

研究ノート

## 致死性自律型ロボット (LARs) の 国際法規制をめぐる新動向

岩本 誠 吾

### 目 次

1. はじめに —— 問題の所在
2. ロボット兵器の現状
3. ロボット兵器規制の動き
  - (1) 2010年8月のオールストン報告
  - (2) 2012年11月のヒューマン・ライト・ウォッチ報告
  - (3) 2012年11月の米国防省指令 3000.09
  - (4) 2013年4月のヘインズ報告
  - (5) 2013年5月の人権理事会及び6月の英国議会
4. 国際人道法上の論点
  - (1) 兵器に関する国際人道法の枠組み
  - (2) 兵器自体の合法性に関する議論
  - (3) 兵器使用の合法性に関する議論
5. まとめにかえて —— LARs の更なる議論に向けて

### 1. はじめに —— 問題の所在

情報 RMA (Revolution in Military Affairs、軍事における革命又は軍事革命) の契機となった 1991 年の湾岸戦争以降、戦争のハイテク化が急激に進展してきた。<sup>(1)</sup> そこでは、遠隔操作による偵察・諜報目的の無人航空機 (Unmanned Aerial Vehicle, UAV) が完全に兵器体系の一部に組み込まれ、2001 年の対テロ・アフガニスタン紛争以降では、無人戦闘航空機 (Unmanned Combat Aerial Vehicle,UCAV) も登場してきた。<sup>(2)</sup> 米国の

無人航空システムは、2003年のイラク紛争を経て、2000年から2008年までに50機以下から6,000機以上に増加し、と同時に、実戦配備された無人陸上車両（Unmanned Ground Vehicle, UGV）も、2001年の100両以下から2007年までに4,400両以上に急増した<sup>(3)</sup>。このように、無人兵器、言い換えれば、ロボット兵器<sup>(4)</sup>が戦闘システムの中で重要かつ不可欠の構成要素となり、現実の戦場に多数投入されている<sup>(5)</sup>。もはや、ロボット兵器のない戦争は、将来、全く想像できない。

言うまでもなく、ロボット兵器は、省人化・省力化だけでなく、人間に代わって、危険な（Dangerous）任務、核・生物・化学兵器などで汚染された環境内での汚い（Dirty）任務、長時間勤務による疲労や精神的弛緩に繋がる単純で単調な（Dull）任務、及び可能な活動範囲を超えた奥深い（Deep）任務を実行することができるという軍事的利点を有するからである<sup>(6)</sup>。

今後、人工知能を含む科学技術の急激な発展を想定すれば、人間の意思が介在することなく標的を選択し攻撃することができる完全自律型兵器（full autonomous weapons）、いわゆる殺人口ロボット（Killer Robots）が20年から30年以内に開発されると予測される<sup>(7)</sup>。これは、1984年公開の米国映画「ターミネーター」のごとく、近未来の戦争では殺人口ロボットがロボット独自の判断で人間を殺傷することを暗示している。もっとも、ここで議論の対象は、あくまで敵対行為・戦闘行為（物体の破壊及び人の殺傷）を実施するロボット兵器<sup>(8)</sup>であって、国連文書では、特に、対人用殺傷兵器という側面を強調するためか、「致死性自律型ロボット（Lethal Autonomous Robotics, LARs）」<sup>(9)</sup>という文言を使用している。

このような状況の中で、幾つかの国際 NGO が集合して、「殺人口ロボット阻止キャンペーン（Campaign to Stop Killer Robots, CSKR）」<sup>(10)</sup>が2013年4月23日に発足した。当該キャンペーンの目的は、国際条約の形式で「完全自律型兵器の開発、生産及び使用の先制的でかつ包括的な禁止」を要請することである<sup>(11)</sup>。本稿は、近未来兵器である致死性自律型ロボット・完全自律型ロボットが、国際人道法上規制すべきという新たな国際的動向

の中で、規制賛成派と規制反対派からどのような議論がなされ、どのような法的評価を受けているのか（当該ロボット自体及びその使用上の合法性問題）を概観する。さらに、未来兵器の開発・生産・使用に関する制限禁止という国際人道法・軍縮法政策が今後どのように模索されるのか、又はされるべきなのかを検討するものである。

## 註

- (1) ステルス戦闘爆撃機 F-117 やトマホーク巡航ミサイルなどハイテク兵器が登場した。参照、防衛研究所編『東アジア戦略概観 2001』2001年 72-73 頁。
- (2) 拙稿「国際法から見た無人戦闘機 (UCAV) の合法性に関する覚書」『産大法学』45 卷 3・4 号 2012 年 1 月 136 頁。
- (3) Philip Alston, *Interim report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, A/65/321, 23 August 2010 (以下、オールストン報告書又は *Alston Report* と略す), par. 22.
- (4) 「ロボット (Robot)」とは、自動的に動作や作業を行う機械装置や目的の作業・操作を自動的に行う機械や装置（『国語大辞典』小学館 1982 年 2506 頁）をいう。一般的には、人間の代わりに、ある程度自律的に又は自動的に作業を行うために作られた機械（用途別に、産業用ロボット、家庭用ロボット、軍用ロボットなどに分類される）を意味する。
- (5) P. W. シンガーによれば、軍用ロボットの利用はイラク戦争で加速的に拡大し、2003 年のイラク侵攻時には米陸軍は無人機を保有していなかったが、2004 年末には 150 両、その 1 年後には 2,400 両、2010 年では 12,000 両を、米空軍も 2010 年現在 7,000 機以上を保有しているという。P. W. シンガー「ロボットが変える戦争」『日経サイエンス』2010 年 10 月号 41 頁。
- (6) 拙稿（注 (2)）137 頁。
- (7) Human Rights Watch, *Losing Humanity: The Case against Killer Robots*, November 2012 (以下、ヒューマン・ライト・ウォッチ報告書又は *HRW Report* と略す), pp. 1 and 7-9.
- (8) 敵対行為の支援活動（情報収集や物資輸送など）を行うロボット兵器は、単に軍事用だけでなく、災害対策用、医療看護用及び家庭生活用にも活用されることから、一層、その研究開発が推奨されるであろう。
- (9) Christof Heyns, *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, A/HRC/23/47, 9 April 2013 (以下、ヘインズ報告書又は *Heyns Report* と略す)。

- (10) HRW, “Arms : New Campaign to Stop Killer Robots”, April 23, 2013, <http://www.hrw.org/news/2013/04/23/arms-new-campaign-stop-killer-robots>
- (11) CSKR, “Urgent Action Needed to Ban Fully Autonomous Weapons”, April 23, 2013, [http://www.stopkillerrobots.org/wp-content/uploads/2013/04/KRC\\_LaunchStatement\\_23Apr2013.pdf](http://www.stopkillerrobots.org/wp-content/uploads/2013/04/KRC_LaunchStatement_23Apr2013.pdf)

## 2. ロボット兵器の現状

国際人道法上の議論に入る前に、まず、ロボット兵器の現状を把握しておく必要がある。オールストン報告<sup>(12)</sup>によれば、ロボット技術は、監視、偵察、検問所の安全確保、即席爆発装置 (Improvised Explosive Device, IED) の無害化、生物・化学兵器の検知、破壊物の破片除去、捜索・救助、市街パトロールなど幅広い任務のために使用されている。現在使用されているそのほとんどは、人間が遠隔操縦するシステムである。

具体例として、ボムボット (爆弾ロボット、BomBot) は、IED を遠隔操作で爆破処理する。特殊兵器監視偵察探知システム (Special Weapons Observation Reconnaissance Detection System, SWORDS) は、機関銃、ライフル、擲弾発射機やロケットランチャーを搭載し、遠隔操作のビデオカメラで2マイル遠方まで操縦できることから、市街パトロール及び検問所・歩哨所に使用可能である。移動式無人警戒システム (Mobile Detection Assessment and Response System, MDARS) は、重要施設や基地の周辺を自動的にパトロールする歩哨システムとして、車両に搭載される無人小型ロボットパトロール隊というべきものである。当該ロボット兵器は、標的の選定及び発砲が遠隔操作する人間に依存することから、攻撃の意思決定過程の「輪の中に人間がいる (man in the loop)」兵器と言える。現在使用されている半自動化兵器システムの無人戦闘機 (米国のプレデターやリーパー) も、これに含まれる。

イスラエルは、2007年以降、歩哨技術システムとして、ガザ境界線沿いの監視塔に遠隔操作による7.62 mm 機関銃を設置している。同システムが、センサーを介して潜在的標的の位置を確認し、その情報を作戦指揮

所に送信し、作戦指揮所の兵士が標的の攻撃・発砲を決定する。将来の計画として、潜在的標的の確認・標的化・攻撃過程に人的介入を必要としない「閉鎖された輪 (closed loop)」システムが構想されている。同国は、境界監視、戦闘支援その他警備任務 (空港や発電所の周辺警備) 用に、無人偵察車両 Guardium (UGV) も開発中である。

