

## 自律型兵器の倫理問題 ——正戦論に基づく分析——

張開慧

### はじめに

ドローンやその他の自律型兵器システムは、近代戦争の様相を大きく変えつつある。これらのシステムは、既にアフガニスタンやシリアにおける軍事作戦で実戦に投入され、その使用に伴う倫理的な議論が活発に進行している。さらに、深層学習技術の進展により、これらのシステムは今後いっそう高度な自律性を持つようになると考えられる。その際、任務遂行の精度と柔軟性の向上が期待されるが、同時に新たな倫理的懸念も生じてくるだろう。例えば、非戦闘員の保護や攻撃の比例性、責任の所在といった伝統的な戦争倫理の問題についても、自律型兵器の文脈においては新たな解釈や対応が必要となる。

このような戦争倫理の問題には様々なアプローチがあるが、本論文では、自律型兵器の倫理問題を、正戦論を通じて深く分析する。論文は次のように進む。第一章で自律型兵器の特徴と運用の状況に関する基本的な背景を紹介した後、第二章では、正戦論の核心的な概念である非戦闘員免除 (non-combatant) や比例性原則 (Principle of Proportionality) に焦点を当て、自律型兵器の運用がこれらの原則とどのように衝突する可能性があるかを探る。第三章では、特に自律型兵器の高度な自律性がもたらす責任空白の問題を検討し、この問題を解決するためにプログラマー、国家、指揮官、市民などの複数の関係者の責任を詳細に分析し、分散型の責任システムを構築する必要があると主張する。

### 第一章：自律型兵器の特徴と運用

正戦論に関する詳細な議論に入る前に、本論文で扱う自律型兵器の範囲とその特徴について明確にする。本論文では、ある程度の意思決定の能力を持つ致死性人工知能兵器を対象とし、特に「自律」の概念に焦点を当てる。自律型兵器に関する見解は多様であるが、本稿では、アメリカ安全保障センターの Paul Scharre が提供する「自律」の三つの次元を、自律型兵器の自律性を検討し評価する上での基盤とする。

Scharre は、自律型兵器の「自律」を以下のように説明している。

機械が行うタスクのタイプ、そのタスクを行うときの人間と機械の関係、タスクを行うときの機械の意思決定がどれほど精巧であるか。つまり、三つの異なる次元の自律がある。

これらの次元は独立していて、機械はこの三つの範囲のいずれかで自律を高めることで、「より自律的」になれる<sup>1</sup>。(Scharre, 2018 : 翻訳 59 頁)

これらの三つの次元の中でも、特に「タスクを行う際の機械の意思決定の精巧さ」と「タスクを行う際の人間と機械の関係」の二つの次元はしばしば混同される傾向にあり、それぞれの特徴を明確に理解し区別することが、自律型兵器に関する深い理解を得る上で不可欠である。

まず、「タスクを行う際の機械の意思決定の精巧さ」という次元は、技術の「知能レベル」を示している<sup>2</sup>。この次元は、機械がどれだけ予測不可能かつ複雑なタスクを達成できるかを評価する指標となる。例えば、イギリス防衛省は自律型兵器を「高いレベルの意図や指示を理解できる」兵器であると定義し<sup>3</sup>、自律型兵器の「知能レベル」の高さを強調し、自律性を評価するためにこの指標を重要視している。

一方で、「タスクを行う時の人間と機械の関係」という次元は、意思決定の際に人間がどのように関与するか、すなわち人間が「ループの内」にいるか、「ループの上」にいるか、また「ループの外」にいるかを強調している。Scharre によれば、人間が「ループ内にいる」とは、機械がタスクを行い、その後で人間が行動を承認または変更してから機械が続行する状況を指し、「ループの上にいる」とは、機械が自律的に感知、決定、行動するが、人間がその運用を監督し、必要に応じて介入する状態を意味し、「ループ外にいる」とは、機械が完全に自律して感知、決定、行動する状況で、人間がタイムリーに介入することが不可能な場合を指す<sup>4</sup>。ドイツ外務省は「自律型兵器とは、その使用に関する決定から人間の要素を完全に排除する兵器システムである」と定義し<sup>5</sup>、この定義では自律型兵器における人間の役割が完全に排除されるシナリオが想定されている。

以上のように、この2つの次元は、似ているように見えるがそれぞれ別の指標である。以下では、自律型兵器の機能を深く理解するために、Scharre が提案する「知能レベル」に基づく異なるカテゴリーを詳細に検討する。これにより、自律型兵器の多様性と複雑さを明らかにすることを目指す。Scharre によれば、自律型兵器は「知能レベル」に基づき大きく三類に分類され、範囲は単純な「自動システム」から高度な「オートメーションシステム」、さらには完全に「自律的なシステム」に至るまで多様である。

まず、「自動システム」は、環境を感知し行動するだけである。感知と行動が、直線的に直に結び付いている。例えば、特定の閾値に達したときにのみ反応する古い形式のサーモスタットは、自動システムの一例である<sup>6</sup>。このような自動システムの定義に従い、1951年から Ryan Aeronautical によって開発された一連のターゲットドローン、ファイヤー・ビーは<sup>7</sup>、その典型的なシステムと言える。このドローンは発射後、予め計画と設定されたルートに従い飛行し、任務を遂行する。

「オートメーションシステム」は、より高い知能度を持ち、行動を起こす前に複数の入力情報を検討し、実行可能な選択肢を比較考慮する<sup>8</sup>。例えば、アメリカのプレデターMQ-1はその一

