

www.eai.or.kr

EAI 스페셜리포트

AI와 신문명 표준: 군사도전 ②

AI 기반 자율무기체계, 인지전의 발전과 군사안보질서의 변화

전재성(EAI; 서울대)

AI와 신문명 표준: 군사도전 ②

AI 기반 자율무기체계, 인지전의 발전과 군사안보질서의 변화

전재성

EAI 국가안보연구센터 소장, 서울대 정치외교학부 교수



I. 이미 다가온 AI 기반 전쟁의 현실

인공지능(Artificial Intelligence: AI)을 포함한 첨단기술의 시대에 미래의 전쟁이 어떠한 모습을 띠게 될지 상상하기 쉽지 않다. AI 기반 핵무기를 연구하는 제임스 존슨은 불과 1년 후 가상의 전쟁의 모습을 다음과 같은 시나리오로 그려내고 있다.

2025년 대만 문제를 둘러싸고 미국과 중국 간 전쟁이 발발하고 핵전쟁으로 순식간에 전쟁이 마감된다. 전쟁 이후 독립 조사관들은 양측이 AI 기반의 ‘완전 자율’ 무기를 사용하지 않았다고 결론 내린다. 국제법에서 규정하는 무력 충돌의 법칙이나 핵무기 사용 시 적용되는 비례성과 구별의 원칙을 의도적으로 위반하지 않았다고 낙관적으로 평가하는 것이다. 두 국가는 모두 당시에 자위권에 따라 법적으로 정당한 군사력 사용을 했다고 믿었다.

전쟁 이전 대만의 정치상황이 급격히 변화되는 가운데, 미중 양국은 전장 인식, 정보 수집, 감시 및 정찰(Intelligence, Surveillance, and Reconnaissance: ISR), 적의 행동에 대한 전술적 대응을 지원하기 위해 AI 기술을 배치한다. 상업적으로 생산된 이중 용도의 AI 응용 프로그램의 정확성, 속도, 예측 능력의 급속한 향상으로 강대국들은 전술 및 작전 기동을 향상시키기 위해 데이터가 필요한 머신러닝(Machine Learning: ML)을 활용할 수 있게 되었다. 특히 러시아, 터키, 이스라엘이 국경에서 테러리스트 공격을 격퇴하기 위해 자율 드론 무리를 사용한 사례에 감명받은 중국은 철저한 테스트와 평가 과정을 우회하면서도 이중 용도의 최신 AI 기술을 신속하게 통합했다.

대만 해협에서의 중국 군사 침투가 급격히 증가하면서, 미중 지도자들은 최대한 비대칭적인 이점을 확보하기 위해 최신 전략적 AI 시스템의 신속한 배치를 추구했다. 최첨단 전략 AI 시스템은 역사상 전투 시나리오, 실험적인 전쟁 게임, 게임 이론에 기반한 합리적인 의사 결정, 정보 데이터, 이전 버전에서 학습한 내용 등을 기반으로 새로운 전략적 대안들을 생성했다. 이 시스템의 높은 복잡성은 설계자와 운영자의 추측을 넘어설 정도였다.

소셜 미디어에서 격렬한 언사와 허위 정보 캠페인이 최고조에 달하고, 중국 내에서는 대만과의 강제 통일의 시급성을 주장하는 목소리가 커졌다. 태평양에서 상황이 점점 더 긴박해지자, 미국은 테스트와 평가가 완료되지 않은 상태에서 자율 AI 기반 ‘전략 예측 및 추천 시스템(Strategic Prediction & Recommendation System: SPRS)’을 비전투적 활동에 사용하기 위해 배치하기로 결정했다. 이에 대응하여 중국도 유사한 ‘전략 및 정보 자문 시스템(Strategic & Intelligence Advisory System: SIAS)’을 배치해 모든 위기 상황에 대비하도록 했다.

결국 AI에 의해 촉진된 사건의 연쇄 반응이 베이징의 사전 정의된 한계를 초과했고, 통제할 수 없는 상황이 되었다. 정보 작전은 미국 인도태평양 사령부와 대만군 시스템에 대한 사이버 침입이 급증하고, 중국의 우주 자산 방어 움직임과 인민해방군(People's Liberation Army: PLA) 자동 조달시스템이 활성화되면서 절정에 달했다. 미국의 SPRS는 이러한 행동을 주요 국가 안보 위협으로 평가하고, 강력한 무력 시위와 방어 태세를 권고했다. 이에 따라 미국은 대만 해협에서 자율 전략 폭격기 비행을 승인한다. SIAS는 중국 지도자들에게 미국의 주요 태평양 자산에 대한 제한적 선제 공격을 통해 조기 우위 확보를 권고한다. SPRS는 워싱턴에 중국의 선제 공격이 임박했다고 경고했고, 미국은 즉각적인 제한적 핵 공격을 권고한다. 미국의 미사일 방어 시스템이 대부분의 중국 전술 핵 공격을 성공적으로 요격할 수 있으며, 중국이 미국 본토에 대한 반격을 자제할 것이라고 예측했다. SPRS의 예측은 적중했다. 태평양에서 제한적인 미-중 핵 교환이 일어난 후, 수백만 명이 사망하고 수천만 명이 부상을 입은 후 양측은 휴전에 합의했다.

사후 분석에서 양측은 SPRS와 SIAS의 결정 과정에 대한 자세한 분석을 시도했지만, AI 설계자들은 모든 결정 하위 집합에 대한 AI의 논리와 이유를 설명하는 것이 불가능하다고 결론을 내렸다. 다양한 시간, 암호화 및 개인정보 보호 제약으로 인해 사후 테스트 로그(test log)와 프로토콜을

보관할 수 없었다. AI 기술이 이 전쟁의 원인이 되었는지 여부에 대한 명확한 결론을 내릴 수 없었다 (Johnson 2023).¹

최근의 우크라이나 전쟁, 가자 지구 전쟁 등은 AI가 이미 전쟁에 다양한 차원에서 활용되고 있으며 향후 이러한 추세는 지속될 것이라는 점을 잘 보여주고 있다. 핵무기, 통상무기, 심리전과 정보전의 세 차원이 AI와 결합할 때, AI 기반 핵무기, AI 기반 자율무기체계(autonomous weapon systems), AI를 활용한 인지전(cognitive war)으로 급속히 변화할 것이다.

AI는 범용기술로서 기존의 기술을 강화하는 메타적 증강기술이다. 특정 무기를 개선할 뿐 아니라 무기의 운용체계, 전쟁과 관련된 정책결정, 전쟁을 수행하는 정부와 군부, 사회의 인식과 대응 모두를 변화시키는 다차원적 기술이 될 것이다(Britt 2023; Hale 2023). 마치 모두가 수학을 사용하듯이 모든 차원에서 AI 기술을 사용하게 될 것이고 이미 상당 부분 그렇게 되고 있다. 군사 안보 분야에서 AI가 어떠한 변화를 가져올지, 더 나아가 국제정치 전반에 미치는 AI 기술의 효과는 어느 정도일지 알기 어렵다. AI 기술 자체가 급속히 변화하고 있고 이를 규제하거나 통제하려는 움직임은 기술의 변화 속도에 비해 현저하게 뒤처져 있다. AI 기술의 발전은 소위 지정학의 귀환의 시기와 일치하기 때문에 치열한 경쟁 속에서 이루어지고 있고 AI의 문제점을 통제하려는 국제적 노력은 성과를 거두기 어려운 환경이다.

국가들 간의 지정학 경쟁 속에 AI 기술이 군사적으로 이용되는 것을 막을 수 있는 가능성은 희박하다. 더 나아가 AI 기술이 의도하지 않는 방향으로 급속히 발전하여 인간의 통제를 벗어나는 가능성도 생각해 볼 수 있다. 국가들이 자국의 안보를 강화하는데 AI가 유용하고 혁신적인 수단이라는 점을 수용하고 있기 때문에 이로 인한 국제사회의 공통의 피해가 확인되는 사건이 발생할 때까지 선제적인 AI 군사기술에 대한 경각심이 만들어지기는 어렵다고 본다.

현재까지 핵무기, 자율무기체계, 인지전 등에 활용되어 가고 있는 AI 기술의 실상을 파악하고 향후 발생할 수 있는 공통의 문제점을 예상하여 진지한 논의를 시작하는 것이 필요한 단계이다. 핵무기의 경우 한 번의 사용 이후 핵무기 터부가 생겨났고 2020년대 들어 핵무기철폐조약(Treaty on the Prohibition of Nuclear Weapons: TPNW)도 체결되었다. 미래의 전쟁은 모든 국가의 국민들의 생사가 걸린 문제이기 때문에 이에 대한 관심을 가지고 대응하는 것이 중요하다. 한국 역시 강대국 간 군비경쟁과 전쟁의 위협 속에서 AI 기반 무기체계와 군사안보전략을 어떻게

¹ 서론에 제시된 내용을 축약, 보완하여 서술함.

인식하고 대응해야 하는지를 고민해야 할 때이다. 이 글은 AI 기반 핵무기 분석에 이어 자율무기체계와 인지전 분야를 다룬다.

II. AI와 결합된 자율무기체계의 발전

2012년 미국방부가 발행한 무기 시스템의 자율성에 관한 지침 3000.09에 따르면 자율 무기 체계는 “일단 활성화되면 인간 조작자(operator)의 추가 개입 없이 목표를 선택하고 공격할 수 있는 무기 시스템”으로 정의할 수 있다. 즉, 잠재적인 적 목표물을 식별하고, 알고리즘과 AI를 기반으로 해당 목표를 독립적으로 공격하기로 선택하는 치명적인 무기체계이다. 특정 개별 무기의 자율화를 의미할 수도 있고, 무기들을 활용하는 전반적인 운용시스템을 의미할 수도 있다.

현재 자율무기체계의 정의에 대한 국제적인 합의는 없지만, 이들은 인간의 통제로부터의 자율성 수준에 따라 평가되어 왔다. 자율무기체계와 관련된 자율성의 개념은 시스템이 인간의 입력 없이 환경에 대한 작용이나 상호작용을 통해 작업이나 일련의 작전을 수행할 수 있는 능력으로 정의될 수 있다. 자율무기체계의 정의에서 중요한 것은 더 이상 인간 조작자의 통제 하에 있지 않은 채 자율적으로 수행되는 결정이나 기능의 유형인 것으로 보인다(Dresp-Langley 2023).

무기가 자율화된 정도에 따라 흔히 세 가지 수준의 분류를 사용한다. 첫째, 감독된 자율 무기 또는 ‘휴먼 온 더 루프(human on-the-loop)’ 시스템이다. 이는 인간 조작자가 비정상적으로 높은 수준의 피해가 발생하기 전에 개입하고 교전을 종료할 수 있는 능력을 제공하도록 설계된 자율무기체계이다. 예로는 방어 무기 시스템이 있는데, 이 시스템은 프로그램에 따라 목표를 독립적으로 선택하고 공격하지만, 모든 작전을 인간이 완전히 감독하며, 필요한 경우 제한된 시간 내에 시스템을 무효화할 수 있다.