韓国は、2010年7月以降、SGR-1という5.5mm機関銃を搭載した無人監視装置を南北朝鮮の非武装地帯にある警戒哨所に設置している<sup>(13)</sup>。現在、当該歩哨システムの致死力行使の決定は、人間の指揮官が行っているが、ロボットが自律的に発砲する能力を装備しているとの指摘もある。

自律型ロボットへの一段階としての自動兵器防御システム<sup>(14)</sup>が、現在、幾つか実戦使用されている。先ず、米海軍のMK15 Phalanx 接近防空システム (Close in Weapons System, CIWS) は、対空脅威 (対艦ミサイルや航空機) を自動的に探知・追尾・迎撃するものであり、20mmガトリング砲、搜索・追跡レーダー及び火器管制システムが一体化した完全自動型防空システムである。米陸軍は、ミニ・ミサイル防衛として、ファランクス<sup>フアランクス</sup>の地上版である対ロケット・野戦砲・迫撃砲システム (The Counter Rocket, Artillery, and Mortar System, C-RAM) を2005年にイラクで初めて実戦配備した。22基の同システムが、100発以上の砲弾を破壊したという。

イスラエルも、米国と同様に、C-RAMに分類される Iron Dome<sup>アイアン ドーム</sup> をガザ境界線などに実戦配備している。2011年4月の配備以降、80%の成功率でイスラエル市民を殺害するミサイルを迎撃しているという。

無人戦闘機に関して、現在運用中であるイスラエルの Harpy<sup>ハービー</sup> は、レーダーサイトなどの敵防空網の攻撃を主眼に開発されたものであり、レーダー発信所を探知、攻撃、破壊するための fire and forget (打ち放し) 自動兵器システムを搭載している。現在開発中の英国の Taranis<sup>タラニス</sup> は、自律的に敵を搜索し、その位置を確認し、司令部の許可を得て攻撃できる。さらに、敵航空機からの攻撃に反撃する防御能力も装備するという。米国が開発中の次世代無人機は、自律的に離発着し、人工知能、センサー情報及び

画像処理により脅威や標的を探知でき、ミサイルを搭載・発射できる能力を持つようになるという。米海軍の無人艦載機の試作機「X47B」が、2013年5月14日に空母からのカタパルト発艦に、そして、7月10日に空母への着艦に成功した事実は、自律型無人戦闘機への大きな一歩となった。

以上、上記のロボット兵器は、現在、最終的な武力行使の意思決定過程に人間が介在して操作されているが、その中には、自動 (automatic) モードを持つロボット兵器もある。

#### 註

(12) *Alston Report*, pars. 17 and 23-28.

(13) 「知能型警戒ロボット開発……来年末にも休戦ライン配置」『中央日報日本語版』2007年9月28日 <http://japanese.joins.com/article/340/80340.html?sectcode=330&servcode=300>、「韓国軍：DMZ内に監視ロボットを配備」『朝鮮日報』2010年7月14日

<http://www.chosunonline.com/news/20100714000034>

(14) *HRW Report*, pp. 9-13. 及び *Heyns Report*, pars. 44-49.

### 3. ロボット兵器規制の動き

NGO「殺人ロボット阻止キャンペーン<sup>(15)</sup>」によれば、完全自律型ロボットの法規制の動きは、2007年8月にロボット学者 Noel Sharkey が、殺傷行為を自ら判断する完全自律型ロボットの開発に懸念を表明し、緊急の国際規制を要請したことに始まる。2009年9月には、「武装自律型無人システムの開発、配備及び使用の禁止」を要請する NGO「ロボット軍備管理国際委員会 (the International Committee for Robot Arms Control, ICRC)」が設立された。ここ2・3年、当該ロボットの規制に関する国際的動きが、以下のように見られる。

#### (1) 2010年8月のオールストン報告

NGOによる自律型ロボットの規制運動が始まったその中で、2010年8

月 23 日に、Philip Alston を国連人権理事会の特別報告者とする「司法外、略式又は恣意的な処刑に関する特別報告者の暫定報告書」<sup>(16)</sup>が、国連総会に提出された。以下は、その概要である。

報告者が指摘するロボットの利点は、戦闘員の犠牲を発生させないことだけでなく、自己保存の意欲がない分、人間より慎重に致死性武力を行使することが可能であり、人間の感情（興奮、恐怖、疲労、復讐心）から生じる判断ミスも回避できる。また、機械が作戦を記録し、法的要件の遵守を監視する能力を有する範囲において、軍事的な透明性及び説明責任を一層果たすことができることである。他方、報告者は、致死性ロボット技術の急激な展開が、その人道上及び人権上の含意の議論の推移を遥かにしのぐ事態を問題視する。

報告者の懸念事項<sup>(18)</sup>の第 1 は、定義について。中核概念の「自律型 (autonomous)」は、機械が道徳的・倫理的な推理力に従って行動する能力を指すのか、又は人間の統制から独立して行動する能力をさすのか、当該定義の混乱があり、その特徴に関する共通理解の必要性が指摘される。

第 2 に、国際責任及び刑事責任について。遠隔操作によるロボットの場合、致死性武力行使の最終決定は、人間が行うので、結果として生じた危害に関する個別責任や上官責任は、一般的に容易に決定することが可能である。他方、自動化 (automation) の急増により国家責任や個人責任の枠組みが適用困難になってきているという。具体的には、ロボットが国際法違反により文民を殺傷した場合に責任を負う者は、ロボットの行動を規定したプログラム設計者か、プログラムを承認した軍の官僚か、当該ロボットの配備を受けた部隊指揮官か、監視を怠ったロボット担当兵士なのか。ロボットの誤作動による殺傷の場合の責任は、製造者なのか。人間がロボットの行動に対して責任を負うためには、どの程度の監視レベルを実施する必要があるのか（又は絶対責任なのか）。国際責任に関する様々な問いが示されている。また、報告者は、透明性や説明責任の補足的観点から、現在の無人システムが事後調査を支援するよう設計（例えば、記録装置の設置）が施されていない点を問題視する。

第3に、安全装置及び配備基準<sup>(19)</sup>について。国際人道法が要請する予防措置として、どのような技術的な安全装置・プログラミングが組み込まれる必要があるのか。また、米軍が認めているように、ロボット兵器が厳格な又は標準化された検査や実験を受けることなく配備されていた事実を問題視する。そして、ロボットが戦闘員と文民との区別原則、軍事的必要性和人道的考慮の比例原則及び予防原則を判断することができるか否かを判定するための基準は、どのようなものが不明確である、と指摘する。人間による最終的な統制が排除されるのを防止するためにどのような安全装置が取付けられるべきか、テロリストによるハッキングや悪用を防止するためにどのような安全装置が取付けることができるのか、問題提起する。

第4に、文民支援<sup>(20)</sup>について。重要な政治的考慮は、広範なロボット使用によって支援しようとする対象の文民が離反しないのか、政治的に考慮すべき重要な点であると指摘し、過度な技術への依存を戒めている。ロボットが感情に基づいて行動しない利点がある反面、戦闘員や指揮官の行動抑制要因となる同情、憐み、感情移入を持たない欠点が指摘される。

第5に、武力行使の敷居と *jus ad bellum* の考慮<sup>(21)</sup>について。戦争への決断は、軍人の死傷や高額な戦費の見積りによって制約されるが、ロボット軍によって、政策決定者が武力紛争を容易に選択する可能性が高まるといふ。また、ロボット技術の広範な使用の結果として攻撃側軍隊の被害軽減化が、被攻撃側文民の被害増大化に繋がるのか否かが検討されるべきであると指摘する。

最後に、人権の方法論として、従来の「巻返し感覚 (catch-up mentality)」ではなく、「先制的アプローチ (proactive approach)」を採用すべきと主張して、次の勧告<sup>(22)</sup>を提示する。国連人権高等弁務官事務所に作業部会を招集して、戦争に限定しないでロボット技術の開発及び使用の法的、倫理的、道徳的含意を考察する。特に、1) 無人・ロボット兵器システムが有人システムと同等かそれ以上の安全基準を持つこと、2) 実戦配備の前に技術の信頼性を検査するという要件、3) 間違った武力行使を効果的に事後調査できるようにする記録システムの装着の3点を含めて、人権法



及び人道法の要件の遵守を確保するためにどのようなアプローチが採択されるのかを検討すべきであるという。

## (2) 2012年11月のヒューマン・ライト・ウォッチ報告

2012年11月19日に、Human Rights Watch (HRW) が、ハーバード・ロースクール国際人権クリニック (IHRC) と共同で『失われつつある人間性：殺人ロボットに反対する理由<sup>(23)</sup>』を発表した。本報告書は、完全自律型兵器が国際人道法を遵守できるのか、文民殺害における他の抑制要因を保持するのかを分析するものであり、結果として、完全自律型兵器は法的基準を満たすことができないだけでなく、文民に対する非法律的な安全装置をも害すると分析する。<sup>(24)</sup>以下、その概要である。