例である<sup>9</sup>。20世紀末から導入されたプレデターMQ-1は地上に配置された誘導ステーションを通じて衛星通信などを媒介して遠隔操作されるが、航行する際の挙動の多くは、オペレーターが直接指示することなく機体そのものの判断で行われる。現在戦闘で使用されているシステムの多くはこのタイプに属すると考えられる。

最後の「自律システム」は、ユーザーが設定した目標に基づいて自己決定的に行動する<sup>10</sup>。1990年代にイスラエル航空宇宙産業（IAI）によって開発された無人機のハーピーは、敵のレーダーサイトからの照射を検知すると自動的に突撃し、自己犠牲的な攻撃を行う能力を持っている<sup>11</sup>。ただし、ハーピーのようなシステムは、特定の任務における自律的機能を持つ高度なオートメーションシステムと見なすことも可能である。将来の完全自律型兵器システムは、ハーピーよりもさらに進化すると考えられる。

現在、自律型兵器はスワーム（群体）として運用されることが可能で、多数の無人機が協調して任務を実行する機能を有している。つまり、これらの無人機は分散しながらも、集中的な意思決定を共有し、目標に対する連続的な反応や集団行動を行うことができる。龐によれば、スワーム内の各無人機は小型で、従来の防御システムを回避する能力を持ち、一部が破壊されても全体の行動に影響を及ぼさない特性を持っている<sup>12</sup>。アメリカ国防総省は2016年に103機の「ペルディクス」ドローン・スワームの編隊飛行実験に成功し、半年後、中国は119機の編隊飛行で記録を更新した<sup>13</sup>。このようなスワーム作戦の多様な利点があるため、戦争の様々なシナリオでの運用が見込まれている。

本論文では、以上のような低い知能から高い知能に至るまでのさまざまな知能レベルを有する自律型兵器システムを対象にしている。重要なことは、自律型兵器は特定の装備により、高度な致命的な兵器システムへと変貌する可能性があることである。この致死性は、自律型兵器の運用とその影響において重要な考慮点となる。また、本稿では「ドローン」や「無人機」といった用語をも使用する。これらは一般に自律型兵器と同義で使われることがあるが、本研究ではこれらを自律型兵器の広範なカテゴリーの一部として捉える。

## 第二章：正戦論の枠組みにおける自律型兵器の倫理問題

正戦論は、国家及びその他の政治共同体による民間人の保護、国際的な安全の保障、及び戦争の正当性を評価し、軍事的行動を指導する枠組みを提供するものである。自律型兵器がどのようにこの枠組みに適合するかを検討することは、これらの兵器が法的及び倫理的な基準を満たしているかどうかを判断する上で不可欠である。

正戦論はもともと宗教的枠組みから発展し、主に政権の道具として使われ、戦争を正当化するための手段と見なされていたが、第二次世界大戦及びベトナム戦争中のミライ虐殺事件などの衝撃的な出来事によって、戦争行為を正当化するだけでなく、批判や反省を促す枠組みとしても有効であるという再評価に繋がった<sup>14</sup>。正戦論は「戦争の正義（*jus ad bellum*）」、「戦争

における正義（*jus in bello*）」と「戦争後における正義（*jus post bellum*）」という三つの領域に分けられる。Michael Walzerはこの三つの領域が戦争における倫理的評価の異なる段階を表し、それぞれ独自の評価基準があると主張する<sup>15</sup>。本稿ではWalzerの考え方に従い、これらの領域を独立した関係にあると見なす。

本章では、主に「戦争における正義（*jus in bello*）」に焦点を置き、自律型兵器がこの原則にどの程度適合しているかを深く探求する。眞嶋俊造によれば、この原則の核心は「非戦闘員免除の原則」と「比例性原則」にある。非戦闘員免除の原則では、戦闘員と非戦闘員（民間人）の区別を明確にし、後者は攻撃の対象として扱ってはならないと規定し戦争中の行動の倫理的な制約を提供する。一方、比例性原則は、行動がもたらす害と達成される軍事的利益との間に適切なバランスを求める<sup>16</sup>。自律型兵器がこの二つの原則をどの程度遵守しているかが本章の焦点である。

以上を要約すると、正戦論は戦時中の暴力行為を正当化する手段として用いられる時もあるが、Walzerの言うように、政治的および軍事的行動に対する倫理的な制約を提供し、政治指導者や軍事指導者に自己の行動に対する責任を持たせ、非戦闘員の被害を可能な限り減少させるよう促すこともできる<sup>17</sup>。本章では、Walzerの見解に基づきながら、自律型兵器の使用に伴う倫理的問題にアプローチする。

## 2. 1 ドローンが変える戦争に関する思考：Schwarzの分析

Elke Schwarzは、武器技術、特にドローンなどの遠隔技術が道徳的決定に与える影響について示唆に富む議論を行っている。彼女はGerhard Øverlandから火炎放射器のケースを引用し、次のように論じる。火炎放射器が唯一の自衛手段と見なされる特定の状況下では、その火炎放射器を用いて無実の者にリスクを負わせても、その行為が正当化されてしまう危険性がある<sup>18</sup>、と。これは、技術介入（この場合は火炎放射器）が本来不釣り合いな手段（無実の者へのリスク）を妥当な手段に見せかける可能性があることを示し、新たな武器が導入されることによって比例性に対する私たちの認識が変わる可能性を示唆している。Schwarzはまた、ドローン技術も同じように、特定の状況下では比例的かつ唯一の選択肢であると誤って見なされる可能性があることを指摘している<sup>19</sup>。これは、他のより少ない被害をもたらす代替案が十分に考慮されていないことを意味する。

