないことが強調されている。

追跡可能性に関して、AWSの「開発」規制では、デジタル・ログなど追跡可能性を容易にするAWSの技術的特徴を組み込むこと、意思決定過程に関与する1人または複数の要員にまで遡って特定行為を追跡する作業を容易にするメカニズム、そして、国際法に基づいてその開発・使用の結果に対する責任を個人や国家に帰属させることができるようにAWSを設計することが規定された。AWSの「使用」規制では、国際法に基づいて使用結果に対する責任が個人や国家に帰属するように、AWSのパラメーターを設定すると規定された。

最後に、パレスチナは、法文書に関連して、非拘束的文書であれば、受容可能性の明確な法的・倫理的境界がないので、AWSの広範な開発・使用を促すリスクがあり、また、明確な法的基準が普遍化しなければ、非拘束的なプラクティス（実行例）が、ある国では受容可能としても他の国では受容不可能といった深刻な不一致が生じるという。それゆえ、非拘束的な原則やプラクティスは、拘束力のある国際法文書が確立する以前に合意されるべきではないという（53項）。

(ii) オーストリア改訂作業文書⁽¹⁰⁾

オーストリアは、多数のGGE参加国間では2層構造（two-tier struc-

(10) 「改訂作業文書」CCW/GGE.1/2023/WP.1/Rev.1, 3 March 2023.

ture) に従って AWS を規律する見解に収斂しつつあるという認識を示す (1 項)。そして、AWS に対する有意な人間による制御を築き、人間の尊厳を保護し、そして、国際人道法を含む現行国際法の遵守を可能にするために、標的を選択しそれに武力行使するという重要な機能に自律性を統合する兵器システムの使用許可権限者は、次の 3 点をすべて満たさなければならないと規定する (2 項(a))。① 検討対象のシステムの機能を十分理解すること (意図せざる交戦を引き起こす条件を含み、どのような状況または条件でシステムによる武力行使が引き起こされるのかを理解すること)、② 当該システムが使用される状況を適切に評価すること (それには、特定の環境やその使用の時刻と期間に関連して、武力行使を引き起こす状況または条件の評価を含む)、③ 法的義務に従って武力行使の予期される結果について情報に基づいた判断ができるようにするために必要な範囲内で、システムが機能する期間や地理的範囲およびシステムが実行できる交戦回数を制限すること (使用許可権限者は、システムが機能し続けて情報に基づいた法的判断が下される文脈から外れてしまう場合には、システムを再調整、中断または不活性化できなければならない)。これらの要件を満たさないシステムは、受容不可であり、禁止されるべきであるという。さらに、人間の尊厳や価値、人道性原則または公共の良心の命令を侵害する方法で人間を標的として選択し攻撃する AWS も受容不可であり、禁止されなければならない (2 項(b)) という。

(iii) パキスタン作業文書⁽¹¹⁾

パキスタンによれば、CCW の議定書は若干の分類の兵器の製造および使用を禁止し (議定書 I と IV)、他の分類の兵器の開発、配備および使用を明確に規制しており (7 項)、以上のことから、CCW の趣旨および目的を達成するために LAWS の開発、配備および使用に関する禁止と規制を想定する国際法文書の策定のための明確な論拠がある (8 項) という。そ

(11) 「条約の趣旨および目的：自律型致死兵器システム (LAWS) の国際法文書提案」
CCW/GGE.1/2023/WP.3, 6 March 2023.

して、パキスタンは、透明性・信頼醸成措置を含む非法文書は国際法規則の補完物であって代替物ではない（25 項）という立場から、LAWS に関する国際法文書を提案した。

いかなる事態においても、以下の AWS の開発、配備および使用が禁止される（27 項(a)）として、具体的に、① 人間による制御なく武力行使の決定をするもの、② 文民、戦闘外の戦闘員、戦闘員との間で区別できないかまたは無差別的に民用物を標的化するもの、③ 不必要な苦痛または過度の傷害をもたらすもの、④ 攻撃中に文民および民用物を保護するためのあらゆる実行可能な措置を取ることができないもの、⑤ 予期される直接的な軍事的利益を超える文民および民用物への付随的危害を及ぼすもの、⑥ 適切に予期し理解し説明することができない効果を有するものが規定された。

兵器の開発、配備、使用に関する制限および規制に関して（27 項(b)）、武力行使の決定に人間による制御のある兵器は、① 人間による明確な承認がなければ、その使用状況を変更できないし、② その標的化のパラメーターも変更できない。また、③ 本質上、軍事目標である対象物だけに標的化能力を限定すること、④ 時間的・領域的制限を含む使用の空間や規模を制限するよう設計されていること、⑤ 潜在的標的の同定および選択において、透明で、疑う余地もなく、予測可能であるように設計されていること、⑥ 文民や民用物が存在しない事態においてしか作動しないように設計されていること、⑦ 人間が作動・使用のすべての段階で中断できること、が規定された。

(iv) 6 か国作業文書⁽¹²⁾

豪・加・日・韓・英・米の 6 か国は、前文と全 7 か条の条文および 3 つの付属文書⁽¹³⁾を含む作業文書を提出した。それによれば、AWS とは「一旦

(12) 「自律兵器システムに関する条文案—— 国際人道法に基づく禁止とその他の規制措置」
CCW/GGE.1/2023/WP.4, 6 March 2023.

(13) 国際人道法の関連する中核的諸原則に関する付属文書 I、GGE が承認した指針原則に関する付属文書 II および 2017 年から 2022 年までの GGE でコンセンサスを得た結論・勧告

起動すれば、オペレーターによる更なる干渉なく標的を選択し致死力でそれを攻撃することができる兵器システム」と定義し（前文2項）、禁止と他の規制措置を規定したAWSの開発、配備および使用に関する7か条の条文案を提案した（前文8項）。

性質上、国際人道法に従って使用できないAWSの防止（1条）として、本質上、不必要な苦痛を与えるか、性質上無差別的なAWSは禁止されると規定する。そして、如何なる事態においても、国際人道法に従って使用できないシステムの開発および使用を防止するために、以下のようにAWSを設計してはならない（1条1項）。(a) 文民もしくは民用物を標的にしまたは文民たる住民の間に恐怖を広めること、(b) 予期される具体的かつ直接的な軍事的利益との関連で、文民に対する偶発的な殺傷または民用物の損害が結果として常に過度となる攻撃を行うこと、(c) システム使用の指揮官やオペレーターが責任を負わない攻撃を行うこと。そして、AWSの開発は、攻撃でのその効果が区別や比例の諸原則によって、その使用事態において求められるように、予期され制御され得るようにしか認められない（1条2項）。そのために開発段階で、次の措置が含まれるべきであるという。(a) 予期される使用状態において兵器システムがどのように作動するのかについての査定を知らせるための厳格な実験・評価、(b) AWSが区別や比例の諸原則に従って使用可能か否かを特に考察する法的審査、(c) 文民や民用物への危害リスクを軽減するためにそのシステムの設計や使用において実施される可能な予防措置または機能の考察、たとえば、i. そのシステムが攻撃する標的タイプを制御、制限または他の方法で影響を及ぼす措置、ii. 弾薬またはそのシステムに自己破壊、自己不活性化もしくは自己中立化の装置の組込みなど、兵器システム作動の期間、地理的範囲および規模を制御、制限または他の方法で影響を与える措置、iii. システム・オペレーターに内在する自動化バイアス⁽¹⁴⁾（偏向）の軽

、告に関するテーマ別編集の付属文書Ⅲである。

(14) 自動化バイアス (automation bias) とは、人間が自動化された意思決定システムからの提案を優先し、自動化なしで作成された矛盾した情報を、たとえそれが正しい場合でも無

減措置、iv. 兵器システム使用に関連して依存する AI 能力における意図せざるバイアスの軽減措置、v. その他、タイミング、精密性および精確性を含めて、武力行使に対する制御を向上させ、その意思決定を改善する措置。

AWS 使用を規制する中核原則（2 条）として、攻撃時の区別、比例および予防の諸原則が確認されている。当該原則を詳述した規定として、攻撃実施時の区別確保のための規制措置（3 条）、攻撃実施時の比例確保のための規制措置（4 条）および攻撃時の予防確保のための規制措置（5 条）がある。加えて、説明責任を確保するための規制措置（6 条）と当該条文案草案の地位と法（国際法—著者注）との関係性（7 条）⁽¹⁵⁾が規定された。

（v）ロシア作業文書⁽¹⁶⁾

2022 年 7 月 26 日にロシア国防相が承認した本文書は、ロシア軍における AI 技術搭載の兵器システムの開発および使用の諸原則を統合し、国際人道法を含む国際法上の義務の遵守を確保することを目的として策定された。その基本原則として、安全性、透明性、技術主権（Technological sovereignty）、責任、制御を列挙している（2 項）。安全性に関して、AI 技術搭載の兵器システムの開発または取得では、物理的な安全性要件、ハッキングやデータ・スプーフィングに対するサイバー・セキュリティを含む非物理的な保護措置およびテロ集団による取得を含む拡散リスクが考慮されるべきである。透明性に関して、AI の働きや AI による結果達成過程は説明可能とすべきである。技術主権に関して、国内の AI 技術の優先的な使用を含めて、AI 分野におけるロシアの必要なレベルの独立性が

↘ 視する傾向を指す。

(15) 7 条は、「これらの条文案草案は国際人道法の要件を明確にし、それらの要件を効果的に満たす措置を特定しようとするものであって、如何なる国際法的義務を創設するものでも、それからの逸脱を認めるものでもない」と位置付ける。

(16) 「ロシア作業文書：AI 技術搭載の兵器システムの開発および使用におけるロシア連邦軍の活動概念」CCW/GGE.1/2023/WP.5, 7 March 2023. なお、本作業文書はロシア語版しか配布されていないので、ここでは非公式の英訳版を参照した。https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/documents/Russia_March2023.pdf, accessed 14 February 2024.

確保されるべきである。責任に関して、人間は、AI 技術搭載の兵器システムのライフサイクルを通して、そして、適用される国際法規範に従ってその開発と使用に関する決定に責任を負うべきである。制御に関して、AI 技術搭載の兵器システムに対する人間による制御を維持する必要性を認めつつ、その制御の具体的形式と方法は国家の自由裁量に委ねるべきであり、システムへの効果的な人間による制御は直接的な制御以外でも達成できるという。

ロシアは、国連憲章や国際人道法を含む国際法が、AI 技術搭載の兵器システムの開発および使用を含むすべての兵器システムに引き続き完全に適用されることを認めており、現行の国際法規範の遵守を維持するために、以下の方法で機械に対する人間による制御を維持する必要があるという(3項)。信頼性と耐障害性⁽¹⁷⁾の向上、標的タイプの制限、運用期間・地理的範囲・使用規模の制限、適時の介入と不活性化、実際の運用環境での AI 技術搭載の兵器システムの実験、AI 技術搭載の兵器システムに関する責任ある使用手順を十分に修得した者に管理(統制)を認めること、個別の要素の生産過程や生産品全体の管理、個別の要素と生産品全体の事前仕分けや廃棄の管理、という方法である。そして、国際法義務の遵守や当該兵器使用の責任は、当該システムに任務を割り当て、その使用の命令を下した当局者にあると規定した。

ロシアは、AI 技術搭載の兵器システムのリスク対策として、軍事上の意思決定を支援するために必要な情報の収集と処理における AI 技術の使用を優先させること、オペレーターが継続的に当該システムの動作を監視し、オペレーターの命令による戦闘任務の緊急終了を保証することが規定された(5項)。

結論として、ロシア軍の優先事項の一つは、AI 技術搭載の兵器を含む兵器の責任ある使用であり、そして、諸国家間で関連するベスト・プラク

(17) 耐障害性 (Fault tolerance) とは、構成部品の一部が故障しても、バックアップシステムを使うなどして正常に処理し続けるシステムをいう。

ティスや最先端技術の成果を自発的に交換することも重要であると指摘している。

(b) GGE 第 1 会期での議論

本会期でも、LAWS の法規制プロセスを推進する諸国家、それに反対する諸国家、その中間に位置する諸国家の解決し難い三つ巴状態が、依然、見られた。イスラエルは、兵器に対する如何なる形態の人間による制御要件は存在せず、人間による制御は国際人道法上の義務ではないと強硬に主張した。⁽¹⁸⁾ 他方、パレスチナは、国際人道法が初めて制定化された時には自律性を議論する必要性がなかったので、GGE 参加国全てが重要と承認するすべての要件（予測可能性、説明可能性、追跡可能性など）は現行国際人道法に法典化されていないと指摘し、新技術が登場した時には、新次元が生じ、法典化が必要とされるのはごく自然であると反論した。⁽¹⁹⁾

3月6日には、GGE で一般発言と上述の5つの個別提案のプレゼンがあった。⁽²⁰⁾ 一般発言の中で、インドは政治宣言または非拘束的合意を支持し、イスラエルは、新たな法規則を創設する必要はないと主張しつつ、コンセンサスで合意が可能であれば、6か国の条文案を支持できると発言した。3月7日午前会合で「議題1：LAWSの特性——定義と範囲」が議論された。そこでは、LAWSの包括的定義は必要か、その定義内容は何か、定義よりも特性に焦点を当てるべきか、「有意な人間による制御」はLAWSの中心的特性か否か、さらに、LAWSかAWSかという「致死性 (lethality)」の問題が改めて議論された。

3月7日午後と8日午前、「議題2：国際人道法の適用：可能な禁止と規制」が議論された。⁽²¹⁾ 国際人道法を遵守できないAWSの開発、製造、使

(18) Ray Acheson, "Editorial: Taking on the War-Builders," *CCW Report*, vol. 11 No. 2, 14 March 2023, p. 1. <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/reports/CCWR11.2.pdf>, accessed 21 February 2024.

(19) *Ibid.*, pp. 2-3.

(20) Ray Acheson and Laura Varella, "General Statements and Presentation of Proposals," *ibid.*, pp. 5-9.

(21) Ray Acheson, "Application of International Humanitarian Law," *ibid.*, pp. 9-12.

用はすでに禁止されており、AWSの新規則は不要と主張する一握りの国とは対照的に、ほとんどの国は新規則を求めた。「禁止と規制」という2層アプローチに関連して、NGOのArticle 36は、国家は法に従って使用できない兵器システムが違法であるというトートロジー（同語反復）に陥らないように注意喚起した。AWSの利点に関して、日本は、AIがより迅速にデータを分析・処理するので、AWSは標的化の精確性や速度を改善することで文民への危害を軽減・阻止できると主張した。他方、ベネズエラは、AWSは逆効果となると反論した。また、日本や英国は、AWSの禁止により技術的發展が妨げられるとの懸念を表明した。それに対して、オーストリアは、AWSを巡る法的明確性の欠如が技術的發展を阻害するのであって、AWSの規制は技術を禁止することでなく、技術の適用を制限することであると反論した。

責任、説明責任および人間による制御に関して、ロシアは、国際人道法の義務を遵守する責任は将校（officers）にあり、武力行使の決定者は常にヒト（people）であると発言した。フランスも、AWSの使用責任は人間の指揮官が負うことに同意した。さらに、人間による制御は多様な形態を取り得るが、AWSに割り振る任務を明確に定め認可するのは人間でなければならないと発言した。

3月8日午前に、「議題3：人間と機械の相互作用/有意な人間による制御（MHC）/人間による判断および倫理的考慮」が議論された。⁽²²⁾ 米国は、国際人道法は兵器の要件として「人間による制御」という用語を使用しておらず、ほとんどすべての兵器や弾薬は、発射後に制御不能であるように、ある時点から人間による制御外にあると発言した。しかし、他のほとんどの代表は、人間による制御の支持を表明し、アルゼンチンは、兵器システムの合法性を決めるのは人間による制御であり、要求される人間による制御のレベルに議論を絞るべきであると反論した。

オーストリアがMHCの支持を表明し、たとえば、フィリピンは収斂し

(22) Laura Varella, "Human Control and Ethical Considerations," *ibid.*, pp. 12-14.