定義<sup>(25)</sup>として、ロボットとは、基本的には、プログラムされた方法で感知し行動する能力を持った機械であり、人間の監視なく機械が作動する能力を意味する自律性 (autonomy) をある程度保有しているという。その自律性の程度によって、1) 人間の指令でしか標的を選択し武力を行使できない「Human *in the Loop* Weapons (人間が輪の中にある兵器)」、2) ロボットの行動を停止できる人間のオペレーターが監視する下で標的を選択し武力を行使できる「Human *on the Loop* Weapons (人間が輪の上にある兵器)」、3) 人間の入力又は相互作用なく標的を選択し武力を行使することができる「Human *out of the Loop* Weapons (人間が輪の外にある兵器)」に分類される。本報告書では、ロボット兵器は、遠隔操作の無人機から完全自律型兵器までの3分類の無人兵器を包含するという。そして、完全自律型兵器とは、「輪の外」兵器及び「輪の上」兵器 (監視が限定されているので、実際は「輪の外」兵器) を意味するという。

先ず、すべての国家は、慣習国際法としてか又は最善慣行 (best practice) としてかは別として、新兵器及び改良兵器が国際法違反にならないように審査すべきである<sup>(26)</sup>という。国家による兵器の国際法審査は、1977年のジュネーブ諸条約第1追加議定書36条のように、早期の開発段階から、生産決定の最終段階まで継続して行われるべきである。また、兵器化

(weaponize) される潜在可能性があるロボット技術について、国家は、当該技術が兵器化された段階で、直ちに兵器審査過程を開始すべきであるが、国家は開発過程において後になればなるほど兵器の放棄に一層抵抗が生じるので、兵器化の前に当該技術を審査する方がより良いという。法審査にかかわる国際法規則として、文民と戦闘員との区別規則、文民被害と軍事的利益との比例性、軍事的必要性及びマルテンス条項が挙げられる。

国際人道法の遵守メカニズムについて、完全自律型兵器の推進派も当該新兵器が国際人道法を遵守しなければならないことは認識している<sup>(27)</sup>。その中でロボット学者 Ronald Arkin が、遵守メカニズムとして「倫理的統治者 (ethical governor)」を提唱する。それは、第 1 段階で、完全自律型兵器が、収集した情報を評価して、国際人道法違反か否かを決定し、もし国際法違反がないとしても、作戦命令において標的の攻撃が要求されているか否かを判断する。第 2 段階で、技術的データに基づいて、軍事的効果と文民被害との比例性審査を行う。Arkin は、「倫理的統治者」を通じて完全自律型兵器は、人間以上に国際人道法を遵守ことができると主張する。また、英国防省が認識しているように、国際人道法の諸原則を完全に遵守する自律型兵器を確保するためには、何らかの形態の人工知能 (Artificial Intelligence, AI) が必要となるだろうし、人間の知能を有する「強い人工知能 (strong AI)」を装備するロボット兵器の開発を提唱する学者もいる。

計算力が人間の認識力に近づくであろうという彼らの想定に対して、HRW は否定的に評価しており、完全自律型兵器が、広範に採用されれば、法的諸問題を惹起し、文民への脅威となると予測する。

国際人道法の遵守に対する難題として、HRW は、完全自律型ロボット<sup>(28)</sup>が国際人道法の中核的原則 (区別規則、比例規則、軍事的必要性、マルテンス条項) を遵守できないと推測する。区別規則について、国家対国家の戦争から住民がいる所での市街戦という非対称戦に武力紛争の性格が変化している中で、標的と非戦闘員とを区別することがますます困難となっている。文民と武装反徒の区別は機械の認識を超え、テロリストは武器を隠すことでロボットを容易に騙せる。さらに、完全自律型兵器は、個人の意

図を評価するのに必要な人間の資質を有していないことから、戦闘員か否かを判断するための基準となる個人の意図（感情）が判断できないと指摘する<sup>(29)</sup>。

国際人道法の最も複雑な規則の一つである比例性は、予期する文民被害が予期する軍事的利益を上回る場合に攻撃が禁止される。その決定は、極めて状況に依存する。ロボットは直面するかもしれない無数のシナリオを処理するよう事前にプログラムすることができそうにないという。赤十字国際委員会（ICRC）による注釈書<sup>(30)</sup>に規定されるように、比例性の審査は、主観的であって、「軍指揮官にとって常識及び誠実の問題」である。また、国際裁判や軍隊では、「合理的な軍指揮官」基準が採用されているという。このように、ロボットは、比例性の評価に必要な人間の判断における心理的過程を再生するようプログラムすることができない。それらに対応可能な「強い人工知能」の開発がなされるとの意見もあるが、それは論争されており、ロボット技術の開発が人工知能の開発を凌駕する脅威もあるという。

軍事的必要性も、比例性と同様に、主観的分析が必要とされる。完全自律型ロボットの歩哨は、一度撃った侵入者が爆風で地面に叩き付けられただけなのか、軽傷なのか、重症なのか判断できず、不必要に2度撃ちするかもしれないという。

マルテンス条項は、人道の法則及び公共良心の要求に従って戦争手段を評価するが、公共良心の要求を決定する場合、多数派の見解が考慮される。HRWは、Ronald Arkinの調査結果を基に、自律型兵器は、人間に統制されなくなればなるほど、多くの人々にとって受諾不可能となり、彼らは、兵士の人命を救う能力を有する致死性自律型ロボットによる文民殺害など付随的損害を懸念しているという。従って、当該兵器は、マルテンス条項に規定された原則から、多くの人々に許容されないことを認識すべきであるという。

国際法の遵守問題を要約すれば、完全自律型兵器は、相手側の意図を理解する能力を欠き、複雑で刻々と変容する状況を処理することが困難であり、そして、主観的評価を必要とする人間の判断が適用できない。さらに、

以前人間の手にあった生死の決定を機械が下すことは人々の良心に反することから、自律技術の開発は、人間を意思決定過程の輪の外に置く時点に到達する前に、停止させるべきである、と結論付ける。

HRW は、文民保護に関する国際法的視点からだけでなく、非法律的視点<sup>(31)</sup>からも、完全自律型兵器の問題点を分析する。ロボット推奨者が人間の兵士よりもロボットが優れていると主張する特徴は、1) 感情の欠如と、2) 軍事的被害の軽減能力である。1) に関して、ロボットは、恐れや怒りの感情がなく、また自己保存への執着がない分、任務のために自己犠牲をいとわぬ。他方、人間の感情（憐みや感情移入）が殺害意思の抑止要因となるが、ロボットはその要因がない<sup>(32)</sup>。また、感情のないロボットは、人民を弾圧する抑圧的独裁者の道具になり得る。感情は、理性に対する障害物と解するのではなく、戦時での制約要因とみなすべきであるという。2) に関して、ロボットによる攻撃側戦闘員の被害軽減により、戦争抑止の心理的障害が除去されることで戦争への決断の敷居が低下する。さらに、「戦闘の非人格化」<sup>(33)</sup>により戦闘行為が容易となる。他方、それらが相手側文民の被害が増加することにつながるという。

最後に、HRW は、完全自律型ロボットに関する説明責任問題<sup>(34)</sup>を取り上げる。当該兵器が国際法に違反して文民被害を惹起した場合に、誰がその国際責任を負うのか。軍指揮官か、プログラマーか、製造者か、ロボット自身か。ある状況では、「上官責任」原則から、指揮官が部下の戦争犯罪について責任を負う。しかし、指揮官は、ロボットをプログラムしていないし、犯罪の実行後にロボットを処罰もできず、完全自律型兵器の行動について説明責任を負えないという。プログラマーは、完全自律型ロボットが複雑な戦場の場面で下す決定を完璧に予測できないことから、過失の場合を除き、プログラマーに責任を負わすのは、不公平であり効果的でないという。製造物責任制度から製造者に責任を負わせるという意見について、民間製造者は兵器使用の方法について一般的に処罰されず、製造物責任制度は犠牲者に負担を強いる民事訴訟を必要とするので、製造物責任の追及は非現実的であるという。上記の者達に責任追及することは、公平で効果

的とは思えず、ロボットによる国際人道法の侵害を抑止できず、犠牲者に意味ある応報の正義を提供できないという。

結論<sup>(35)</sup>において、HRWは、オールストーン報告による問題提起型の勧告をさらに進展させ、明確な方向性を含む具体的な勧告を提出した。先ず、すべての国家に対して、完全自律型兵器の開発、生産及び使用を拘束力ある国際法文書によって禁止すること。当該禁止の国内法及び国内政策を採択すること。完全自律型兵器に至る技術及び構成部分の審査を開始すること。次に、ロボット兵器の開発に携わるロボット研究者その他の者に対して、武力紛争時の完全自律型ロボット兵器の使用に関する法的及び倫理的な懸念事項が、技術開発の全段階において適切に検討されることを確保するために、当該兵器、特に完全自律型となる可能性のある兵器の研究開発に関する専門的な行動規範を確立すること、である。