Schwarzはさらに、Bradley J. Strawserが提起した「不必要なリスク原則（PUR: Principle of Unnecessary Risk）」に注目している。この原則は、正義を実現するためや善を追求する際に、他者に不必要な致命的リスクを負わせることは誤りであると主張している<sup>20</sup>。特に戦争の文脈では、軍人に不必要な致命的リスクを強いるべきではないと示唆している。中国古代の軍事書『孫子』にも述べられるいくつかの原則にも、迅速な勝利を通じて兵士の損傷と資源の消耗を最小限に抑えることを強調する同様の思想が見られるという。一方で、Strawserによると、ドロー

ンの使用は PUR の原則に則り、行動中の兵士のリスクを減らすために「倫理的に強制される」かもしれない<sup>21</sup>。それにも関わらず、Schwarz はこの立場に懸念を示している。彼女は、ドローンが表面上は非難される余地のない殺戮技術として容易に受け入れられると指摘し、このような技術の使用が軍事的な安全を追求する過程で、実際には民間人の安全を脅かす可能性があると主張している<sup>22</sup>。

現実の戦場において PUR の適用が広がっている中で、Schwarz の指摘は特に重要である。ドローンの使用はしばしば他の兵器と比較せずに使用され、その副次的な被害が軍事目標の達成に必要な程度を超える危険性がある。もしそうなってしまえば、「戦争における正義(jus in bello)」の「非戦闘員免除の原則」と「比例性の原則」に違反する可能性がある。しかし Schwarz の仮定が正しかった場合だけでなく、現実の自律型兵器の使用が正戦論の原則に背いているかどうかも問題である。これを明らかにするために、二つの問題に答える必要がある。第一は、現代の戦争において自律型兵器は民間人への被害を増大させる形で比例性に反して使用されているかという問題である。第二に、現在の技術が不十分であっても、将来の技術の発展が一つ目の問題に対応できるかという問題である。次に、順を追ってこの二つの問題に答える。

## 2. 2 ドローン使用の実態分析：非戦闘員免除と比例性の評価

まず、Schwarz が例に挙げる火炎放射器の事例をもう一度考えよう。火炎放射器は、犠牲者に不必要な苦痛を与えるという理由のため特定通常兵器使用禁止制限条約 (CCW) によって原則的に禁止される<sup>23</sup>。しかしドローンの使用が正戦論の「比例性の原則」に適合するか否かは、具体的な軍事作戦の文脈によって異なると考えられる。例えば、精密攻撃の下でのドローン使用は、比例性の原則に適う可能性がある。つまり、使用の地域に注意を払うことで、民間人への被害を最小限に抑えることができるのである<sup>24</sup>。しかし、これはあくまで理想的な状況を述べたものにすぎず、実際の戦場での使用はこの理想から逸脱している場合がある。

ドローンの使用に関する Grégoire Chamayou の分析によれば、アメリカ政府はアフガニスタンのような軍事衝突地域だけでなく、ソマリア、イエメン、特にパキスタンのような平和とされる国々でもドローンを頻繁に展開している<sup>25</sup>。対テロ作戦におけるこれらのドローン攻撃が敵対勢力の排除には一定の効果を持つものの、同時に民間人の不必要な犠牲を引き起こすリスクがある。特にオバマ政権時代には、CIA によるパキスタンでのドローン攻撃が顕著に増加し、平均して4日に1回という高い頻度に達している<sup>26</sup>。しかし、テロ対策としてのこれらの攻撃と、それによって生じる民間人の犠牲との間には、適切な比例関係が欠けている可能性がある。これは、ドローンの攻撃の規模や頻度が目的に対して過剰であり、民間人への配慮が不足している可能性があることを示唆している。

一方で、ドローン戦争が伝統的な戦争に比べて民間人の死傷数が多いということを直接的に証明することは困難であるが、遠隔操作による戦闘が民間人に与える影響については、一部の

例が示唆に富む。例えば、1999年のコソボ戦争ではパイロットの生命を優先することを目的に米軍は遠隔作戦を多用することで、民間人を躊躇なく負傷させたり殺害したりする蓋然性を増大させたことが指摘されている<sup>27</sup>。このような事例は、ドローンや他の遠隔兵器が民間人への被害を十分に考慮せずに使用されるリスクを示している。さらに、Walzerの研究は、ドローン戦争における民間人の被害に関して重要な洞察を提供している<sup>28</sup>。アメリカ軍とCIAがパキスタンで行ったドローン戦争に対する彼の分析における例を見ると、ドローン戦争による民間人の被害が比較的少ないというアメリカ政府の主張する見解とは対照的に、独立研究者のデータは、殺害された人々の中で民間人の割合が7%から40%の間であることを示している。このように、民間人の被害についてはアメリカ政府公式の報告と外部の観察者の間で顕著な違いがあり、実際には公式に認められているよりも多くの民間人が死傷している可能性があることを示唆している。

以上の分析と事例から、ドローン作戦と伝統的な戦争における民間人の被害を具体的に比較することは困難であるが、ドローン戦争における民間人の被害への配慮を改善する必要があることは確かである。ドローンが特定の地域や方法に限定して使用されることが理想的であるが、実際には慎重さの欠ける使用がなされ、民間人の死傷リスクを増大させる傾向にある。したがって、ドローンや他の自律型兵器の使用が正戦論の原則に適合するかどうかは、事例ごとに詳細に分析する必要がある。自律型兵器の使用に伴う倫理問題について、更なる慎重な議論が求められる。