つつある2層アプローチ構造にMHC概念を含めるべきと主張した。他方、米国は、本用語に反対して、新たな法的基準として、MHCに焦点を当てるよりも兵器システムが人間のオペレーターの意図を予期し、意図せざる交戦を回避するよう確保することが重要であると反論した。

多くの提案には、中核的概念として、予測可能性、理解可能性、説明可能性、追跡可能性及び信頼可能性といった概念、標的のタイプ、運用期間、地理的範囲の特定の制限、そして、人間が兵器システムを不活性化・中断・停止できる適時の介入（timely intervention）要件が含まれていると、様々な国家から指摘された。なお、パレスチナが提出した作業文書4項で指摘した「名ばかりの人間による入力」⁽²³⁾は実質的にプロセスに影響を及ぼさないので、パレスチナは、これは人間による介入に該当しないと指摘した。オーストリアもこの意見に賛同して、人間による制御は名ばかりの入力以上のものでなければならぬと発言した。

倫理的考慮に関して、オーストリアやフィリピンは、国際人道法に加えて、倫理的考慮を重視した。他方、日本は倫理的考慮が国際人道法に含まれていると言い、フランスはCCWの中心に国際人道法があり、GGEの文脈でもそうあるべきだと述べた。

3月8日午後、「議題4：責任（responsibility）と説明責任（accountability）」⁽²⁴⁾⁽²⁵⁾が議論された。事実上すべての代表が、人間が責任を負うこと、責任は機械に委ねられないこと、そして、国家は国際違法行為について説明責任を負うことに合意した。AWSの責任および説明責任に関して、米国は、説明責任を確保するために、AWSの配備前に、国家は、オペレーターに確実に教育・訓練を受けさせ、国際法に違反した場合に確実に個人に責任を負わせる指針を持たなければならぬと明言した。他方、チリや

(23) それは、「人間のオペレーターが更なる道義的・法的熟慮なく単純にボタンを押す行為」を指す。Ibid., p. 13.

(24) Responsibilityは任務の遂行責任を指し、現場担当者レベルの責任の果たし方を指すのに対して、Accountabilityは過去の行為や結果に対する責任であり、監督者や上司の責任の果たし方を指す。

(25) Ray Acheson, "Responsibility and Accountability," *ibid.*, pp. 14-16.

メキシコは、攻撃開始の決定は、事態を分析し、すべてが国際人道法に完全に準拠して実施されていることを保証できる人間に帰属すべきであると示唆した。

3月9日午前に、「議題5：法的審査」が議論された⁽²⁶⁾。多くの諸国家が、AWSの法的審査が国際人道法の遵守を確保するための重要な措置であると認めていた。オランダは、その義務は、追加議定書の法的義務だけでなく、全ての国家の義務（＝慣習法上の義務—筆者注）であるという。諸提案間での共通性として、フィリピンは、ほとんどの提案で兵器システムは法的審査の専門家に説明可能で、かつ、理解可能でなければならないことが確立しているという。ロシアも、法的審査によってアルゴリズムが予測可能であることを確実にする必要があるという。幾つかの国は、機械学習能力により兵器システムが変容する可能性を懸念して、ソフトウェアに変更がある場合にそのシステムは新たな法的審査を受ける必要があるかもしれないと強調した。

法的審査は国内当局に任せるべきか、多国間のコミットメント（約束）に基づかせるべきかについて、意見の対立が見られた。日本は、法的審査の実施方法は各国に委ねられているが、諸国家が自国の法的審査を確立するために、グッド・プラクティスやリスク軽減措置の編集を支持した。オランダは、諸国家が法的審査の結果を公表するよう提案した。

会議の終了前に、議長が、以下の質問を提示した。1. 自国による法的審査で国際人道法の遵守が確保できるのか。2. 現行の枠組みでは、国家は共通の基準なく独自にLAWSの法的審査を行うことになるが、ある国の基準で合法でも他国の基準で違法となる「規範の分裂」リスクを回避できるのか。3. AWSの国際的に合意された定義がない状態で効果的な法的審査をどのように促進するのか。質問1について、エクアドルは、自身による法的審査は遵守が確保できないが、重要なツールとなり得ると発言した。質問2の規範の分裂問題について、オランダは、誤解や解釈の分散

(26) Laura Varella, "Legal Reviews," *ibid.*, pp. 17-19.

化を回避するために、理想的な手段は何かを議論し続ける必要があると述べた。質問3に関して、オランダは、第1追加議定書36条は国家が兵器や戦闘の手段・方法を評価しなければならないことを確立したが、それは技術的定義それ自体が必ずしも必要とされているわけではないことを明確にしているという。

3月9日午後に、「議題6：リスク軽減と信頼醸成措置」が議論され⁽²⁷⁾た。すべての提案でAWSの軽減すべきリスクが規定されている。たとえば、非国家行為者（テロリスト）による取得、技術的不具合による制御喪失、サイバー攻撃に対する脆弱性、AIのバイアスなどが指摘された。パレスチナは、GGEが注視すべき5つのリスクとして、偶発事故や誤作動、人種その他の差別の悪化、拡散、地球規模的な不安定性の悪化、環境や生物多様性への危害を列挙した。フィリピンは、リスク軽減措置を国内措置に制限すべきでなく、多国間合意の枠組みに委ねるべきであると繰り返し主張した。英国は、リスク軽減措置や信頼醸成措置が法文書の代替物とはならないという発言に答えて、その立場の変更を提案しつつ、法文書は信頼醸成やリスク軽減の必要条件を満たし得ないと強調した。

3月10日午前に、非公式協議がGGE議長により招集された⁽²⁸⁾。今週の主な収穫として、2層アプローチの進展ならびに責任・説明責任およびリスク軽減措置に関する意見の収斂が見られたという。

3月10日午後の閉会の辞においては、ほとんどの国は、2月の自律兵器に関する社会的・人道的影響に関するラ米・カリブ会議で採択されたベレン・コミュニケに注目し、AWSの法文書の要求を繰り返した。2層アプローチに関する意見が収斂している状況において、フィリピンは、本アプローチに対する2つの異なる出発点があると指摘した。一つは、禁止と規制は専ら国際人道法の操作化（operationalization、実際に運用可能な形に

(27) Laura Varela, "Risk Mitigation and Confidence-Building Measures," *ibid.*, pp. 20-21.

(28) Laura Varela, "Informal Discussion," *ibid.*, p. 20.

(29) Ray Acheson, "Closing Remarks," *ibid.*, pp. 23-24.

すること)に基づくべきであると考え。多数派であるもう一つの出発点は、武力行使に対する国際人道法の遵守の確保と有意な人間による制御の確保の両方であるという。フィリピンは、人間による制御の欠如が禁止基準となるべきであるという。対照的に、オーストラリアは、人間による制御は禁止のための「第1層」基準ではなく、AWSの合意における「第2層」制限に含まれるべきであると繰り返した。

議長は、5月15-19日にGGEの公式第2会期が開催されるが、第1会期と第2会期の間に2回のオンライン非公式協議(4月20日と5月2日予定)も行われると告知した。

(2) GGE 第2会期

(a) 第2会期前の非公式協議

前回のよう、GGEの公式会期前の4月20日に、GGEの第1回オンライン非公式協議が行われた。そして、GGE議長作成の2023年報告書に挿入する結論部分の草案が⁽³⁰⁾4月25日に回覧され、5月2日の第2回オンライン非公式協議で議論された。⁽³¹⁾議長案で国際人道法に焦点を絞っている点について、国際人道法だけでなく、国際人権法や倫理的考慮にも言及するようにとの批判がある(メキシコ)一方で、国際人道法に焦点を絞る議長案を擁護する意見(ロシアやフランス)も見られた。

LAWSの特性(characterization)について、議長案は「一旦起動すれば、人間の介入なしに標的を特定し、追跡し、交戦し、武力を加えることができる兵器システム」(5項)と規定した。その条項に対して、LAWSの機能性を重視すべき、実用的定義にすべき、目的を明確化すべきなどの

(30) *Draft conclusions for inclusion into the 2023 report of the GGE Status date: 25 April 2023*, <https://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/documents/draft-conclusions-april2023.pdf>, accessed 23 February 2023. 結論案は、LAWSの特性、国際人道法の適用、人間と機械の相互作用、責任と説明責任、法的審査・リスク軽減・信頼醸成措置の5節から構成される。

(31) Laura Varela, "Report on the Second Virtual Informal Meeting," *CCW Report*, vol. 11 no. 3, pp. 3-5.

自律型致死兵器システム（LAWS）規制議論の動向と今後の展望

意見が出された。国際人道法の適用に関して、議長案の6項⁽³²⁾および7項⁽³³⁾に規定された禁止と規制の2層アプローチに賛同する発言が幾つかの代表から聞かれた。前述の如く、国際人道法に焦点を絞るか否かについて意見対立が見られた。また、オーストリアは、LAWSの禁止対象を「開発、配備または使用」と規定している6項と比較して、その規制対象が兵器システムの「使用」に限定されている7項について、その「開発および配備」も取り上げるべきと述べた。米国は、6項と7項は当該兵器の開発における措置に焦点を当てるべきで、その使用は8項⁽³⁴⁾と9項⁽³⁵⁾で焦点を当てるべきと勧告した。オーストリアは、7項(d)が2層アプローチの第1層の一部となるべきで、6項に位置する方がよいと述べた。フランスなどは、国際人道法と両立できないシステム（第1層）と国際人道法と両立できる方法で開発および配備できるシステム（第2層）との間でのみ区別する2層アプローチを支持した。米国は、7項(d)の用語が現行の国際人道法にみられないので、その定義が必要であると述べた。他方、フィリピンは、7項(d)の要素は殆どの提案に反映されていると反論した。

人間と機械の相互作用に関して、日本は、それがまだコンセンサスに至っていないと発言し、オーストリアは但書の「実行可能な場合に

(32) 6項「……攻撃時の当該兵器システムによる影響が、その使用の事態において求められるように、国際人道法に従って予期し制御することができない場合には、当該兵器システムは開発、配備、または使用してはならない。」

(33) 7項「……LAWS分野での新興技術に基づく兵器システムの使用は、適切な場合には、次の措置を含まなければならない、(a) システムの交戦可能な標的タイプの制御、制限その他の作用、(b) 自己破壊、自己不活性化または自己中立化の装置を含めて、兵器システムの期間、地理的範囲、作動規模の制御、制限その他の作用、(c) 状況において国際人道法の遵守に必要な場合には、制御の向上または武力行使に関する意思決定の改善をするために兵器システムの機能を起動または解除するための習熟した人間のオペレーター用の明確な手順、(d) システムが十分に予期可能、信頼可能、理解可能、説明可能、追跡可能であるように確保すること。」

(34) 8項「LAWS使用での人間と機械の相互作用は、攻撃時における区別、比例、予防の要件・諸原則の履行と一致しなければならない。」

(35) 9項「LAWS分野での新興技術に基づく兵器システムの使用責任者は、実行可能な場合に、国際人道法の遵守に求められるように、システムまたはシステム機能を制御し、中断または無能化できる立場にいないなければならない。」

(where feasible)」が不要であると述べた。責任および説明責任に関して、インドは、議長案の10項⁽³⁶⁾(b)の帰属概念および10項(d)・(e)の内容に反対を表明した。フィリピンは、10項(c)をリスク軽減の節に置く方がいいと提案した。フランスと日本は、10項(e)でのメカニズムはLAWSに特化する必要がなく、報告・調査の一般的な国内メカニズムが当該兵器の事件にも同様に対処できると述べた。法的審査・リスク軽減・信頼醸成措置に関して、オーストリアは、11項⁽³⁷⁾が第1追加議定書36条にもっと準拠すべきであるという。インドは、第1追加議定書の署名国でないので、但書の「適切な場合(as appropriate)」を追加するよう提案した。オーストリアとパキスタンは、12項⁽³⁸⁾は既に6項と7項で取り扱われている概念と重複するので、その削除を提案した。ベルギーは、法的審査・リスク軽減・信頼醸成措置の節に機械学習能力の事項が含まれ得ると言った。日本は、13項⁽³⁹⁾に標的のタイプだけでなく、期間や地理的範囲なども含まれ得ると述べた。

(36) 10項「国家は、特に、以下のものを通して、LAWS使用に対する説明責任を確保すべきである、(a)責任ある指揮統制系統内での当該システムの運用、(b)国際法に基づいて、兵器システムの使用結果についての責任を個人や国家に帰属させることができるように、指揮系統の特定を可能にする兵器システムの設計、(c)センサーの制限や武力適用を引き起こす状況を含む、システム機能に関する使用者への適切な訓練、(d)関連の法律・手順の実施を通じての個人の責任の確保、(e)国際人道法違反を含むかもしれないLAWS特有の事件の報告・調査に関するメカニズムの確立。」

(37) 11項「LAWS分野での新興技術に基づく兵器システムの法的審査は、それらが適用される国際人道法および国際法に従って使用できるか否かを評価するよう努めなければならない。法的審査は、技術的性能、使用目的、あり得る任務と標的のタイプを考慮すべきである。この文脈で国家間での関連するベスト・プラクティスの交換によって遵守のための適切な評価が強化できる。」

(38) 12項「リスク軽減措置には、運用環境での兵器システムの予期される機能に関する不確実性を回避するために、システムが十分に予測可能で、信頼性があり、理解可能かつ説明可能であり、追跡可能であるメカニズムが含まれるべきである。」

(39) 13項「リスク軽減措置には、意図せざる交戦のリスクを考慮し、システムが交戦できる標的のタイプに影響を与える制御、制限その他の措置を検討するべきである。さらに、それらは、自動化バイアスや兵器システムの使用に関連するAI能力における意図せざるバイアスの軽減を含むべきである。」

(b) GGE 第2会期提出の作業文書

GGE 第2回会期の会合が5月15日から19日まで開催された。当会合にグループ14⁽⁴⁰⁾による作業文書と51カ国共同声明が提出された。

(i) グループ14の作業文書⁽⁴¹⁾

本作業文書は、CCWの新議定書として「自律兵器システムに関する議定書（第6議定書）案」を提案するものであった。前文で、AWSのリスクや諸問題が認識され、その対処のため新議定書の締結が必要であると認めた。1条（一般条項）は、本議定書がAWSに適用される（1項）が、他方で、本議定書の如何なる規定も、AIを含む平和目的での新興技術のアクセス・開発・研究・製造・調達・移譲および使用に関するすべての国家の固有の権利を害するものではない（3項）と明記した。

2条（特性）は、「AWS」とは「標的を選定し、交戦し、そして武力を行使するという重要な機能に自律性を組み込んだ兵器システム」（1項）と規定した。また、「有意な人間による制御」とは、「武力行使に対する人間による判断および干渉の保持を含む、人間による作用の維持（maintenance of human agency）に関連する」（2項）。特に、次の要件がそれに含まれるという。2.1 兵器システムの目標もしくは任務を再定義や修正その他当該兵器システムを環境に適合させる能力、必要に応じてその運用および使用を不活性化、中止、終了または中断させる能力ならびにその機能が強制的に自己起動する能力、2.2 時間的および空間的な制限を含む、兵器システム使用の範囲および規模を制限する能力やその標的化パラメーターや標的化能力を制限する能力、2.3 責任および説明責任の帰属を含む、兵器システムの運用に関する法的およびその他の要件を満たす説明を振り返ってするために兵器機能を理解し説明する能力。

(40) アルゼンチン、コロンビア、コスタリカ、エクアドル、エルサルバドル、グアテマラ、カザフスタン、ナイジェリア、パナマ、ペルー、フィリピン、パレスチナ、シエラレオネ、ウルグアイの14カ国グループである。

(41) Draft Protocol on Autonomous Weapon Systems (Protocol VI), CCW/GGE.1/2023/WP.6, 10 May 2023.

