

http://dx.doi.org/10.17703/JCCT.2023.9.4.375

JCCT 2023-7-44

치명적 자율무기체계의 도덕적 책임 문제 연구 : 로버트 스페로우의 ‘책임간극’ 이론에 대한 고찰

A Study on the Moral Responsibility of Lethal Autonomous Weapon Systems (LAWS): Focused on Robert Sparrow's “Responsibility Gap” Theory

문현영*, 김상수**

Hyunyoung Moon*, Sangsu Kim**

요약 미래 전장 양상이 초연결성에 기반한 전장 네트워크 발전에 기반할수록 그 결과 또한 예측할 수 없는 불확실성이 더욱 커질 것으로 예상된다. 이런 환경에서 인공지능에 기반한 자율무기체계 도입은 그 불확실성을 더욱 증대시킨다. 불확실성이 증대된 환경에서 공동의 작업으로 인해 도출된 결과에 대한 책임을 평가하는 것은 그 중요성에도 불구하고 선명한 방향을 정하지 못하고 있으며 누구도 충분한 책임을 지지 않는 책임간극(responsibility gap)에 대한 우려도 있다. 이에 따라 이 글에서는 자율무기체계의 운용양상을 분석하고 그 과정에서 발생할 수 있는 윤리적 문제로서 책임간극의 문제를 로버트 스페로우의 이론을 통하여 고찰해볼 것이다. 그의 주장을 면밀히 살펴보고 치명적 자율무기체계(lethal autonomous weapon system, LAWS)의 운용에서 책임간극은 극복 가능한 것임을 밝힌다. 간극이 아닌 책임의 프레임이 앞으로 다가올 전장 환경에 더 적합함을 보여줌으로서 신중한 무기체계 개발 및 사용의 기반을 마련하고자 한다.

주요어 : 자율무기체계, 인공지능, 책임간극의 문제, 도덕적 책임, 로버트 스페로우

Abstract In an increasingly hyperconnected battlefield, the reliance on battlefield networks and AI-based autonomous weapons systems creates uncertainty and raises ethical concerns. This article explores the responsibility gap in operating autonomous weapons systems, using Robert Sparrow's theory. By analyzing Sparrow's arguments, we propose overcoming the responsibility gap in lethal autonomous weapon systems (LAWS). Our objective is to establish a framework of responsibility that aligns with the evolving battlefield, promoting the development and use of responsible weapon systems.

Key words : Autonomous Weapon System, Lethal Autonomous Weapon System, Artificial Intelligence, Responsibility Gap, Moral Responsibility, Robert Sparrow

1. 서론

로봇 전투원, 즉 인공지능에 기반한 자율무기체계는 오랫동안 인간에게 호기심과 두려움의 대상이었다. 그

런 이유로 많은 영화 및 대중문학 작품이 로봇 전투원을 주요 소재로 다루었다. 인류를 멸종시키기 위해 스카이넷이 파괴한 터미네이터에서부터 인간의 귀여운 친구인 스타워즈 속의 C-3PO에 이르기까지 인간은 다

*정회원, 육군사관학교 영어과 조교수 (제1저자)

**정회원, 육군사관학교 국어철학과 부교수 (제2저자)

접수일: 2023년 6월 19일, 수정완료일: 2023년 7월 5일

게재확정일: 2023년 7월 10일

Received: June 19, 2023 / Revised: July 5, 2023

Accepted: July 10, 2023

*Corresponding Author: hymoon2016@gmail.com

Dept. of English, Korea Military Academy, Korea

가을 미래의 자율로봇과 인간의 관계에 대해 끊임없는 상상을 해왔다. 다양한 모습의 로봇 중에서도 상상력을 갖춘 로봇은 특히 인류를 공격하는 모습으로 그려져 대중이 가지는 자율무기체계에 대한 두려움을 엿볼 수 있다.