### (3) 2012年11月の米国防省指令 3000.09<sup>(36)</sup>

前述のヒューマン・ライト・ウォッチ報告が発表された2日後の2012年11月21日に、米国防省は、指令3000.09「兵器システムにおける自律性」を公布した。当該指令によれば、その目的は、その目的は、<sup>(37)</sup>有人・無人を含む「兵器システムにおける自律的及び半(semi-)自律的機能の開発並びに使用に関する国防省の政策を確立し、その責任を割り当てる」ことであり、「意図せざる交戦に至る自律的及び半自律的システムでの故障の蓋然性及び結果を最小限にするための指針を確立する」ことである。その適用範囲は、自律型及び半自律型兵器システムの設計、開発、取得、実験、配備並びに展開であり、そして、自律型及び半自律型兵器システムによる致死性の・非致死性の、運動学的・非運動学的な武力に及ぶとする。

米国防省の定義<sup>(38)</sup>によれば、自律型兵器システムとは、「一旦起動すれば、人間のオペレーターによる更なる介入がなくとも標的を選択し攻撃することができる」ものを指す。これには、「人間のオペレーターが兵器システムの運用を停止させることができるよう設計された人間監視・自律型兵器」が含まれる。また、半自律型兵器システムとは、「一旦起動すれば、

人間のオペレーターが選択した個別の標的又は特定の標的群しか攻撃しないよう意図された」ものを指す。

米国防省の政策は、「自律型及び半自律型兵器システムは、指揮官及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の判断を行使できるように設計される」(4. a) とする。

そのために、ハードウェアやソフトウェアの検証及び確認 (V&V)、実践的な開発及び運用の試験評価 (T&E) を実施することで、当該システムが指揮官やオペレーターの意図と一致した時間枠内で交戦を完了する。もしできなければ、交戦を停止するか、交戦の継続前に人間のオペレーターによる追加的入力三要請する (4a (1) (b))。また、意図せざる交戦や制御不能に備えて、安全装置や誤作動防止装置の装備や人間と機械の連接 (interface) 及び制御を組み込んだ設計が行われる (4a (2))。オペレーターが標的を攻撃する際に情報に基づき適切な決定をするために、当該システム用の人と機械の接続が「訓練済みオペレーターにとって容易に理解可能なものであり、システムの現状に追跡可能なフィードバックを提供し、訓練済みオペレーターがシステム機能を活性化・不活性化させる明確な手続きを提供する」(4a (3)) ものでなければならない。

半自律型兵器システム (4. c. (1)) は、「有人又は無人プラットフォームを含めて、致死性・非致死性の、運動学・非運動学の武力を行使するために使用できる。」また、「無人機に搭載された半自律型兵器システムは、通信の低下又は喪失の場合に、権限ある人間のオペレーターによる事前に選択されていない個別の標的又は特定の標的群を自律的に選択し攻撃しないように設計されなければならない。」

人間監視・自律型兵器システム (4. c (2)) は、「標的として人間を選択することを除き、有人施設の陣地防御や有人プラットフォームの搭載防御のために集中攻撃を阻止する局地防御用として標的を選択し攻撃するために使用できる。」

そして、自律型兵器システムは、「例えば、電子攻撃のような非致死性の、非運動学の武力を物理的標的に作用させるために使用することができ

る (4. c (3))。]

本指令は、2012年11月21日に発効し、再公布や破棄を含めて5年以内に見直しがなされる。もし見直しがなければ、2022年11月21日(10年間)<sup>(40)</sup>まで有効となる。

米国の政策をまとめると、致死性武力を行使する意思決定に、「適切なレベルでの人間の判断 (appropriate levels of human judgment)」が必要であり、それが介在しない人間監視・自律型兵器を含む自律型兵器及び半自律型兵器システムを設計、開発、取得、実験、配備並びに展開することは禁止される。もっとも、非致死性・非運動学的武力を対物使用する自律型兵器システムは、人間の判断が介在しなくとも合法となる。また、人間監視・自律型兵器システムは、人間を攻撃対象としなければ、陣地防御や局地防御のために使用できる。具体的には、対空脅威 (対艦ミサイルや航空機攻撃) に対抗する局地防御用の MK15 Phalanx<sup>フアラックス</sup>、陣地防御用の対ロケット・野戦砲・迫撃砲システム C-RAM 並びにイスラエルのミサイル防御用の Iron Dome<sup>アイアン ドーム</sup>、海上の Aegis システム<sup>イージス</sup> 及び陸上の Patriot システム<sup>ペイトリオット</sup> が当該定義に該当する。同指令は、当然のことながら、それらが合法兵器であることを前提としている。また、自律型兵器システムは、意図せざる交戦や制御不能に備えての予防措置 (安全装置や誤作動防止装置等) が義務付けられている。

#### (4) 2013年4月のヘインズ報告

2013年4月9日に「特別報告者クリストフ・ヘインズによる司法外、略式又は恣意的執行に関する報告書」<sup>(41)</sup>が、年次報告書として国連人権理事會に提出された。当該報告書は、正面から「致死性自律型ロボット (LARs)」を議論している。以下、その概要である。

火薬、核爆弾に次ぐ軍事革命であると位置付けられるロボットは、国際人道法及び国際人権法を遵守できず、倫理的に人の生死を決定する権限を持つべきではないとのロボット反対論 (開発、製造及び使用の一切禁止) と、技術的進歩が紛争の人道化や全当事者側の人命救助に役立つとのロ



ロボット推進論が対立している。その中で広範に受容されているのは、国家によるロボット技術の利用について何らかの形態の制御が必要であるという意見である<sup>(42)</sup>。

そして、新しい戦闘手段の出現の後になって深刻な批判が生じた他の軍事革命（火薬及び核爆弾）と対照的に、先を見越してLARsによる危険性を停止させ対処する機会は今であり、本報告は、オールストン報告の勧告を強化して、LARsに関する若干の活動を一時的に国内停止（国内モラトリアム）するよう諸国家に要求する<sup>(43)</sup>。

定義について<sup>(44)</sup>、ヘインズも、米国防省やHRWと同様に、LARsを「一旦起動すれば、人間のオペレーターによる更なる介入がなくとも標的を選択し攻撃することができる」ロボットと理解する。監視された自律性は、「人間が輪の上」にいて、ロボットの決定をモニターし、それを無効にすることができる。しかし、ロボットの決定過程はナノ秒で測定されるので、人間はその決定過程に介入できず、事実上「人間が輪の外」にいることになるという。自動的（automatic）と自律的（autonomous）を区別して、前者は「構築された予想可能な環境の中で作動する」ことを示し、後者は「オープンな環境の中や構築されずダイナミックな状況の下で機能することができる」ことを指す。ヘインズは、人間のような自律的行動は、特に武力紛争のような混乱状態では、最終的には予想不可能とみなしている。それゆえ、ロボットに「自律型」という文言を使用することは、誤解を生じさせる可能性があり、現在、知覚力のあるロボットや「強い人工知能」の実現可能性はないという。

従来から指摘されているように、ヘインズも、LARsの促進要因として<sup>(45)</sup>、自軍兵士の人命の保護、兵力投入の増加、3D（dirty, dull, dangerous）任務の遂行及び人間の欠点（復讐、狼狽、怒り、遺恨、偏見、恐怖）からの解放を列挙し、阻害要因として、人権法や人道法で必要な人的判断（常識、相手側の行動意図の理解、価値の理解、事態の展開予測）及び戦闘行為での抑止効果を持つ憐みや直視力の欠如を指摘する。そして、LARsの使用による武力行使への影響について、人的被害の減少による「武力紛争の常態



化 (normalization)」や殺傷行為に対する敷居の低下により、武力衝突ごとの被害が減少したとしても、総計の被害者数はより高くなり得るという。

国際人道法の区別原則及び比例原則<sup>(46)</sup>の遵守に関して、兵士が負傷し戦闘外に置かれたか、又は降伏過程にあるかをロボットが確定することは困難である。非対称戦争、非国際武力紛争、市街戦又は文民の「敵対行為への直接参加」の場合、なおさらである。戦闘員か文民かが不明である事態に直面した兵士は、生存本能から即発砲する。他方、ロボットであれば、接近し発砲されるまで反撃しない戦術が取れることから、人間の兵士のために「戦雲 (fog of war)」を除去し誤射を防止することができる長所も指摘する。比例原則の法解釈は、常識、誠実及び合理的な軍指揮官の基準に依拠するが、ロボットは、当該概念がプログラムできる範囲でしか比例原則を理解できない。

国際法の累積的基準は、例えば、ロボットを反撃しかできないよう設定すれば、満たし得るが、そうでない場合には、満たさない場合も生じる。それゆえ、比例原則のような分析は少なくとも当初指揮官に委ね、それ以外の側面を LARs に任せるといった意見があることも指摘する。

国家や個人の国際責任制度<sup>(47)</sup>は、違反行為の抑止や防止のために重要である。誰がロボットの国際法違反の責任を負うのか不明で責任追及が不可能ならば、その使用は非倫理的で不法とみなされるべきである。法的責任の新たな確立方法として、責任の事前割り当て、記録装置の装着とそれに基づく強制的事後審査又は責任候補者間の責任分担システムを LARs の使用条件にするよう提案する。