3条（禁止）は、「選定し、標的化し、武力行使するという重要機能において有意な人間による制御により使用できないAWSをいかなる事態においても設計、開発、生産、保有、取得、配備、移譲または使用は禁止される」。また、「それらには、予測、説明、予期、理解または追跡できない方法で作動するものも含まれる」と規定した。

4条（規制）によれば、各締約国は、次の措置を制定する（1項）。「1.1 兵器システムに対する効果的な人間による監視を確保し、および如何なる時でも干渉や不活性化を可能にする要件、1.2 人間が、標的タイプ、期間、地理的範囲および使用規模を制限する能力を確保するための要件、1.3 人間がAWSに対する制御を行使するための情報と権限を確実に入手するための明確な手順、1.4 兵器システムが使用される多様な事態においてどのようにそれが機能するかを評価するための厳格な実験、1.5 人間が兵器システムの運用を確実に命令し、撤回し、そして、制限できるようにするために、兵器システムの理解可能性、説明可能性および予言可能性を保証するデータ処理方法の複雑さの制限」。そして、「各締約国は、兵器システム運用での自動化バイアスを回避し、そして、兵器システムの使用に関連して依拠するAI能力でのアルゴリズムのバイアスを排除する措置とメカニズムを制定することを約束する」（2項）。さらに、「各締約国は、AWSの使用における国際人道法の完全な遵守を確保し、説明責任を維持管理するための追加的な規制措置とメカニズムを制定することを約束する」（3項）。その他に、5条（兵器審査）、6条（リスク軽減）、7条（遵守）および8条（締約国協議）が規定されている。

(ii) 51カ国共同声明⁽⁴²⁾

本声明は、禁止と規制のアプローチを支持した。また、人間による判断、

(42) *Joint Statement Translating the Progress at the GGE LAWS into a Substantive Outcome*, 15.5.2023, https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/statements/15May_51states.pdf, accessed 24 February 2024. 51カ国の中に、英・米・仏・独・伊・日・加・豪・韓・蘭が含まれている（中・露・印・イスラエルが含まれず）。

制御、関与（human judgement, Control or involvement）と用語は異なるが、人間が中心に留まらなければならないという幅広い認識が共有され、関連する倫理的観点がこの枠組みの考慮の指針となるべきであるという。

⁽⁴³⁾
一般発言では、その 51 か国を含む多数の国家および EU が禁止・規制の 2 層アプローチの支持表明をした。赤十字国際委員会（International Committee of the Red Cross, ICRC）は、本アプローチは出発点であり、設計上どのタイプの AWS が禁止カテゴリーに該当するか、AWS により生じる特定の懸念に対処するために設計や使用においてどのような要件が必要なのか、より詳細に規定する必要があると指摘した。アラブ連盟やラ米・カリブ諸国などは、AWS の禁止・制限の法的文書交渉の支持を表明した。グループ 14 は、その法文書が CCW の新議定書の形式をとることに賛成した。中・印・米その他の国は、CCW が引き続き AWS を議論する適切なフォーラムであると発言した。

法的枠組みに関して、国際人道法だけでなく、広く、国際法や倫理的観点への言及を望む意見（51 개국提案の如く）もあれば、国際人道法の要件だけに限る意見（インド）もあった。人間による制御に関して、51 개국共同声明が示すように、人間中心主義が歓迎され、中国も兵器システムが人間による制御の下で使用されなければならないと述べた。問題は、人間による関与をどのように条文案に挿入するのか、であるという。個別提案のプレゼンに関して、グループ 14 は、法文書交渉の開始を目的とした前述の第 6 議定書案を提案した。加・日・米は、GGE の 3 月会期に提出した 6 개국共同提案を想起させた。外部のフォーラムに関して、51 개국共同声明やその他多くの国が、本稿第 3 節(2)項「国際会議および地域会議」で後述する CCW 枠外での発議（オランダの REAIM 会議、自律兵器の社会的人道的影響に関するラ米カリブ諸国会議、ルクセンブルグ自律兵器システム会議）を認識していた。

(43) Laura Varella, "General Statements," *CCW Report*, vol. 11, no. 3, pp. 5-7.

(c) GGE 第 2 会期での議論

5 月 15 日午前から 16 日午後まで、GGE は報告書案第 3 節：結論部分の 16～20 項までの第一読を実施した⁽⁴⁵⁾。多くの国家が報告書草案 18 項に賛同したが、フィリピンは、技術中立アプローチに関して、自律性によるリスクや課題の理解とバランスを取るべきであるという。他方、中国もロシアも、LAWS は技術の平和利用へのアクセスを妨げないという立場である。日本と英国は、技術そのものを汚名化しないための用語を支持し、英国は、AI は本質的に兵器よりもむしろ一つの機能性を指すと言って、この技術不可知論的アプローチ (technology-agnostic approach) を繰り返した⁽⁴⁷⁾。

19 項に関して、中国は、LAWS の定義および特性が実質的な措置を採択する前提条件であり、GGE で次年度に引き続き議論することを提案した。ロシアも、GGE は LAWS の定義に関して合意に達するほど十分議論していないので、本報告書に定義を挿入することに反対した。19 項は、LAWS の定義を試みるものではないことを明確にして本条項を支持する国 (ノルウェー) や、本条項に含まれた機能的アプローチを支持する国 (スイスなど) もあった。パレスチナは、本条項に「名ばかりの人間による入力」の追加を提案した。

(44) 2023 年 GGE 報告書案、CCW/GGE.1/2023/CRP.2, 06 May 2023.

(45) Laura Varella and Emma Bjertén, “First Reading: Paragraphs 16–20,” *CCW Report*, Vol. 11 No. 3, p. 7–9.

(46) 18 項「国際人道法の諸規則および諸原則は、すべての兵器および戦闘方法に適用され、従って、使用される軍事技術から独立している。このように、技術によって実行される機能を考慮する技術中立アプローチ (technology-neutral approach) が、LAWS 分野での新興技術の含意を考慮する場合に適用されるべきである。」

(47) 不可知論とは、「哲学で、ある種の主題の認識が人間には不可能であると主張する認識説」(『国語大辞典』小学館 1981 年 2089 頁) を指す。従って、技術的不可知論とは、AI の本質や全体を技術的に認識することは不可能であり、AI を経験や現象として捉える認識手法を意味すると考えられる。

(48) 19 項「可能な措置およびその他の選択肢は、一旦起動すれば、人間による介入なく標的を識別し、選択し、追跡し、武力行使することができる LAWS 分野での新興技術に基づく兵器システムに対処すべきである。」

(49) 20 項に関して、14 か国が、国際法や倫理的考慮に言及してその範囲を広げるべきであると主張する一方で、インドや英国は、報告書草案での国際人道法の諸規則および諸原則の優先を支持した。エクアドルは、予測可能で信頼可能な兵器の必要性に関する文言の挿入を提言した。オーストリアは、予測可能、信頼可能、理解可能・説明可能および追跡可能でない兵器は禁止されるべきであることを明確にするために、21 項(d) の内容を 20 項に移行する意見に賛同した。米国は、それらの用語は現行国際人道法の諸要件の用語法ではないと主張した。また、本条項の用語「致死 (lethal)」に反対し、報告書はすべて AWS に言及すべきとの見解 (パナマ) がある一方で、GGE のマンデートの用語法「LAWS 分野での新興技術 (emerging technology in the area of LAWS)」の維持を支持する見解 (イスラエル) があった。インドは、国際人道法は兵器の開発ではなく使用だけに適用されることから、第 1 文での「開発」の削除を提案した。しかし、多くの国家は、「開発」の用語は第 1 文で維持されるだけでなく、第 2 文でも追加されるべきであると反論した。⁽⁵⁰⁾

5 月 16 日午後に、GGE は報告書草案第 3 節の 21・22 項に集中して議論した。⁽⁵¹⁾ 議長によれば、21 項は、本年に提出された諸提案の共通項であると⁽⁵²⁾

(49) 20 項「特に攻撃時における区別、比例および予防を含む国際人道法の諸規則ならびに諸原則は、LAWS 分野での新興技術に基づく兵器システムの開発、配備および使用において遵守しなければならない。当該兵器システムは、その使用事態において国際人道法が要求するように、攻撃時のその効果を予測し制御できない場合には、配備または使用してはならない。」

(50) Note 45, p. 9 によれば、開発の用語は、「本条項第 2 文で維持されるだけでなく、第 1 文で追加されるべき」と記述されている。これは第 1 文と第 2 文が逆であり、筆者 (Laura Varella and Emma Bjertén) の誤記と思われる。

(51) Emma Bjertén, “First Reading: Paragraphs 21–22,” *CCW Report*, Vol. 11 No. 3, pp. 10–11.

(52) 21 項「攻撃時の効果とその使用事態において要求されるように予期でき制御できることを含めて、国際人道法の遵守を確保するために、LAWS 分野での新興技術に基づく兵器システムの使用、開発および配備は、適切な場合、システムの設計および使用において、以下の実行可能な措置を含まなければならない (must)。(a) システムが交戦する標的タイプを制御、制限またはその他の方法で影響を与えること、(b) 自己破壊、自己不活性化、自己中立化または同等のメカニズムの組み込みを含めて、兵器システム運用の期間、地理的範囲および規模を制御、制限またはその他の方法で影響を与えること、(c) 事態に

説明された。オーストラリアは、「must」は現行国際人道法の要件を記述するためにのみ使用されるべきであり、それを「should」に置き換えるように提案した。スリランカは、21 項の内容は常に適用可能とすべきであり、「where appropriate」の削除を提言した。オーストリアその他の国も、それに賛同した。フィリピンは国際人道法の中心性に異議を唱え、ベルギーも国際人道法の尊重は重要だが、国際刑事法や国連憲章の尊重の重要性も強調した。他方、米国は国際人道法を超えて前に行くコンセンサスは存在しないと主張した。それに対して、オーストリアは、テキストの目的は、現行国際法の繰り返しではなく、それを発展させることであると強調した。

21 項の「開発」の維持か削除かで意見が分かれた。インドは、国際人道法は兵器の使用のみに適用されその開発には適用されないと発言し、他方、スリランカは、兵器システムの制限はその開発ではなくその使用のみに関連することに同意しなかった。21 項(d) に関して、ロシアは「予測可能、信頼可能、理解可能、説明可能、追跡可能」の用語が法的用語でなく、異なる解釈のある主観的な用語であると強調し、21 項(d) の削除を提言した。米国も当該措置は法的要件を反映していないと論じた。人間のオペレーターによる不活性化システムの要件に関して、スリランカは、オペレーターが兵器システムの継続的監視を行使している場合にのみシステムの不活性化を許可されると明記されるべきであると強調した。さらに、21 項(d) に関して、スリランカは、「文民や民用物が存在しない事態においてしか作動しないように設計されること」の追記を提案した。

⁽⁵³⁾ 22 項に関して、米国は 22 項が GGE の作業に貢献していないと懸念を

ㄨ より国際人道法を遵守するために必要な場合、武力行使の制御を向上させ、武力行使の意思決定を改善するために、熟練した人間のオペレーターが兵器システムの機能を起動または停止させる明確な手順を提供すること、(d) 兵器システムがその予期された使用事態においてどのように作動すると予想されるかを評価することによって、そして、その使用事態において国際人道法の遵守を確保するためにシステムの計画された使用を見直すことによって、システムが十分に予測可能、信頼可能、理解可能・説明可能および追跡可能であることを保証すること。]

(53) 22 項「LAWS 分野での新興技術に基づく兵器システムの使用における人間と機械の相互作用は、国際法、特に、攻撃時の区別原則、比例原則および予防要件の履行を含む国際

表明し、インドは人間と機械の相互作用が国際人道法の要件と合致すべきであるという。トルコは、「must」と「should」の置き換えや、国際法の前に「applicable」の挿入を示唆した。ロシアは、技術開発を阻害しないために、軍事的必要性和人道的考慮とのバランスの達成に関する文言の追記や可能な措置が AI もしくは自律技術での進展またはそれへのアクセスを阻害すべきでないとの明記を提言した。

5月17日に、GGEは報告書草案の23～28項までの第1読を実施した。⁽⁵⁴⁾ 23項⁽⁵⁵⁾に関して、当該条項の削除を提案するロシアは、代替案として「LAWS分野での新興技術に基づく兵器システムを使用するかもしれない場合、国際人道法の遵守を強化するために学術研究で確認されまたは産業界で開発されたグッド・プラクティスを含む、人間と機械の相互作用の自発的なグッド・プラクティスを識別し共有することが有益である」を提案した。米国は、国際人道法は人間が「中断できる立場」にいることまで要求していないと発言して、本条項に懸念を示した。本条項は現行兵器システムに要求するよりも高い基準をLAWSに設定しているという。それに対して、オーストリアは、低い基準の現行兵器類によって、GGEが前進し自律兵器といった新システムの問題に対処するのを妨げられるべきではないという。オーストラリアやイスラエルも本条項に懸念を示し、インドは「must」の代わりに「should」または「may」の使用を主張した。ベルギーは本条項の「feasible (実行可能)」の削除を提案し、幾つかの国もそれに賛同した。

24項⁽⁵⁶⁾に関して、ロシアは、手続きを標準化する必要がないと述べて、

ㄨ 人道法と両立しなければならない (must)。」

(54) Laura Varela, “First Reading: Paragraphs 23-28,” *ibid.*, pp. 12-16.

(55) 23項「LAWS分野での新興技術に基づく兵器システムの使用責任者は、実行可能な場合に、国際人道法を遵守するために求められるように、システムまたはシステム機能を中心、無効化またはその他の方法で制御できる立場にいないなければならない (must)。」

(56) 24項「各国は、特に、以下のことを通じて、LAWS分野での新興技術に基づく兵器システムの使用に対する説明責任を確保すべきである。(a) 責任ある指揮統制系統内で当該システムを運用すること、(b) 国際法上の義務に従ってその使用結果に対する責任が個人および国家に帰属することを確保するために兵器システムを設計すること、(c) システ

24 項 (b) の削除と GGE で以前合意された文言の追加を提言した。他方、パキスタンは単に以前合意された文言の記述という考えに反対であった。米国は、(b) の削除、(c) の文言に反対、(d) の「ensure」の「enable」への置き換えを提言した。インドも、(c) に強い懸念を表明し、国家責任の基準は普遍的支持を得ていないと述べた。そして、(b)・(d)・(e) の省略または修正を、(c) をリスク評価・軽減措置の節への移行を主張した。パキスタンは、(d) や (e) の文言を支持し、ドイツは「適切な訓練」を含む (c) を歓迎した。

⁽⁵⁷⁾25 項に関して、ドイツや日本およびその他の国は本条項を支持し、ベルギーは法的審査が自己学習能力のある兵器システムに関して特に重要であると発言した。スリランカは、第 1 文の「seek to (するよう努める)」の削除を提案した。さらに、法的審査の場合に、兵器システムの使用が標的の生命に対する権利または尊厳に対する権利の侵害を構成するか否かの評価が含まれることから、兵器が国際人道法だけでなく国際人権法も遵守しているか否かを考慮することが重要であると付言した。ブルガリアなどは、本条項の第 1 文が以前合意した文言を薄めるかもしれないと懸念表明した。ロシアは、「must」を「should」への変更、第 2 文の削除、第 3 文の冗長さを指摘した。

⁽⁵⁸⁾26 項に関して、ロシアは、リスク評価や軽減措置といった範疇は国際

ムが交戦できる標的のタイプや武力行使を引き起こす事態を含む、システムの機能、能力および制限に関して使用者に適切な訓練を実施すること、(d) 個人の責任および適切な説明責任の評価や帰属を確保する関連の法規および手順、(e) 国際法上の国家の義務に従って、国際人道法の違反を含むかもしれない LAWS 分野での新興技術に基づく兵器システムに関連する事案を報告し調査するためのメカニズム。」

(57) 25 項「LAWS 分野での新興技術に基づく兵器システムの法的審査は、適用される国際人道法および現行の受容力や権能を考慮する国際法に従って使用可能かどうかを評価するよう努めなければならない (must)。法的審査は、適切な場合、特に、技術的性能、使用意図、可能な任務および標的タイプを考慮すべきである。この文脈で、諸国家間で関連するベスト・プラクティスを交換することで、遵守に関する適切な評価が強化され得る。」

(58) 26 項「リスクの評価および軽減措置は、意図せざる交戦を回避し攻撃時の区別、比例および予防の要件や原則を履行するのに必要となるかもしれない場合にシステム使用の予期される結果の理解を考慮しつつ、システムが十分に予測可能で信頼性が高まることを目