최근 세계 각 군에서 개발 중이거나 개발 완료된 무기체계를 보자면 로봇 전투원은 더 이상 상상 속이 아닌 현실에 존재한다는 것을 알 수 있다[1]. 이스라엘의 하피(Harpy) 드론이나 미국 팔랑스(Phalanx) 근접방공 체계의 경우 인간의 통제나 감독 없이 스스로 목표물을 식별 및 결정하고 타격까지 수행한다. 일련의 과정에서 인간 운용자의 역할이 배제되어 있다는 점에서 기존의 무기체계와 구분된다. 인공지능과 결합된 이러한 무기체계는 주변 환경의 변화를 기민하게 감지하고, 감각 입력을 기반으로 행동을 결정하며, 때로는 인간의 개입 없이 행동할 수 있는 역량을 보인다. 이런 무기체계를 치명적 자율무기체계(lethal autonomous weapon systems, LAWS)라고 부른다. LAWS는 전투에서 신속성을 보장하고 인간 전투원의 손실 및 노동을 줄여준다는 점에서 여러 국가에서 도입하는 추세이다[1], [2].

LAWS는 인간의 개입이 최소화 되었거나 부재하다는 점에서 법적 및 도덕적 논란의 대상이 된다. 이런 우려를 해소하고자 LAWS의 운영에 대한 법적 토대를 설계하고 이를 정당화하기 위한 다양한 도덕적 성찰이 이루어지고 있다[3]. 특히 국제인도법(International Humanitarian Law, IHL)에서 보호의 대상으로 규정하고 있는 비무장 민간인 및 전투능력을 상실한 군인에 대한 살상 행위에 대한 도덕적 책임의 평가에 대한 기초논의로서 '책임간극(responsibility gap)'의 문제가 제기되었다[4].

전통적으로 책임간극의 문제는 집단적 행동의 맥락에서 논의되었다. 이 문제는 여러 개인이 관계하는 집단 행위의 결과로 도덕적으로 정당화될 수 없는 결과가 발생하면 이에 대한 도덕적 책임을 평가하고 그 책임을 귀속시키는 것이 집단 구성원 중 특정 개인이나 소수에 게만 하는 것이 제한된다는 점에서 출발한다. 만약 인공지능에 기반한 LAWS가 개입한다면 도덕적 책임에 대한 평가의 문제는 더욱 난해해진다. 그 이유는 자율무기체계가 개발이 다수의 노력이 결집된 산물이며 그 기술에 내재되어 있는 메커니즘은 기존의 기술체계에 비해 더욱 복잡하기 때문이다. 특히 인간의 개입이 미미한 자율성을 보유한 LAWS의 행위는 인간의 행위라

기보다 인공지능 무기의 자율적 행위로 인식될 수 있는 여지가 있기에 그 결과에 대한 도덕적 책임을 누구에게 부여해야할지 논란이 있는 것이다.

따라서 이 글에서는 LAWS의 운용양상 및 분류를 간략히 소개하고 그 운용과정에서 도덕적 책임 평가의 어려움을 초래하는 책임간극의 문제를 로버트 스페로우의 주장을 중심으로 분석할 것이다. 스페로우 개념의 분석을 토대로 책임간극 문제를 극복할 수 있는 가능성을 제시하도록 하겠다.

II. 자율무기체계와 도덕적 문제

1. 자율무기체계(AWS)의 정의

자율무기체계의 정의는 국가별 및 학자별로 차이를 보인다. 하지만 가장 널리 통용되며 영향력이 있는 정의는 미 국방부의 정의이다. 미 국방부가 발간한 지침서 DODD 3000.09는 자율무기체계를 “일단 활성화되면, 사람의 추가 개입 없이도 목표물을 선택하고 교전할 수 있는 무기 시스템(a weapon system that, once activated, can select and engage targets without further intervention by an operator)”이라고 정의한다[5]. 미 국방부의 정의는 국제적십자사 및 국제인권감시기구의 정의와 맥을 함께 하는데 이 두 기구는 각각 “독자적으로 목표물을 선택하고 공격할 수 있는 무기”와 “인간의 개입없이 목표물을 선택 및 타격할 수 있는 무기”라고 자율무기체계를 정의하고 있다[6], [7]. 그 운용양상보다 구체적으로 이해하기 위한 예시로 무인 항공기