LARs 非保有国との関係<sup>(48)</sup>について、LARs は、技術保有側の人命尊重を、他方、技術非保有側の人命軽視を意味し、究極の非対称的状況 (殺傷ロボット vs 通常の人間) を提示することになる。LARs 保有の優越性も永久ではなく、技術拡散や兵器拡散の可能性や非国家行為者による取得の潜在的危険性もあると指摘する。

また、意思決定過程から人間を外すことは、人道性を外すことになる<sup>(49)</sup>。殺傷の意思決定権は人間にあり、LARs は、スポーツの線審のような機械

的監視の場合しか使用できず、価値判断を含む事態では使用できない。地雷が禁止されたように、機械が人間の生死を決定する権限を持つべきではない。

その他の懸念事項として、ハッキングに対する脆弱性や大災害となりうる誤作動の危険性が指摘される<sup>(50)</sup>。

では、LARsの規制体制をどうするかという問題について、兵器関連条約を参照して、兵器自体の規制（禁止又は禁止に至らない制限）や活動の規制（取得、保有、貯蔵、研究、開発、実験、配備、移譲、使用など）の組合せが考えられる。条約形式だけでなく、ソフト・ロー形式（行動準則、政府間対話、情報共有及び信頼醸成措置など）も考えられる。さらに、ジュネーブ諸条約第1追加議定書第36条のように、兵器の国内審査手続きが規定されている。2012年の米国防省指令は、一種のモラトリアムを課していると評価される。

結論<sup>(52)</sup>として、現在、LARsが国際人道法などの要件をどの程度満たすか不明である。LARsに殺傷を認めると、人命軽視となる可能性がある。特定の使用が特別の事態で許容されるとの立証責任は、LARs使用側にある。後になって撤回することが困難な段階に進むのを防止するために、モラトリアムが必要であるという。

ヘインズは、勧告として、特に以下のことを要請する。国連人権理事会に対して、すべての国家に、LARsに関する国際法的枠組みが将来確立するまで、LARsの実験、生産、組み立て、移譲、取得、展開及び使用に関するモラトリアムの宣言及び履行を要請すること。人権高等弁務官に対して、法学、ロボット学、軍事作戦、外交、倫理学、哲学等の多分野の専門家から構成されるLARsハイレベル・パネルを招請すること。国家に対して、LARsに関する国内モラトリアムを実施すること。

#### (5) 2013年5月の人権理事会及び6月の英国議会

運営委員会には9の国際NGOが結集して、国際的な議論を活性化するために「殺人ロボット阻止キャンペーン」が2013年4月23日に発足した。

上述した一連の国際的な動きの中で、2013年5月30日に、ヘインズ報告の後に完全自律型兵器についてどうすべきか、国連人権理事会において初めて議論が行なわれた。<sup>(54)</sup> 22か国が発言し、どの国も更なる議論に反対しなかった。唯一英国だけが、完全自律型兵器のモラトリアム又は禁止の要請に反対した。ブラジル及びフランスは、更なる議論の場として特定通常兵器条約（CCW）を示唆した。パキスタンは禁止を支持したが、他の諸国はモラトリアムの要請に賛成した。

英国は、LARs に対する政策として、従来、「指揮の意思決定過程で人間の介入なく標的を迫及する能力を持つ攻撃的システムの計画はない<sup>(55)</sup>」し、「兵器システムの運用は、常に人間の統制下にあることは絶対に明白である<sup>(56)</sup>」と主張する。人権理事会での議論を受けて、アリストター・バート外交連邦政務次官<sup>(57)</sup>は英国議会において次のように発言している。人権理事会はLARs に関する討議のための適切なフォーラムではない。英国は国際的議論に参加する用意があり、この重要事項についての議論の場として、特定通常兵器条約が適切であろう。英国だけがモラトリアムの要請に反対した理由を問われて、英国は法の自由が付与する以上に制限的な政策を採択し、国際人道法を遵守しているので、国内モラトリアムとして形式化する意図はない。そして、幾つかの国は何らかの規制や統制を支持して発言したが、多くの国はそうではなく、モラトリアムの普遍的支持があると理解すべきではないと指摘する。

## 註

(15) <http://www.stopkillerrobots.org/chronology/>

(16) 注 (3)、*Alston Report*.

(17) *Ibid.*, pars. 29-31.

(18) *Ibid.*, pars. 32-36.

(19) *Ibid.*, pars. 37-42.

(20) *Ibid.*, par. 43.

(21) *Ibid.*, par. 44.

(22) *Ibid.*, pars. 45-49.

- (23) 注 (7)、*HRW Report*.
- (24) *Ibid.*, pp. 1-2.
- (25) *Ibid.*, pp. 2-3.
- (26) *Ibid.*, pp. 21-26.
- (27) *Ibid.*, pp. 27-29.
- (28) *Ibid.*, pp. 30-36.
- (29) 具体例として、母親が二人の子供の後を追いかけて、兵士の近くで玩具の鉄砲で遊ぶのをやめなさいと叫ぶ状況では、人間の兵士は彼らの意図が無害であると判断し発砲しないが<sup>8</sup>、完全自律型ロボットは、近づく人間と2人の武装した個人としか見えず、発砲するかもしれないという。*Ibid.*, pp. 31-32.
- (30) International Committee of the Red Cross, *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949*, 1987, par. 2208.
- (31) *HRW Report*, pp. 37-41.
- (32) 例えば、戦闘地帯内のロボットは、それに鉄砲を向ける子供に発砲するかもしれないが<sup>9</sup>、人間の兵士であれば、自分の子供を思い出し、発砲を停止し、より慈悲深い解決策（捕獲）を模索するかもしれないという。*Ibid.*, p. 38.
- (33) P. W. シンガー『ロボット兵士の戦争』2010年572頁、P. W. Singer, *Wired for War: The Robotics Revolution and Conflict in the 21<sup>st</sup> Century*, 2009, p. 396.
- (34) *HRW Report*, pp. 42-45.
- (35) *Ibid.*, pp. 46-48.
- (36) US Department of Defense, Directive 3000.09, *Autonomy in Weapon Systems*, November 21, 2012.
- (37) *Ibid.*, p. 1.
- (38) *Ibid.*, pp. 13-14.
- (39) *Ibid.*, pp. 2-3.
- (40) *Ibid.*, p. 4.
- (41) 注 (9)、*Heyns Report*.
- (42) *Ibid.*, pars. 28-32.
- (43) *Ibid.*, pars. 33 and 35.
- (44) *Ibid.*, pars. 38-43.
- (45) *Ibid.*, pars. 51-61.
- (46) *Ibid.*, pars. 66-74.
- (47) *Ibid.*, pars. 75-80.
- (48) *Ibid.*, pars. 86-88.
- (49) *Ibid.*, pars. 89-93.

- (50) *Ibid.*, par. 98.
- (51) *Ibid.*, pars. 100-108.
- (52) *Ibid.*, pars. 109-112.
- (53) *Ibid.*, pars. 113-126.
- (54) <http://www.stopkillerrobots.org/2013/05/nations-to-debate-killer-robots-at-un/>
- (55) 2013年3月7日のアスター国防省政務次官の答弁書、*Official Report, House of Lords*, 7 Mar 2013, Vol. 743, Column WA 411.
- (56) 2013年3月27日の英国上院でのアスター政務次官の答弁、*Official Report, House of Lords*, 26 Mar 2013, Vol. 744, Column 960.
- (57) 2013年6月17日の英国下院でのパート外交連邦政務次官の発言、*Official Report, House of Commons*, 17 Jun 2013, Column 736.

#### 4. 国際人道法上の論点

##### (1) 兵器に関する国際人道法の枠組み

LARsに関連する国際人道法は、まず兵器自体が合法か否か、そして、兵器の使用が合法か否かの2つの次元で議論されるので、まず、その全体図について概説する。<sup>(58)</sup>

兵器自体について、無差別的兵器禁止原則（ジュネーヴ諸条約第1追加議定書51条4項b参照）が適用される。LARsは、十分に信頼できる詳細なデータが提供され軍事目標を指向することができるならば、本質的に無差別的兵器とならないことになる。そして、不必要な苦痛を与える兵器禁止原則（同議定書35条2項参照）が適用され、当該原則を満たす弾薬を使用しなければならない。国家は、これら2つの法原則を基に、LARsの法的審査を実施する義務を負う（同議定書36条）。

兵器自体が合法であるとしても、兵器の使用方法が合法であるかを議論する必要がある。その際の中核的要件が、1) 区別原則、2) 比例原則及び3) 予防義務である。第1の要件は、戦闘員及び軍事目標だけを攻撃し、文民及び民用物を攻撃してはならないという区別原則（同議定書48・51・52条）である。正規軍同士の高強度紛争や遠隔地域（水中、砂漠又は非