人道法のそれではなく単なる国家の特権であると述べ、LAWSの軍事利用に関する条項の挿入を提言した。それに対して、アイルランドは、CCWの任務は兵器を促進する機関として行動することではないと述べて、その提案に反対を示した。スイスは、責任、説明責任およびリスク軽減措置といった概念を支持して、報告書草案に維持されることを希望した。パキスタンやメキシコは、リスク軽減措置条項は20項や21項で規定された措置を補完すると強調した。中国は、21項、26項および27項の間でかなり重複があり、それらの統合か削除を示唆した。オーストリアなどはどのようなリスクを軽減する必要があるのかを精確に規定する合意された文言がなく軽減措置を議論することに疑念を持ち、そして、幾つかの国はそれを草案から取り除くことを提案した。米国は、本条項の焦点を意図せざる交戦リスクに置くべきと発言した。

⁽⁵⁹⁾ 27項に関して、カナダは、バイアス概念をエスニシティ（民族性）、ジェンダー（性差）、年齢およびディザビリティ（障害）などに拡大することを示唆し、多くの国がそれに賛同した。日本は、本条項に期間と地理的範囲が追加されたことを歓迎した。

⁽⁶⁰⁾ 28項に関して、パキスタンは、28項の最後の文が国際人道の要求よりも低い基準を無意識に設定しているため、その削除を提案した。中国とフランスは、「ethical reviews」を「relevant ethical considerations」に置き換えることを提言した。米国も、28項の最後の文が弱い表現となってい

、指すべきである。」

(59) 27項「リスク評価および軽減措置は、意図せざる交戦リスクや当該リスクを軽減するための制御・制限・その他の以下のような措置を考慮することも含めるべきである。システムが交戦できる標的のタイプ、システムの運用期間、地理的範囲および規模に影響を与える措置；システム・オペレーターの自動化バイアスを軽減する措置；兵器システムの使用に関するAI能力の意図せざるバイアスを軽減する措置。」

(60) 28項「国家は、LAWS分野での新興技術に基づく兵器システムの設計、開発、実験および配備においても、リスク評価および軽減措置を考慮すべきである。当該評価および軽減措置は、学際的観点を通じて研究・開発に組み込まれ、そして、倫理的審査を含めることができる。さらに、リスク評価および軽減措置に、文民の偶発的な死傷リスクや民用物への損害を最小限にするのを支援するために、文民の犠牲リスクや予防の考慮が含まれるべきである。」

るのでその削除を提案し、ニュージーランドもそれに賛同した。

5月17日に、GGEは報告書草案のマンドートに関する30項および1~14項を議論した。⁽⁶¹⁾ GGEは、30項に含まれた次年度のGGEマンドート(委任事項)を審議した。ロシアは本年度の成果を知る前に次年度のマンドートの内容を議論することは時期尚早であると主張した。米国は、現在は諸提案により議論の終了ではなく開始の段階にあると見なし、報告書案には諸提案を引き続き考慮するという言及が含まれていないと指摘した。インドとスイスは、引き続き諸提案の考察を支持した。オーストリアは、GGEが行うべき作業を完了したとの立場から、現在のマンドートを継続しても、それは必ずしも包括的成果に至ると考えていなかった。そして、別のGGEマンドートの設定を主張した。キューバも、過去2年の繰り返しではGGEに進展はもたらされまいだろうと考えた。

パキスタンは、3段階に分けてGGEの作業の進展を説明した。2019年までの第1段階で、GGEはその趣旨と目的を確立して作業を開始した。2019年からの第2段階で、GGEは法的、技術的および軍事的側面を考慮するよう要請され、指針原則を策定した。2021年からの第3段階で、現行のCCW議定書を斟酌しながら諸提案を考慮し精巧に練り上げた。この発展を念頭において、パキスタンは、マンドートの自然的な展開は2つの方向性(2層アプローチの収斂と法的拘束力のある文書交渉)が反映されることであると述べ、禁止と規制を含む法文書交渉の開始を30項に追加するよう提案した。多くの国家がそれを支持したが、米国は、法文書交渉のマンドートへの移行を受け入れる用意がなく、この方向性にコンセンサスが存在すると考えていなかった。米国の見解に賛同した日本は、全ての見解のバランスを反映した文書の必要性を強調した。ベルギーは交渉開始に合意は必要ないと主張し、パナマもごく少数の国家の政治的意思の欠如によって多数の国家の意思が害されることは受容できないと述べた。さらに、パナマは確立した実行としてコンセンサスを引き続き要請するが、そ

(61) Emma Bjertén, "First Reading: Paragraphs 30 and 1-14," *ibid.*, pp. 16-18.

のコンセンサスを拒否権として使うことは許されないと付言した。

GGE の会合日数に関して、幾つかの国家は、GGE に勢いをつけるために 10 日間以上を希望する考えを表明した。会合の最後に、GGE は報告書草案第 1 節（序説、1-3 項）、第 2 節（GGE の組織と作業、4-14 項）を議論したが、特に問題はなかった。

5 月 18 日午後に、GGE 報告書草案の第二読（20・21 項）⁽⁶²⁾ が行われた。多数の国が、特に、AWS と国際人道法との関係性および GGE 報告書の枠組みを国際人道法だけに基礎づけるのか、それとも国際法や倫理的考慮といったより広いものに基礎づけるのかを議論した。アイルランドは GGE 報告書が現行基準を害してはならないと再確認する一方で、ベルギーなどは、現報告書草案での文言が以前の報告書での文言より軟弱だけでなく、現行の国際人道法の諸原則や基準を害していると強調した。

多くの国家が、AWS の全サイクルで人間による判断・制御および AWS の人間のオペレーター側の説明責任・責任を確保する重要性を強調した。他方、米国は、人間による制御は現行法を反映しているという点に同意しておらず、文書（報告書 — 著者注）は、代表たちが法的要件を反映していると合意していないものを法的要件と認定してはならないと主張した。

5 月 19 日の午前に、GGE は報告書草案の第二読（22~28 項と 20 項）⁽⁶³⁾ を開始した。午後に、議長が諸国家との二者協議を行った。会合が公式モードでその日の終り頃に再開された時に、新たな 20 項草案⁽⁶⁴⁾ を議論した。公式会合の終了後に、非公式協議（市民社会は参加できず）が続けられた。

22 項に関して、字句の修正が議論された。23 項に関して、ロシアと米国はその削除を提案し、パナマやフィリピンその他はそれに反対した。24

(62) Audrey Kelly, "Second Reading: Paragraphs 20-21," *ibid.*, pp. 18-19.

(63) Laura Varela, "Second Reading: Paragraphs 22-28 and paragraph 20," *ibid.*, pp. 20-24.

(64) 報告書第 3 節：結論部分（15-29 項）の修正草案、*Draft Conclusions (18 May 2023)*, <https://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/documents/conclusions-draft.pdf>, accessed 14 March 2024.

項に関して、ロシアと中国は24項(e)の削除を、パキスタンは24項(c)の削除を提案した。25項に関して、幾つかの字句の修正が提案された。26項と27項に関して、パキスタンは、議長の26・27項統合案に規定される意図せざるバイアス（ジェンダー、エスニシティ、年齢、障害）に人種（race）の挿入を提言した。パレスチナはこのコメントに賛同するとともに、「宗教」の追加も示唆した。「avoid」か「prevent」か、「should aim to」か「must」か、など字句の修正が議論された。28項に関して、パキスタンやロシアは新たな条項案を提案した。報告書草案29項⁽⁶⁵⁾は、結論部分の修正草案の段階で削除された。

20項に関して、「開発」か「製造」の追加の提案があったが、米国は、それらが規定されたら、その文は「must」から「should」に変更せざるを得ないとして、いずれの追加の提案にも反対した。フィリピンによる代替案が提案され、議長も、諸国家との協議後の午後に、新たな提案を行った。議長は公式会合を閉会しても、非公式会合を継続した。その非公式会合は、NGOのオブザーバーが退場させられた。

(3) GGE 報告書と CCW 締約国会議

GGE 第2会期での議論を総括すれば、AWSの開発、保有および使用の明確な禁止・規制を規定することにほとんどの代表から圧倒的な支持を得たにも関わらず、最終のGGE報告書⁽⁶⁷⁾は、21項(b)のように、AWS使用において国際人道法を遵守しなければならないと主張するに過ぎなかったという。本報告書は、22項のように、「必要な場合、国家は…すべき(should)」などの表現を使って、公約（commitments）を提案（sugges-

(65) 29項「議長の責任で準備されたグループ会合中の議論の要約は、付属書〔II〕として本報告書に添付する。」

(66) Ray Acheson, “Editorial: Time to Leave the CCW Chatbox,” *CCW Report*, vol. 11 no. 3, 22 May 2023, pp. 1-2.

(67) *Report of the 2023 session of the Group of Governmental Experts on Emerging Technologies in the Area of Theal Autonomous Weapons Systems*, CCW/GGE.1/2023/2, 24 May 2023.

tions) に替える但書が挿入された。さらに、AWS の使用に関して自主的（非強制的）な制限が幾つかあり、それも非常に曖昧であると指摘された。その例として、人間を標的化できる兵器システムの禁止という広く支持された要求に言及することなく、「システムが交戦（engage）できる標的のタイプを制限する」という規定（22 項(a)）がそうある。また、国家が国内の審査過程を通じて、兵器システムの配備が国際法によって禁止されると決定しても、そのシステムの開発または使用をしないように求められず、23 項のように、単にベスト・プラクティスの交換を奨励されるだけであるという。加えて、報告書は、GGE の会議日数、その会議内容のマンデートを決定することなく、それらをそのまま 11 月の CCW 締約国会議に一任する形となった。

具体的には、報告書草案の第 3 節結論は 15 項から 29 項までの計 15 項にわたり詳細に規定されていたが、採択された報告書は 16 項から 23 項までの計 8 項に削減された。実質的な内容を規定する条項は、21、22、23 項だけであり、以下の通りである。

「21 項：国際法、特に、国連憲章や国際人道法および関連の倫理的観点に引き続き従う専門家会合（GGE を指す一著者注）の将来の作業を害することなく、専門家会合は次のように結論付けた。(a) 国際人道法は、LAWS の潜在的な開発と使用に全面的に引き続き適用される、(b) LAWS 分野での新興技術に基づく兵器システムは、国際人道法に従って使用できない場合には使用してはならない、(c) LAWS 分野での新興技術に基づく兵器システムに関する制御は、攻撃時での区別、比例性、予防措置の原則と要件を含む、国際法、特に、国際人道法の遵守を維持するために必要である。」

「22 項：国家は、LAWS 分野での新興技術に基づく兵器システムのライフサイクル全体を通じて、国際法、特に国際人道法に基づく義務の遵守を確保しなければならない。国家は、必要な場合、特に次のことを行うべきである。(a) システムが交戦できる標的のタイプを制限すること、(b) 兵器システム運用の期間、地理的範囲、および規模を制限すること、(c)

人間のオペレーターに適切な訓練と使用説明書を提供すること。]

「23 項：国際法上の国家の義務に従って、新しい兵器、戦闘の手段もしくは方法の研究、開発、取得または採用において、その使用が一定の場合またはすべての状況に国際法によって禁止されているか否かを決定しなければならない。この文脈で、国家安全保障上の考慮や専有情報に対する商業上の制限に留意しつつ、国家間での関連するベストプラクティスの自発的交換が奨励される。」

そして、CCW 締約国会議の年次会合が、11 月 15 日から 17 日までジュネーブで開催された⁽⁶⁸⁾。初日に、ロシアが国連諸機関・ICRC・NGO などのオブザーバー参加に反対して、1 日が過ぎた。2、3 日目に、各国代表は CCW 関連の議題（地雷、焼夷兵器、人口密集地での爆発性兵器）について駆け足で自国の立場を表明し、最終日に、オブザーバーを排除して、報告書案が議論された。

AWS に関して、ラ米・カリブ諸国（後述するベレン・コミュニケの署名国）、アフリカ諸国、非同盟諸国その他 100 以上以上の国家が、AWS の法文書交渉を要求し、中には CCW の議定書を支持する国家もあった。中国は、LAWS の定義や特性の問題について共通理解が生れ条件が熟したら、完全自律型致死兵器システムを禁止する法文書の締結を支持すると発言した。EU、カナダ、米国などを含む幾つかの国家は CCW 枠内で AWS 問題に取り組むことの支持を、そして、ドイツ・イタリアなどは 2 層アプローチの支持を表明した。EU・ドイツ・日本などは、後述する 2023 年 10 月の国連総会第 1 委員会での AWS 決議を歓迎した。ラ米・カリブ諸国やアフリカグループなどは、兵器や武力行使に対する有意な人間による制御の維持の必要性を強調した。さらに、予測可能性、比例性および説明可能性といった基準を保持し、かつ、確実に国際人道法や国際人権

(68) Laura Varela, "CCW Operates in the Dark for Lowest Common Denominator Outcomes," *CCW Report*, Vol. 11 No. 4, 20 November 2023, pp. 1-2, <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/hcp-meeting/reports/CCWR11.4.pdf>, accessed 20 March 2023.

法の義務が確実に遵守されるだけでなく、強化されるというアプローチを擁護した。しかしながら、コンセンサス方式が及ぼす頑固な障害や10年以上のCCW枠内での成果のない議論を見ると、国連総会などの他のフォーラムで作業を継続する必要があると指摘される。

結果として、2023年のCCW締約国会議の最終報告書は、第4節：結論と勧告において、以下のように規定された。⁽⁶⁹⁾ 次年度のGGEマンドートは、法文書交渉は明記されず、例年と同様の規定振りとなった（20項）。ただし、GGEのマンドートは1年ごとではなく、3年間分が設定された。GGEは、コンセンサス方式で報告書を採択し、それを2026年開催予定の第7回再検討会議（Review Conference）に付託しなければならないのである。GGEは、できる限り早期（願わくは2025年末まで）に、その作業を完了すべきであるという。GGE議長はGGEの作業に関して締約国の年次会合を更新することになる。GGEの会合は毎年10日間と規定された。

LAWS関連のGGE会合は、2024年3月4日から8日までの第1会期と8月26日から30日までの第2会期の計10日間開催される（24項(c)）。CCW締約国会議は、2024年11月13日から15日までの開催予定とされた（24項(f)）。

3. CCW枠外での議論

LAWS規制に関する議論は、ここ10年間、専らCCW枠内で我慢強く実施されてきたが、遅々として進まず、具体的な成果物の達成までに至らなかった。そのために、以下のように、CCW枠内の議論と同時並行して、CCW枠外のプラットフォームで議論が活発に行われるようになってきた。それらの動きを概観する前に、LAWS規制の議論の中で重大な影響を及ぼす米国の国内動向を把握しておく必要がある。

(69) Final report, CCW/MSP/20213/7 (Advanced version), 23 November 2023.

(1) 米国防総省指令の改訂

2023年1月に、米国防総省指令「兵器システムにおける自律性⁽⁷⁰⁾」が公布から10年余りを経て改訂された。改訂版は、不明瞭だった審査プロセスを「第4節：若干の自律兵器システム審査のガイドライン」の「図1：上級審査と承認が必要か否かの判断を手助けするためのフォロー・チャート」のように詳細かつ明瞭に改善した。さらに、「第1節：一般的発行情報」の「2項：方針 (Policy)」に幅広いAI倫理原則が追加されているが、いわゆる殺人口ロボットの開発や配備を実際に禁止しているわけではないという⁽⁷¹⁾。そして、管理や規制の層が厚くなり、自律型兵器の開発が難題のように思えるかもしれないが、他方、満たすべき基準が明確になったことで容易に手続きを進めることができることから、指揮官やプロジェクト・マネージャーにとっては「青信号 (a green light)」が灯されたとも言えるという。実際に、2023年8月28日に、ヒックス米国防副長官より「レプリケーター (Replicator) 構想」が発表された⁽⁷²⁾。それによれば、米国は、今後、18か月から24か月以内 (2025年2月～8月—著者注) にマルチ・ドメイン (陸、海上、海中、空・宇宙—著者注) で数千の使い捨て可能 (attribution⁽⁷³⁾) な無人の自律システムを配備するという。このように、米国防

(70) DoD Directive 3000.09, *Autonomy in Weapons Systems*, January 25, 2023, <https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf>, accessed 26 March 2024. 本指令の発効により、2012年11月21日付の旧指令は失効した。2012年の国防総省指令について、岩本誠吾「致死性自律型ロボット (LARs) の国際法規制をめぐる新動向」『産大法学』47巻3・4号2014年1月342-344頁。

(71) Sydney J. Freedberg JR, “DoD’s clarified AI policy Flashes ‘green light’ for robotic weapons: Experts,” *Breaking Defense*, February 09, 2023, <https://breakingdefense.com/2023/02/dods-clarified-ai-policy-flashes-green-light-for-robotic-weapons-experts/>, accessed 26 March 2024.