## 2. 3 技術進歩と倫理的限界：技術の発展が問題解決に役立つか？

自律型兵器の将来的な進歩が正戦論の原則に関連する倫理問題を解決する可能性については、見解が分かれる。一部の支持者は、現在の自律型兵器の倫理問題が、将来の技術進歩によって克服されると主張している。例えば、人工知能の性能向上に伴い、自律型兵器が戦争法をより厳密に遵守できるようになり、非戦闘員の被害やその他の偶発的損害を減らす可能性がある。このため、人間の兵士よりも、これらの兵器の方が効果的かつ倫理的に行動できるだろう、と<sup>29</sup>。しかし、技術進歩がすべての問題を解決するわけではない。というのも眞嶋が指摘するように、比例性の原則は紛争における民間人死傷者の絶対数と全死傷者の相対比率を制限するが、具体的な値や基準がないため、その解釈や適用においては人間の主観に依存することが多く、このような曖昧さは恣意的な操作のリスクを孕み、軍事作戦において民間人への危害を正当化する政治的な利用につながる恐れがある<sup>30</sup>。

一方、Ronald C. Arkinは、機械は人間の主観や感情に影響されず、客観的なデータに基づいて判断を下すことができると主張し<sup>31</sup>、正戦論の原則の適用をより透明かつ一貫したものにするだろうと期待している。理論上、これは自律型兵器を使用した戦争において、より正確な戦闘員の識別と比例性の原則に適った判断につながる可能性がある。しかし、たとえArkinの主張

が正しかったとしても、機械による決定プロセスには技術的な限界があり、それに伴う倫理問題は残り続ける。この問題を解明するために、二つの論点を詳細に検討する必要がある。

### 2. 3. 1 正戦論の理論の実装：区別の原則の解釈の難しさ

Arkinのようなロボット工学者は、国際人道法や現在の戦争で使われる交戦法規の明確なルールを機械にプログラムしようと試みていた<sup>32</sup>。しかし、この試みは理論上は有望に見えたものの、実は戦闘員と非戦闘員を区別する明確な基準が欠けているため、その区別基準をコード化して機械に教えることは困難であると判明した。

区別の原則に基づく判断は、しばしば状況に応じて行う必要がある。たとえば、「合法的な目標」と「非合法的な目標」の識別は文脈によって変わってくる<sup>33</sup>。合法的な目標は国際人道法や交戦法規などに基づき、軍事行動の対象として認められる一方、非合法的な目標は攻撃から保護される対象である。たとえその攻撃対象が戦闘員であっても、攻撃が過度に破壊的である場合や、民間人の死傷者が多くなる可能性がある場合、または標的が降伏の意思を示しているか戦闘不能である場合、そのような攻撃は合法ではなくなる。さらに、民間人や非戦闘員が間接的に戦争に関与している場合、そのような人々に対する攻撃の正当性は、「合法的目標」と「道徳的目標」の間でまた葛藤を生じさせる。この葛藤は、Scharreの戦争経験を通じても検証されうる。Scharreはアフガニスタンとパキスタン国境地帯で、5～6歳の少女がタリバン戦士のために偵察活動を行っていた事例を述べている<sup>34</sup>。戦時国際法では、戦闘員の年齢は定められておらず、戦闘行為に関与する者は適法な交戦対象となり得る。しかし、この少女を攻撃することは法的に許容される可能性がある一方で、道徳的には別の問題が生じる。

これらの事例は、正戦論における区別の原則が根本的に曖昧であることを明らかにしている。しかしその曖昧さは原則の欠陥ではなく、文脈によって区別する基準の解釈が必要だからである。そのためこれらの原則を自律型兵器に導入するとしても、単純にコード化するだけでは不可能である。

### 2. 3. 2 自律型兵器における区別と比例性の判断の限界：技術的現状

自律型兵器が区別と比例性の判断を行う能力に関する疑問を解明するためには、現在の技術が実際に可能とすることと、そうでないことを区別することが必要である。一部の論者は、技術の進歩によりドローンなどの自律型兵器が高度に精緻化され、無実の民間者の被害を減少させると主張している。しかし、この「精緻化」とは、主に「射撃の精確さ」を指すものであるため、標的を選択する際の「識別能力」と混同されてはいけない<sup>35</sup>。技術の進歩が命中率を向上させる可能性はあるが、それが自律型兵器による戦闘員と非戦闘員の正確な区別・識別に直接寄与するわけではない。

現在、戦闘員を識別する技術は主にカメラやセンサーなどの監視装置を通じて収集されたデータに基づく生活パターンの分析に依存している。このアプローチは、個人の日常的な行動から戦闘員である可能性を推測することを目的とする。しかしながら、この方法は誤認識のリスクを孕んでいる<sup>36</sup>。機械による分析は、特定のパターンに偏ることがあり、したがって必ずしも包括的であるとは限らない。このため、現在の戦闘員の識別技術には限界が存在する。これは将来的にも妥当である。技術の発展によって別の分析ができたとしても、依然として人間の社会的、文化的、政治的コンテキストを深く理解することには大きな限界がある<sup>37</sup>。Scharreの指摘のように、将来において使用される兵器がいかに高度な知能を持っていたとしても、背景の理解や意味の解釈においては人間に比べて劣ると考えられる<sup>38</sup>。

以上、2. 2と2. 3で、戦争における自律型兵器の使用実態と自律型兵器の技術の限界という二つの課題の検討を通じ、機械による戦闘員と非戦闘員の区別を正確に判断することにおいて困難さが生じることが明白になった。現実の使用状況を見ると、ドローンを含む自律型兵器はしばしば区別の原則と比例性の原則を考慮せずに使用され、時には濫用されて多くの民間人の被害をもたらす実例がある。技術の発展が進んでも、このような問題が解決される保証はない。そのため、自律型兵器の開発と使用には慎重なアプローチを取る必要がある。