武装地帯)での戦闘では、低レベルでの区別原則を満たす事態はあり得るが、複雑な対反乱型戦闘や市街戦では、区別原則の履行が困難となる。第2の要件は、攻撃から予期される付随的損害が予期される軍事的利益との比較で過度であるか否かを判断する比例原則(同議定書51条5項b及び57条2項iii)である。標的の軍事的利益は、極端に状況に依存し、戦場の展開によっては急激に変化し得るので、攻撃前の評価は、文民被害の予測と同様に、容易ではない。第3の要件は、攻撃時における実行可能(feasible)な予防措置を取る義務(同議定書57条2項i及びii)である。これは、標的が軍事目標であることを確認するための実行可能な措置をとること、同程度の軍事的利益を得るために幾つかの軍事目標から選択が可能な場合に文民被害の最も少ないと予期される標的を攻撃すること、及び付随的損害を回避し少なくとも最小限にするための攻撃手段の選択における実行可能な措置をとることである。これらの実行可能な措置が取られたか否かの価値判断は、極めて難解な課題となる。

上記の実体規定とは別に、手続き規定としての前述の兵器の法審査(同議定書36条)及び国際責任(例えば、上官責任、同議定書87条)の議論がある。

## (2) 兵器自体の合法性に関する議論

Schmitt 及び Thurnher<sup>(59)</sup> (以下、シュミットらとする)は、合法的標的を指向できないので無差別的である兵器システムの禁止と、区別可能な兵器を無差別的方法で使用することの禁止とが混同されていると注意喚起する。例えば、スカッド・ミサイルは、不正確ではあるが、区別的に使用できる事態(砂漠のような開放地域での軍隊への使用)が存在し、それ自体不法ではないという。

同様に、自律型兵器の批判に対する反論として、区別原則が全く適用できない自律型兵器でさえ、ある環境では、合法的に使用可能であることが見落とされていると指摘する。<sup>(60)</sup>

さらに、LARs 批判者の主張は事実と反するという。軍事技術が単に個

人か物かを見分ける以上に進化している。特に、最近のセンサーは、物の形状や大きさを評価し、その速度を判定し、使用されている推進力のタイプを識別し、その材質を決定し、対象物やその周辺の音を聴取し、通信を傍受することができる。彼らは、自律的特徴だけで自律型兵器システム自体が不法であるとの特徴付けに反対する。

LARs が兵器の隠蔽や動作の工夫で騙され、文民と戦闘員の区別が困難であるがゆえに、LARs の禁止が必要であるとの主張に対して、非対称的で不利な敵は、数世紀にわたり、通常兵器の攻撃を回避するために文民その他の保護される地位の振りをしてきた。そのような欺瞞テクニックが成功した事実は、当該兵器を無差別兵器と区分し違法化することに値しない。<sup>(61)</sup>もし LARs を無差別兵器として違法化すれば、敵の戦術（兵器を隠蔽し文民の振りをする）の利用を奨励することになり、非生産的であるという。

LARs が人間の意図を認識できないという批判に対して、人間が操作する兵器システムは、感情的な感受性によることなく、すでに標的を攻撃している。<sup>(62)</sup>実際、人間の判断が戦場の興奮により技術的な指示器ほど信頼できない事例として、1994 年の米軍の友軍撃墜事件及び 1988 年の米軍ミサイル巡洋艦ヴィンセンスによるイラン航空機撃墜事件が指摘される。<sup>(63)</sup>これらの事件は、輪の中の人間が軍・民の区別困難な事態において万能薬ではないことを示しているという。

人間行動の抑制要因となっている感情がないロボットは、抑圧的独裁者の道具として役立つという批判に対して、感情は、人間を抑制するとともに、ルワンダ、バルカン、ダルフル及びアフガニスタンでの悲劇的事例のように、本能の最も根源的な部分を開放する。<sup>(64)</sup>それ故、LARs の感情の欠如は、徹底的な禁止の正当化事由にならないという。

LARs は、文民と戦闘員とを明確に区別できないが、文民のいない場所での使用が計画されている場合、それ自体不法ではない。<sup>(65)</sup>その場合に LARs を使用するために、戦闘員が文民のいる領域に進駐するのを阻止するために地理的制約が可能でなければならないし、文民のいない地域がほとんどないので、時間的制約が可能でなければならないという。

無差別的兵器の第2の形態として、標的を精確に攻撃できるにも拘らず、統制不可能な効果を持つ兵器システムは禁止される。それ故、自律型プラットフォームに、例えば、生物兵器を合法的に搭載できない。

### (3) 兵器使用の合法性に関する議論

シュミットらは、LARs 自体が不法である見込みは極めて低く、問題は、その使用であるとみなして、以下の論点を取り上げ、国際人道法の視点から分析している。

#### (a) 区別性<sup>(66)</sup>

区別可能な兵器の無差別的使用の禁止には、文民が存在するかもしれない事態において区別能力を向上させることができる LARs と連携したセンサーの使用が要求される。

#### (b) 比例性<sup>(67)</sup>

比例性の中心概念は「過度性 (excessiveness)」であるので、攻撃から生じる合理的に予期される軍事的利益が大きければ大きいほど、武力紛争法は、予期される付随的損害もより大きく許容される。このことは、LARs が比例性の計算ができるか否かの問題を棚上げにする。すでに、標的付近の物や人への付随的損害の可能性を決定するための効果的なシステム、「付随的被害見積法 (Collateral Damage Estimate Methodology, CDEM)<sup>(68)</sup>」がある。標的領域内の文民被害の可能性を決定するために CDEM のような分析をするようプログラムできる。文民被害の予測だけをするこの分析は、予期される軍事的利益に照らして付随的損害を考慮していないので、これだけでは、特定攻撃が比例規則を遵守しているか判断できない。

問題は、LARs が状況 (文脈) によって変化する軍事的利益を測定することができるかである。比例規則における軍事的利益の要素は、一般的に事例ごとの決定が必要である。比例性の決定は、必然的にかなり広範な判断領域を含み、付随的損害に対する標的の軍事的価値の数量化は、あまり精密である必要はないという。



比例性の遵守には、基盤となる最大限の付随的損害の敷居が極めて低い<sup>(69)</sup>か、交戦状況によりその調整が可能であるか、<sup>(70)</sup>どちらかが必要となる。前者は、最大限の付随的損害の基準を超えるような状況にLARsを投入しなければ、比例性の問題は生じない。後者も、状況に応じて最大限の付随的損害基準が調整できるので、比例性の問題は生じないことになるという。

ロボットは比例性の評価に必要な人間の判断における心理的過程を再生するようにプログラムできないとの批判に対して、人間以上のことを機械に求めることは不適切である。人間も機械も比例性に関する完全な基準を遵守できず、武力紛争法の基準は常に合理的なものであると述べつつ、近未来でも、現実の比例性の決定は人間によって行われ続けると付け加える。<sup>(71)</sup>

#### (c) 攻撃時における実行可能な予防措置

標的が軍事目標であることを検証するための実行可能なすべてのことをするという要件は、LARsの場合、標的識別の信頼性を増大することができる搭載又は外部のセンサーを最大限使用することを要求する<sup>(72)</sup>。そこでの「実行可能な」とは、「人道上及び軍事上の考慮を含むその時点におけるすべての事情を勘案して実施し得る又は実際に可能と認められる」<sup>(73)</sup>ことである。軍事的考慮には、技術上及び作戦上双方の要素、特に兵器・関連システム又は運用関連担当者の残存性が含まれる<sup>(74)</sup>という。

文民被害を最小限にするために、軍事目標の中で選択するという要件は、LARsにも適用される<sup>(75)</sup>。そして、軍事的利益を犠牲にしないで文民被害をあまり引き起こさない戦闘手段を選択する必要がある。LARsが合法的に使用できる唯一の事態は、付随的損害を軽減させる利用可能なシステムがあっても、それでは軍事目標が達成できないが、LARsではその軍事目標が実現できる場合であるという。

#### (d) 比例性及び予防措置に共通の論点 —— 疑惑と主観性

文民か否かについて疑惑がある場合には文民とみなすという文民推定規則（同議定書 50 条 1 項）は、比例性や予防措置を考慮する場合に適用され、LARsにも適用される<sup>(76)</sup>。LARsの搭載するセンサーが、身長によって児童であると判断する場合、潜在的標的が戦闘員である蓋然性が低くなる。

他方、当該標的が武器を携行し敵対行為に従事していることが確認されれば、それが戦闘員である見込みが高くなる。いずれにせよ、LARsは、比例性規則を侵害する危険性があるので、攻撃を差し控える疑惑の敷居を設定する必要があるという。疑惑が特別なレベルに達すれば、それは攻撃しない。もっとも、落ち着いた環境よりも激しい戦場でのほうが多くの疑惑が黙認される。将来の進化したLARsは、おそらく遠隔操作で疑惑の敷居を調整することができるだろうと予測する。

また、LARsは、自己保存概念がないので、発砲されるまで発砲しないよう自己犠牲を強いるプログラム設定が可能である。この戦術は、標的を攻撃する前に、潜在的標的の疑惑の程度を低減化できるといふ。