(72) U. S. Department of Defense, *Speech Deputy Secretary of Defense Kathleen Hicks Keynote Address: ‘The Urgency to Innovate’ (As Delivered)*, Aug. 28, 2023, <https://www.defense.gov/News/Speeches/Speech/Article/3507156/deputy-secretary-of-defense-kathleen-hicks-keynote-address-the-urgency-to-innov/>, accessed 3 April 2024. 牧田純平「レプリケーター——中国に対抗するアメリカのUAV開発イニシアティブ——」先端技術安全保障研究所2023年11月16日 <https://giest.or.jp/contents/briefs/4076/>, 2024年4月3日閲覧。

(73) Attributable とは、「低価格で、再利用可能で、最終的に使い捨て可能な (low-cost, reus-

は、AI を活用した自律兵器システムを積極的に導入しようとしている。

では、本指令を、Paul Scharre による条文の逐条解釈⁽⁷⁴⁾および NGO のヒューマン・ライト・ウォッチ（Human Rights Watch, HRW）とハーバード大学法科大学院国際人権クリニックによる法的分析を参照しながら概観してみよう。本指令は、その目的で⁽⁷⁵⁾「武装プラットフォーム」に限定されていることから、兵器を搭載していない（非武装）偵察ドローンは本指令の適用外となる⁽⁷⁶⁾。

1.2(a) では、「自律および半自律兵器システムは、指揮官およびオペレーターが武力行使に対する適切なレベルの人間による判断（appropriate levels of human judgment）を行使することができるように設計される（will be designed）」と規定された。HRW と国際人権クリニックは、人間による判断の「適切なレベル」が旧版も改訂版も不明確であると批判している⁽⁷⁸⁾。そして、改訂版が「制御（control）」の用語を使用していないが、control は、精神的判断（mental judgment）と物理的行為（physical act）の両方を包含するために、control の方が適切であると主張する⁽⁷⁹⁾。また、

able, and ultimately expendable)」を意味する。Thomas Hamilton, David Ochmanek, *Operating Low-Cost, Reusable Unmanned Aerial Vehicles in Contested Environments*, Rand Corporation, 2020, p. 1.

(74) Paul Scharre, “Noteworthy: DoD Autonomous Weapons Policy,” Center for a New American Security, February 06, 2023, <https://www.cnas.org/press/press-note/noteworthy-dod-autonomous-weapons-policy>, accessed 26 March 2024.

(75) Harvard Law School/International Human Rights Clinic, “Background Briefing: Review of the 2023 US Policy on Autonomy in Weapons Systems,” February 2023, https://www.hrw.org/sites/default/files/media_2023/02/DoDrobots_2.13.2023_Final_0.pdf, 26 March 2024.

(76) 「目的：本指令は、遠隔操作または搭乗員による操作の武装プラットフォームを含めて、兵器システムにおける自律および半自律機能を開発し使用するための政策を確立し、その責任を割り当てる」。旧版では、「manned and unmanned platforms」であったが、これらのジェンダー関連用語は意図的に削除され、「armed platforms that are remotely operated or operated by onboard personnel」に変更された。

(77) 本指令の 1.1 適用 b. (2) でも、その適用除外に非武装プラットフォームが規定されている。

(78) Note 75, p. 4.

(79) Note 75, p. 5.

旧版では法的義務を示す「shall be」が使用されていたが、それが、ある事象が確実に発生することを示唆する「will be」に変更された。Paul Scharre は特に文言に変更なし⁽⁸⁰⁾としているが、HRW と国際人権クリニックは、法的義務が解消されたのかと、それに疑念を抱いている⁽⁸¹⁾。

1.2a(1)(b) で要求される自律・半自律兵器システムは、「指揮官およびオペレーターの意図に従って、時間的制約、地理的領域、その他関連する環境上および作戦上の制約の中で交戦を完了する」と規定された。旧版では時間的制約しか言及されなかったが、改訂版で地理的制約に加え、その他の環境上・作戦上の制約も追記された。本条項は、国防総省が時間的にも地理的にも制限されない自律兵器の配備ができないことを意味する⁽⁸²⁾という。

1.2b で、「自律または半自律兵器システムでの AI 能力の利用は、国防総省 AI 倫理原則と一致する」と規定された。本条項は、2012 年以降の重要な国防総省の政策発展の一つである 2020 年 2 月に国防総省が採択した「AI 倫理原則⁽⁸³⁾」との関係性を明確にする新規定であった。AI 倫理原則（責任、衡平、追跡可能、信頼可能、統制可能）は、そのまま 1.2.f. で再現されている。

正式な兵器の開発や配備の前の上級審査を必要としない兵器システムに、1.2.d(2)「企図された緊急攻撃または飽和攻撃をインターセプトする領域防御用に物的標的 (material targets) を選択し交戦するために使用されるオペレーター監視型自律兵器システム」が挙げられた。「オペレーター監視型」は旧版の「人間監視型」の置き換えであり、「物的標的」は

(80) Note 74, 1.2. Policy.

(81) Note 75, p. 4.

(82) Note 74, 1.2. (1)(b).

(83) US Department of Defense, "DOD Adopts Ethical Principles for Artificial Intelligence," Feb. 24, 2020, <https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/>, accessed 26 March 2024. 米国防総省の AI 倫理原則について、岩本誠吾「自律型致死兵器システム (LAWS) 規制 —— 多層的・多角のアプローチからの規制の試み ——」『軍縮研究』11 卷 1 号 2022 年 17 頁。

「標的として人間を選択することを除き」の置き換えであった。本条項は、人間ではない物体（たとえば、航空機、車両、ロケット、ミサイル）を標的化する監視型自律兵器を認可しているという⁽⁸⁴⁾。

HRW と国際人権クリニック⁽⁸⁵⁾は、改訂版は旧版を劇的に変更しておらず、多くの諸国が主張する有意な人間による制御がなくとも作動する AWS や人間を標的にする AWS を禁止する新提案を十分反映していないと批判している。さらに、正式な兵器の開発や配備の前の上級審査の要件は、「緊急の軍事的必要な場合に」特定の高官職員によって免除が付与されるという「ウェーバー条項」(4.2) が旧版にも改訂版にも規定されている。そのため、それが深刻な抜け穴になっていると指摘し、ウェーバー条項の規制を勧告している。

また、旧版の指令は LAWS 交渉開始の 2013 年 11 月より以前の 2012 年に策定されたもので、不完全なものであったとしても、提示された問題を認識させ対処する方向への重要な措置であったと評価している。他方、10 年後の改訂版の指令は AWS の禁止・規制に関する新条約に関する広く支持された国際的な提案と乖離している⁽⁸⁶⁾と批判した。

(2) 国際会議と地域会議

(a) オランダ・韓国共催「軍事領域での責任ある AI (REAIM) サミット」

(i) REAIM 行動要請 2023 年 2 月 15・16 日にオランダ・ハーグにおいてオランダ・韓国共催の「軍事領域での責任ある AI サミット」⁽⁸⁷⁾が開催された。ロシアはウクライナ侵攻のために招待されなかったが、米・中を含む 50 カ国余りが参加した⁽⁸⁸⁾。本会議は、AI が偵察、監視、状況分析の軍

(84) Note 74, 1.2. d(2).

(85) Note 75, pp. 1-3.

(86) Note 75, p. 7.

(87) The Summit on Responsible AI in the Military Domain, REAIM 2023.

(88) REAIM 2023 については、The Brussels Times with Belga, "International conference on the use of AI in the military opens in the Netherlands," in *The Brussels Times*, 16, February 2023; Toby Sterling and Stephanie van den Berg, "Duch host first summit on

事領域において既に利用され、将来、ドローン・スウォームを含む自律標的選択や核指揮統制システムでの利用の可能性も指摘される状況下で、初めて軍事領域での AI 利用規制に向けた第 1 段階となった。

本会議に参加した政府代表者は、会議の成果として、軍事領域での責任ある AI の開発、配備および利用に関する「REAIM 行動要請 (REAIM Call to Action)⁽⁸⁹⁾」に合意した。オランダ国防相の発言のように、AI は将来の軍隊にとって極めて重要なものである。適切な枠組みや立法を導入すれば、AI 利用により我々の作戦上および兵站上のプロセスがより簡便でより能率的となり、さらに、より迅速でより適切な情報に基づく意思決定が可能となる。そのことにより、自軍の保護だけでなく、最大限可能な範囲で相手側への危害を制限することができるという。なお、本行動要請に賛同した諸国は、豪・中・仏・独・伊・日・蘭・韓・土・英・米を含む 60 か国 (後に、57 か国) であった (ロシアは招待されず、ウクライナは出席せず、イスラエルは署名せず⁽⁹⁰⁾)。

前文と本文 25 項から構成される本行動要請 (添付資料 1: 「軍事領域での責任ある AI 行動要請 (REAIM Call to Action)」参照) は、法的義務を設定するのではなく、あくまで AI の軍事利用に関する現状認識および国家が今後取るべき行動を奨励する非法文書である。具体的な内容は、次の

、 ‘responsible’ use of AI in the military,” in *Reuters*, February 15, 2023.

(89) 関連記事については、Government of the Netherlands, “Call to action on responsible use of AI in the military domain,” 16-02-2023, <https://www.government.nl/ministries/ministry-of-foreign-affairs/news/2023/02/16/reaim-2023-call-to-action>, accessed 18 January 2024. 軍事領域での責任ある AI 行動要請 (REAIM Call to Action) の文書については、<https://www.government.nl/ministries/ministry-of-foreign-affairs/documents/publications/2023/02/16/reaim-2023-call-to-action>, accessed 18 January 2024. 日本外務省は、本行動要請を「軍事領域における責任ある AI 利用 (REAIM) 宣言」と称している。外務省「軍事領域における責任ある AI 利用 (REAIM) イニシアチブ」2023 年 2 月 28 日 https://www.mofa.go.jp/mofaj/gaiko/arms/page23_004201.html, 2024 年 1 月 19 日閲覧。

(90) *REAIM 2023 Endorsing Countries and Territories*, <https://www.government.nl/ministries/ministry-of-foreign-affairs/documents/publications/2023/02/16/reaim-2023-endor-sing-countries>, accessed 19 January 2024. 後のリストには、ボリビア、グアテマラ、インドネシア、カザフスタンの 4 か国が削除され、ギリシャが追記されたために 57 か国となった (accessed 3 April 2024)。

通りである。① 軍事領域での責任ある AI 利用が国際法上の義務に従っていること（前文 5 項）、② 人間が軍事領域での AI 利用時の決定の責任を負うこと（6 項）、③ AI の軍事利用に関するグッド・プラクティスや教訓の情報交換・共有を奨励していること（8、14、18、23、24 項）、④ AI の軍事利用に関するすべてのステークホルダー（国家、民間部門、市民社会、学界、研究所、シンクタンクなど）の交流・協力・情報交換が重要であること（9、15、16、17、19、20、21、22、23、24、25 項）、⑤ AI システム利用での適切な安全措置および人間による監視（human oversight）の確保が重要であること（12 項）。

なお、本会議では、オールラウンドの認識を高め、軍事領域での AI の定義方法を明確にし、その技術が責任をもって開発し製造し配備する方法を決定するために、軍事領域での責任ある AI グローバル委員会（Global Commission on Responsible AI in the Military Domain）の設立が予定された。また、軍事 AI の規制に関する国際討議の定期的行事となったと言えるかは断言できないが、2024 年に韓国・ソウルで第 2 回 REAIM サミットが予定された。

（ii）米国イニシアチブ 上記の REAIM 行動要請とは別に、米国は、2 月 16 日の REAIM2023 において「AI および自律性の責任ある軍事利用に関する政治宣言」イニシアチブ（構想）を公表した⁽⁹¹⁾。その宣言の目的は、如何に軍が責任をもって AI および自律性をその作戦に組み込むことができるかを巡り国際的コンセンサスを構築すること、そして、本技術が国際法、安全保障や安定性の尊重を確実に促進するするために防衛目的用の本技術を国家が開発し、配備し、そして、使用する際の指針の助けとなることであるという。当該宣言は、防衛文脈（defense context）での責任ある

(91) U.S. Department of State, “Building Consensus on the U.S. Framework for a Political Declaration on the Responsible Military Use of Artificial Intelligence and Autonomy,” February 16, 2023, <https://www.state.gov/building-consensus-on-the-u-s-framework-for-a-political-declaration-on-the-responsible-military-use-of-artificial-intelligence-and-autonomy/>, accessed 10 October 2023.

AI 利用のためのベスト・プラクティスを記述する一連の法的拘束力のない指針から構成されている。その後公表された政治宣言⁽⁹²⁾でも、同様に、責任ある行動をめぐる国際的コンセンサスを構築し、軍事 AI の国家による開発、配備および使用の指針となることを目的として位置付けている。そして、宣言はベスト・プラクティスの情報交換や国家の能力向上のための基盤を規定しており、それによって賛同国（endorsing states）は経験や理念を共有することができるという。米国は、今後、その政治宣言の賛同国間の定期的な対話を招集し、これらの責任ある実行の国際的支援や履行をさらに促進する考えであり、2024 年の第 1 四半期に会合の開催が予定された。

本政治宣言は、ファクト・シートによれば、確認事項として、政治宣言は自衛権および責任をもって AI を開発し軍事領域で利用する国家の権能を認めている⁽⁹⁵⁾。以下、注釈を参照しながら、本政治宣言（添付資料 2：米

(92) U.S. Department of State, “Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy,” November 9, 2023, <https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy-2/>, accessed 19 January 2024. 2024 年 2 月 12 日現在、52 か国（加、仏、独、伊、日、韓、土、英、米を含む）であり、中、露、イスラエルは含まれない。2023 年 11 月 9 日公表の政治宣言は、同年 2 月 16 日の政治宣言案から大幅に文言修正が行われた。たとえば、旧 B 項は、「国家は、核兵器使用（nuclear weapons employment）に関する主権的決定を知らせ、実行するために重要なすべての行動について、人間の制御および関与（human control and involvement）を維持すべきである」と規定され、核兵器使用に言及していた。しかし、11 月の新版ではその部分が削除された。もっとも、2023 年 11 月 13 日の同政治宣言の初会合（ニューヨーク）では、日本は、核兵器のリスク軽減に関して、米・英・仏が人間の制御および関与を維持するとのコミットメントを歓迎する発言をした。国際連合日本政府代表部「『AI と自律性の責任ある軍事利用に関する政治宣言』に関する式典における石兼大使ステートメント」https://www.un.emb-japan.go.jp/itpr_ja/ishikane111323.html, 2024 年 3 月 18 日閲覧。

(93) 日本の参加に関して、外務省「『AI と自律性の責任ある軍事利用に関する政治宣言』への我が国の参加」2023 年 11 月 14 日、https://www.mofa.go.jp/mofaj/press/release/press5_000156.html, 2024 年 1 月 20 日閲覧。

(94) U.S. Department of State, *Fact Sheet: The Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy*, November 1, 2023, <https://www.state.gov/political-declaration-on-the-responsible-military-use-of-artificial-intelligence-and-autonomy/>, accessed January 20 2024.