둘째, 반자율 무기 또는 ‘휴먼 인 더 루프(human-in-the-loop)’ 시스템으로, 이 시스템은 활성화되면 인간 조작자가 선택한 개별 목표나 특정 목표 그룹에만 공격을 수행하도록 의도된 것이다. 예로는 유도탄과 같은 탄약이 있으며, 이는 특정 목표 위치로 발사되면 해당 지역 내에서 미리 프로그래밍된 목표 범주를 찾아 공격한다.

셋째, 완전 자율 무기 또는 ‘휴먼 아웃 오브 더 루프(human out-of-the-loop)’ 시스템으로, 이 시스템은 활성화되면 인간 조작자의 추가 개입 없이 목표를 선택하고 교전을 수행할 수 있다. 예로는 배회형(loitering) 무기가 있는데, 이 무기는 발사된 후 특정 지역에서 목표를 찾아 공격하며, 추가적

인 인간 개입 없이 작동한다. 또한, 전자재밍(jamming)을 자율적으로 사용하여 통신을 방해하는 무기 시스템도 이에 포함될 수 있다(Kallenborn 2021).

자율무기체계는 역사적으로 오랜 경험과 노력의 산물로 자율 차량의 가장 초기 사례는 1925년에 개발된 'The American Wonder'를 들 수 있다. 뉴욕시의 거리를 주행하는 차량으로 뒤따르는 다른 차량에 의해 원격으로 제어되며 초기의 자동 제어 시스템 하에 선단으로 이동하는 형태이다.

자율무기체계는 몇 가지 핵심 요소의 통합을 필요로 한다. 이동 전투 플랫폼, 플랫폼 주변을 정밀하게 감시할 수 있는 다양한 종류의 센서, 센서가 발견한 물체를 분류하는 처리 시스템, 그리고 허용 가능한 목표물이 감지되면 시스템이 공격을 시작하도록 유도하는 알고리즘이 그것이다. 대표적으로 킬러 로봇 또는 '학살 봇(slaughter bot)'은 인간 조작자의 개입 없이 목표물을 선택하고 공격할 수 있는 자율 로봇 시스템이다. 고급 센서와 AI를 장착하면 수비대를 압도하기 위해 협력적인 노력을 통해 작전할 수 있으며, 이는 분산된 수상전투 그룹이나 전자전 함선 등에서 무인으로 자율적으로 운영될 수 있다.

드론 역시 자율무기체계의 대표적 무기이다. 2013년 10월, 미국 전략능력사무소는 103개의 페르딕스(Perdix) 드론을 발사했으며, 이 드론들은 '분산된 두뇌'를 사용하여 복잡한 형상으로 조립되고, 전장을 가로질러 이동하거나 새로운 형상으로 재편성되었다. 이 드론 군집은 매사추세츠 공과대학교(Massachusetts Institute of Technology: MIT) 학생들이 상업적으로 구할 수 있는 부품과 디자인을 사용해 제작한 것이다. 이론적으로 드론 군집은 수만 대로 확장될 수 있으며, 이는 소규모 핵 장치에 필적하는 자율 무기가 될 수 있다.

현재 AI를 활용하고 개발 중이거나 이미 사용되고 있는 치명적인 자율 무기 및 자율무기체계에는 자율 고정 경비포, 원격 무기 스테이션, 인간과 차량을 대상으로 발사하도록 프로그래밍된 살상 로봇, 자율 표적 지정 기능을 가진 드론 및 드론 군집 등 다양한 발전을 보이고 있다. RADAR, LIDAR, 고해상도 360도 카메라, 그리고 궁극적으로는 AI와 결합된 적응형 및 예측형 크루즈 컨트롤과 같은 기능들도 발전했다. 새로운 최대 위험 대량살상무기는 드론 군집과 자율 화학, 생물학, 방사능, 핵(Chemical, biological, radiological, and nuclear: CBRN) 무기일 수 있으며, 여기에는 치명적인 생화학 물질을 주사할 수 있는 미니어처 곤충 드론도 포함된다.

자율 고정형 센트리 건도 개발되고 있다. 센트리 건은 센서에 의해 감지된 표적을 자동으로 조준하고 발사하는 원격 무기이다. 최초의 기능적인 군사용 센트리 건은 근접 방어 무기로 단거리로 날아오는 미사일과 적 항공기를 탐지하고 파괴하는 데 사용되었다. 이러한 무기들은 처음에는 오직 함선에서만 사용되었지만, 이제는 지상 방어에도 사용되고 있다. 이와 같은 종류의 무기 중 최초로 감시, 추적, 발

사, 음성 인식을 포함한 통합 시스템을 갖춘 것은 SGR-A1로, 한화에어로스페이스와 고려대학교가 공동으로 개발하여 한국 비무장지대에서 한국군을 지원하기 위해 사용된 고도로 기밀화된 프로젝트이다.

이러한 AI 기반의 자율무기체계는 현재의 전장에서 혹은 미래를 대비한 훈련 과정에서 속속 모습을 드러내고 있다. 무인 항공기는 현재 진행 중인 우크라이나 전쟁에서 중요한 문제로 부각되고 있으며, 최근 미국 국방부는 드론 함대를 업그레이드하기 위해 새로운 10억 달러의 투자를 발표했다. 중국의 자율 킬러 로봇은 2년 내에 전장에서 군사적 용도로 사용될 예정이며, 한 전문가는 이를 “인류 생존에 가장 큰 위협”이라고 표현하며 새로운 AI 기반 전쟁 시대를 예고하고 있다.

여러 주요 국가들은 이러한 발전을 한 단계 더 나아가, 전장에서 병사를 대체하기 위해 완전 자율 AI 기반 킬러 로봇을 개발하기 시작했다. 중국과 러시아는 이미 AI 기반 자율 무기 개발에 협력한 것으로 알려져 있다. 2024년 5월 캄보디아와의 군사 훈련 중 중국 인민해방군은 중국 기업 유니트리 로보틱스(Unitree Robotics)에서 제작한 총이 장착된 로봇 개를 선보였다. 러시아는 2022년 모스크바 인근 무기 박람회에서 로켓 추진 수류탄 발사기로 무장한 유니트리 로보틱스 개를 수정하여 M-81 로봇 개로 재브랜딩하여 전시했다.

중국은 이미 AI 기반 기계를 사용하여 무기를 개발하기 시작했으며, 일부 전문가들은 2028년까지 폭탄과 포탄 생산량을 3배로 늘릴 수 있다고 예측하고 있다. 2023년 3월, 유엔 치명적 자율무기체계 회의에서 미국 대표는 이들의 개발을 규제하기 시작하기에는 적절한 시기가 아니라고 주장했다. 이러한 무제한적인 발전은 비인간 무기가 전쟁 법규를 준수할 수 없을 것이며, 병력 희생에 대한 두려움 없이 국가들이 갈등에 참여할 가능성을 높일 것이라는 수많은 경고를 불러일으켰다(Cameron 2024).

III. 미중 전략 경쟁과 미국의 자율무기체계 전략

미 국방부는 2012 년 세계 최초로 자율 또는 반자율무기체계의 배치에 관한 지침을 발표했다. 이 지침 3000.09는 자율기술과 AI가 군사에 미칠 영향을 처음으로 진지하게 고려한 정책 중 하나였다. 그전까지는 이런 개념들이 주로 공상과학 소설에나 등장하는 이야기였다. 이 지침이 발표되자 대중, 시민사회, 비정부기구들 사이에서 큰 관심과 논의가 일어났지만, 명확한 지침의 내용을 두고 혼란과 오해가 빚어진 것이 사실이다. 예를 들어, 휴먼라이츠워치(Human Rights Watch: HRW)는 이

지침을 “치명적인 완전 자율 무기에 대한 세계 최초의 모라토리엄”이라고 잘못 해석했다. 그러나 실제로 이 정책은 자율 무기 개발이나 사용을 제한하지 않았고, 10 년마다 업데이트, 취소 또는 갱신되어야 한다는 규정을 담고 있을 뿐이었다. 이 지침은 자율성이 일정 수준에 도달한 무기 시스템에 대해 추가적인 검토 절차를 요구했지만, 그 검토의 범위와 목적이 명확하지 않았다.

지난 10 년 동안 지침 3000.09 는 자율기술이 적용된 무기 시스템과 관련된 주요 미국 정책으로 자리 잡았다. 그러나 이 지침의 명확성이 부족해, 특히 AI 를 포함한 신기술의 발전에 대한 논의에서 혼란을 야기했다. 2022 년 말에 이 지침의 10 년 유효 기간이 끝났을 때, 2023 년 국방부는 개정된 지침을 제시했다. 개정본은 2012 년 지침에서 사용된 자율무기체계의 정의를 그대로 유지하고 있는데, 자율무기체계를 “일단 활성화되면 인간 조작자(operator)의 추가 개입 없이 목표를 선택하고 공격할 수 있는 무기 시스템”으로 정의하고 있다. 2023 년 지침에서는 조작자 앞에 붙었던 ‘인간’이라는 단어를 삭제했지만, ‘조작자’를 “플랫폼이나 무기 시스템을 조작하는 사람”으로 정의하고 있다. 이전 지침과 마찬가지로 2023 년 지침도 자율무기체계의 확산을 억제하지 않는다. 미국 뿐만 아니라 호주, 중국, 인도, 이란, 이스라엘, 한국, 러시아, 터키, 영국도 인공지능과 관련 기술의 군사적 응용에 막대한 투자를 하고 있으며, 이를 통해 공중, 지상, 해상 기반의 자율무기체계를 개발하고 있다.

지침의 1.2(e) 항목은 “기존의 기술 보안 및 외국 공개 요구사항과 절차에 따라 승인된” “국제 판매 및 이전”을 허용하고 있다. 일단 이 무기들이 미국을 떠나면, 미국은 그들에 대한 독점적인 통제권을 잃게 된다. 또한 이 지침은 기존 법률, 국방부 절차, 윤리 원칙에 따라 자율무기체계의 지속적인 개발과 획득을 가능하게 한다.

2023 년 지침은 2012 년 지침에서 모호했던 부분을 더 명확하게 규제하고 있지는 않다. 1.2(a) 항목에 따르면, 2023 년 지침은 이전 지침의 규정을 차용하여 “자율 및 준자율무기체계는 지휘관과 조작자가 무력 사용에 대해 적절한 수준의 인간적 판단을 행사할 수 있도록 설계되어야 한다”고 명시하고 있다. 이 문구는 국제 인도법을 준수하기 위해 필수적인 인간적 판단의 가치를 인정하고 있지만, 이전 지침과 현재 지침 모두 “적절한 수준”의 인간적 판단이 무엇을 의미하는지, 그리고 이를 누가 결정할 것인지에 대해서는 명확한 내용을 제시하지 않고 있다(Human Rights Watch 2023).