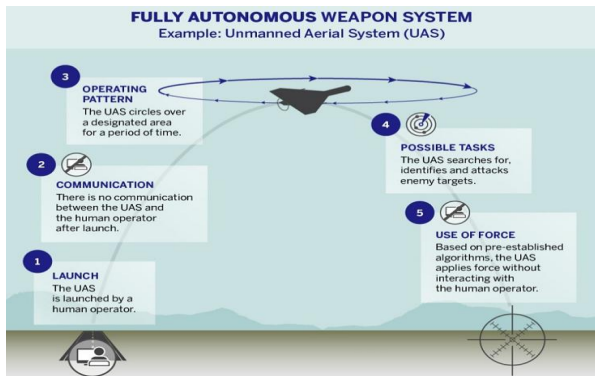
출처: Autonomous Weapon Systems: Selected Implications for International Security and for Canada(2019)

그림 1. 비 자율무기체계 UAV 예시
Figure 1. Non-autonomous weapon system example

(Unmanned Aerial Vehicle, UAV)와 무인항공시스템(Unmanned Aerial System, UAS)의 차이를 살펴보겠다. 그림 1은 전투원이 비 자율무기체계를 운용하는 모습을 제시한다.

전쟁에서 전투원은 UAV를 활용한 정찰 활동을 통해 표적을 식별하고 공격 여부를 결정한다. 그 과정을 보다 구체화하면 전투원은 전장 상황을 인식(cognition)하고 자신에게 주어진 임무를 수행하기 위한 판단(judgement)을 한 후 이를 자신이 보유하고 있는 무기를 사용하여 행위(action)로 옮긴다. 이 과정에서 효과적인 임무 수행을 위해 UAV와 이를 운용하는 전투원은 지속적인 정보교류를 하며 전투원은 그 과정에서 판단에 필요한 정보를 획득한다. 또한, 행위에 있어 식별된 가능한 과업을 전투원이 가용한 무력수단을 통해 직접 실행, 즉 행동한다.

이와 달리 자율무기체계는 인간의 개입이 (거의) 없이 운용된다. 그림 2는 전투원의 개입 없이 운용되는 UAS의 활용모습이다.



출처: Autonomous Weapon Systems: Selected Implications for International Security and for Canada(2019)

그림 2. 자율무기체계 UAS 예시
 Figure 2. Fully autonomous weapon system example

자율무기체계는 전투원-기계(UAS) 사이의 소통이 없이 UAS가 자율적으로 판단 및 행위를 한다. 그 과정에서 전투원은 단순히 UAS를 전장에 투입하지만 할 뿐 그 이후의 과정은 오로지 UAS의 독립적 운용으로 이루어진다. 자율무기시스템은 시스템에 내재하여 있는 센서를 통해 전장 환경을 인식하고 세팅된 소프트웨어와 알고리즘에 따라 판단 및 결정을 한 후 최종적으로 행위를 한다.

2. 자율무기체계의 분류

앞서 살펴본 UAV와 UAS의 차이처럼 자율무기시스템은 그 자율성(autonomy) 및 인간이 무기체계에 관여하는 방식 및 정도에 따라 표 1과 같이 세 단계로 나눌 수 있다[5], [7]. 용어사용은 상이하지만 그 내용에서는 대동소이함으로 이 글에서는 미 국방부 용어를 사용하도록 하겠다.