(95) Shawn Steene, Chris Jenks, “The Political Declaration on Responsible Military Use Of

国提案の「AI および自律性の責任ある軍事利用に関する政治宣言」参照)を概観する。

序説部分では、AI の軍事利用は適用される国際法に従っていなければならないと、特に、武力紛争での AI 利用は、国際人道法の諸原則 (慣習法) に従わなければならないと指摘する。なお、本宣言での「軍事 AI 能力 (military AI capabilities)」とは、AI 活用兵器 (AI-enabled weapons) または自律兵器システムと同義ではなく、兵器のみならず防衛指導者がすべてのレベルでより良好でより適時な決定を下すのを助ける意思決定支援システムが含まれるという。そして、本宣言は、法的拘束力はなく、賛同国の国際法に基づく義務を追加したり変更したりするものでもないという。

本政治宣言は、A 項から J 項まで 10 項目にわたり、国家がなすべき (should) 行動を宣言している。A 項に関して、それは、政治レベルでの賛同だけでなく、軍事防衛組織も実際に一定の措置の採択や履行を確保するための作業を伴うことを示している。B 項に関して、本国軍隊が武力紛争法に従って確実に軍事 AI 能力を利用するための適切な措置を取る必要性が強調されている。適切な措置として、法的審査や、武力紛争法違反の報告メカニズム、武力紛争法教育⁽⁹⁶⁾などが含まれるという。C 項に関して、それは軍事 AI 能力の使用だけでなく、開発や配備にも言及している。そして、賛同国の防衛官庁の一般的な監督を超えて政府高官に対して、より具体的に⁽⁹⁷⁾より詳細に軍事 AI 能力の開発や配備を監督するよう求めている。D 項に関して、それは国防総省の AI 倫理原則の「衡平 (Equitable)」に⁽⁹⁸⁾

↘ Artificial Intelligence and Autonomy,” in *Lieber Institute(West Point) Articles of War*, Nov 13, 2023, <https://lieber.westpoint.edu/political-declaration-responsible-military-use-artificial-intelligence-autonomy/>, accessed 20 January 2024.

(96) 米軍では、米国防総省指令「戦争法プログラム (DoD Directive 2311.01, DoD Law of War Program, July 2, 2020)」がある。

(97) その実行例として、米国防総省指令「兵器システムの自律性 (DoD Directive 3000.09, Autonomy in Weapon Systems, January 25, 2023)」では、自律兵器システムの正式な開発前と再度実践配備前に、その上級審査が要請されている。

(98) C. Todd Lopez, “DOD Adopts 5 Principles of Artificial Intelligence Ethics,” Feb. 25, 2020, *DOD News*, <https://www.defense.gov/News/News-Stories/article/article/2094085/dod->

基づいており、北大西洋条約機構（NATO）の責任ある使用原則⁽⁹⁹⁾の「バイアス軽減化（Bias Mitigation）」に類似している。意図せざるバイアスを完全に排除することは達成不可能な基準であるため、排除（eliminate）ではなく、最小化（minimize）の用語が基準として使われている。また、意図せざるバイアスには、人種、肌の色、出身国、宗教もしくは信仰、性別もしくはジェンダーを含むが、それらに限られないという。

E 項に関して、特定の軍事 AI 能力の開発、配備および利用において適切な注意を払う「関係する」者は、軍隊構成員や国防職員の様々な要員を含んでいる。F 項に関して、これは、一定の国防機関以外の職員がそれらの方法論、設計手順および文書化を理解し、監査できるようにすることを提案または要求することを意図していないという。G 項に関して、これは、軍事 AI 能力を使用する要員（オペレーター）またはその使用を承認する要員（指揮官）が当該能力の使用に関して文脈に基づいた判断を下さなければならないことを強調している。それには、当該要員がシステムの能力と限界を十分に理解しておく必要があり、その理解があれば、自動化バイアスのリスクも軽減するという。H 項に関して、明確で十分定義された用途があることを軍事 AI 能力に保証することで、その意図された機能を満たすように設計し製作することが促進される。

I 項に関して、実験と保証は、異なる軍事 AI 能力全般にわたる異なるレベルのリスクや共通点のない機能不全の結果を考慮するために、「適切」で「厳格」であるべきだという。そして、軍事 AI 能力の機能不全の結果が重大であればあるほど、その実験や保証はより厳格になる。J 項に関して、それは、兵器に限定されず、そして、攻撃に使用するかまたは攻撃に

↘ adopts-5-principles-of-artificial-intelligence-ethics/, accessed 16 March 2024. 「衡平：国防総省は、AI 能力での意図せざるバイアスを最小限にするための慎重な措置を取る。」

(99) Zoe Stanley, Edward Hunter Christie, “An Artificial Strategy for NATO,” 25 October 2021, in *NATO REVIEW*, <https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html>, accessed 16 March 2024. 「バイアス軽減化：AI アプリの開発および使用ならびにデータ・セットで意図せざるバイアスを軽減するために事前措置を取る。」

関連する能力でない可能性があるかもしれない、すべての軍事 AI 能力に適用される。J 項は、システムが作戦中に使用されている場合だけでなく、システムのライフサイクルの様々な段階で機能することができる適切な安全装置⁽¹⁰⁰⁾を備えている。そして、「適切な措置」は、機能不全から起こり得る結果を念頭に、重大な結果が発生する場合には、それに応じた厳格な保護措置が必要になるという。最後の部分で、様々な解釈があるとしても、本宣言での「人工知能（AI）」と「自律性（Autonomy）」が定義されている。

注釈の結論は、本宣言が重要な第一歩であるとしても、まだ最初の一步にしかすぎないということである。その後、賛同国は 2024 年初め頃の会合（実際には、3 月に第 1 回全体会議が開催⁽¹⁰¹⁾）で、責任ある方法で合法的に軍事 AI 能力を開発、配備、使用する能力を相互に向上させるために、ベスト・プラクティスや会得した教訓を共有し始めると締めくくられた。

(iii) REAIM サミットのその後

第 2 回 REAIM サミットは、オランダと共同議長国であった韓国のソウルで 2024 年 9 月 9、10 日に開催される予定である。その前段階で、韓国主催の会議が 2 つ開催された。その一つとして、国連と韓国の共催で 2023 年 12 月 4、5 日にジュネーブで「第 22 回軍縮・不拡散合同会議」が、「軍事領域での人工知能（AI）のガバナンス」という議題の下で開催された⁽¹⁰²⁾。本会議は、2024 年開催予定の第 2 回 REAIM サミット（ソウル）

(100) Shawn Steene と Chris Jenks は、具体例として、宇宙ロケットの AI 活用モニタリング・自己破壊システムを例示している。Cf., Note 95.

(101) 2024 年 3 月 19・20 日に、「AI・自律性の責任ある軍事利用政治宣言に関する第 1 回全体会議が開催された。その時点での同政治宣言賛同国数は、54 か国となっている。U.S. Department of State, Office of the spokesperson, “Inaugural Plenary Meeting for the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy,” March 29, 2024, <https://www.state.gov/inaugural-plenary-meeting-for-the-political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/>, accessed 1 April 2024.

(102) UN Press Release, “UN, Republic of Korea Host Twenty-Second Joint Conference on Disarmament and Non-Proliferation, Focusing on Military Artificial Intelligence,” DC/3865, 4 December 2023, <https://press.un.org/en/2023/dc3865.doc.htm>, accessed 19 March 2024.

の準備段階として位置付けられた。その初日に、リー国際安全保障担当韓国大使および中満国連事務次長兼軍縮担当上級代表は、軍事 AI のツールやアプリの安心・安全な設計、開発および使用を確保するために必要な信頼性を構築するために国際規範の確立の重要性を強調した。本会議は、国連事務総長「平和に対する新たな課題」⁽¹⁰³⁾の中での AI に関する要請に共鳴していると評価された。

もう一つの国際会議として、韓国は、オランダとシンガポールと共同主催国として、地域の諸国間で出現しつつある軍事領域の AI 開発について理解を深め、公開の議論を促進するために、シンガポールで 2024 年 2 月 26、27 日に「REAIM 地域協議（アジア）」⁽¹⁰⁴⁾を開催した。共同主催国 3 か国に加えて、12 か国の外務・防衛職員が参加し、意見交換が行われた。⁽¹⁰⁵⁾

(b) コスタリカ主催「CELAC 地域会議」⁽¹⁰⁶⁾

2023 年 2 月 23、24 日に、ラ米諸国とカリブ諸国の 33 か国がコスタリカのベレンに参集して、「自律兵器の社会的・人道的影響に関するラ米・カリブ会議」を開催した。本会議で採択された「自律兵器の社会的・人道的影響に関するラ米・カリブ会議コミュニケ」⁽¹⁰⁷⁾は、a から h までの前文 8 項と本文 3 項から成る（添付資料 3：「自律兵器の社会的・人道的影響に

、 2024.

(103) 国家が「国際人道法や国際人権法での加盟国の義務と一致して、AI の責任ある設計、開発および使用に関する国家戦略を早急に作成すること」を要請している。後述「3. (3)

(a) 国連事務総長イニシアチブ」参照。

(104) Korian Ministry of Foreign Affairs, Press Release, “REAIM Regional Consultations (ASIA) Held on February 26-27,” 2024-02-29, https://overseas.mofa.go.kr/eng/brd/m_5676/view.do?seq=322428, accessed 19 March 2024.

(105) オーストラリア、中国、インド、インドネシア、日本、マレーシア、モンゴル、ニュージーランド、パキスタン、フィリピン、タイ、ベトナムの 12 か国。

(106) ラテンアメリカ・カリブ諸国共同体 (Comunidad de Estados Latinoamericanos y Caribeños, CELAC) は、33 か国から構成される地域的国際機構である。

(107) Communiqué of the Latin American and The Caribbean Conference of Social and Humanitarian Impact of Autonomous Weapons, February 23 & 24 2023,

<https://conferenciaawscostarica2023.com/wp-content/uploads/2023/02/EN-Comuniquue-of-La-Ribera-de-Belen-Costa-Rica-February-23-24-2023.pdf>, accessed 20 March 2024.

関するラ米・カリブ会議コミュニケ（バレン共同声明）」参照）。その中で、「戦闘の更なる非人間化を防止するために有意な人間の制御（meaningful human control）を維持すること」（d項）の重要性を明記するとともに、「兵器システムの自律性に関する禁止と規制のある国際法文書の緊急交渉」（1項）を要請する地域的コミュニケとなった⁽¹⁰⁸⁾。なお、本会議には、オブザーバーとして13か国（豪、仏、独、日、蘭、露、米など）も参加した。

(c) ルクセンブルグ主催国際会議

2023年4月25、26日に、ルクセンブルグ主催の「ルクセンブルグ自律兵器システム会議」が開催された⁽¹⁰⁹⁾。本会議には、5つのパネル（政治的見解、軍事適用、国内・国際規制、技術的課題と規範・規制、倫理的側面）があり、そこでは、グローバル・サウスでの自律兵器の使用可能性から軍隊内での自律性の履行方法のより良き理解を主張する議論まで様々な意見が交換された。

AWSの法的枠組みを確立する意欲の程度は様々であったが、ほとんどの参加者は運用上または法的な枠組みに向けた合意が早期に実現することへの激励と決意を表明した。さらに、会議では、システムが文民をより良く保護するかもしれないという主張、自律兵器システム使用の境界問題、様々な「自律性レベル」などの検討を含む、より技術的な議論が掘り下げられた。本会議は、それほど大掛かりなものではなかったが、国連CCWとは別に、本問題に取り組む意欲が高まっていることを改めて示したという。

(108) Ousman Noor, “Significant act of political leadership as 33 Latin American and Caribbean States deliver Communiqué on AWS,” Stop Killer Robots, 27/02/2023, <https://www.stopkillerrobots.org/news/significant-act-of-political-leadership-as-33-latin-american-and-caribbean-states-deliver-communication-on-autonomous-weapons-systems/>, accessed 20 March 2024.

(109) Ullduz Sohrabi, “Luxembourg Autonomous Weapons Systems Conference 2023: A global gathering explores legal, ethical and technological aspects of autonomous weapons,” Article 36, 27.04.23, <https://article36.org/updates/luxembourg-autonomous-weapons-systems-conference-2023-a-global-gathering-explores-legal-ethical-and-technological-aspects-of-autonomous-weapons/#:text=On%20April%2025%2D26%2C%202023,%2C%20research%20institutes%2C%20and%20academia,> accessed 20 March 2024.

(d) トリニダード・トバゴ主催「CARICOM⁽¹¹⁰⁾地域会議」

2023年9月5、6日に、トリニダード・トバゴ主催の「CARICOM 会議：自律兵器の人的影響」⁽¹¹¹⁾が開催された。その会議で採択された「自律兵器システムに関する CARICOM 宣言」⁽¹¹²⁾（添付資料4：AWSに関する CARICOM 宣言（CARICOM 会議「自律兵器の人間による影響」参照）によれば、武力行使に対する有意な人間による制御が不可欠であり、AWSの禁止と規制を組み合わせた2層アプローチの国際法文書の追求が奨励される（1項）。その2層アプローチに関して、有意な人間による制御なしに武力行使ができる予測不可能または制御不可能な AWS および対人用武力行使に設計されるかそのように使用される AWS を禁止する一方で、その他の形式の AWS を規制する国際法文書の交渉を目指す（2項）というコミットメント（公約）が宣言された。また、人種、民族、国籍、階級、宗教、性差、年齢およびその他の地位に基づくデジタルによる差別を含めて、AWSに関連する拡散の危険性、意図せざる拡大、倫理的考慮、デジタルによる非人間化およびその他人間的・社会的含意を認識するように諸国家に要請するコミットメントが規定された（7項）。

(e) フィリピン主催「インド太平洋地域会議」

2023年12月13、14日にフィリピン主催「LAWSに関するインド・太

(110) ^{カリコム}CARICOM とは、the Caribbean Community（カリブ共同体）の略称で、14カ国1地域から構成される地域的国際機構である。

(111) Isabelle Jones, “Historic Declaration Adopted at CARICOM Conference,” Stop Killer Robots, 06/09/2023, <https://www.stopkillerrobots.org/news/historic-declaration-adopted-at-caricom-conference/#:~:text=The%20CARICOM%20Declaration%20on%20Autonomous%20Weapons%20Systems%2C%20adopted%20on%206,international%20legally%20binding%20instrument%20that>, accessed 20 March 2024. INPS Japan「カリブ諸国、『殺人口ロボット』禁止」2023年10月21日 <https://inpsjapan.com/news/un-civil-society/caribbean-moving-against-killer-robots/>, accessed 20 March 2024. なお、CARICOM とは、Caribbean Community（カリブ共同体）を意味し、14カ国1地域が加盟している域内の経済統合を目指す地域的国際機構である。

(112) CARICOM Declaration on Autonomous Weapons Systems, https://www.caricom-aws2023.com/_files/ugd/b69acc_4d08748208734b3ba849a4cb257ae189.pdf, Accessed 20 March 2024.

自律型致死兵器システム（LAWS）規制議論の動向と今後の展望

平洋地域会議」がフィリピン・マニラで開催された。⁽¹¹³⁾ その会議には、インド太平洋地域 24 개국、オブザーバー 5 개국および国際市民社会機構 16 団体が参加した。主催者国フィリピンは、LAWS などの新興技術に対応するために、インド太平洋諸国の声をもっと挙げるよう要請した。第 1 パネルは最近の AI 技術の発展を概観し、第 2 パネルは国際人道法の遵守や倫理的、社会のおよび人道的含意に関する課題について議論した。そして、最後のパネルは、フィリピン作成の CCW 議定書 VI 案⁽¹¹⁴⁾ を含め、自律兵器のリスクや課題に対処するための措置について議論をした。

(3) 国連事務総長と国連総会

(a) 国連事務総長イニシアチブ

2023 年 7 月 22 日に、国連事務総長は、「我々の共通課題：政策概要 9 平和のための新課題」⁽¹¹⁵⁾ を公表した。第 4 議題「平和および紛争の潜在的領域に対する新アプローチ」の「行動 11：新興領域の兵器化を防止し、責任あるイノベーションを促進する」において、自律型致死兵器システムの禁止について次のように述べている。「完全自律兵器システムは、戦争を劇的に変更する潜在力があり、現行の法的枠組みを曲げてしまい、浸食することさえあり得る」と指摘する。そして、「人間による関与なく命を奪う力や裁量を持つ機械は、道義的に不快であり、政治的にも受け入れられず、国際法によって禁止すべきである」と主張する。さらに、「多国間交渉での進展を踏まえて、2026 年までに、人間による制御または監視なく機能し、国際人道法を遵守して使用することができない LAWS を禁止し、

(113) Philippines Department of Foreign Affairs, “Philippines calls for Indo-Pacific voices to Address Lethal Autonomous Weapons Systems Risks,” 15 December 2023, <https://dfa.gov.ph/dfa-news/dfa-releasesupdate/33834-philippines-calls-for-indo-pacific-voices-to-address-lethal-autonomous-weapons-systems-risks>, accessed 20 March 2024.