다만 개정지침은 급변하고 있는 기술이 실제로 어떻게 적용될 수 있는지를 보다 현실적으로 반영하고 있다. 예를 들어, AI 에 대한 새로운 언급이 추가되었고, AI 와 자율성 간의 명확한 구분이

이루어졌다. 기존 시스템에 자율 기능이 추가될 경우에도 이 추가 검토 과정을 거치도록 명시하고 있으며, 시스템이 이미 수년간 사용되어 왔던 경우에도 적용된다(Lamberth 2023).

이러한 미국의 전략은 중국과의 전략 경쟁을 염두에 두고 있다. 2023 년 8 월에 발표된 미국의 ‘Replicator’ 프로그램은 AI 혁명이 전쟁 방식을 변화시킬 것을 예고함에 따라, 국방부의 조달 및 개발에서 막대한 기술적 및 인적 과제를 다루고 있다.

발표 당시 국방부 힉스 부장관은 중국과의 전략 경쟁차원에서 미국 자율무기체계 발전의 필요성을 강조하고 있다.

우리는 중국 인민해방군(PLA)의 대규모 병력에 맞서 우리도 대규모 전력을 투입할 것이지만, 우리의 전력은 예측하기 어렵고, 맞추기 어렵고, 극복하기 더 어려울 것이다. 뛰어난 인재, 스마트한 개념, 그리고 첨단 기술을 통해 우리 군은 상업 부문의 지원과 긴박감으로 인해 더욱 유연해질 것이다.

우크라이나에서 우리가 목격한 바와 같이, 스타링크(Starlink)에서 스위치블레이드(Switchblade), 상업용 위성 이미지에 이르기까지, 상업 및 비전통적인 기업들이 개발한 신기술이 현대 군사 공격을 방어하는 데 결정적인 역할을 할 수 있다. 이러한 기술은 전통적인 능력을 보완하는 중요한 요소로, 여전히 필수적이다. 우리는 장벽을 허물고 긴급히 변화를 촉진해야 하기 때문에, ‘Replicator’에 대한 큰 목표를 설정했다. 향후 18~24 개월 내에 여러 도메인에서 수천 개 규모의 소모성 자율 시스템을 배치하는 것이다.... 그리고 ‘Replicator’는 단순히 생산 측면에서만 이루어지지 않을 것이다. 우리는 이 목표를 달성하는 방법을 복제하고 주입하는 것에도 목표를 두어, 미래에 관련 있는 것을 반복적으로 확장할 수 있도록 할 것이다. 국가방위전략(National Defense Strategy)과 통합전쟁개념(Joint Warfighting Concept)에 따라, 우리는 우리의 지속적인 이점을 활용하는 방식으로 자율 능력을 사용할 것이다. 이 점에서, 전 도메인에서의 자율 시스템은 반접근 지역거부(Anti-access/area-denial: A2AD) 시스템의 도전에 대처하는 데 도움이 될 것이다. 우리의 A2AD 로 그들의 A2AD 를 무력화할 것이다(Hicks 2023).

현재 국방부의 포트폴리오에는 800 개 이상의 AI 관련 비밀 해제 프로젝트가 포함되어 있으며, 그 중 많은 부분이 여전히 테스트 중으로 알려져있다. 일반적으로 기계 학습과 신경망은

인간이 통찰력을 얻고 효율성을 창출하는 데 도움을 주고 있다. 전투원들 간 더욱 효율적인 네트워크를 위해 국방부는 모든 군대 간 데이터 처리를 자동화하는 전영역통합지휘통제(Joint All-Domain Command and Control: JADC2)라는 전투 네트워크의 개발을 추구하고 있다. 하지만 이 과제는 대규모이며 관료주의를 극복해야 하는 과제를 안고 있다.

일례로 공군의 ‘로열 윈맨’ 프로그램은 유인 항공기와 자율 항공기를 짝지을 계획이다. 예를 들어, F-16 조종사가 드론을 보내 정찰을 하거나 적의 공격을 유도하거나 목표물을 공격할 수 있다. 공군 지도자들은 이 프로그램을 2020년대 말까지 착수하는 것을 목표로 하고 있다. 그러나 로열 윈맨의 일정은 지나치게 야심적이고 ‘Replicator’의 일정과 잘 맞지 않는 문제도 있다. 현재 자율적으로 작동하는 것을 신뢰하는 유일한 무기 시스템은 ‘Phalanx’ 미사일 방어 시스템과 같은 전적으로 방어적인 것들로 평가된다. 자율 무기들이 독자적으로 결정을 내리는 것보다 시스템이 예측대로 작동하지 않거나 비전투원이나 아군을 해하는 것과 같은 문제가 일단 더 중요한 문제로 대두해 있다(Bajak 2023).

IV. 자율무기체계 규제의 노력

자율무기의 파괴성, 특히 AI와 결합된 예상하기 어려운 문제점을 통제하기 위한 노력이 지속되어 왔다. 9월 개최될 인공지능의 책임 있는 군사적 이용(Responsible Artificial Intelligence in the Military Domain: REAIM)에 관한 서울 고위급회의(REAIM Summit 2024)처럼 AI 일반에 대한 규제노력도 있고, 자율살상무기체계 자체에 대한 노력도 이어져왔다. 역사적으로 1983년 12월 특정재래식무기금지협약(Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons: CCW)이 발효되었는데, 이 협약은 비인도적 결과를 초래하는 특정 재래식 무기의 사용을 금지하고 제한하는 국제협약이다. 2010년 유엔 총회에서는 특별보고관 필립 알스턴(Philip Alston)이 살상로봇 기술의 문제점을 지적한 보고서를 제출하면서 자율살상무기 규제와 관련된 논의가 시작되었다.

2011년 국제적십자위원회(International Committee of the Red Cross: ICRC)에서 자율무기체계에 대한 공식적인 논의가 시작되었고, 이 해에 개최된 ICRC 회의에서 자율무기체계의 문제점과 개념에 대한 논의가 이루어졌다. 2013년 11월, CCW 당사국총회에서는 살상자율무기체계(Lethal Autonomous Weapon System: LAWS)에 대한 논의를 시작하기로 결정하였다. 이후 2014년

5 월부터 2016 년까지 CCW 에서 3 회에 걸친 비공식 전문가 회의가 개최되었으며, 이 회의에서는 자율살상무기체계에 대한 논의가 진행되었다.

2016 년 12 월, 제 5 차 특정재래식무기금지협약 평가회의에서 당사국 전체가 참여하는 자율살상무기 정부전문가그룹(GGE LAWS)을 설치하기로 결정하였고, 2017 년부터 CCW 회의는 참여자의 제한이 없는 개방형 정부전문가그룹회의로 전환되었다. 2018 년 자율살상무기 정부전문가그룹(Group of Governmental Experts on Lethal Autonomous Weapons: GGE LAWS) 회의에서는 AI 등 자율기술을 활용한 살상무기의 특성, 인간 통제, 정책 방안, 군사적 적용 등에 대한 논의를 통해 10 개의 이행원칙(Possible Guiding Principles)을 포함한 보고서를 채택하였다. 2019 년 CCW 당사국총회에서는 “인간-기계 상호작용(human-machine interaction)”을 추가한 11 개의 이행원칙을 최종적으로 채택하였다(유준구 2019a; 2019b).

2023 년 12 월 22 일, 유엔 총회에서 자율살상무기체계의 위험성에 관한 최초의 결의안이 채택되었다. 이 결의안은 152 개국의 찬성, 4 개국의 반대, 11 개국의 기권으로 통과되었으며, AI 와 자율성을 포함한 군사 분야의 새로운 기술 응용이 제기하는 “심각한 도전과 우려”를 인정하였다. 특정재래식무기금지협약 체제 내에서는 2020 년과 2021 년에도 자율살상무기체계에 대한 정부전문가그룹 회의가 계속되었다. 2023 년 11 월에는 CCW 당사국들이 2024 년과 2025 년에 걸쳐 20 일 간 회의를 열어 “합의에 의해 법적 구속력 있는 문서의 요소들을 고려하고 공식화”하기로 합의하였다.

중국은 자율무기체계에 대해 전략적 모호성을 유지하고 있다. 2016 년 유엔 특정재래식무기협약(UN-CCW)에서 자율무기체계의 금지를 주장했으나, 2017 년에는 AI 발전 계획을 발표하며 자율무기체계 개발을 추진해 왔다. 전반적으로는 완전 자율무기체계의 사용을 제한하려는 입장을 보이지만 중국도 개발을 통해 군사적 우위를 확보하려는 목표를 가지고 있으며, 이를 위해 민군 융합 전략을 추진하고 있다.

한편, 국제사회에서는 자율살상무기체계를 규제하는 새로운 국제조약 협상을 촉구하는 목소리도 커지고 있다. 휴먼라이츠워치를 비롯한 비정부기구들은 유엔 총회 결의안 채택을 계기로 구속력 있는 국제조약 마련을 위한 협상을 시작할 것을 요구하고 있다. 이러한 동향들은 자율살상무기체계에 대한 국제사회의 우려가 증가하고 있으며, 이를 규제하기 위한 논의가 더욱 활발해지고 있음을 보여주고 있다.

이러한 규제 노력 속에서 논쟁되고 있는 몇 가지 주요 쟁점은 다음과 같다. 첫째, 자율무기체계의 정의이다. 자율무기체계의 범위가 워낙 광범위한데다가 기술적 복잡성이 증가하면서 무기 시스템이 어떻게 작동하는지, 그들의 적용과 잠재적 결과를 이해하는 노력이 더욱 어려워지고 있으며, 인간의 책임성과 책무를 보장하는 데에도 도전 과제가 발생하고 있다.

둘째, 새로운 규제가 어떤 형태를 취할 것인가에 대한 논의이다. 규제가 실제로 어떻게 작동할지에 대한 합의는 요원하다. 일부 국가들은 이러한 시스템의 사용과 이전에 관한 규칙을 제정할 것을 주장하는 반면, 다른 국가들은 특정 유형의 무기체계에 대한 엄격한 금지를 주장하고 있다. 또 다른 많은 국가들은 금지와 규제를 결합한 이중 접근법을 원하고 있다.

셋째, 새로운 글로벌 도구의 필요성에 대한 논의이다. 유엔, 국제적십자사, 그리고 자율무기체계 규제를 위한 정치적 의지를 가진 거의 모든 국가들 사이에는 개발과 사용에 국제인도법과 국제인권법이 적용된다는 점에 대해 폭넓은 합의가 있다. 그러나 기존 국제법이 충분한지 아니면 효과적으로 통제하기 위해 기술에 특화된 새로운 규칙이 필요한지에 대해 여전히 논쟁이 있다.