표 1. 자율성 정도에 따른 LAWS 용어사용
 Table 1. Levels of AWS autonomy

	미 국방부	국제인권기구
낮음	반자율 (semi-autonomous)	human-in-the-loop
	인간감독 (operator-supervised)	human-on-the-loop
높음	완전자율 (autonomous)	human-out-of-the-loop

무기의 자율성이 가장 떨어지는 ‘반자율’ 단계는 인간이 의사결정의 주체로서 무기체계에 표적을 설정하고 교전 실시여부를 판단하는 경우로 이 단계에서 무기체계는 인간의 승인 없이는 어떠한 단독행동도 하지 않는다. 이 단계에 해당하는 무기체계는 재블린(Javelin)을 포함한 발사 후 망각(fire and forget) 방식을 가진 유도미사일이 있다. ‘반자율’ 단계보다 한 단계 높은 자율성을 가진 무기체계 분류는 ‘인간감독’ 단계인데 이 단계에서 무기체계는 탐지, 결정, 행동 단계를 무기체계가 단독으로 수행하지만 인간 운영자가 언제든지 작동을 멈출 수 있다. 현재 약 30개 국가에서 ‘인간감독’ 단계의 무기체계를 사용하고 있는데 주로 포화된 공격(saturated attacks)과 같이 인간이 모든 행동을 인식하고 통제할 수 없는 경우를 대비한 자율 무기체계라고 할 수 있다. 미군의 팔랑스 근접방공체계, 패트리엇 미사일 방어체계, 그리고 독일의 반포병 시스템인 맨티스 등이 여기에 해당한다. 현재 전력화되어있는 ‘인간감독’ 체계는 앞선 예시에서도 볼 수 있듯이 함정이나 부대 등을 방어하는 목적으로 사용되며 인간에 대한 공격보다는 인공물(미사일, 로켓 및 항공기) 등을 그 대상으로 한다.

반면 ‘완전자율’ 단계의 무기체계는 인간 운영자가 완전히 결여된 상태인데 이 단계가 바로 DODD 3000.09가 정의하고 있는 ‘자율무기체계’에 해당하며 가장 많은 윤리적 논란이 동반된다[8]. 이 단계에서 무기체계는 활성화 상태에서 탐지, 결정, 행동의 단계를 인간의 개

입 없이 실행하며 인간과 어떤 교신도 하지 않고 단독으로 목표물에 대한 공격을 실시한다. 현재 전력화 되어있는 무기체계 중 가장 대표적인 ‘완전자율’ 체계는 이스라엘의 하피 드론이다. 하피는 흔히 ‘배회하는 무기(loitering munition)’라고도 불리는 드론으로서 최대 500km의 반경 안에서 스스로 적 레이더를 탐지 및 식별하고 파괴까지 실시한다. 이 경우 활성화 여부는 인간이 결정하지만 어떤 목표물을 공격할지에 대한 결정은 하피 스스로 한다.

3. 자율무기체계의 도덕적 쟁점

인간의 개입이 없는 완전한 자율성을 보유한 자율무기체계는 작전목적을 달성하는데 효과적이다. 특히 미래전에서는 전장이 다영역 환경(multi-domain environment)으로 확장됨에 따라 전장에서 활용할 수 있는 무기체계의 자율성 향상은 선택이 아닌 필수적 요구라 할 수 있다. 하지만 자율무기체계의 수많은 이점에도 불구하고 인간의 개입이 부재한 상태에서 상대 인간에 대한 살상을 한다는 점은 많은 도덕적 논란이 있다. 특히 살상에 대한 의사결정을 기계에게 위임하는 것은 상대의 생명권 및 존엄성을 침해한다는 견해가 지배적이다[9], [10]. 또한 자율살상무기가 오류나 오인으로 인하여 국제인도법에 벗어난 살상행위를 할 경우 책임의 주체가 사라지는 ‘책임간극’이 발생한다는 점도 쟁점이 된다[4], [11].

전통적으로 도덕적 책임은 도덕적 행위자(moral agent)에게 귀속시킬 수 있는 것으로 받아들여져 왔다. 도덕적 책임의 귀속 대상으로서 도덕적 행위자의 개념이 등장하는 것은 책임이라는 그 가치를 평가하기 위해서는 어떤 대상이 자신으로부터 발생한 행위에 대한 자율적 통제(control)와 인식(knowledge)을 할 수 있어야 하는 점이 필요조건(necessary condition)으로 제시되기 때문이다. 성인의 경우 도덕적 행위자의 조건을 대개 만족하고 있으며 그런 까닭에 만약 성인이 도덕적으로 정당화될 수 없는 행위를 할 때 이에 대한 책임을 귀속시키는 것은 보편적으로 수용되어온 관점이다. 하지만 자율무기체계의 경우 행동을 한 주체가 다수의 참여자가 생산한 기계이며 생산자와 운용자가 상이함에서 오는 ‘많은 참여자의 문제(the problem of many hands)’에 직면하게 된다.