### 第三章：自律型兵器と責任のジレンマ：戦後正義の視点

「戦争後における正義（*jus post bellum*）」は、戦争終結後の和解、復興と責任などに関する問題に焦点を当てるが、この領域の内容については、学者らの間でもまだ広範な合意には至っていない。しかし、典型的な問題はいくつか存在するため、ここではWalzerの議論に沿ってそれらを考察する。

まずWalzerは侵略罪と戦争犯罪を区別する。侵略罪は戦争全体の正義性に関する議論であり、戦争犯罪は戦争遂行中に発生する非戦闘員への不当な攻撃や国際法・人道法違反といった行為を指す。そして、戦争犯罪において、Walzerは軍人と指揮官の責任（*responsibility*）について論じている。軍人は戦争のルールに従い、民間人への加害を避けるべきである一方、指揮官には、自分が合法的な指令を出すことを保証するだけでなく、部下の行動を監督し、彼らが国際法や倫理規範に従って行動するように指導する責任を持つ。部下による戦争犯罪が発生した場合、指揮官はその行為に対して間接的な責任を負う可能性がある<sup>39</sup>。

しかし、従来の戦争とは違い、自律型兵器によって戦争犯罪になった場合、その行為における人間の判断が希薄化し、行為の責任の所在が曖昧になる可能性がある。自律型兵器による責任問題にどのようにアプローチすればよいのだろうか。

#### 3. 1 事故は避けられない—自律型兵器による責任空白



まず自律型兵器の使用に際し、非戦闘員への攻撃や国際人道法の違反といった戦争犯罪のリスクが避けられないと考える。これは主に自律型兵器システムの二つの特徴に起因する。一つ目は、高度に複雑なこれらのシステムでは、全ての状況や組み合わせを完全にテストすることは不可能であるため、予期せぬ相互作用や故障が発生する可能性がある<sup>40</sup>。二つ目は、深層ニューラルネットワークの利用により、予期せぬデータや些細な差異による誤動作が発生するリスクが伴う<sup>41</sup>。この脆弱性は敵によって利用される可能性があり、ハッキングを通じてミスが発生することもある。

ドローンが遠隔操作される場合、意思決定は人間によって行われるため、最終的な責任は人間であるオペレーターにあると考えられる。ただし、オペレーターと目標との間に心理的距離が生じることで、敵を人間ではなくて抽象的な存在や単なる目標としてしか認識しなくなり、オペレーターの責任感の低下といった別の倫理問題を招く可能性が指摘されている<sup>42</sup>。だが、その場合でも責任の所在は明確である。しかし完全自律型兵器の場合、責任の所在はより複雑になる。自律型兵器による誤動作や事故が発生した際、オペレーターは直接的な意思決定を行っていないため、機械の行動に対する直接的な責任を負うことが難しい。一方、Robert Sparrowによれば、システムの自律性が高まるほど、人間の予測を超える決定が生じる可能性があり、そのため機械の設計者やプログラマーに責任を問うことは適切ではない<sup>43</sup>。つまり、自律型兵器の導入により、従来の戦争犯罪に関連する責任の枠組みが変化し、事故が発生する際に責任の所在が不明確になる。その結果、責任の空白が生じる可能性がある。このような Sparrow の指摘は適切であり、現在の倫理的責任の枠組みで自律型兵器が引き起こす被害の責任を位置づけるのは困難だろう。

### 3. 2 自律型兵器における多面的な責任

正戦論の観点から自律型兵器の使用における責任問題を検討すると、直接的な責任者が不明確な場合がある。しかしその場合でも、プログラマーと国家は、システムの開発と運用の段階で重要な責任を担うという指摘がある。そこでこの論点について、いくつかの論者の指摘を参考にしたい。

まず、Merel Noormanによれば、システムが自律的であればあるほど、人間が予期せぬ方法で動作する可能性が高まるとしても、自律型兵器の意思決定の条件と方法の設定は人間に依存する<sup>44</sup>。したがって、プログラマーは自律型兵器の設計段階からシステムの動作を適切に規制し、システムの信頼性と安全性を保証する責任を負う。これには、慎重なプログラミング、安全性向上のための継続的な取り組みが含まれる。プログラマーがこの責任を適切に果たすことは、自律型兵器の倫理的な開発において不可欠である。しかし、自律型兵器の全責任がプログラマーに押し付けられるのも間違っているだろう。そのため、この枠組みのみで自律型兵器の責任の問題を解決することはできないと考えられる。

次に、自律型兵器システムに関連する戦争犯罪の責任は国家が負うべきであるという提案を考えよう。Rebecca Crootofの提案によると、自律型兵器の開発と使用において国家が重要な役割を果たし、結果として国家レベルでの責任が強調されることになる。国家の責任は個々の兵士や指揮官の責任を超える強力な抑止力として機能し、自律型兵器システムによる戦争犯罪行動に対する責任を個人から国家に移行させることが提案されている<sup>45</sup>。しかしながら、このアプローチにも限界がある。犠牲者への補償能力や戦後の賠償や復興における国家の役割を考えると、国家の責任は確かに重要である。しかし、たとえ国家が自律型兵器の開発と使用において重要な役割を果たすとしても、戦争犯罪の責任が国家のみに帰属するわけではない。人間の兵士が行った戦争犯罪が、国家の責任であると同時に犯罪行為を行った個人の責任でもあるように、自律型兵器システムによる戦争犯罪に関して国家だけが責任を負うわけではない。