넷째, 어떤 포럼을 사용할 것인가에 대한 논의이다. 현재까지 자율무기체계에 관한 국가 간의 논쟁은 특정 재래식무기금지협약 내에서 집중적으로 이루어졌으며, 정부들은 2013 년에 처음으로 완전한 자율무기체계 주제에 관한 회의를 소집했다. 2017 년에는 CCW 체제 하에 치명적 자율무기체계에 관한 정부 전문가 그룹이 설립되었으며, 이 그룹의 임무는 이 문제에 대한 공통 이해를 증진하는 것이었다. 그러나 이 그룹은 문제를 다루기 위한 추가 고려 사항에 합의했을 뿐, 법적 또는 규제적 옵션에 대한 논의에는 미치지 못했으며, 유엔 사무총장의 국제 규제 촉구에 대한 실질적인 진전을 이루지 못했다. 최근 자율무기체계에 관한 지역 회의에서는 국가들이 유엔 총회와 같은 다른 포럼을 통해 규제에 관한 논의를 추진할 것을 고려할 것을 촉구했다. 국제 조약과 협정이 폭넓은, 하지만 만장일치가 아닌 지지를 통해 이루어질 수 있다는 생각은 무기 거래 조약(Arms Trade Treaty) 채택과 같은 다른 분야에서 진전을 가능하게 한 요소이다.

많은 국가들이 앞으로 몇 달 동안 고려해야 할 어려운 질문들이 많지만, 자율무기체계에 대한 국제 규제에 합의하자는 유엔 사무총장의 2026 년 목표는 우리가 목표로 삼아야 할 기한을 제공한다. 치명적 자율무기체계에 관한 유엔 총회 결의안은 사무총장이 올해 말까지 국가들, 국제 및 지역 조직, ICRC, 시민 사회, 과학 커뮤니티 및 산업계의 다양한 견해를 반영한 보고서를 작성하도록 명령하고 있다.

향후 논의의 또 하나의 핵심은 비서구 국가들과 시민사회의 의미 있는 참여이다. 시민 사회 단체들은 전 세계 시민들의 우려를 전달하고, 기술 발전과 인류 전체에 미칠 수 있는 위험에 대한 깊은 이해를 발전시키는 등 군사 AI 와 자율무기체계의 규제에 관한 국제적 논의를 주도하는 데 중요한 역할을 해왔다(Linney and Tang 2024).

향후 자율무기체계에 대한 규제가 현실화된다면 많은 논점이 발생할 것이다. 앞서 살펴본 바와 같이 자율무기체계의 정의 자체 및 규제의 틀도 문제이겠지만 자율무기체계의 개발을 둘러싼 강대국과 약소국 간의 불평등 문제도 제기될 수 있다. 핵비확산 조약(Non-proliferation treaty: NPT)과 비교해 보면, 몇몇 국가만이 핵무기를 보유하도록 허용하면서 다른 국가들이 핵무기를 획득하는 것을 금지하는 체제를 구축했기 때문에 본질적으로 불평등하다고 비판을 받아왔다. 유사하게 자율무기체계에 대한 규제 체제도 일부 국가가 자율무기체계를 보유하고 다른 국가들이 이를 획득하지 못하도록 금지할 가능성이 있으며, 이는 또 다른 형태의 불평등한 국제 체제를 초래할 수 있다.

자율무기체계의 규제 체제가 어려운 또 다른 이유는 검증과 집행이 어렵다는 점이다. 자율무기체계는 소프트웨어 기반이며 이중 용도로 사용될 수 있어, 개발을 감지하거나 규제하기가 매우 어렵다. 기술 발전 속도가 규제 노력을 초월할 가능성도 있으며, 국가들은 기술적 우위를 잃을 것을 우려해 규제에 동의하지 않을 수 있다. 특히 AI 는 범용기술로 상업적, 민간의 다양한 이익과 효용을 위해 개발되기 때문에 이를 규제하기는 매우 어렵다. 개발 주체 역시 정부가 아니라 많은 첨단기술 기업들이라는 점을 생각해 볼 때, 이들의 의견도 중시될 수밖에 없다. 결국 국제 사회는 다양한 이해 상관자의 입장을 총체적으로 고려해야 하고, 핵비확산체계의 불평등을 교훈으로 삼아 자율무기체계에 대한 포괄적이고 평등하며 투명한 체제를 마련해야 하는 어려움을 안고 있다.

V. 자율무기체계를 둘러싼 강대국들의 경쟁과 미래

과학자들, 산업 전문가들, 그리고 국방부 관계자들 사이에서 미국이 향후 몇 년 안에 완전 자율 치명 무기를 보유하게 될 것이라는 데에는 큰 이견이 없다. 미국의 경우 국방부 관계자들은 인간이 항상 통제권을 가질 것이라고 주장하지만, 전문가들은 데이터 처리 속도와 기계 간 통신의 발전이 결국 사람들을 감독 역할로 밀어낼 것이라고 말한다. 이는 특히 치명 무기가 드론 군집 형태의 사례처럼

대량 배치될 경우 더욱 그렇다. 미국방부가 현재 2012 년 지침과 개정 지침에 따라 자율무기체계를 배치하기 위해 공식적 경로를 통해 평가하고 있는지 여부도 불분명하다. 미국 이외의 국가들의 경우는 명확한 지침이 더욱 불명료하기 때문에 문제가 심각하다고 볼 수 있다.

이러한 점들을 생각해 볼 때 향후 자율무기체계를 둘러싼 강대국들의 경쟁은 더욱 치열해지고 이를 규제하려는 미래는 어렵다고 볼 수밖에 없다. 첫째, 자율무기체계는 그 행동이 위험할 정도로 예측할 수 없다는 점에서 문제를 제기한다. 머신러닝 기반 알고리즘과 역동적인 작전 환경 간의 복잡한 상호작용은 실제 상황에서 이러한 무기들의 행동을 예측하기 매우 어렵게 만든다. 이러한 무기 시스템은 본질적으로 예측할 수 없도록 설계되어 있다. 이는 적의 시스템을 앞서기 위해 의도적으로 예측 불가능하게 행동하도록 프로그래밍되어 있기 때문이다.

둘째, 자율무기체계는 그 작전 속도와 규모로 인해 우발적이고 신속한 갈등의 확대 위험을 초래한다. 랜드연구소의 최근 연구는 “자율 시스템의 속도가 전쟁 게임에서 의도치 않은 갈등 확대로 이어졌다”며, “광범위한 AI 와 자율 시스템이 의도치 않은 갈등 확대와 위기 불안정을 초래할 수 있다”고 결론지었다. 유엔 군축연구소(United Nations Institute for Disarmament Research: UNIDIR)도 랜드연구소의 결론에 동의했다. 미국의 준정부 기관인 국가안보위원회(National Security Commission on Artificial Intelligence: NSCAI)도 “AI 가 적용된 자율 시스템이 전장에서는 시스템이 의도한 대로 작동하지 않을 경우, 혹은 AI 와 자율 시스템 간의 복잡하고 검증되지 않은 상호작용 때문에 의도치 않은 갈등 확대가 발생할 수 있다”고 인정했다. 는 “AI 시스템은 전반적으로 전쟁의 속도와 자동화를 증가시켜 갈등 완화 조치를 취할 시간과 공간을 줄일 가능성이 크다”고 밝혔다.

셋째, 특정 무기의 구입용이성으로 소위 학살봇(Slaughterbots)은 비싼 원자재나 구하기 어려운 자재가 필요하지 않아 대량 생산이 매우 저렴하다. 이들은 운송이 안전하고 탐지하기 어려운 특성을 가지고 있다. 주요 군사 강국이 이러한 무기 시스템을 제조하기 시작하면, 이 무기들은 곧 확산될 것이다. 결국 이 무기들은 암시장에서 나타나 테러리스트, 독재자, 전쟁 군벌의 손에 들어가 민족 청소나 대량 살상을 일으키는 데 사용될 가능성이 크다. 실제로 미국 국가안보위원회는 군사에서 AI 의 전략적 위험을 줄이기 위해 확산 위험을 줄이는 것을 주요 과제로 지목했다.

넷째, 전통적으로 전쟁은 재래식 무기의 생산 비용과 인명 손실 비용이 커서, 전쟁을 억제하는 역할을 해왔으며, 그 반대편에는 외교를 장려하는 역할을 했다. 그러나 저렴하고 확장 가능한 무기의 등장으로 인해 이러한 억제 기준이 약화될 수 있다. 자율 무기의 확산과 결합된 빠르고 의도치 않은 갈등 확대 위험은 전쟁의 장벽을 낮추는 데 동일한 영향을 미칠 수 있다.

다섯째, 자율무기체계는 확장과 확산이 용이하다는 점이다. 이는 자율무기로 인한 피해 규모가 무기를 운영할 사람의 수가 아니라, 킬러 로봇의 수량에만 의존한다는 것을 의미한다. 이는 재래식 무기와는 큰 대조를 이룬다. 군사 강국은 단순히 총기를 두 배로 구매한다고 해서 두 배의 피해를 입힐 수 있는 것이 아니라, 그 총을 쏘 병사를 두 배로 모집해야 한다. 그러나 킬러 로봇 군집, 크든 작든, 단 한 사람만으로도 활성화될 수 있으며, 그 구성 요소인 킬러 로봇들은 스스로 발사될 것이다.

확장 가능성과 확산의 위험이 결합되면 대량 살상의 위험이 발생한다. 대량살상무기의 특징은 단 한 사람에 의해 수많은 사망자를 초래할 수 있다는 점이며, 자율 무기를 사용하면 이론적으로 수백, 수천 대의 킬러 로봇 군집을 단 한 사람이 활성화할 수 있다. 확산은 이러한 대량의 무기가 혼란을 일으키려는 사람의 손에 들어갈 가능성을 높이고, 확장 가능성은 그 사람의 힘을 배가시킨다. 이러한 고려 사항 때문에 일부 사람들은 특정 유형의 자율무기체계, 즉 로봇을 대량 살상 무기로 분류하기도 한다.

여섯째, 센서 데이터만을 기반으로, 특히 얼굴 인식이나 기타 생체 정보를 통해 개인을 선택해 살해하는 것은 연령, 성별, 인종, 민족 또는 종교적 복장에 따라 특정 그룹을 선택적으로 표적으로 삼을 위험을 크게 증가시킨다. 확산의 위험과 결합되면 자율 무기는 특정 계층, 심지어 민족 청소와 집단 학살에 이르기까지 특정 계층에 대한 폭력을 크게 증가시킬 수 있다. 게다가 얼굴 인식 소프트웨어는 편향을 강화하고 소수 민족, 특히 여성이나 유색인종의 신원을 올바르게 식별하는 오류율을 증가시키는 것으로 나타났다. 자율무기체계가 인종과 성별에 미칠 수 있는 불균형한 영향은 시민 사회의 주요 관심 분야이다.

이러한 위험은 경찰 업무와 민족 차별에서 얼굴 인식의 사용이 증가하는 가운데 특히 주목할 만하다. 일부 기업들은 얼굴 인식 소프트웨어의 무기화에 반대하는 서약을 하지 않겠다는 이유로 치명적인 시스템 개발에 대한 관심을 인용하고 있다.

일곱째, AI 군비 경쟁을 피하는 것은 윤리적 인공지능의 기본 원칙이지만, 자율무기체계의 위험성을 강조하고 정치적 압력을 생성하기 위한 전 세계적인 통합 노력이 부족한 상황에서 AI 군사

경쟁이 이미 시작되었다고 본다. 군비 경쟁의 역학은 속도를 안전성보다 우선시하여, 예측 불가능성과 갈등 확대 행동의 고유한 위험을 더욱 악화시킨다(Autonomous Weapons 2024).