4. 로버트 스페로우의 ‘책임간극’

철학자 스페로우는 자율무기체계를 ‘킬러 로봇(killer robot)’이라고 칭하며 책임을 귀속시킬 수 없는 무기를 전쟁에서 사용하는 것은 범죄이며 부도덕한 행위라고 비판하였다[4]. 스페로우는 LAWS 사용은 정의롭고 의미있는 책임귀속을 불가능하게 한다고 주장하며 무기를 개발한 프로그래머, 무기를 사용하는 지휘관, 그리고 기계 그 자체 모두 책임을 묻기에는 불충분한 주체라고 주장하였다. 그는 LAWS가 자율적으로 목표물을 설정하고 타격한다는 점을 들어 프로그래머와 지휘관은 운용 간 충분한 통제권을 가지고 있지 못하다고 주장하였으며 LAWS는 기계임으로 근본적으로 처벌이 불가능하다고 주장하였다. 다음 표 2는 LAWS의 독단으로 행해진 전쟁범죄 책임에 대한 스페로우의 주장을 보여준다.

표 2. LAWS 독단 전쟁범죄에 대한 스페로우의 주장
 Table 2. Sparrow's arguments regarding LAWS' full autonomous war crimes

	스페로우의 주장
프로그래머	“시스템 자율성이 높아질수록 프로그래머가 예측하거나 의도하지 않은 선택을 할 수 있는 능력이 더 커진다... 시스템이 자율성을 갖게 된 경우, 프로그래머에게 창작물의 행동에 대한 책임을 묻는 것은 부모가 자신의 보호를 떠난 자녀의 행동에 대한 책임을 묻는 것과 비슷할 것이다”[4, p. 70].
지휘관	“따라서 자율무기를 사용하면 군인은 자신이 통제할 수 없는 기계의 행동에 대해 책임을 져야 할 위험이 있다... 기계가 정말 스스로 표적을 선택한다면, 그 결과 발생하는 살상에 대해 지휘관에게 책임을 물을 수 없다”[4, p. 71].
무기체계	“기계의 행동에 도덕적 책임을 묻기 위해서는 기계에 대한 처벌이나 보상이 가능해야 한다... [그리고] 기계가 고통을 받을 수 있어야 한다... [또한] 충분한 내부 복잡성과 표현 능력을 가지고 있어야 한다... 역설적이게도, 그러한 기계의 창조는 애초에 개발 동기를 무색하게 한다... 기계는 우리의 전우가 될 것이고, 우리는 기계가 파괴될 때 도덕적 슬픔을 느끼게 될 것이기 때문이다”[4, pp. 72-73].

스페로우는 책임간극을 논함에 있어서 ‘완전자율’ 단계의 무기체계를 그 논의의 대상으로 하고 있다. ‘반자율’ 및 ‘인간감독’ 단계의 무기체계는 현재 이미 전력화 되어 있으며 기존의 재래식 무기체계에 비교하여 별 다른 도덕적 귀속책임 논란이 없기 때문이다. 앞서 살펴본 바와 같이 ‘완전자율’ 무기체계가 책임간극의 우려대상이 되는 이유는 높은 자율성에서 오는 목표물의 ‘선택’

이 기계에게 달려있다는 점이다. 인간 운용자의 역할이 배제된 목표물 선정과 공격이 쟁점이라고 할 수 있다. 하지만 스페로우의 주장대로 LAWS가 저지른 전쟁범죄에 대해서는 누구에게도 책임을 물을 수 없는 것일까?