以上のように、自律型兵器の責任は開発者だけでも、国家だけでもなく、もっと広い範囲で考えなければならない。そこで、以下から、Walzerの理論に基づき、自律型兵器の使用に関連するさらなる二つの主体の責任を考慮することが合理的であると考えられる。これらは指揮官と市民の責任である。Walzerが直接に自律型兵器について言及していないとしても、彼の一般的な原則は新しい軍事技術、特に自律型兵器にも適用される。

まず、指揮官は関係者のなかでも比較的大きな責任を負う必要がある。これには反論もあるだろう。例えばCrootofは、自律型兵器の自律性に基づき、指揮官にはシステムの意図を知らず、また犯罪行為を止める時間も手段もないので間接的な責任を問われないと主張し、指揮官の責任に反対する<sup>46</sup>。しかし、Crootofの見解は自律型兵器の複雑さについても、指揮官の役割についても過小評価している。彼女の主張は、指揮官が適切な訓練と事前のリスク評価を通じて自律型兵器の使用を十分に監督できるという可能性を見落としている。一方、Walzerの指揮官の責任の論点に基づく、指揮官は自律型兵器に関連する犯罪行為に対して、より大きな責任を負うべきである。指揮官は、これらの技術が倫理的および法的規範に従って使用されることを確保する責任を持つ。特に、自律型兵器の特性とリスクを理解し、適切な監督と指導を行うことが、自律型兵器の使用における責任とリスクの全体像を考慮する上で重要である。そのため、この問題については、Walzerの主張には大きな説得力があると思われる。

次は、Walzerの正戦論によると、「民主的市民」もまた侵略戦争に対して一定の責任を負う。民主政治の下で、市民は国の政策に大きな影響を及ぼすからである<sup>47</sup>。この理論によれば、自律型兵器の開発と使用においても市民は監督の責任を持つ。民主政治の下では、市民は自国の軍事政策、特に新技術の採用に関して影響力を持ち、それに基づいた監督を行うべきである。これには、政治的議論や民主的プロセスを通じて自律型兵器の使用に関する倫理的および法的基準に基づいて運用することを促し、監督する役割が含まれる<sup>48</sup>。

本章では、自律型兵器の責任問題を、従来のプログラマーと国家の責任だけでなく、正戦論に基づいて指揮官と市民の責任も含めて検討した。さらにその自律型兵器の運用に伴うさまざまなシナリオにおける多様な責任の分担も少し議論したい。例えば、システムの誤認識により

民間人に被害をもたらした場合、この事故はプログラマーのアルゴリズム設計、国の監督方針、指揮官の攻撃の監督の問題として責任を問うことが適切である。また、自律型兵器が過剰に使用されて大規模な被害をもたらした場合、その責任はプログラマーと指揮官の意思決定、国家と市民の「過剰な力の使用」に対する監督に関連している。つまり、どのような事故に対しても一律に同一の基準で関係者に責任を負わせるアプローチは非効果的であり、各シナリオに応じた責任の分担を慎重に検討する必要がある。本稿では、このようなシナリオごとに責任の所在を多元的に捉える考え方を「分散型責任システム」と呼びたい。このアプローチは、自律型兵器の使用において関係者間の責任を効果的に管理し分担する枠組みとして機能する。事故発生時の責任割り当てを詳細に規定することは、各状況における関係者の責任を適切に評価するために不可欠である。具体的なシナリオと責任のあり方を慎重に検討する必要があり、この枠組みの構築と実際の運用にはさらなる議論と具体的な計画が必要である。これにより、自律型兵器の開発と運用はより責任あるものとなるだろう。

#### 第四章：結論

本稿では、自律型兵器の倫理問題を正戦論の観点から検討し、特に第二章と第三章でこの問題に焦点を当て、自律型兵器が正戦論の基準に適合することの困難さを明らかにした。具体的には、自律型兵器の使用が民間人への被害が増加させるリスクが高いこと、そして技術的進歩が必ずしもこれらの問題を解決するわけではないことを論じた。第三章では、自律型兵器の使用に伴う責任問題を掘り下げ、戦争犯罪における責任の所在が不明瞭になる可能性を考察した。しかし自律型兵器の意思決定に直接的な責任を負う主体が存在しない場合でも、少なくともプログラマー、国家、指揮官、市民という四つの主体が責任を負う可能性があることを主張した。本論文全体の分析を通じ、自律型兵器の使用が正戦論の基準に違反する可能性があること、また正戦論が自律型兵器の時代の課題に十分に対応できていないことが明らかになった。

さらに本論文では触れられなかったが、機械による戦争における意思決定が適切であると仮定しても、人間の命に関する決定を機械に委ねることが倫理的に正当化されるかどうかは重要な疑問である。Walzerの権利理論に基づく正戦論では、戦士間の平等性を尊重することが重要視される。その観点から、自律型兵器による意思決定が戦士間の平等性を損なうかどうかは、今後の研究において重要な問題である<sup>49</sup>。

#### 註

<sup>1</sup> Scharre (2018) 翻訳 59 頁。

<sup>2</sup> 同上、翻訳 62 頁。Scharre によれば、「タスクを行う際の機械の意思決定の精巧さ」は、機械の「知能レベル」を示す次元である。この概念は本稿全体で重要な役割を果たすため、以降は便宜上「知能レベル」という用語で参照する。