軍事的利益と巻き添えによる文民被害との比例性及び実行可能な予防措置を考慮する場合に、主観的判断が必要であるが、LARsはその主観的評価が不可能であることから国際人道法を遵守できないという意見には根拠がないと、シュミットらは反論する<sup>(77)</sup>。

LARsの攻撃を含めて、標的化過程での主観的判断は人間によって引き続き行われており、違いが生じるのは、標的化過程における段階においてである。LARsの標的化過程に人間の主観性が介在するのは、その開発時、投入時又は遠隔操作によりプログラムする時である。主観性との決定的な連結点は、指揮官が戦場にLARsを投入する命令を発する時である。最終的には、LARsを使用する指揮官の決定が合理的であったか否かが、法的要件によって評価される<sup>(78)</sup>。

以上のことから、LARsではなく人間が、国際人道法に内在する主観的決定を引き続き行う。それ故、LARsを「人間が輪の外に置く」と称するのは誤解を生じさせ、LARsが主観的に決定できない事実からLARsを不法とするのは間違いであるという。

#### (e) 手続き規定上の論点 —— 兵器の法審査と国際責任

兵器の法的審査（ジュネーヴ諸条約第1追加議定書36条）では、兵器それ自体が合法か否かは、その可能な使用ではなく、意図された使用での状況において決定される。そのため、兵器の使用に関して適用される比例

性や予防措置は、兵器審査の要因とはならない。<sup>(79)</sup>

法審査でマルテンス条項の適用を主張する意見について、マルテンス条項は法の欠缺に対応する多重安全装置 (failsafe mechanism) であり、すべての場合に考慮すべき包括的原則ではない。<sup>(80)</sup> 兵器関連の条約法が充実する中で、自律型を含む将来の兵器システムが適用可能な条約に違反しないが、マルテンス条項に基づき不法であるという可能性は、極めて低くなっていると指摘する。

法的責任に関して、LARs による不法行為の法的責任を課する公平で効果的な方法が存在しないので、LARs による国際人道法の侵害を抑止できないという批判について、その結論の前提が間違っていると反論する。<sup>(81)</sup> 人間が特定の交戦行為の統制の中にいない単なる事実は、いかなる人間も当該兵器システムの行動に責任を負わないことを意味しない。人間は、当該システムにプログラムする方法やその投入時期を決定する。個人は、戦争犯罪に達する行動に従事するようプログラムすれば、当然、責任を負う。指揮官や文民の上官は、当該システムがそのようにプログラムされていることを認識していたにもかかわらず、その使用を停止するために何もしなかった場合、又は当該システムが戦争犯罪を構成する方法で使用されたことを認識していたが、関連の個人に説明責任を課すために何もしなかった場合、戦争犯罪の責任を負うという。<sup>(82)</sup> 米国は、「自律型及び半自律型兵器システムの使用を許可し、命令し又は運用する者は、適切な注意を持って戦争法、適用可能な条約、兵器システム安全規則及び適用可能な交戦規則 (ROE) に従ってそのように行動しなければならない」(米国防省指令 3000.09, 4(b))と規定し、作戦関係者に責任を課している。米国の戦闘司令部司令官も、同一の義務を課している (同指令 Enclosure 4, 10. Commanders of the Combatant Commands)。

最後に、シュミットらは、次の四つの結論を示している。<sup>(83)</sup> 1) 自律型兵器システムは、それ自体として不法ではない。2) ある状況での自律型兵器システムの使用は、武力紛争法上合法である。3) 人間は、実際、「輪の外」にはいない。4) 人間が、自律型兵器システムの使用について常に責

任を負う。もし上記の結論が正しければ、自律型兵器システムの禁止要求は、多くの国家の支持を得ないという。当該兵器の意図しない人的影響の潜在的可能性及び戦闘能力が十分に理解されるまで、国家は、深刻にそれを禁止しよう思わないだろう。<sup>(84)</sup>

自律型兵器システムを禁止することは、若干の場合、文民被害を軽減する価値あるツールを指揮官に否定する効果を持つかもしれない。それ故、開発段階にある当該システムを禁止することは、無責任であり、当該システムの使用について法上、道義上及び作戦上の最終的結論を出すことは時期尚早であると、彼らは主張する。

#### 註

(58) Cf. Jeffrey S. Thurnher, “The Law That Applies to Autonomous Weapon Systems”, *insights*, America Society of International Law, vol. 17, Issue 4, January 18, 2013.

Thurnher は、weapons law (兵器法) と targeting law (標的化法) に区分して、国際人道法との関連性を概説している。

(59) Michael N. Schmitt & Jeffrey S. Thurnher, “Out of the Loop’: Autonomous Weapon Systems and the Law of Armed Conflict”, *Harvard National Security Journal*, vol. 4, 2013, pp. 244-250. 以下の本稿 4 (2) (3) において、LARs 規制派に対する反対論としてシュミットらの論文を紹介する。

(60) 代表例が、砂漠の遠隔地域での戦車部隊への攻撃用及び航路帯から遠く離れた公海上での軍艦への攻撃用の当該システムの使用である。 *Ibid.*, p. 246.

(61) 攻撃側は、文民の振りをする戦闘員を攻撃しないことによって逆に攻撃を受け、それ以降、文民が戦闘員か不明の場合に発砲せざるを得なくなる。その場合に、無差別的攻撃の様相を呈する。当該兵器を無差別的兵器とみなし兵器自体を禁止することは、最初に欺瞞行為をした側にとって有利に働く。しかし、本来は欺瞞行為自体が法的問題なのであって、兵器自体の問題ではないとこをここで指摘している。 *Ibid.*, pp. 247-248.

(62) 例えば、長射程 (110 海里) の空対空ミサイル AIM-54Phoenix のように、人間の操作する「視認距離外 (Beyond Visual Range, BVR)」攻撃は近代戦では一般的である。 *Ibid.*, p. 248.

(63) 前者は、北イラク飛行禁止区域で、米空軍 F-15 のパイロットのエラーで標的を見過ごし、次の友軍ヘリをイラク軍ヘリと誤認し撃墜した。後者は、イラン航空機が上昇していたにもかかわらず、艦船への攻撃態勢で下降して

きたと誤認し、当該航空機を撃墜した。*Ibid.*, p. 248.

- (64) *Ibid.*, p. 249.
- (65) *Ibid.*, p. 250.
- (66) *Ibid.*, p. 253.
- (67) *Ibid.*, pp. 254-257.
- (68) これは、攻撃軍が、攻撃中に引き起こされそうな文民被害数を見積もるために、兵器の精度、その爆風効果、攻撃戦術、標的付近の建物内に文民のいる可能性、建物の構造などの諸要素を考察する手続きである。*Ibid.*, p. 254. P. W. シンガー前掲書（注（33））578頁（ここでは、「副次的被害推定法」と訳されている）、P. W. Singer, *Wired for War*, p. 400. CDEMの具体的な解説として、河野桂子「アフガニスタン戦争と付随的損害：武力紛争法上の評価」『上智大学論集』56巻4号2013年3月参照。
- (69) その具体例として、戦車1両の破壊で文民1人が犠牲になっても許されると見積もることは、高度の紛争であれば一般的に合理的である。1両の戦車の破壊が特定の時期に紛争の強度に照らして文民被害を正当化しなければ、指揮官は、単純に「調整不可能な」LARsを使用しないと決定すればいいという。注（59）, p. 256.
- (70) LARsは、状況に応じて投入前に最大限の付随的損害のプログラムをするか、状況の変化に応じて、索敵している最中でも遠隔操作でその再プログラムをすることは可能であるという。例えば、集合しようとしている個別の戦車を破壊する軍事的利益は、単独で作戦している個別の戦車の破壊より大きいので、受諾可能な付随的損害レベルの引き上げは、合理的であると指摘する。*Ibid.*, pp. 256-257.
- (71) 具体的には、人間が特定の環境の当該システムを投入することの決定、当該システムを事前にプログラムする方法の決定及び遠隔操作で交戦基準を改定することである。*Ibid.*, p. 257.
- (72) 例えば、無人航空システムが、LARsを投入する前に敵軍の位置を絞り込むために利用可能ならば、利用する。*Ibid.*, p. 260.
- (73) 特定通常兵器条約改正第2議定書3条10項。
- (74) 注（72）の無人航空システムの事例では、無人機の投入が、識別能力の向上の程度によっては正当化されないほどの危険にそれを晒す場合には、その使用が「実行可能」ではないことになる。*Ibid.*, 261.
- (75) 例えば、LARsは、送電線の攻撃が軍事的に実行可能であり、同じ軍事目標を達成しそうで、そして、文民や民用物をあまり危険に晒さない場合には、変電所の攻撃のために使用できないことになる。*Ibid.*, p. 261.
- (76) *Ibid.*, pp. 262-265.
- (77) *Ibid.*, pp. 266-267.