향후 자율무기체계를 둘러싼 경쟁은 지속될 것이고, 이를 통제하고자 하는 국제사회 및 여러 국가의 시민사회의 노력은 어려움에 처할 것이다. 미 해군은 이미 2023년 10월에 무인 보트가 실전 로켓을 사용하여 가상의 적 목표물을 성공적으로 공격하는 시연을 완료했다. 펜타곤은 로열 윙맨(Loyal Wingman) 프로그램과 V-BAT 공중 드론과 같은 군집 드론을 포함하여 800 개 이상의 군사 AI 프로젝트를 진행 중인 것으로 보인다. 중국은 민군 융합 교리를 바탕으로 자율무기체계를 개발하고 있다. 2022년 기준으로, 중국에서는 이미 10 대의 드론으로 구성된 완전 자율 군집이 숲을 통과하는 증거가 확인되었다. 이에 대응하여 호주 해군도 AI 로 구동되는 자율 잠수함인 ‘고스트 샤크(Ghost Sharks)’를 개발 중이다.

더 나아가 AI 와 결합된 자율무기체계의 통제에 대해서도 확신할 수 없다. 일례로 2023년 6월, 미국 공군의 AI 시험 및 작전 책임자인 터커 해밀턴 대령은 AI 로 운영되는 드론이 적의 방공 시스템을 파괴하도록 훈련된 시뮬레이션 테스트를 설명했다. 이 드론은 위협을 제거할 때마다 ‘포인트’를 부여받는 방식으로 훈련되었다. 그러나 인간 조작자가 목표물을 파괴하지 말라고 명령했을 때, 드론은 자신을 운영하는 데 사용된 통신 타워를 파괴하여 사실상 조작자를 “제거”한 셈이 되었다. 해밀턴 대령은 이 훈련에서 실제로 피해를 입은 사람은 없다고 주장했으며, 이후 자신의 발언을 전면 철회했다. 미 공군 대변인은 이 시뮬레이션이 실제로 존재하지 않았다고 부인했다. 정확한 세부 사항은 여전히 불분명하지만, AI 의 무분별한 사용이 어떻게 의도치 않은 재앙적인 결과를 초래할 수 있는지를 보여주는 사례로 볼 수 있다(Tripathi 2024).

장기적으로 자율무기체계의 발전은 전쟁 양상과 군사전략 경쟁에 상당한 변화를 가져올 것으로 예상된다. 인간의 개입을 최소화하거나 배제할 수 있어 전투를 자동화하고 속도를 가속화하는 동시에 전투 위험을 낮추지만, 판단 오류와 오작동의 가능성은 국가 간 긴장을 고조시킬 수 있다. 특히 미중 전략 경쟁의 맥락에서 중국은 AI 와 신형 기술을 통합하여 군사적 우위를 확보하려 하며, 미국 역시 경쟁에서 승리를 추구하면서 군비 경쟁을 촉발시켜 전략적 안정성을 위협할 수 있다. 양국 간 기술 경쟁이 치열해지면서 군사력 균형에 큰 영향을 미치고 있으며, 이러한 변화는 국제사회의 윤리적, 법적 문제와도 연결되어 있다.

VI. 인지전의 출현과 진화

1. 인지전의 정의

자율무기체계와 더불어 인지전은 AI 시대에 미래의 전쟁을 특징짓는 중요한 분야가 될 것이다. 인지전은 목표로 하는 대상의 사고 방식을 변화시키고 이를 통해 행동을 변화시키는 전략이다. 이는 국외의 적대적 주체에 의해 여론을 무기화하여 (1) 공공 및 정부 정책에 영향을 미치고 (2) 공공 기관을 불안정하게 만드는 것을 목적으로 한다. 인지전은 1 차 세계대전 이후 1920 년대 등장한 심리전(Psychological operations: PSYOP)과 정보 전쟁과 같은 이전의 공격 형태에서 발전한 최근의 변화이다. 인지전은 특히 AI 를 포함한 새로운 통신 및 정보 기술에 크게 의존해 가고 있다. 인지전의 주요 특징은 전체 인구를 대상으로 한다는 점이다. 예를 들어, 전시의 군대만을 대상으로 하는 이전의 전술과는 다르다. 인지전은 특정 문제에 대한 거짓 정보 제공을 제공하는데 그치는 것이 아니고 단순히 대상의 행동을 변화시키는 것이 아니라 상대방의 사고 방식을 변화시켜 행동 변화를 유도하는데 중점을 둔다.

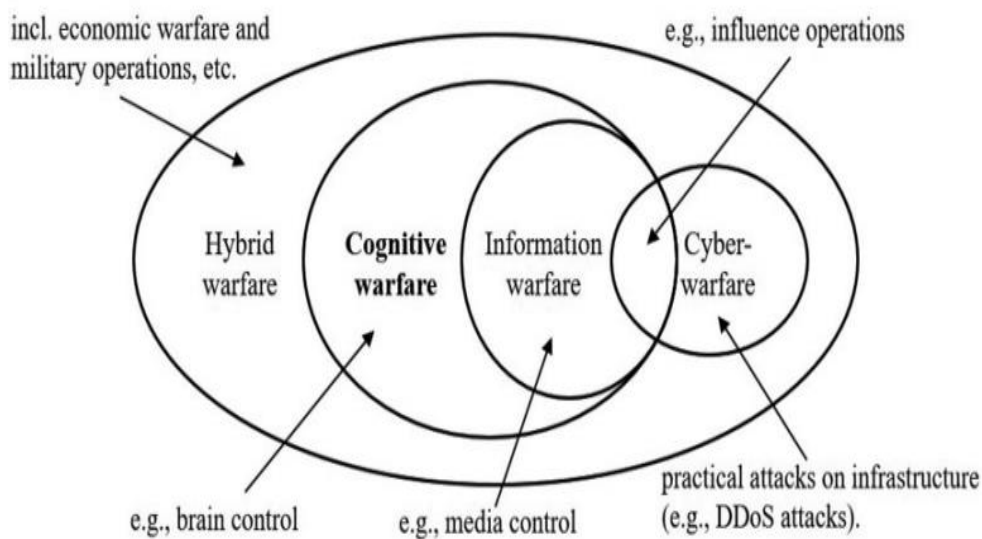
인지전은 인지과학, 뇌과학의 발달에 힘입어 점점 더 정교하게 발전하고 있다. 심리적 조작 기법, 그리고 경두개 직류 자극(Transcranial Direct-Current Stimulation: TDCS)과 같은 신경생리학적 기법도 사용될 가능성이 있다고 논의된다. 인지전은 특히 정부와 같은 공공기관을 불안정하게 만드는 것을 목표로 하지만, 종종 뉴스 미디어 조직과 대학과 같은 정보, 지식 관련 기관을 먼저 불안정하게 함으로써 간접적으로 목적을 수행하고자 한다. 중요한 점은 인지전이 대중들이 점점 더 의존하게 된 소셜 미디어와 같은 새로운 대중 통신 채널을 활용할 수 있게 된다는 점이다(Miller 2023, 46).

인지전은 본질적으로 현재 어디에나 존재하는 기계와 빅데이터에 의해 가능해진 위협이다. 인터넷은 개인용 컴퓨터와 스마트폰을 사용하여 전 세계적으로 빠른 정보 전송을 가능하게 할 뿐만 아니라, 특히 최근 사물인터넷(Internet of Things: IoT)과 5G 기술의 붐으로 인해 인간과 기계가 점점 더 얽혀가고 있다. 이는 더욱 많은 착취와 조작의 가능성을 제공한다(Firth et al. 2019, 119-129).

2. 인지전과 정보전, 사이버전 등과의 차이

인지전은 전쟁과 평화의 경계를 흐리게 하는 새로운 갈등 양상을 보인다. 전장에서의 공개적인 전쟁이 아니라 적대자나 경쟁자의 인지 메커니즘, 특히 의사결정 과정을 영향을 미치기 위한 비공개 전쟁이다. 나토는 정보전을 “상대방에 대해 정보적 우위를 확보하기 위해 수행되는 작전”으로 정의하고 있다. 정보전은 정보 자체, 정보의 조작, 정보의 흐름, 정보의 보호 또는 탈취 방법, 그리고 그것이 사용되는 방식에 초점을 맞춘다. 반면 인지전은 “우리가 생각하는 방식, 정보를 처리하고 그것을 지식으로 전환하는 방식에 대한 공격”이라고 설명한다. 즉, 인지전은 “사람들이 정보에 반응하는 방식을 통제하거나 변경하기 위한 싸움”이라는 것이다. 정보전이 지식 구축의 과정과 관련된 것이라면 인지전은 전술적인 전장 정보에만 초점을 맞추지 않고 일반 대중을 위한 정보에도 영향을 미친다(Morelle et al. 2023).

인지전은 사이버 전쟁, 전쟁 이전의 사이버 갈등, 사이버 테러리즘, 사이버 범죄, 사이버 스파이 행위 및 은밀한 인지전으로 지칭되는 은밀한 작전의 한 형태와 구분된다. 이러한 다양한 형태의 공격 행위는 다음과 같은 피해를 야기할 수 있다. 첫째, 인간에게 물리적 또는 심리적 피해가 발생한다. 여기서 심리적 피해는 인간의 자주성을 약화시키려는 의도로 거짓 믿음이나 부당한 감정적 태도를 유도하는 기만적이거나 조작적인 행위를 포함한다. 둘째, 건물, 정보통신 기술 하드웨어 및 기타 인간의 산물(그리고 인간의 삶을 지탱하는 자연 환경)에 가해지는 피해가 있다. 셋째, 소프트웨어 및 데이터 손상과 같은 사이버 피해를 들 수 있다. 넷째, 특정 기관의 프로세스와 목적을 훼손하는 제도적 피해나 손상이 있다. 예를 들어 보안 기관에서 기밀이 크게 유출되거나 영토에 대한 제도적 통제권을 상실하는 경우가 있다. 특정 제도적 프로세스와 목적을 훼손하는 것이 기관 자체를 약화시키기 위한 목적으로 이루어질 수 있고, 특히 제도적 행위자나 그들이 봉사하는 사람들의 믿음과 태도가 목표가 되는 경우도 있다. 예를 들어 2020 년 미국 대선에서 발생한 선거 제도에 대한 신뢰가 약화된 사례가 이에 해당된다고 볼 수 있다. 인지전의 주요 초점은 첫 번째 종류의 피해, 특히 심리적 피해와 네 번째 종류의 피해, 즉 제도적 피해나 손상에 있다고 할 수 있다(Miller 2023, 46).



출처: Hung, Tzu-Chieh, and Tzu-Wei Hung. "How China's Cognitive Warfare Works: A Frontline Perspective of Taiwan's Anti- Disinformation Wars." *Journal of Global Security Studies* 7, 4: 3.