III. 책임간극의 극복 가능성

1. 프로그래머와 지휘관

프로그래머가 개발한 프로그램과 개발자 사이의 책임 귀속여부는 비단 무기체계에만 국한된 것이 아니다. 프로그래머가 개발한 의도와 다르게 오류를 일으키거나 사용자의 오사용으로 문제를 야기하기도 하고 개발자가 예측할 수 없는 피해를 입히기도 한다. 더욱이 스페로우의 주장대로 프로그램이 인간을 배제한 형태로 모든 의사결정을 자율적으로 하는 경우라면 더욱 프로그래머에게 책임을 물을 수 없다.

다만 스페로우 본인의 주장대로 “프로그래머에게 창작물의 행동에 대한 책임을 묻는 것은 부모가 자신의 보호를 떠난 자녀의 행동에 대한 책임을 묻는 것과 비슷할 것”이다. 이 의미는 프로그래머와 무기체계의 관계는 부모와 성년이 된 자녀 관계와 비슷하다는 뜻이다. 스페로우는 법적 책임이라는 좁은 의미의 책임에서 생각하고 있지만 도덕적 책임이라는 틀에서 보자면 부모는 자녀의 행동에 대한 도덕적 책임에서 자유로울 수 없다. 이것은 자녀의 선행과 악행을 모두 포함한다. 예를 들어 성년이 된 자녀가 큰 업적을 이루거나 극악무도한 행동을 하였을 때 부모가 함께 축하를 받거나 책임감을 느끼는 것에 비유할 수 있다.

부모와 자녀의 관계처럼 프로그래머와 LAWS의 관계는 철학자 트리스턴 괴츠(Trystan Goetze)의 ‘도덕적 얽힘(moral entanglement)’으로 설명할 수 있다[12]. 이 관계는 자신의 에이전시(agency)가 다른 사람이나 대상의 행동에 어느 정도 관여하고 있는지 불확실하거나 모호할 경우에 발생하는데, 대표적인 경우가 부모와 자식 또는 반려동물과 보호자의 관계이다. 부모나 보호자의 과거 양육행동이 자식이나 반려동물의 행동에 크게 영향을 미치기 때문에 아무리 자녀 및 반려동물이 부모로부터 독립되어 자율적으로 행동하고 있더라도 부모의 영향을 받는다고 할 수 있는 것이다. 이렇듯 자신의 에이전시와 상대의 에이전시 사이의 불확실성에서 초래된 도덕적 책임이 ‘도덕적 얽힘’의 관계인 것이다.

이런 대리(vicarious) 책임감은 군대의 지휘책임 개념과 합쳐질 때 그 의미가 더 강해진다. 물론 스페로우의 주장대로 LAWS의 탐지, 결심, 행동의 일련의 과정에 지휘관은 개입의 여지가 없다. 하지만 지휘관으로서 무기체계 사용 결정에 대한 책임에서 자유로울 수 없다. 예시를 통하여 본다면 더 쉬운 이해가 가능하다. 2004년 이라크 아부그라이브 교도소에서 세상을 충격에 빠뜨린 가혹행위 사건이 있었다. 당시 교도소에서 근무하던 육군 및 CIA 소속 구성원은 재소자를 고문하고 강간하였으며 심지어 살인까지 저질렀다. 직접적인 범죄를 저지른 것은 일선 근무자들이었지만 당시 교도소장을 포함한 상급 지휘관도 법적, 도덕적 책임을 졌다. 이것은 직접 고문 및 살인 행위에 가담하지 않았더라도 그런 행위가 일어나지 않도록 사전에 교육 및 감독을 하고 적극적으로 통제하지 않은 것은 물론 예방활동을 소홀히 한 태만에 대한 책임을 지는 것이다.