- (78) 判断する要素は、どれほどのLARsの精密度か、環境が静態的か動態的か、比例性及び予防措置のプログラムされた敷居が当該環境に対して適切か否か、当該環境にどれほど多くの疑惑が存在していたか、などがある。特に考慮すべき重要事項は、時間である。人間による最終的なプログラミングから時間が経てば経つほど、指揮官の予測しなかった攻撃結果の危険性が増大し、LARsの使用決定があまり合理的でなくなる。*Ibid.*, p. 268.
- (79) *Ibid.*, pp. 273-274.
- (80) *Ibid.*, pp. 275-276.
- (81) *Ibid.*, p. 277.
- (82) *Ibid.*, p. 278.
- (83) *Ibid.*, pp. 279-281.
- (84) 歴史的には、戦場に配備されるまでに禁止された唯一の例として、恒常的な盲目化レーザー兵器がある。しかし、その成立は、一時的な盲目化レーザー兵器でも同一の軍事目的を達成することができるので、条約が成立しても軍事的にほとんど何も失わなかったからであると指摘する。*Ibid.*, p. 281。また、一般的に戦闘の方法としての盲目化が禁止されたとは言えず、米国が同意できるほど限定された内容であったから、1995年の特定通常兵器条約第4議定書(盲目化レーザー兵器議定書)が成立したのである。拙稿「盲目化レーザー兵器議定書に対する国際法的評価」『産大法学』38巻2号2004年19-20頁。もう一つ将来の兵器や技術の発達を想定して作成された条約が、1976年の環境改変技術敵対的使用禁止条約である。当該条約は、当初予定されていた環境改変技術の研究開発やその使用準備の禁止ではなく、米ソが合意できる一定の環境改変技術の使用禁止に限定し、さらに、技術の使用禁止の敷居を高く設定した。それ故、当該条約の実効性に疑問が生じる。藤田久一「環境破壊兵器の法規制」『関西大学法学論集』28巻2号1978年13及び47頁。

## 5. まとめにかえて —— LARsの更なる議論に向けて

以上の議論から、まず、自立型兵器システム・致死性自立型ロボットは、米国防省指令が規定したように、「一旦起動すれば、人間のオペレーターによる更なる介入がなくとも標的を選択し攻撃することができる」ものであるという定義が一般的に定着していることがわかる。それをHRWが象徴的に表現した「人間が輪の外にいる兵器」と分類すると、シュミットらが指摘するように、LARsに対する誤解が生じるかもしれない。完全に人

間を武力行使の意思決定過程の外に置くという表現は、人間とLARsの関係が全く存在しないように思える。しかし、LARsの設計及びそのプログラミング、兵器に関する法的審査の評価、投入場面（戦場）の選択、戦場への実戦配備の決定、そして戦場での起動作業は、確実に人間（指揮官又は文民の上官）が実行することになる。その場面では、人間の主観的判断が介入している。人間の介入に関して、米国も認めているように、自律型兵器システムは、「適切なレベルの人間の判断」を行使できるように設計されなければならない。この場合の「適切なレベルの人間の判断」とは何を示すのかについて、今一度、議論を詰める必要がある。

このように、LARsの使用について、それぞれの段階で人間の意思が介入している事実を認識した上で、LARsが果たして国際人道法の諸原則（区別原則、比例原則、予防原則）を遵守できるのか否かが、改めて問われなければならない。現在、人間の認識力を有しないLARsは、国際法の要件を満たせず、制限規制すべきであるとの主張と、人間の主観的判断力を保有しなくとも、ある条件下では国際法の要件を満たすことができる<sup>(85)</sup>ので、制限規制の早期決定に反対する主張とが対立している。これらの解釈論争を見ていると、議論が十分にかみ合っているとは言えない。さらに、現実に存在しない未来兵器を予測しての議論という限界から、LARsの進行しつつある技術的能力及び運用上の問題点に関する両者間での共通認識が成立していないようにも思われる。これらの点を克服した上で、LARsの使用に関する国際法解釈の議論が戦わされることが重要である。現在は、LARsの法的論争の入り口に立ったと言える。その際に、法以外の武力行使の敷居の低下や紛争の常態化の諸問題についても議論されることが不可欠である。

国際法の遵守に関する議論とは別に、LARsの規制時期に関して、オールストン、ヘインズ及びHRWは、兵器が出現してからの「巻返し感覚」ではなく、兵器が出現していない現段階で帰還不能点（point of no return）に至るまでの「先制的アプローチ」を主張する。他方、シュミットらは、議論はしても結論を出すのは時期尚早であると反論する。シュミット



トらの指摘のように、盲目化レーザー兵器が実戦配備前に禁止されたのは、国際的非難も大きな影響を与えたけれども、新兵器開発国にとって新兵器を保有せずとも甘受できないほどの大きな軍事的損失がなかったことが主たる理由といえる。現段階では、LARs 開発国が、規制レベルにより異なるが、法規制による軍事的損失が大きいと判断すれば、当然、法規制の議論自体に消極的な姿勢を示すであろう。

新兵器の実戦配備前に法規制をするか否かは、国際法政策上の問題である。このアプローチ問題に密接に関連して、ロボット技術一般を含めて LARs の理解を深めるための国際会議の要請から、LARs に関する国際法的枠組みを作成する努力をすると同時に、それまでの LARs の実験、生産、組立て、移譲、取得、展開及び使用のモラトリアムの要請、さらに、NGO「殺人口ロボット阻止キャンペーン」の主張する国際条約形式での LARs の開発、生産及び使用の包括的な禁止要請まで、LARs を巡る具体的な目標設定は広範囲に及ぶ。

では、今後の LARs に関する国際的な議論は、どの場で行うのか。オールストン及びヘインズは、国連人権高等弁務官事務所が招集する多分野の専門家から構成される作業部会を提案する。また、ヘインズは、すべての国家に対する勧告要請の主体として国連人権理事会に期待する。他方、英国、フランス及びブラジルは、むしろ兵器規制の条約作成に実績のある特定通常兵器条約（CCW）の枠組みで議論することを推奨する。

確かに、LARs の規制は、人権に関連する部分もある。オールストンもヘインズも、国連人権理事会の特別報告書としての「司法外、略式又は恣意的な処刑に関する報告書」の中で LARs を取り上げたために、人権関連の枠内での議論を主張したのかもしれない。しかし、彼らの報告書は、国際人権法からではなく、国際人道法から LARs を分析している。最初に人権分野のフォーラムで兵器の規制問題を取り上げて、法規制への国際的世論を形成することは十分意義がある。しかし、具体的に兵器規制の条約を作成するための国際的な議論の次の場としては、そのままの人権分野のフォーラムよりも軍事安全保障に直結した CCW 締約国会議内で議論す



る方が、正統的であろう。<sup>(87)</sup>

もっとも、CCW 締約国会議は、「コンセンサス方式」で新たな条約作成過程を進行していくことから、兵器の合法性を前提に<sup>(88)</sup> 穏健な条約規定を形成する傾向がある。他方、1997年の対人地雷禁止（オタワ）条約及び2008年のクラスター弾（オスロ）条約のように、コンセンサス方式による妥協的な条約規定に満足せず、より高度な法規制（全面禁止）を求める「有志連合方式」を選択する国家が出現する可能性も否定できない。後者の条約作成過程の場合、規制レベルが高く設定される利点がある半面、軍事大国が成立した条約に加入しない欠点がある。先ずは、LARsの理解を深め法規制の可能性を探るために、CCW 締約国会議を議論の出発点とすることは、必要であり、かつ順当であろう。

今現在は、議論の最終目的としてLARsの研究開発から使用までのモラトリアム又は包括的禁止を設定することなく、LARsの実態の共通理解を深めるための国際的議論を開始することが先決事項である。その上で、何らかの法規制が必要であるとの国際的なコンセンサスが形成される場合に、CCW 締約国会議において、LARsの条約作成の具体的な議論が俎上に上がってくるだろう。LARsの法的論争の入り口から1歩進んだ第2ラウンドの国際的議論が強く望まれる。

## 註

(85) ヘインズも、反撃しかできないようにロボットにプログラムすれば、国際法の基準を満たしうることを認めている。*Heyns Report*, par. 74.

(86) 赤十字国際委員会（ICRC）は、ロボット分野の技術的發展がもたらす法的その他の問題の多くに決定的な回答を得るまでにまだ時間がかかりそうであり、情報に基づいて国際人道法と新技術の関係についての議論の必要性を主張している。*International Committee of the Red Cross, International Humanitarian Law and the Challenges of Contemporary Armed Conflicts Report, 31<sup>st</sup> International Conference of the Red Cross and Red Crescent, 31IC/11/5.1.2, October 2011, p. 40.*

(87) 兵器規制は、軍事的必要性和人道的考慮のバランスの上に成立するものである。人権分野の会議では、軍事的視点が欠如しているので、兵器の法規制

に関する合意は成立し得ないと思われる。

- (88) 特定通常兵器条約改正第2議定書で対人地雷の合法性基準として、自己破壊装置や自己不活性化装置の装着を義務付けたように、LARsの合法性基準として、故障、制御不能、ハッキング又は誤作動に対する安全装置や誤作動防止装置又は事後調査のための記録装置の装着義務化が考えられる。