3. UK Ministry of Defence, “Unmanned Aircraft Systems”,  
[https://assets.publishing.service.gov.uk/media/5a823670ed915d74e6236640/doctrine\\_uk\\_uas\\_jdp\\_0\\_30\\_2.pdf](https://assets.publishing.service.gov.uk/media/5a823670ed915d74e6236640/doctrine_uk_uas_jdp_0_30_2.pdf), 最終閲覧日：2024年2月11日。
4. Scharre (2018) 翻訳 60–62 頁。
5. German Federal Foreign Office, “German commentary on operationalizing all eleven guiding principles at a national level as requested by the chair of the 2020 Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS) within the Convention on Certain Conventional Weapons (CCW)”,  
<https://documents.unoda.org/wp-content/uploads/2020/07/20200626-Germany.pdf>, 最終閲覧日：2024年2月12日。
6. Scharre (2018) 翻訳 62 頁。
7. Popular Mechanics (1954) August, “Brightly Colored Fish Flies High Over New Mexico’s Sand Dunes”, p. 105.
8. Scharre (2018) 翻訳 62–63 頁。
9. “MQ–1B Predator”, <https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104469/mq-1b-predator/>, 最終閲覧日：2024年2月12日。
10. Scharre (2018) 翻訳 63 頁。
11. “HARPY: Autonomous Weapon for All Weather”, <https://www.iai.co.il/p/harpy>, 最終閲覧日：2024年2月12日。
12. 龐 (2018) 翻訳 219 頁。
13. 「AI 兵器開発，米中が火花-静まりかえった北米攻撃 CG」  
<https://www.asahi.com/articles/ASLDQ7JF6LDQUHBI029.html>, 最終閲覧日：2023年12月6日。
14. Walzer (2002) pp. 925–944. また，十六世紀以来，ヒューゴー・グロティウスやエメリッヒ・ド・ヴァッテルのような法哲学者は，正戦論を国際法の領域に取り入れ，近代国家体系の形成と共に，国際関係におけるその適用範囲を拡大した。この脈絡において，ジュネーブ条約とハーグ条約は，正戦論の人道主義的原則を具体化し，国際法のさらなる進展において重要な役割を果たした。
15. Walzer (1977) Chapter 2. 一方，Jeff McMahan は，これらの領域が完全に独立して考えることはできないと主張している。彼は論文 “The Ethics of Killing in War” の中で，特に「戦争の正義 (jus ad bellum)」と「戦争における正義 (jus in bello)」の間には相互依存関係が存在し，一方の領域での判断が他方に影響を及ぼす可能性がある」と論じ，Walzer の議論を批判した (McMahan (2004) pp. 693–733)。この観点は，Walzer の立場とは異なるものであり，正戦論における複雑な関係性を示唆している。
16. 眞嶋 (2016) 90 頁。ただし，ここで引用されている眞嶋の「戦争における正義 (jus in bello)」に関する記述は，戦争法と国際人道法における一般的な理解を反映している。非戦闘員免除と比例性原則は，多くの国際法の専門家や学者によって受け入れられている基本的な原則であり，特定の個別の見解ではなく，広範な合意に基づいていることを注記する。また，「戦争における正義 (jus in bello)」において，非戦闘員免除原則と比例性原則は時として矛盾することがある。例えば，特定の軍事目標に対する攻撃が，損害と利益を比較した結果，戦略的に有益であると判断されたとする。しかしながら，その攻撃によって一定数の民間人が犠牲になる可能性があるとしたら，どうすべきだろう。非戦闘員免除原則からは，民間人の安全を最優先し，そのような攻撃は避けるべきとされる。一方，比例性原則からは，その攻撃の戦略的利益が民間人の損害を上回る場合，攻撃を実施すべきだと考えられる。このような限界事例において，両原則が矛盾する可能性を認めつつ，本稿ではこの問題について深く探求し，議論を進めたい。
17. Walzer (2002) pp. 925–944.
18. Øverland (2014) pp. 481–506.
19. Schwarz (2018) pp. 280–298.
20. Strawser (2010) pp. 342–368.

21. 同上.
22. Schwarz (2018) pp. 280–298.
23. International Committee of the Red Cross (1977), Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, p. 108. CCW では焼夷兵器の使用を規制している。これには物を燃やし人に火傷を与える目的で設計された火炎、熱を生じる武器や弾薬が含まれ、「火炎放射器」も該当する。CCW では特定の武器種の明示的な禁止はされていないが、民間人に影響を及ぼす可能性がある場合にはその使用に制限が設けられている。
24. Sparrow (2016) pp. 93–116. Arkin の将来的に自律型兵器システムが人間よりも戦争の倫理的行動要件を満たす可能性があるという主張に対して、Sparrow は、自律型兵器システムには合法的な目標と非合法的な目標を区別するのは難しいと指摘するが、この問題は特定の限定された領域ではより扱いやすくなるとも述べている。
25. Chamayou (2015) 翻訳 24 頁.
26. Chris Woods (2011), “Drone Strikes Rise to One Every Four Days”, *The Bureau of Investigative Journalism*, <https://www.thebureauinvestigates.com/stories/2011-07-18/drone-strikes-rise-to-one-every-four-days>.
27. Chamayou (2015) 翻訳 152 頁.
28. Walzer (2016) pp. 12–24.
29. Arkin (2009a) pp. 30–33.
30. 眞嶋 (2009) 57–70 頁. 注記として、眞嶋は別の研究で自律型兵器が武力紛争法や「戦争における正しさ」をどの程度遵守できるかについて論じている。彼によると、現時点では自律型兵器がこれらの法規を完全に遵守することは不確実であるが、技術の進歩と指揮官の倫理教育により、将来的には自律型兵器が人間の兵士よりもこれらの原則を遵守する可能性があるとされている。しかし、この技術が具体的にどのように区別の基準を満たすかについては、彼の議論は詳細ではない。また、本論文のこのセクションでの関心事は、技術的側面よりも、「正戦論」の基本原則の曖昧さに焦点を当てる。
31. Arkin (2009a) pp. 30–33.
32. Arkin (2009b) Chapter 7.
33. Sparrow (2016) pp. 93–116.
34. Scharre (2018) 翻訳 26–27 頁.
35. Chamayou (2015) 翻訳 165–168 頁.
36. 同上, 62 頁.
37. これらの技術が社会的、文化的、政治的コンテクストをどの程度理解し、取り入れることができるかについては、今後の研究でさらに探求する必要がある。
38. Scharre (2018) 翻訳 30 頁.
39. Walzer (1977) Chapter 19. ただし、戦争はその本質において複雑なものであり、Walzer は軍人と将軍の責任を絶対的なものとしてではなく、状況に応じたものとして捉えている。
40. Scharre (2018) 翻訳 212 頁.
41. 同上, 翻訳 248 頁.
42. Royakkers and van Est (2010) pp. 289–296. Lambèr Royakkers たちは、軍用ロボットの操作がゲームのようなインターフェースで行われる現在、オペレーターは敵を非人間的に見る傾向がある。この状況は、オペレーターの責任感を低下させるリスクがあることを示唆している。一方、Chamayou Grégoire (2015) は、従来の遠隔操作と違い、ドローンのオペレーターは高解像度のカメラでターゲットの生活を数週間観察し、生活パターンの情報を蓄積してから攻撃を行うため、心理的な距離はむしろかなり縮まったと主張している。これらの異なる視点は、遠隔操作ドローンの使用が人間の心理に複雑な影響を与えることを示している。どちら