군대라는 특수한 집단에서는 무력사용이 항상 전제되어 있기 때문에 더욱 엄한 잣대로 지휘책임을 부여한다. 비슷한 의미에서 철학자 제프 맥마한(Jeff McMahan)은 군인이 된다는 것은 신체적 위협 뿐만 아니라 도덕적 위협을 항상 내재하고 있다고 말했다[13]. 따라서 지휘관은 부하에 대한 책임에서 쉽게 자유로울 수 없는 것이다. 인간 부하의 행동에 대한 지휘책임을 지는 지휘관이 자율무기체계라고 해서 책임에서 벗어나는 것은 아니다. 인간 부하가 그렇듯 자율무기체계도 어느 정도의 예측 불가능성과 불량의 가능성이 상존한다. 그럼에도 불구하고 지휘관은 무기체계의 사용을 결심한 것이다. 지휘책임은 상하관계 뿐만 아니라 체계 사용 결심에 대한 책임이기도 하다.

2. 무기체계

프로그래머, 지휘관, 무기체계의 세 주체 중에서 스페로우는 전쟁범죄를 저지른 직접행위자인 무기체계에 가장 큰 도덕적 책임이 있다고 인식한다. 다만 인간 행위자에 준하는 처벌을 할 수 없기 때문에 무기체계에 책임을 귀속할 수 없다고 주장하는 것이다. 그는 처벌을 위해서는 괴로움(suffering)을 느낄 수 있다는 전제조건이 필요한데 기계는 내적 복잡성과 감정표현 능력이 인간만큼 없기 때문에 처벌 대상이 될 수 없으며, 기계가 인간 수준에 도달할 때에만 충분한 처벌을 할 수 있다고 설명하였다. 다만 로봇이 인간 수준에 도달

할 경우, 그래서 인격을 갖출 경우, 인간을 대체한 그 초기 목적을 상실하는 것이기 때문에 자기모순에 빠지는 것이다.

여기서 스페로우가 인간에게 적용 가능한 처벌방식을 로봇에 적용할 수 있어야 함을 전제하고 있으며 형벌적 책임만이 국제인도법 및 전쟁윤리에 부합한다고 가정하고 있다. 그는 책임과 처벌을 동일시하는 동시에 처벌과 괴로움을 동일시하고 있다. 처벌을 한다는 것은 억제(deterrence), 징벌(retribution), 갱생(rehabilitation), 무력화(incapacitation)의 네 가지 기능을 가진다[14]. 이 중 갱생과 무력화는 LAWS에도 충분히 적용 가능하다. 재프로그램 및 프로그램 업데이트를 통하여 갱생이 가능하고, 그럼에도 오류와 부작용이 계속 된다면 프로그램을 폐기함으로써 무력화가 가능하기 때문이다.

반면 억제와 징벌은 서로 긴밀히 연결되어 있는 동시에 LAWS에 쉽게 적용할 수 없는 것으로 보인다. 하지만 억제는 프로그래머 및 지휘관의 책임 범위 내에서 본다면 충분히 달성 가능한 특성이라고 할 수 있다. 프로그래머 및 지휘관에게 책임의 주체성을 부여할 경우 그들은 더 면밀하고 신중하게 무기체계를 개발 및 운용할 것이다. 의도된 목적에서 벗어난 행동을 하는 자율 무기체계는 적극적인 개입을 받을 것이며 갱생 불가능한 경우 폐기될 것이다. 그렇게 된다면 억제는 충분히 달성할 수 있는 목표가 된다.

그럼에도 징벌은 여전히 달성할 수 없는 기능으로 남는다. 하지만 면밀히 따져본다면 징벌은 처벌에서 가장 중요한 기능이 아니다. 오히려 억제효과 달성을 위한 장치로서 기능을 하는데 억제가 징벌 없이 가능하다면 징벌은 꼭 필요한 기능이 아니라고 할 수 있다. 어떤 의미에서는 징벌이 없는 처벌이 가장 국제인도법에 부합한다고 할 수도 있다. 억제효과를 달성할 수 있는 다른 방법이 있다면 징벌의 목적은 이미 달성되었다고 볼 수 있기 때문이다. 이렇듯 스페로우가 지적한 징벌적 책임을 지우지 않더라도 무기체계의 책임간극은 극복 가능하다고 할 수 있다.