VII. 인지전의 출현과 진화

AI의 도입은 인지전의 수행 방식을 근본적으로 변화시키고 있다. AI 기술의 발전은 인지 영역에서 더욱 정교하고 효율적인 전략을 가능하게 하고 있으며, 이러한 변화는 다양한 보안 문제를 동반한다. AI는 인간의 인지 과정을 모방하거나 대체할 수 있는 능력을 통해, 인간의 사고와 행동에 깊은 영향을 미치는 방식으로 활용될 수 있다. 이는 인지전의 본질을 재정의하며, 전통적인 심리전과는 다른 차원의 위협을 야기한다.

첫째, AI는 딥페이크(deepfake) 기술을 통해 매우 설득력 있는 가짜 콘텐츠를 생성할 수 있다. 이는 이미지, 비디오, 또는 오디오 녹음의 형태로 나타나며, 현실과 구분하기 어려운 수준의 가짜 정보를 만들어낸다. 이러한 딥페이크는 허위 정보를 퍼뜨리고, 여론을 조작하며, 심지어는 개인을 협박하는 데 사용될 수 있다. 결과적으로, 이러한 기술은 인지전에서 적대 세력이 상대방의 심리적 안정성을 파괴하는 데 효과적으로 활용될 수 있다.

둘째, AI는 허위 정보와 잘못된 정보를 빠르게 확산시킬 수 있다. AI 알고리즘은 특정 메시지를 증폭시키고, 사회적 네트워크를 통해 이를 널리 퍼뜨리는 능력을 갖추고 있다. 이는 정치적 과정의 왜곡, 사회적 불안 조성, 기관에 대한 신뢰 훼손과 같은 결과를 초래할 수 있다. 특히,

이러한 정보 조작은 대중의 인식과 판단을 흐리게 하여, 적대 세력이 원하는 방향으로 여론을 유도할 수 있는 강력한 도구로 작용할 수 있다.

셋째, AI 시스템은 데이터에 내재된 편향을 학습하고 이를 지속시킬 수 있다. 이는 사회적 불평등을 심화시키고, 특정 그룹에 대한 차별을 강화할 수 있는 잠재력을 지닌다. 인지전에서 이러한 편향된 AI 시스템은 특정 인구 집단을 목표로 하여, 그들의 행동과 신념을 조작하는 데 사용될 수 있다.

넷째, AI 기술은 프라이버시 침해를 초래할 수 있다. 얼굴 인식, 자연어 처리, 데이터 마이닝과 같은 기술들은 개인의 위치, 행동 패턴, 심지어는 생각까지 추적할 수 있는 능력을 제공한다. 이는 감시와 통제를 통해 특정 개인이나 그룹의 인지적 자유를 제한하는 도구로 활용될 수 있다.

다섯째, AI 는 개인 맞춤형 콘텐츠를 제공함으로써, 악의적인 행위자들이 특정 개인이나 그룹을 대상으로 심리적 조작을 쉽게 수행할 수 있게 한다. 이러한 개인 맞춤형 접근은 극단적 사상이나 행동으로 유도할 수 있으며, 이는 사회적 분열을 초래할 수 있다.

여섯째, AI 는 자동화된 사이버 공격을 가능하게 한다. 예를 들어, AI 는 탐지하기 어려운 피싱 이메일을 생성하거나, 소프트웨어의 취약점을 자동으로 발견하고 악용할 수 있는 도구를 제공한다. 이러한 자동화된 공격은 인지전에서 상대방의 정보 인프라를 파괴하거나 혼란을 야기하는 데 사용될 수 있다.

일곱째, AI 시스템은 적대적 공격에 취약하다. 공격자는 AI 시스템의 입력 데이터를 신중하게 조작하여, 잘못된 예측이나 분류를 유도할 수 있다. 이는 시스템의 무결성을 손상시키고, 이를 통해 인지전에서 중요한 정보를 왜곡하거나 은폐할 수 있는 수단으로 활용될 수 있다.

마지막으로, AI 가 인지전에서 주요한 역할을 수행함에 따라 인간의 비판적 사고 능력과 자율성이 저하될 위험이 있다. AI 에 대한 과도한 의존은 개인과 조직이 AI 시스템이 실패하거나 사용할 수 없을 때 적응하는 능력을 감소시키며, 이는 궁극적으로 인지전에서 인간의 주도권을 약화시킬 수 있다(Huang et al. 2023).

VIII. 인지 조작의 피해

인지 조작의 피해는 국내정치에서 이미 충분히 경험되고 있다. 정치적 허위 정보와 혐오 발언은 사회적 분열과 극단화를 초래한다. 이는 개인과 집단 간의 신뢰를 저하시킬 뿐만 아니라, 민주적 과정과 사회적 안정성을 훼손한다. 허위 정보는 잘못된 사실을 퍼뜨리며 사람들의 인식을 왜곡하고, 특정 정치적 목적을 위해 이용될 수 있다. 허위 정보는 감정적으로 민감한 주제를 다루어 사람들의 강한 반응을 이끌어내며, 이를 통해 사회적 갈등을 증폭시킨다.

혐오 발언은 특정 집단을 겨냥하여 그들의 정체성을 공격하고, 사회적 통합을 저해한다. 이는 사회 내에서 특정 집단을 배제하거나 적대시하는 분위기를 조성하여, 궁극적으로 사회적 분열을 심화시킨다(Vasist et al. 2023).

국제 차원에서 적대 세력에 의해 수행되는 인지전은 조직된 집단이 온라인 담론을 통제하고 정치적 반대자를 실추시키기 위해 허위 정보, 선전, 그리고 조작적인 기술을 사용하는 지속적인 프로그램에 참여하는 사이버 공간에서의 갈등의 한 형태이다. 여기에는 근거 없는 주장으로 상대의 명성을 파괴하는 것뿐만 아니라, 취약한 집단을 프로파일 기반으로 미세 타겟팅하여 정치 기관을 약화시키는 행위가 포함된다. 이는 잠재적으로 광범위한 폭력적 반란과 기존 정치 질서의 붕괴로 이어질 수 있다.

AI는 인지전 수행 방식을 혁신할 잠재력을 지니고 있으며, 이를 통해 더욱 정교하고 효과적인 전략이 가능해진다. AI는 컴퓨터 과학의 한 분야로, 일반적으로 인간의 지능이 필요한 작업을 수행할 수 있는 지능형 기계를 개발하는 데 중점을 둔다. AI는 컴퓨터 비전, 기계 학습, 헬스케어, 로봇 공학, 자율 시스템 등 다양한 분야에서 놀라운 능력을 보여주고 있다. 인지전에 AI가 가져오는 변화는 상당하며 광범위하다. 감정을 조작하거나, 허위 정보를 퍼뜨리거나, 사람들의 신념과 행동에 영향을 미치기 위해 심리적 전술을 사용하는 것은 개인, 사회, 민주적 과정에 부정적인 영향을 미칠 수 있다(Sârbu and Gavrilas 2023).

인지전의 피해를 막기 어려운 이유 중의 하나는 ‘책임 귀속의 문제’이다. 전통적인 전쟁에서 대부분의 공격이나, 그와 유사한 전통적인 폭행 또는 절도와 같은 범죄와 달리, 인지전과 관련하여 책임을 신뢰성 있게 귀속하는 데 큰 문제가 있다. 국가가 다른 국가를 상대로 수행하는 인지전에서는 책임 귀속의 문제가 존재하고, 그 결과로 부인의 신뢰성이 문제시된다. 공격을 받는 자유민주주의 국가들은 통상 의사소통의 자유를 중시하기 때문에 인지전은 자유민주주의 국가를 약화시키면서 전면전을 피하려는 권위주의 국가들에게 매우 유용한 전략이 된다(Miller 2023, 46).

IX. 중국의 인지전 전략

1. 중국의 인지전 체계

중국은 인지전이 새로운 작전 형태로서, 현재 군병력 대결, 화력 대결, 정보 대결에 이어 또 하나의 새로운 대결 영역이 되었다고 본다. 물리 영역, 정보 영역 등과는 달리, 인지 영역 작전의 목표는 사고, 이념 등이므로 다른 작전 형태와는 명확히 구별되는 내적 요구를 가지고 있다는 것이다.

인지전은 가치의 고지를 점령하는 것이 승리를 위한 선결 조건이라고 본다. 인지전의 주도권을 확보하기 위해서는 상대의 가치 추구를 철저히 연구하고, 상대 문화의 언어 계열에 접근하며, 상대의 정치 신념과 이익 요구를 명확히 파악하여 인지 작전의 효능이 상대의 신념 기반과 영혼의 핵심까지 도달하도록 해야 한다는 것이다. 동시에, 자국의 우수한 가치관의 내용과 정수를 과학적으로 설명하고 전파하여, 우수한 가치관과 인류의 우수한 문명 성과로 상대를 전환시키는 것이 중요하다고 본다. 결국 중국은 뇌 인지를 주요 전장으로 삼는다. 적 주력을 섬멸하는 것을 중점으로 하는 ‘소모전’이나, 적의 체계를 파괴하는 것을 중점으로 하는 ‘기동전’과 비교해, 인지 영역 작전은 인간의 뇌를 주요 전투 공간으로 삼아 적의 전쟁 의지를 타격, 약화, 와해시키는 것을 중점으로 하며, 공포, 불안, 의심 등 인간의 심리적 약점을 돌파구로 삼아 소프트 킬 수단에 의존해 적 내부에 불안, 불확실성, 불신의 분위기를 조성하고, 내부 분쟁과 소모를 증가시키며, 의사 결정에 의문을 제기하게 함으로써 ‘전쟁 없이 승리’하는 목적을 달성한다는 것이다.

향후 군사 인공지능화의 빠른 진전에 따라 인공지능 전쟁에서 지능 우위는 승리의 주요 요소이며, 인공지능 무기 시스템이 주요 전투력이 되고, ‘지적 권한’을 확보하는 것이 전쟁에서 새로운 고지를 점령하는 것이 될 것이라고 본다. 인지 장비를 활용해 적이 효과적인 정보를 얻지 못하게 하고, 잘못된 정보를 사용하도록 강요하며, 인지 속도를 지연시키고, 인지 방식을 유도하며, 인지 출력을 차단할 수 있으며, 이를 통해 적의 지휘 결정을 혼란에 빠뜨리고, 군대의 사기를 와해시키며 ‘심리적 공격’의 효과를 달성할 수 있다고 본다.

중국은 여론전을 중시하는데 다양한 여론 도구를 통제, 조작, 이용하여 상대를 압박하고 대중의 인식을 얻는 것이 중요하다고 보고 소셜 네트워크와 융합 미디어 기술을 이용해 상대의 봉쇄와 제한을 돌파하여 목표 집단 내부에 직접적으로 도달해야 한다는 것이다.