の場合でも、人間の操作員に責任を負わせるのは合理的である。

<sup>43</sup> Sparrow (2007) pp. 62–77.

<sup>44</sup> Noorman (2014) pp. 51–62.

<sup>45</sup> Crootof (2016) pp. 1347–1402.

<sup>46</sup> 同上.

<sup>47</sup> Walzer (1977) Chapter 18.

<sup>48</sup> 注記として、ここでの市民の「責任」は、Walzer (1977) の議論に基づき、民主制度の下での市民が国の戦争の決定に持つ道徳的・政治的責任を指す。法的な責任ではなく、市民を軍事攻撃の合法的な対象とするものではない。国際人道法により、非戦闘員（市民）は直接軍事行動に関与しない限り合法的な攻撃対象にあたらぬ。

<sup>49</sup> 本研究は JST 科学技術イノベーション創出に向けた大学フェローシップ創設事業 JPMJFS2129 の支援を受けたものである。

## 参考文献

Arkin, Ronald, C. (2009a). “Ethical Robots in Warfare”, in *IEEE Technology and Society Magazine*, Vol. 28, No. 1, pp. 30–33.

——— (2009b). *Governing Lethal Behavior in Autonomous Robots*, Routledge.

Chamayou, Grégoire. (2015). *A Theory of the Drone*, Penguin. (渡名喜庸哲 (翻訳), 2018 年『ドローンの哲学—遠隔テクノロジーと「無人化」する戦争』明石書店.)

Crootof, Rebecca. (2016). “War Torts: Accountability for Autonomous Weapons”, in *University of Pennsylvania Law Review*, Vol. 164, No. 6, pp. 1347–1402.

McMahan, Jeff. (2004). “The Ethics of Killing in War”, in *Ethics*, Vol. 114, No. 4, pp. 693–733.

Noorman, Merel and Deborah G. Johnson. (2014). “Negotiating autonomy and responsibility in military robots”, in *Ethics and Information Technology*, Vol. 16, pp. 51–62.

Overland, Gerhard. (2014). “Moral Obstacles: An Alternative to the Doctrine of Double Effect”, in *Ethics*, Vol. 124, No. 3, pp. 481–506.

Royakkers, Lambèr and Rinie van Est. (2010). “The cubicle warrior: the marionette of digitalized warfare”, in *Ethics and Information Technology*, Vol. 12, pp. 289–296.

Scharre, Paul. (2018). *Army of None: Autonomous Weapons and the Future of War*, Tantor Media Inc. (伏見威蕃 (翻訳), 2019 年, 『無人の兵団—AI, ロボット, 自律型兵器と未来の戦争』早川書房.)

Schwarz, Elke. (2018). “Technology and moral vacuums in just war theorizing”, in *Journal of International Political Theory*, Vol. 14, pp. 280–298.

Sparrow, Robert. (2007). “Killer Robots”, in *Journal of Applied Philosophy*, Vol. 24, No.1, pp. 62–77.

——— (2016). “Robots and Respect: Assessing the Case Against Autonomous Weapon Systems”, in *Ethics & International Affairs*, Vol. 30, No. 1, pp. 93–116.

- Strawser, Bradley J. (2010). “Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles”, in *Journal of Military Ethics*, Vol. 9, No. 4, pp. 342–368.
- Walzer, Michael. (1977). *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, Basic Books.
- (2002). “The Triumph of Just War Theory (and the Dangers of Success)”, in *Social Research*, Vol. 69, No. 4, pp. 925–944.
- (2016). “Just & Unjust Targeted Killing & Drone Warfare”, in *Daedalus*, Vol. 145, No. 4, pp. 12–24.
- 眞嶋俊造 (2009). 「民間人保護を巡る正戦論への建設的批判」, 『応用倫理』, 1, 57–70 頁.
- (2016). 『正しい戦争はあるのか?: 戦争倫理学入門』, 大隅書店.
- 龐宏亮 (2018). 『21 世紀戦争演變與構想: 智能化戦争』上海社会科学院出版社. (安田淳, 上野正弥, 金牧功大, 御器谷裕樹 (翻訳) (2021) 『中国軍人が観る「人に優しい」新たな戦争—智能化戦争』五月書房新社. )