IV. 결 론

본 연구에서는 치명적 자율무기체계의 사용 양상 및 구분을 살펴보고 로버트 스페로우의 '책임간극' 이론을 바탕으로 자율무기체계의 전투참여와 그 책임귀속에

대하여 고찰하였다. 이를 통해 스페로우가 책임의 주체라고 했던 프로그래머, 지휘관, 무기체계 모두에게 책임을 귀속할 수 있으며 책임간극이 극복될 수 있음을 밝혔다. 전쟁의 범위가 확장되고 그 속도가 빨라짐에 따라 인공지능에 기반한 자율무기체계 도입은 더 이상 피할 수 없는 현실이다. 많은 국가가 자율무기체계를 이미 도입하였으며 LAWS의 도입도 고려하고 있다. 이런 상황에서 '책임간극' 틀로만 사고하는 것은 발전적이지 못하다. 오히려 간극보다 책임을 논해야 더욱 신중하고 보수적인 사용이 가능하기 때문이다. 그런 의미에서 이 글에서는 '책임간극'의 극복가능성을 LAWS를 접근하는 새로운 사고의 틀로 제안하고자 한다.

References

- [1] P. Scharre, *Army of None: Autonomous Weapons and the Future of War*. W. W. Norton & Company, 2018.
- [2] M. C. Horowitz, "When speed kills: Lethal autonomous weapon systems, deterrence and stability," *J. Strateg. Stud.*, vol. 42, no. 6, pp. 764-788, 2019, doi: 10.1080/01402390.2019.1621174.
- [3] D. Amoroso and G. Tamburrini, "Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues," *Curr. Robot. Rep.*, vol. 1, no. 4, pp. 187-194, 2020, doi: 10.1007/s43154-020-00024-3.
- [4] R. SPARROW, "Killer Robots," *J. Appl. Philos.*, vol. 24, no. 1, pp. 62 - 77, 2007, doi: 10.1111/j.1468-5930.2007.00346.x.
- [5] "DoD Directive 3000.09 Autonomy in Weapon Systems." Department of Defense, Jan. 25, 2023. [Online]. Available: <https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf>
- [6] "ICRC Position on Autonomous Weapon Systems." ICRC, May 12, 2021.
- [7] Human Rights Watch, "Losing Humanity: The Case Against Killer Robots." Human Rights Watch, Nov. 19, 2012.
- [8] S. Kim and Y. Yoon, "A Study of Cognitive Warfare and Ethical Issues: Focusing on Just War Theory," *JCCT*, vol. 9, no. 3, pp. 451-459, 2023, doi:10.17703/JCCT.2019.5.4.413
- [9] P. Asaro, "On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making," *Int. Rev. Red Cross* 2005, vol. 94, no. 886, pp. 687-

- 709, 2012, doi: 10.1017/S1816383112000768.
- [10]E. Rosert and F. Sauer, “How (not) to stop the killer robots: A comparative analysis of humanitarian disarmament campaign strategies,” *Contemp. Secur. Policy*, vol. 42, no. 1, pp. 4 - 29, 2021, doi: 10.1080/13523260.2020.1771508.
- [11]A. Matthias, “The responsibility gap: Ascribing responsibility for the actions of learning automata,” *Ethics Inf. Technol.*, vol. 6, no. 3, pp. 175 - 183, 2004, doi: 10.1007/s10676-004-3422-1.
- [12]T. S. Goetze, “Moral Entanglement: Taking Responsibility and Vicarious Responsibility,” *The Monist*, vol. 104, no. 2, pp. 210 - 223, 2021, doi: 10.1093/monist/onaa033.
- [13]J. McMahan, “The Moral Responsibility of Volunteer Soldiers,” *Boston Review*. Accessed: May 27, 2023. [Online]. Available: <https://www.bostonreview.net/forum/moral-wounds-ethics-volunteer-military-service/>
- [14]N. V. Demleitner, “Types of Punishment,” in *The Oxford Handbook of Criminal Law*, in *Oxford Handbooks in Law*. Oxford University Press, 2014. doi: 10.1093/oxfordhb/9780199673599.013.0041