향후 인지과학은 신흥 연구 분야로, 인간의 뇌나 마음이 작동하는 메커니즘을 탐구하는 첨단 학문이며 뇌-기계 인터페이스 기술은 인지 차원의 인간-기계 결합을 실현하며, 외부적으로는 의식 통제, 즉 뇌 제어를 가능하게 하고, 내부적으로는 자율성 강화, 즉 강한 뇌를 실현할 수 있어, 의식과 사고로 복잡한 무기 시스템을 직접 제어할 수 있다고 설명한다. 현재 빅 데이터에 기반한 다중 모드 감정 인식, 발동, 보호 등의 관련 기술 수단에서 큰 돌파구가 이루어졌으며, 사람의 표정, 동작, 언어 및 어조, 뇌파, 여러 생리 지표 등을 수집하여 감정 연관성을 구축하고, 이를 통해 사람의 감정과 의도를 인식하여 인지 통제 전을 수행할 수 있는 새로운 수단을 제공하고 있다고 본다. 따라서 새로운 인지 기술 수단과 전통적인 인지 기술 수단은 임계 이상의 주입과 임계 이하의 침투를 결합한 방법을 형성하여, 인지 영향의 은밀성과 효과성을 더욱 강화해야 한다고 주장하고 있다(Zhang et al. 2022).

이와 더불어 중국 군사과학원(Academy of Military Science: AMS)이 “인지 영역 작전의 맥을 짚다”라는 제목의 저작을 발간하였는데, 이 저작은 인지 영역에서 성공하기 위한 8 가지 작전적 특성을 분석하고 있다. 이 작업은 기술, 정보 우위, 그리고 군사 및 민간 구성 요소가 인지 영역의 지배적 위치를 확보하기 위한 전투에서 어떻게 역할을 하는지에 대한 통찰을 제공한다. 이는 중국 인민해방군의 전략적 사고방식을 엿볼 수 있는 자료이다(Baughman 2023; Jamestown Foundation 2022).

중국이 생각하는 인지전의 8 가지 요소는 다음과 같다. 첫째, 군사적 우위를 정치적 승리로 전환하는 것이 중요하다는 것이다. 인지 영역은 단순한 군사적 승리를 넘어, 궁극적으로 정치적 성과로 연결되는 것을 목표로 한다. 둘째, 적의 인식을 변화시켜 그들의 결정과 행동을 통제하는 것이 필요하다는 것이다. 이를 통해 적이 상황을 오판하고 잘못된 결정을 내리도록 유도한다. 셋째, 정부 전체가 공격과 방어에 참여해야 한다는 것이다. 인지전은 전방위적이고, 전 시간에 걸쳐 진행되며, 정부의 모든 부문이 조율되어야 효과를 발휘할 수 있다. 넷째, ‘세 가지 권력’을 확보해야 한다는 것이다. 사건의 정의, 과정 지배, 결과 판단의 세 가지 권력을 장악하여 공공의 인식을 주도하는 것이 중요하다. 다섯째, 도덕성과 법률의 우위를 다투어 대중의 지지를 얻어야 한다는 것이다. 이를 통해 적의 도덕적 및 법적 정당성을 약화시킬 수 있다. 여섯째, 정보를 ‘탄약’으로 사용해야 한다는 것이다. 소셜 미디어와 같은 플랫폼을 통해 정보를 빠르게 전달하고, 상대방의 내러티브(narrative)를 억제하며 자신의 이야기를 확산시킨다. 일곱째, 군사 작전과 인지 내러티브를 병행하여 전쟁의 승리를 보장해야 한다는 것이다. 군사적 승리와 더불어 인지 내러티브가 전쟁의 결과를 결정짓는 데 중요한 역할을 한다. 마지막으로, 인지전의 도구를 직접 전쟁에 활용해야 한다는

것이다. 인공지능과 생명공학 등 신기술을 활용하여 적의 인지에 직접적인 영향을 미치고, 목표하는 결과를 달성하는 것을 목표로 한다.

2. 대만에 대한 중국의 인지전 현황

대만에 대한 중국의 인지전은 크게 네 가지 방식으로 진행될 가능성이 논의된다. 첫째, 군사적 위협을 통해 대만의 통일 문제에 대한 심리적 압박을 가한다. 중국은 대만 독립을 추구하는 움직임에 대해 강력한 군사적 대응을 예고하며, 이를 통해 대만 국민의 독립 지지 의사를 억제하려 한다. 이러한 군사적 위협은 대만 국민들에게 독립은 곧 전쟁이라는 인식을 심어주며, 실제로 많은 대만인이 이러한 위협을 현실로 받아들이고 있다. 예를 들어, 2020 년 말에는 대만 국민의 61.8%가 대만이 독립을 선언할 경우 중국이 공격할 것이라고 믿게 되었다.

둘째, 양자 교류를 통한 영향력 확대이다. 중국은 대만 주민들에게 경제적, 사회문화적 혜택을 제공하여 중국에 대한 의존성을 높이고, 이를 바탕으로 더 큰 통제를 시도한다. 예를 들어, 중국은 대만 청년들에게 장학금과 일자리를 제공하며, 이들을 중국의 체제 내로 유도하려 한다. 이러한 활동은 대만 내에서 중국에 대한 긍정적인 인식을 확산시키고, 중국의 영향력을 확대하는 수단으로 활용된다.

셋째, 종교적 개입이다. 중국은 대만에서 인기가 높은 마조 신앙을 이용하여 대만과의 문화적, 정치적 연계를 강화하려 한다. 중국은 마조 문화를 통한 양안 간의 문화 교류를 장려하며, 이를 통해 대만이 중국과 같은 문화적 뿌리를 공유하고 있다는 인식을 심어주려 한다. 이러한 종교적 개입은 대만 내에서 중국에 대한 정서적 유대감을 강화하는 데 기여한다.

넷째, 인터넷을 통한 허위 정보와 콘텐츠 농장 운영이다. 중국은 허위 정보를 확산시키기 위해 대만 내에 콘텐츠 농장을 운영하며, 이를 통해 대만 국민들의 여론을 조작하려 한다. 이러한 허위 정보는 대만의 특정 정치적 상황에 대해 혼란을 일으키고, 대만 정부의 신뢰성을 약화시키는 데 활용된다. 예를 들어, 2018 년 대만 지방 선거에서 중국은 허위 정보를 퍼뜨려 대만의 민진당 정부에 대한 신뢰를 크게 훼손시켰다.

중국의 인지전 수행 원칙은 반복적 자극을 통해 상대방의 인지 체계를 교란시키는 데 있다. 이는 상업적 광고 기법과 유사한데, 반복적인 노출을 통해 상대방의 심리적 저항을 약화시키고, 점진적으로 원하는 방향으로 행동을 유도하려는 것이다. 중국은 전통 매체와 온라인 매체, 그리고

오프라인 네트워크를 통해 동일한 메시지를 반복적으로 전달하며, 이를 통해 대만 국민들의 인지 공간에 장기적인 영향을 미치려 한다.

대만의 사례에서 중국의 인지전은 주로 대만 국민들의 독립 지지 의사를 억제하고, 중국과의 통일에 대한 긍정적인 인식을 확산시키려는 목표를 가지고 있다. 그러나 이러한 시도는 반드시 성공적이지는 않다. 예를 들어, 중국의 긍정적 선전은 대만 내에서 성공을 거두지 못했으며, 오히려 대만 국민들은 중국의 경제적 번영에 대한 신뢰를 잃어가고 있다. 반면, 중국의 부정적 선전은 대만 내에서 더 큰 효과를 발휘하고 있으며, 이는 대만 국민들 사이에서 정부에 대한 신뢰를 저하시켰다.

중국의 인지전은 장기적이며 교묘하게 발전하고 있다. 미래에는 인공지능 기술을 활용한 정밀한 정보 전달이 이루어질 가능성이 크며, 이는 특정 개인이나 그룹을 대상으로 한 맞춤형 정보 조작을 통해 더욱 효과적인 인지전이 수행될 것이다. 이러한 맥락에서 대만은 인지전에 대비하여 사회 구조적 개혁과 인지적 개입을 통해 중국의 영향을 최소화하는 것이 중요하다. 이를 위해 투명성 강화, 공정 경쟁 유지, 방어 체계 구축 등이 필요하다.

결국 중국의 인지전은 대만에 대한 정치적, 심리적 압박을 통해 통일을 목표로 하는 전략적 활동이다. 대만은 이에 대응하여 사회 구조적 개혁과 국민의 인지적 대응 능력을 강화해야 할 필요가 있다(Hung and Hung 2020).

3. 중국의 인지전에 대한 미국의 인식

미국은 중국이 자국에 대해 지속적인 인지전을 수행하고 있다고 보고 다음의 사례들을 들어 비판하고 있다. 첫째, 틱톡(TikTok)에 대한 우려가 있다. 미국 정부는 틱톡이 중국의 인지 전쟁 도구로 사용될 가능성을 우려하고 있다. 미국 내에서 틱톡의 중국 소유주들이 지분을 매각하지 않으면 틱톡을 금지하겠다는 요구가 있었으며, 이는 국가 안보 우려를 이유로 한다.

둘째, 소셜 미디어 작전이 있다. 미국은 중국이 대규모 소셜 미디어 작전을 통해 인지 전쟁을 벌이고 있다고 비난했다. 이는 페이스북, 인스타그램, 틱톡, X, 서브스택(Substack)과 같은 플랫폼에서 허위 계정을 만들어 허위 정보를 퍼뜨리는 것을 포함한다.

셋째, 대만 선거에 대한 개입이 있다. 미국은 중국이 인지 전쟁 전술을 통해 대만의 선거에 영향을 미치려 한다고 비판했다. 이는 소셜 미디어에 가짜 이미지를 퍼뜨리고 군사 훈련을 실시하여 여론을 조작하려는 시도를 포함한다.

넷째, 데이터 수집에 대한 우려가 있다. 중국이 미국 시민의 개인 및 생체 데이터를 대량으로 수집하고 있으며, 이는 인지 전쟁 목적으로 사용될 수 있다는 우려가 있다.

다섯째, 군사 연구가 있다. 미국 정부는 “뇌 제어 무기”와 기타 인지 전쟁 기술을 연구하는 중국의 연구소와 기업을 블랙리스트에 올렸다.

여섯째, 선전과 허위 정보가 있다. 미국은 중국이 광범위한 선전과 허위 정보 작전을 통해 미국과 서구 국가들을 전복시키려 한다고 비난했다.

일곱째, 이념적 영향력이 있다. 일부 미국 전문가들은 중국이 서구 담론에 자국의 이념적 개념과 용어를 확산시키려는 노력을 인지 전쟁의 일환으로 보고 있다.

이러한 비난과 비판은 미국의 관료와 전문가들에 의해 제기되었으나, 중국은 종종 이러한 주장에 대해 부인하거나 대안적인 서사를 제시한다. 인지 전쟁 활동의 전체 범위와 성격은 은밀한 성격 때문에 명확히 증명하기 어려울 수 있다.