(114) フィリピンを含む 14 国が共同提案した作業文書「自律兵器システムに関する議定書案（第 6 議定書案）を指していると思われる。CCW/GGE.1/2023/WP.6, 10 May 2023.

(115) United Nations, *Our Common Agenda Policy Brief 9 A New Agenda for Peace*, July 2023, p. 27, <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-new-agenda-for-peace-en.pdf>, accessed 30 January 2024.

そして、それ以外のすべてのタイプの自律兵器システムを規制するための法的文書を締結せよ」と勧告している⁽¹¹⁶⁾。

さらに、10月5日に、「諸国家が自律兵器システムの新たな禁止と制限を確立するように国連事務総長と赤十字国際委員会総裁による共同要請⁽¹¹⁷⁾」が発表された。その共同要請によれば、「我々は、武力行使に対して人間による制御を保持するために今行動しなければならない。人間による制御は、生死の決定において保持されなければならない。機械によって人間を自律的に標的化することは、我々が越えてはならない道義的一線である。人間による関与なく生命を奪う力や自由裁量権を持つ機械は、国際法によって禁止されるべきである」と訴えた。そして、その効果が予測できない方法で機能する自律兵器システムは禁止される。たとえば、機械学習のアルゴリズム（基本的にはそれ自体を書き込む予測不可能なソフトウェア）に制御される自律兵器を認めることは、危険な提案であり受け入れ不可能であるという。他方、それら以外の自律兵器には、国際法の遵守や倫理的な受容可能性を確保するために、明確な制限が必要であるという。それには、使用の場所、時間、期間、標的タイプ、使用される武力の規模を制限すること、そして、効果的な人間の監視、適時な干渉および不活性化（effective human supervision, and timely intervention and deactivation）の能力を確保することが含まれるという。

また、国際人道法は、若干の兵器を禁止し、その他すべての兵器の使用に関して一般的な制限を設けているが、しかし、AWSを規律する特定の国際協定がなければ、一般的な規則がどのように適用されるかについて国家によって異なる解釈が生まれる可能性がある。それゆえ、自律兵器に関する新たな国際規則は、現行法を明確化し強化するために必要であるとい

(116) *Ibid.*, p. 27.

(117) ICRC, *Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems*, 05 October 2023, <https://www.icrc.org/en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems>, accessed 30 January 2024.

う。それらは予防措置であり、当該兵器によって影響を受ける者を保護する機会にもなるだろうと指摘している。

そして、最後に、国連事務総長と赤十字国際委員会総裁は、世界の指導者に、自律兵器システムの明確な禁止と制限を規定する新たな法文書の交渉を開始し、2026年までにその交渉を終えるよう訴えた。

(b) 国連総会での議論

国連では、2022年10月に国連総会第1委員会⁽¹¹⁸⁾で70カ国共同声明が提出されたが、それに関連する決議が採択されたわけではなかった。翌年の2023年12月に初めて、国連総会⁽¹¹⁹⁾はLAWS決議を採択した。本決議は、自律兵器システムに国際法、特に国連憲章、国際人道法および国際人権法が適用されることを確認し（前文1項）、兵器システムにおけるAIと自律性に関するものを含めて、軍事領域での新技術の適用が、人道的、法的、安全保障上、技術的および倫理的見地から生じる深刻な挑戦ならびに懸念に留意している（前文3項）。さらに、LAWSに関連する最近の動きとして、2022年10月7日にコンセンサスで採択された人権理事会決議51/22⁽¹²⁰⁾

(118) *Joint Statement on Lethal Autonomous Weapons Systems First Committee, 77th United Nations General Assembly Thematic Debate-Conventional Weapons, 21 October 2022.* https://estatements.unmeetings.org/estatements/11.0010/20221021/A1j8bNfWGIL/KLw9WYcSnnAm_en.pdf, accessed 22 March 2024. 70カ国の中には、米・英・日・仏・独・伊・豪・韓・加が含まれ、他方、中・印・露・イスラエル・イラン・トルコはそれに参加しなかった。そこでは、国際人道法に従って使用できない自律兵器システムの禁止とその他の種類の自律兵器の規制という2層アプローチおよび人間による判断、制御または関与（human judgement, control or involvement）といった用語の異なる理解があるかもしれないが、人的要素（human element）が武力行使の中核であり、中核で居続けなければならないという共通認識が存在する。そして、人間が兵器システムの使用に関連して適切な制御、判断および関与を行使する必要性が強調された。

(119) A/RES/78/241, 28 December 2023 (adopted 22 December 2023). 総会第1委員会では（第1委員会報告書A/78/409, 10 November 2023）、賛成164（後にガンビア賛成165）、反対5（印・露など）、棄権8（中・北朝鮮・土・イラン・イスラエルなど）、欠席16で採択された。本会議では、賛成152、反対4（印・露など）、棄権11（中・北朝鮮・土・イラン・イスラエルなど）、欠席26で採択された。

(120) 2022年10月7日の人権理事会決議「軍事領域での新技術および新興技術の人権上の含意」(A/HRC/RES/51/22, 11 October 2022). ここでは、特に、データ・セット、アルゴリズムによるプログラミングおよび機械学習プロセスに依存する軍事領域での新興技術の

(前文 6 項)、英・韓共同主催サミット 2023/2/15-16、コスタリカ主催 2/23-24、ルクセンブルグ主催 4/25-26、トリニダード・トバゴ主催 9/5-6 などの国際会議や地域会議（前文 7 項）および「平和の新課題」イニシアチブでの事務総長の努力（前文 9 項）に言及している。

そして、国際社会が、特に、GGE を介しての自律兵器システムが提示する課題および懸念に対処する焦眉の急を強調し（本文 1 項）、事務総長に加盟国やオブザーバー国家（パチカン市国とパレスチナ）に LAWS に関する見解を聴取し、次年度の 79 会期の総会に当該見解を含む実質的な報告書を付託するよう要請した（本文 2 項）。加えて、国際機構、地域機構、ICRC、市民社会団体、学術団体や産業界の見解も求めるように事務総長に要請した（本文 3 項）。そして、79 会期の暫定議題に「Lethal Autonomous Weapons Systems」を挿入することが決定された（本文 4 項）。

4. まとめにかえて——2026 年に向けて

(1) 三つ巴対立から 2 項対立への移行

CCW 枠内で大勢を占める法文書作成交渉のマンデート挿入は、コンセンサス方式で 2023 年度も例年と同様に実現できなかったが、その中でも、若干の議論の変化が垣間見られた。3 月 6 日の一般発言の中で、インドが政治宣言または非拘束的合意を支持し、イスラエルもコンセンサスで合意が可能であれば、現行国際人道法の明確化を目的とする西側 6 か国の作業文書を支持できると発言した。また、5 月 17 日にロシアが、その作業文書の結論で示したように、人間と機械の相互作用に関する自発的なグッド・プラクティスの共有を提案した。従来、追加的な法規制は不要であると主張していたロシア、イスラエルおよびインドが、西側諸国の主張する

ㄨ 使用が人権の侵害や濫用および国際人道法の違反を引き起こし推進していることを認識し（前文 8 項）、人権理事会諮問委員会に、軍事領域での新興技術の人権上の含意を研究し、第 60 会期（2025 年 8 月か 9 月か）の人権理事会に研究報告書を提出するよう要請した（本文 1 項）。

非法文書（政治宣言やベストプラクティス編集）派に接近してきたように思える。その意味では、LAWS 規制論争は、前述した三つ巴の対立構造から、法文書（ハード・ロー）による AWS 規制派か非法文書（ソフト・ロー）による AWS 規制派かの 2 項対立構造へと緩やかに移行してきたのかもしれない。

従来、米国等のソフト・ロー派は、行動準則やグッド・プラクティスなどの自発的措置が将来の法規制に繋がる可能性を指摘していた⁽¹²¹⁾。NGO を含めてハード・ロー派の諸国は、ソフト・ローの成果物が達成されれば、その時点で国際法文書（条約）作成へのモメンタム（勢い）が低下し、国際立法まで行き着かなくなることを強く恐れていた。しかし、AI 技術の急激な発展により AI リスクの脅威を放置できない事態を迎えて、追加的法规制不要派の主張が説得力を失ってきたことから、論争の軸は、むしろ、国際規制・立法に向けて漸進していると解釈できよう。

(2) CCW 枠外での議論の活性化

2023 年に活性化した CCW 枠外での議論は、2 系統に区分できる。それは、① CCW 枠内での議論の延長上にあり LAW/SAWS 規制に関するものと、② LAWS/AWS という兵器システムを超えた AI・自律技術の軍事利用に関するものである。①は、さらに、(i) CCW 枠内での特定の議論（AWS 規制の法文書交渉促進派）を外部から支援する応援団的な地域会議（CELAC 地域会議や CARICOM 地域会議など）の議論と、(ii) CCW 枠とは別の議論の受け皿（フォーラム）の提供を暗示する国連総会の議論に細分される。特に、注目すべきは、後者の (ii) の国連総会のものである。

CCW 締約国会議は、「LAWS 分野での新興技術の問題を取り扱うための適切な枠組み⁽¹²²⁾」と評価され、今まで LAWS 規制議論を独占してきた。

(121) 注 1「岩本前掲資料」114 頁。

(122) 2019 年に GGE が策定した指針原則 (k)、CCW/GGE.1/2019/3, 25 September 2019, Annex IV Guiding Principles, p. 13.

しかし、CCW 枠内での議論が 10 年間も成果なく時間が過ぎた状況にシビレを切らした多くの諸国家が LAWS/AWS 規制に関する別の議論の受け皿を模索し始めた。その動きの表れとして 2022 年の国連総会第 1 委員会に 70 カ国の共同声明であり、2023 年の LAWS 規制問題に関する総会決議および次年度（2024 年度）の LAWS 問題の継続審議であった。

この国連総会の動きは、CCW 締約国会議の議論を促進する最大の圧力になるとともに、硬直化した LAWS 規制問題の突破口を開く画期的な出来事のように思われる。というのも、CCW 枠内でその専管事項のような LAWS 規制問題について何らかの成果物が達成されなければ、それは当該問題が国連総会側に移行する可能性を示唆し、ひいては、CCW 締約国会議の存在意義すら問われかねないからである。

なお、これに関する留意点は、国連総会は、コンセンサス方式を採用する CCW の議論とは異なり、多数決で議論を進めることができる利点がある一方で、少数の見解を含め、妥協点を追求する姿勢が弱まるリスクもある。国連総会の議論であっても、あくまで AI 研究開発国を含む LAWWS/AWS 規制内容のコンセンサスを追求することが望まれる。もはや、AI 研究開発国は CCW 枠内でコンセンサス方式を濫用して法文書の作成を阻止するだけでは済まされず、その意味において CCW 枠内での議論の深化が期待される。

もう一つの系統の②は、既に実戦で活用中のパランティア・テクノロジーズの「ゴサム」といった AI 意思決定支援システムを議論の対象とする AI の軍事利用に関するものである。その例として、オランダ・韓国共催の「REAIM サミット」と米国の「AI・自律性の責任ある軍事利用政治宣言構想」が挙げられる。それぞれの試みの中で、すでに非法文書（ソフト・ロー）形式の「REAI 行動要請」（60 か国→ 57 か国賛同）および「AI・自律性の責任ある軍事利用政治宣言」（52 → 54 か国賛同）が策定されている。今後、開催される関連の国際会議を介して、これら二つの非法文書の行動準則がより精緻化され、そして、その賛同国が一層増加することが期待される。その延長上に、政治的コミットメントのソフト・ローか

【図2：AI利用領域と規制状況】

| | |
|--|--|
| AI 民生利用： 倫理原則・行動規範（ソフト・ロー）から 法規範（EUのAI規則案や米大統領令などハード・ロー）へ | |
| AI 軍事利用： 軍事利用政治宣言（ソフト・ロー） | * 人間中心主義 * リスク・ベース・ アプローチ（禁止と規制） |
| AI 自律兵器：規制なし CCW・国連総会で審議中 * 人的制御・判断・関与 * 2層アプローチ（禁止と規制） | |

著者作成

ら法的義務のハード・ローへの転換も視野に入ってくる。

責任あるAIの軍事利用規制問題は、AWS規制問題それ自体とは言えないが、標的（人的・物的軍事目標）を選定し軍司令官に軍事作戦案を提言する「ゴサム」のように、自律的な攻撃命令の一手前段階でのAIの軍事利用であり、LAWS問題とは極めて密接な関係性にある。別稿に委ねるが、AIを巡る規制問題は、現在、民生用AI、軍事用AIおよびAI兵器の3系統で同時並行的に議論が進行している（【図2：AI利用領域と規制状況】参照）。規制対象は異なるが、AIという共通の新興科学技術を取り扱っている点から、3系統の規制動向は相互関連性がある。そして、共通課題は、巨視的に言えば、人類とAIの共存・共生の在り方の追求とも言える。LAWS/AWS規制の在り方は、AI規制の全体的な観点を踏まえて検討する必要がある。

(3) 2026年が勝負年

今後の予定として、LAWS規制問題は、2024年にCCWのGGE（3月会期と8月会期に計10日間）と11月のCCW締約国会議で議論する。それと同時並行して、国連事務総長は各加盟国やオブザーバー国から聴取したLAWS関連の見解を纏めた報告書を国連総会に提出し、それを基に総会（第1委員会および本会議）で議論する。同様に、2025年もCCWと国連総会の2系統の議論が見込まれる。そして、2026年に、CCWのGGEは、5年毎に開催されるCCW第7回再検討会議に報告書を付託する。

その報告書の作成作業は、2025 年末までに完了することが望まれている。この 2026 年とは、国連事務総長が強く要望してきた LAWS 禁止・規制の法文書を交渉し締結し終えるデッドラインでもある。

核兵器に次ぐ「第 3 の軍事革命」または「戦場のゲーム・チェンジャー」と称される LAWS/AWS の登場が急迫している状況の中で、国際法規制問題は、時間との戦いであり、この 3 年間で大きな山場を迎える。それは、近未来を占う、人類の叡智が真剣に問われる 3 年間とも言える。人類と AI の共存・共生のためにも、LAWS/AWS 規制問題は国際法上、早急に解決されなければならない喫緊の課題であることは言うまでもない。

＊「添付翻訳資料（全文または部分翻訳）」

資料 1：「軍事領域での責任ある AI 行動要請（REAIM Call to Action）」

人工知能（AI）は、我々の世界に影響を与え、それを根本的に変えようとしている。

我々は、AI が、我々の働き方や生活様式に影響を与えているように、将来の軍事作戦に劇的な影響を及ぼすと認識している。軍事は、一連のアプリやコンテンツを横断してその AI 利用を増やしつつある。

AI は、実現可能な技術として顕著な機会を与え、桁違いの潜在能力を有しており、他の利点の中でも特に、我々が以前では想像できないほどのデータ量を有効に活用することができ、意思決定を改善できる。しかしながら、我々は、現在までその多くを予見できない関連リスクが存在することも認識している。

軍事領域での AI 利用をめぐり、そして、AI システムの信頼性の潜在的低さ、人間による関与問題、責任に関する明確性の欠如や意図せざる潜在的な結果、その他潜在的な影響の中でも特に、武力の一連の範囲内で意図せざる段階的な拡大リスクが世界的に懸念されている。

我々は、軍事領域での AI の責任ある利用が国際法上の義務に完全に従い、かつ、国際的な安全保障、安定性および説明責任を害しない方法で援用されることが最も重要であると強調する。

行動要請によって、我々は、政府、企業、研究機関、国際機構その他が以下のことをするのを支援するように奨励する。

1. 我々は、国際的な安全保障や安定性において、軍事領域での AI システムの急速な採用の結果として機会や課題を含めて、潜在的な影響を認識する。
2. 我々は、武力紛争での文民や民用物への被害リスクを軽減するための AI 活用を含めて、幅広い目的や人類のために、軍事領域での AI 活用の潜在的な能力を認める。
3. 我々は、軍事領域での幅広い活用に亘り AI の導入から結果として生じる含意や課題を十分に理解せず、予測せず、そして、できないことを認める。
4. 我々は、軍事領域での AI の影響について我々の理解を高める必要があると理解する。それは、神話を打破し、軍事領域での AI の利益、リスクおよび限界に関する知識やリテラシーを改善することを含む。
5. 我々は、関連する国家戦略、AI 原則や国際的イニシアチブを含む、軍事 AI の責任ある開発、配備および利用について多くの行為者が行った作業や、軍事領域に AI を組み込むことで生じる課題に効果的に対処するために異なるステークホルダー集団が構築した専門知識を認識する。
6. 我々は、AI が意思決定を形成しそれに影響を与えるために利用できることに留意する。そして、我々は、人間が軍事領域で AI を利用する際に決定の責任および説明責任を負い続けることを確保するために努力する。
7. 我々は、軍事領域や AI が適用される異なる軍事的文脈において多様な AI 技術の現在および将来の適用の様々なタイプに含まれるリスクを評価する必要性を認識する。
8. 我々は、十分な研究、実験および保証なく尚早の採用をすることは、結果として不注意からの損害となり得るが、適宜に AI を採用しないことで結果として軍事的不利益となり得ることも認識する。我々は、

リスク軽減の実行や手続きに関して会得した教訓の交換を増やす必要を感じる。

9. 我々は、軍事領域での AI 利用のあり得る影響、機会および課題に対処する上で全体的、包括的、包摂的なアプローチの重要性ならびに国家、民間部門、市民社会および学界を含むすべてのステークホルダーが軍事領域での責任ある AI に関する情報を協力して交換する必要性を強調する。
10. 我々は、AI システムのデータが国際法および関連の国内的・地域的・国際的な法的枠組みやデータ基準と両立する方法で、収集、利用、共有、文書化、適切な場合には削除されるべきであることを確認する。適切なデータ保護やデータ品質ガバナンス・メカニズムが、AI 訓練データの確保や利用における場合も含めて、設計の初期段階から確立し確保すべきである。
11. 我々は、軍事的意思決定の分散した性質や AI システムの複雑さにより、軍事領域での AI の開発、配備および利用のすべての段階で細心の注意を払うことが必要であると認識する。我々は、官民の協力を奨励し、軍事領域での AI の開発、配備および利用に関わる複数のステークホルダーとの関係を継続するよう努める。
12. 我々は、時間と能力の制約による人間の限界に留意しつつ、AI システムの利用に関する適切な安全措置と人間による監視を確保することの重要性を繰り返す。
13. 我々は、AI を活用する軍人は、データにおける潜在的バイアスといった制約から生じる結果を含めて、AI システムの特徴、当該システムの利用の潜在的結果を十分理解すべきであり、それ故、望ましくない効果を回避するために AI システムでのユーザー相互作用や信頼性の方法について研究、教育および訓練が必要であると認識する。
14. 我々は、軍事領域での AI の利用に関する国内の枠組みや政策の相互理解を高めるために、国家間で学んだグッド・プラクティスや教訓の交換を推奨する。我々は、また、規範、政策、原則および技術的な専

門知識に関して、民間部門が学んだグッド・プラクティスや教訓の重要性も肯定する。

15. 我々は、軍事領域での AI の履行が国ごとに異なることを認識する。軍事領域での AI の責任ある利用は、すべての国家、特に発展途上国がその機会から利益を受け、課題やリスクに対処するために、国際的でマルチなステークホルダーの交流が必要である。
16. 我々は、軍事領域での AI 利用から生じる利益、ジレンマ、リスクおよび課題についてバランスの取れた国際的でマルチなステークホルダーの議論の継続が必要であると理解する。我々は、国際的な民間部門、学界、市民社会および他の関連ステークホルダーがマルチレベルで議論に積極的に貢献するように奨励し、軍事領域での責任ある AI を促進する。

我々の行動要請

17. AI 分野での技術的開発は、主として民間領域で発生している。それ故、我々は、軍事領域での AI 導入がマルチなステークホルダーの課題であると認識している。我々は、マルチステークホルダーにおいて、そして、包括的な方法で軍事領域での責任ある AI に関する地球規模の対話を継続し、すべてのステークホルダーに国際法に従って国際的な安全保障や安定性に貢献することにおいて、その責任を果たすよう要請することをコミット（約束）する。
18. 我々は、国家が研究、訓練課程および能力構築活動を通じての知識形成によって軍事 AI の一般的な理解を高めるよう奨励する。我々は、国家が共同作業し、会得したグッド・プラクティスや教訓を交換し、自らの能力を高めることで知識を共有するよう奨励し、そして、民間部門、市民社会および学界が軍事領域での責任ある AI を促進するようかかわらせる。
19. 我々は、国家が軍事領域での責任ある AI の国内的枠組み、戦略および原則を開発するよう奨励する。
20. 我々は、軍事領域での責任ある AI を促進する、国家、学界、市民社

会、産業界その他のステークホルダーによるイニシアチブを歓迎する。

21. 我々は、軍事領域、特に、人間と機械のチームング領域で、異なる軍事的文脈で異なる AI システムの多面的な利用事例を認識しつつ、世界中の学界、研究所およびシンクタンクが急速に AI を採用することの影響、機会および課題をより良く理解するために追加的な研究を行うことを支援する。
22. 我々は、学界、研究所およびシンクタンクが国際的な安全保障や安定性に貢献するために軍事領域での AI 利用が提示する課題に手順を提案し、そして、現実的な解決策に貢献するよう奨励する。
23. 我々は、民間部門に軍事領域での責任ある AI 利用を支援し促進するよう要請する。我々は、国家安全保障の考慮や商業専有情報の制限に留意しつつ、他のステークホルダーとともに企業による責任ある AI 利用に関するグッド・プラクティスや政策、特に、軍事領域での AI 利用に関連するかもしれないグッド・プラクティスや政策の共有を奨励する。
24. 我々は、軍事領域での責任ある AI 利用に関するグッド・プラクティスおよび政策を通じて学際的議論を確保するために、軍事領域での AI の開発、配備および利用を導くベスト・プラクティスに関するマルチステークホルダーの対話を奨励する。
25. 我々は、世界中のすべてのステークホルダーがこの行動要請に参加するよう求める。

資料 2：米国提案の「AI および自律性の責任ある軍事利用に関する政治宣言」

自律的な機能やシステムを可能にするために AI 利用を含むことがある軍事 AI 能力を開発する国家が増加している。AI の軍事利用は、倫理的であり、責任を負い、そして、国際安全保障を高めることができるし、そうすべきである。AI の軍事利用は、適用される国際法に従わなければな

らない。特に、武力紛争での AI 利用は、国際人道法の基本原則を含めて、その国家の義務に従っていなければならない。AI 能力の軍事利用は、責任ある人間の指揮統制（a responsible human chain of command and control）内で軍事作戦中の当該利用を含み、責任を負う必要がある。AI の軍事利用に対する原則アプローチは、リスクと利益の注意深い考察を含むべきであり、また、意図せざるバイアスや偶発事故を最小限にすべきである。国家は、自律的な機能やシステムを可能にするものを含めて、軍事 AI 能力の責任ある開発、配備および利用を確保するための適切な措置を取るべきである。これらの措置は、軍事 AI 能力のライフサイクルを通して関連する段階で適用されるべきである。

賛同国は、以下の措置が自律的な機能やシステムを可能にするものを含めて、軍事 AI 能力の開発、配備または利用において実施されるべきであると信じる。

- A. 国家は、自国の軍事組織が AI 能力の責任ある開発、配備および利用のためのこれらの原則を採択し実施するように確保すべきである。
- B. 国家は、自国の軍事 AI 能力が国際法、特に国際人道法の下での各自の義務と両立して使用されるように確保するために、適切な措置、たとえば、法的審査を行うべきである。国家は、また、国際人道法の実施を高め、武力紛争での文民や民用物の保護を改善するための軍事 AI 能力の利用方法も考慮すべきである。
- C. 国家は、高級官吏がそのような兵器システムを含み（それに限定されないが）、重大な結果をもたらすアプリを持つ軍事 AI 能力の開発および配備を効果的に、そして、適切に監督することを確保すべきである。
- D. 国家は、軍事 AI 能力で意図せざるバイアスを最小限にするための先行的措置を取るべきである。
- E. 国家は、関係する者が軍事的な AI 能力を装備した兵器システムを含み、当該能力の開発、配備および利用において適切な注意を払うよう確保すべきである。

- F. 国家は、軍事 AI 能力が、関係する防衛要員に平明で、そして同要員による検査が可能な方法論、データ源、設計手順、文書化をもって開発されるよう確保すべきである。
- G. 国家は、軍事 AI 能力を利用しまたはその利用を承認する者が、当該システムの利用に関して適切な文脈に基づく判断を下し、自動化バイアスのリスクを軽減化するために、当該システムの能力や限界を十分理解するように訓練を受けることを確保すべきである。
- H. 国家は、軍事 AI 能力に明確で十分定義された用途があり、それらが意図した機能を満たすよう設計され工作されるよう確保すべきである。
- I. 国家は、軍事 AI 能力の安全性、保安性、実効性がその十分定義された用途内でそのすべてのライフサイクルを通じて適切で厳格な実験および保証に従うよう確保すべきである。国家は、自己学習するかまたは継続的に更新される軍事 AI 能力のために、重要な安全性の特徴がモニタリングのような過程を通じて低下しないように確保すべきである。
- J. 国家は、軍事 AI 能力における機能不全リスクを軽減するために適切な安全措置、たとえば、意図せざる結果を検知し回避する能力や、具体的には配備済みシステムが意図せざる行動を示した際にそれを解除するか不活性化することで対応する能力を満たすべきである。

本宣言の目的を促進するために、賛同国は、以下のことを行う。

- ・自律的な機能やシステムを可能にするものを含め、軍事 AI 能力を開発、配備または使用する際に、これらの措置を取る。
- ・本宣言に対するコミットメント（約束）を明らかにし、これらの措置の実施に関する適切な情報を提供する。
- ・軍事 AI 能力が責任をもって合法的に使用されるよう確保する他の適切な努力を支援する。
- ・如何に軍事 AI 能力が責任をもって合法的に開発、配備および使用されるかを賛同国間で継続して議論する。
- ・これらの措置の効果的な実施を促進し、これらの措置を改善しまたは賛

同国が適切と判断する追加措置を確立する。そして、

- ・関連主題に関する他のフォーラムを含め、他のフォーラムでの関連主題の進行中の議論を害することなく、その他の国際社会がこれらの措置を促進するように一層働きかける。

賛同国は、AI や自律性の概念は様々な解釈が可能であることを認識している。本宣言の目的のために、AI とは、人間の知能を必要とする任務を実行する能力を指すと理解され得る。これには、パターンの認識、経験からの学習、結論の導き、予測の見立て、提言の生成が含まれる。あるAI アプリは、自律的な物理システムの行動を導きもしくは変更させ、または、純粋にデジタル領域に留まっている任務を実行することができた。自律性は、一つのスペクトルとして、そして、起動後に更なる人間による干渉なく作動するシステムを含むものと理解され得る。

資料3：「自律兵器の社会的・人道的影響に関するラミ・カリブ会議コミュニケーション」（ベレン共同声明）

- a. 新興技術により、国際の平和と安全に具体的な課題が提示され、戦闘での人間の役割について新たな疑問が生じている。兵器システムにおける自律性の効果および有意な人間による制御へのその影響についての全体論的な理解が、倫理的、法的、人道的および安全保障上の含意を十分評価するために必要とされる。
- b. 新たな兵器技術が開発され武力紛争で使用され続けているので、人道性の諸原則、公共の良心の命令および倫理的視点に由来する国際人道法ならびに当該原則を含めて、国際法による新たな禁止や規制が早急に必要とされている。
- c. 国連憲章、国際人権法および国際人道法に確立された目的や原則を含めて、国際法に従って、国際共同体は、文民の避難、死傷および民用物の破壊という結果を防止するために、国際法的枠組みを発展させ強化することにより、これら脅威に対処するよう求められている。
- d. 戦闘の更なる非人間化を防止し、個人の説明責任や国家の責任を確保

するために、有意な人間による制御を維持することが重要である。

- e. ラ米・カリブ諸国は、積極的に CCW 枠組み内の LAWS 分野での新興技術に関する政府専門家の議論に積極的に参加している。ある国家集団は、将来の規範的枠組みのための諸要素や CCW 枠組み下の第 6 議定書草案を提示した。当該地域の幾つかの国家は、自律兵器に関する 2022 年 10 月の国連総会での 70 カ国による声明にも参加していた。
- f. さらに、2023 年 1 月 24 日の第 7 回 CELAC（ラテンアメリカ・カリブ諸国共同体）サミットでのプエノスアイレス宣言は、2014 年 1 月（29 日）にハバナでの第 2 回 CELAC サミットで署名された平和地帯にとしてのラ米・カリブ諸国の宣言を強調した。
- g. 国連事務総長、司法外の略式または恣意的な執行に関する特別報告者、障害者の権利に関する国連特別報告者、赤十字国際委員会（ICRC）、市民社会、科学団体および学術団体は、兵器システムでの自律性についての懸念を表明し、諸国家がそれに関する法文書の交渉をするようにと繰り返し要請している。
- h. 本地域の諸代表は、兵器システムでの自律性が提示するリスクや課題とともに、有意な人間による制御を保障するために、禁止と規制を確立する国際法的文書の必要性を承認した。

ラ米・カリブ諸国の諸代表は、以下のような更なる行動が必要であると認識している。

1. 国際人道法や倫理的視点を含む国際法の遵守および兵器システムの自律性に伴う社会的・人道的影響の防止を確保するために、兵器システムの自律性に関する禁止と規制を有する国際法文書を緊急に交渉するように共同して促進すること
2. さらなる当該交渉に対する共通の立場を促進するために、積極的に関与することをコミットすること
3. それらの専門意見や支援を引き出すために、国際機構、市民社会組織、学界その他関連のステークホルダーとの協力やパートナーシップを

継続し強化すること、そして、

4. CCW の枠組みその他の多国間フォーラムで推進できる可能な勧告を作成するために、ラ米・カリブ諸国間での議論を継続すること

2023年2月23、24日

コスタリカ・エレディア州ラ・リベラ・デ・ベレン

資料4：「AWSに関するCARICOM宣言」（CARICOM会議「自律兵器の人間による影響」トリニダード・トバゴ、ポート・オブ・スペイン、2023年9月5、6日）

我々、カリブ共同体（CARICOM）代表は、ここに以下のことを確認するために、2023年9月4、5日にトリニダード・トバゴのポート・オブ・スペインでの「CARICOM会議『自律兵器の人的影響』」に参集した。

前文

人工知能（AI）を含む新興技術が、包括的経済成長を開く力を有し、人間の生命を高める並外れた潜在力を持ち、深刻な社会的課題に対処していることを認識しているだけでなく、新興技術が、特に、自律兵器システム（AWS）の場合に、地球規模の安全保障の含意を含めて、国家や地域の安全保障に対するリスクを引き起こしていることも認識して、

第1節：基本的原則および懸念（略）

第2節：倫理的・法的命令（略）

第3節：協調的努力と多角的協力（略）

第4節：共同行動と将来のコミットメント（略）

第5節：最終的なコミットメントおよび勧告

それ故、我々は以下のコミットメント（公約）を宣言する。

1. 武力行使に対する有意な人間による制御の不可欠性を支持し、それ故、AWSの禁止と規制を組み合わせた国際法文書の追求を奨励する決意である。
2. 有意な人間による制御なしに武力行使ができる予測不可能または制御不可能なAWSを禁止し、そして、対人用の武力行使に設計されるか

そのように使用される当該兵器を禁止するが、他方で、その他の形式の AWS には規制を導入する国際法文書の交渉を目指す努力に協力する。

3. 他の適切なイニシアチブの支持および AWS の懸念事項に対処する多国家間のプラットフォームへの積極的な参加を誓う。

4～12 (略)