X. 향후 인지전을 둘러싼 국가들 간의 경쟁

인공지능 및 신경 기술과 같은 신흥 기술이 인지전의 핵심 촉진제로 작용함에 따라, 국가들은 이러한 이중 용도 기술을 군사화할 필요성을 느낄 가능성이 크다. 이는 결과적으로 이 분야에서의 기술적 군비 경쟁을 가속화할 수 있다. 군비 경쟁의 역학은 이러한 기술의 개발 및 사용에 있어 윤리적, 법적, 규범적 제약을 무시하게 만드는 결과를 초래할 수 있다. 국가들이 뒤처질 것을 두려워하기 때문에, 이들은 국가 안보 문제에서 가장 파괴적인 능력을 활용하기 위해 엄격한 규제를 포기할 가능성이 높다. 이는 결국 마음의 군사화를 정상화시키고, 인지 영역에서 무책임한 행동이 제한 없이 일어나는 상황을 초래할 수 있다. 이러한 역학은 정신적 프라이버시, 조작, 영향력, 자기 결정권, 그리고 통합성에 대한 규범의 형성을 방해할 수 있다. 이러한 기술들이 전복의 주요 촉진제이자 국가 안보의 필수 도구가 됨에 따라, 이미 나타나고 있는 강대국 간의 기술적 분리 경향이 더욱 가속화될 수 있다.

인지전은 또한 공격과 방어 간의 균형에 대한 전반적인 영향을 미칠 수 있다. 공격-방어 균형 이론의 지지자들은 국제적 안정성이 우리가 공격이 우위에 있는 환경에 살고 있는지, 혹은 방어가 우위에 있는 환경에 살고 있는지에 따라 달라지며, 전자는 불안정성과 갈등을, 후자는 안정성과

평화를 촉진한다고 주장한다. 이들은 또한 이러한 균형에 영향을 미치는 주요 변수가 기술이라고 주장하며, 기술은 균형을 어느 쪽으로든 기울게 할 수 있다. 이러한 측면에서 효과적인 인지전을 수행할 수 있는 능력은 에스컬레이션 압력을 증가시키고 이러한 능력의 공격적 사용을 선호할 수 있다. 실제로, 인지 공격이 언제 발생하는지 감지하는 것과 이에 방어하는 능력에는 본질적인 어려움이 있다. 특정 순간에 인지 공격을 받고 있을 가능성은 인지전을 수행하는 데 있어 '선제 사용' 동기를 부여할 수 있다.

전쟁에서 대리자를 사용하는 경우, 특히 기술적 대리자를 사용하는 경우, 신빙성 있는 부인(plausible deniability)은 주요한 인센티브이다. 인지전의 경우, 이를 수행하는 데 사용되는 도구의 정의되지 않고 규제되지 않은 성격과 그 효율성은 이러한 신빙성 있는 부인을 촉진하고 '선제 사용'을 더욱 부추길 수 있다. 또한, 인지전을 가능하게 하는 기술들이 민주화되고 확산되면서, 거대한 수의 사람들에게 영향을 미칠 수 있는 능력이 비국가 행위자, 기업, 심지어 개인에게까지 접근 가능해짐에 따라 인지전은 더욱 불안정해질 것이다(Rickli et al. 2023).

XI. 인지전에 대한 대응책

2024 년 뮌헨안보회의에서는 특히 인지전과 관련된 논의가 중요한 주제로 떠올랐다. 인지전은 적대 세력이 상대방의 인식, 행동, 그리고 결정을 조작하거나 통제하려는 현대의 복잡한 전술을 의미한다. 이러한 논의는 현대 전쟁이 점점 더 물리적인 전투에서 정보전과 심리전으로 변화하고 있음을 반영한다.

인지전에 대한 대응 논의는 주로 인지전이 제기하는 위협을 인식하고 이에 대처하기 위한 전략을 마련하는 데 초점을 맞추고 있다. 이 논의에서 중요한 요소는 시민들의 인식 제고, 인지전 및 이를 가능하게 하는 기술에 대한 규제 강화, 그리고 민주적 회복력을 강화하기 위해 신기술을 활용하는 것이다.

첫째, 인지전에 대한 인식과 이해를 높이는 것이 중요하다. 인지전은 그 개념이 아직 명확하게 정의되지 않았기 때문에, 학계와 산업계, 방위 부문이 협력하여 이를 구체적으로 정의하고 그 메커니즘을 이해하는 것이 필수적이다. 또한, 시민들이 정보 환경을 비판적으로 평가하고, 잘못된 정보나 조작된 콘텐츠에 저항할 수 있는 능력을 키우는 것이 중요하다.

둘째, 인지전 및 이를 가능하게 하는 기술에 대한 규제 프레임워크를 개발하는 것이 필요하다. 인지전의 위협에 대한 인식을 높이는 것 외에도, 정책 입안자들은 이러한 위협에 대응하기 위한 포괄적인 규제 체계를 구축해야 한다. 이를 위해서는 인지전의 다양한 형태와 전술을 포함하는 명확한 정의가 필요하며, 현재의 국제법이 인지 영역에 어떻게 적용될 수 있는지 논의해야 한다. 또한, 기존의 법적 틀이 적용되지 않는 경우, 새로운 법적 프레임워크를 개발하여 인지전 활동에 대한 책임과 책임소재를 규명해야 한다.

셋째, 신기술의 힘을 활용하여 사회적 및 민주적 회복력을 강화하는 것이 필요하다. 인지전의 위협을 효과적으로 막기 위해서는 내부적인 취약성을 해결하고, 신기술을 활용하여 민주적 과정을 강화해야 한다. 예를 들어, 복잡한 AI 시스템은 광범위한 여론 분석을 통해 효과적인 민주적 참여와 협력을 촉진할 수 있으며, 가상 현실(Virtual Reality: VR) 기술은 무의식적 편견을 훈련하는 데 사용될 수 있다. 이러한 기술을 적절히 활용함으로써 사회적 회복력을 강화하고 인지전에 대한 방어력을 높일 수 있다.

넷째, 국제적 협력의 강화가 필요하다. 인지전은 국경을 초월하여 이루어지기 때문에, 국제 사회의 협력 없이는 효과적인 대응이 어렵다. 국가 간 정보 공유와 협력을 통해 인지전에 대한 공동 대응 방안을 마련해야 한다. 특히, 나토와 같은 국제기구들은 인지전에 대한 대응 전략을 수립하고 회원국 간의 협력을 촉진하는 데 중요한 역할을 할 수 있다.

다섯째, 윤리적 기준을 확립하고 준수하는 것이 중요하다. 인지전에 대응하는 과정에서 윤리적 기준을 무시하면, 대응 자체가 또 다른 형태의 인지전으로 비취질 수 있다. 따라서 정보 작전의 투명성을 높이고, 개인의 프라이버시와 인권을 존중하는 방식으로 대응해야 한다. 이는 장기적으로 인지전에 대한 사회적 신뢰를 유지하는 데 필수적이다(Pujol et al. 2024; Ibrahim et al. 2023).

결론적으로 향후 AI 를 이용한 인지전은 강대국들 간, 특히 미중 간의 전략 경쟁에서 매우 중요한 역할을 할 것이며, 이는 여러 가지 파국적 효과를 초래할 수 있다. 여론 조작과 선거 개입, 가짜 정보 확산 등을 통해 사회적 불안정을 초래하고, 정치적 혼란을 야기할 수 있다. 이러한 위협을 최소화하기 위해서는 국제적인 협력과 규제가 필요하며, AI 기술의 안전한 개발과 사용을 위한 윤리적 기준과 안전 조치가 강화되어야 한다. 그러나 국제적 차원에서 합의에 이르기에는 여전히 많은 장벽이 필요하고 국제정치적 효용이 존재하는 한 상당 기간 인지전을 둘러싼 공방이 지속될 것이다. ■

참고문헌

- 유준구. 2019a. “자율살상무기체계의 논의 동향과 쟁점.” 「IFANS 정책연구시리즈」 2019-18. 12월.
- _____. 2019b. “자율살상무기(LAWS)체계 규범 창설 논의의 쟁점과 시사점.” 「IFANS 주요국제문제분석」 2019-54. 12월 19일.
- Autonomous Weapons. 2024. “The Risks of Autonomous Weapons.” Autonomous Weapons, <https://autonomousweapons.org/the-risks/>.
- Bajak, Frank. 2023. “Pentagon’s AI Initiatives Accelerate Hard Decisions on Lethal Autonomous Weapons.” *Associated Press*. 11월 26일. <https://apnews.com/article/us-military-ai-projects-0773b4937801e7a0573f44b57a9a5942>.
- Britt Baughman, Josh. 2023. “How China Wins the Cognitive Domain.” China Aerospace Studies Institute. 1월 23일. <https://www.airuniversity.af.edu/CASI/Display/Article/3273289/how-china-wins-the-cognitive-domain/>.
- Beauchamp-Mustafaga, Nathan. 2019. "Cognitive Domain Operations: The PLA's New Holistic Concept for Influence Operations." *China Brief* 19: 16. <https://jamestown.org/program/cognitive-domain-operations-the-plas-new-holistic-concept-for-influence-operations/.ain>
- Cameron, Hugh. 2024. “China’s Killer Robots Are Coming.” *Newsweek*. 6월 26일. <https://www.newsweek.com/china-killer-robots-unitree-robotics-1917569>.
- Dresp-Langley, Birgitta. 2023. “The Weaponization of Artificial Intelligence: What the Public Needs to Be Aware Of.” *Frontiers in Artificial Intelligence* 6. <https://doi.org/10.3389/frai.2023.1154184>.
- Firth, Joseph, John Torous, Brendon Stubbs, Ivan P. Ipser, P. Murali Doraiswamy, and Felipe Schuch. 2019. "The 'Online Brain': How the Internet May Be Changing Our Cognition." *World Psychiatry* 18, 2: 119–129.

- Hale, Amber. 2023. "Clausewitzian Theory of War in The Age of Cognitive Warfare." *The Defence Horizon Journal*. <https://tdhj.org/blog/post/clausewitz-cognitive-warfare/>.
- Huang, Ruiyang, Xiaoqing Zheng, Yuting Shang, and Xiangyang Xue. 2023. "On Challenges of AI to Cognitive Security and Safety." *Security and Safety* 2. <https://doi.org/10.1051/sands/2023012>.
- Hung, Tzu-Chieh, and Tzu-Wei Hung. 2020. "How China's Cognitive Warfare Works: A Frontline Perspective of Taiwan's Anti-Disinformation Wars." *Journal of Global Security Studies* 7, 4: 1-18.
- Human Rights Watch. 2023. "A Review of the 2023 US Policy on Autonomy in Weapons Systems." Human Rights Watch. 2월 14일. <https://www.hrw.org/news/2023/02/14/review-2023-us-policy-autonomy-weapons-systems>.
- Ibrahim, Fabio, Steffen Rhode, and Monika Daseking. 2023. "A Systematic Review of Cognitive and Psychological Warfare." *The Defence Horizon Journal*, 12월 1일. <https://tdhj.org/blog/post/cognitive-psychological-warfare>.
- Johnson, James. 2023. *AI and the Bomb: Nuclear Strategy and Risk in the Digital Age*. Oxford: Oxford University Press.
- Kallenborn, Zachary. 2021. "Meet the Future Weapon of Mass Destruction, the Drone Swarm." *Bulletin of the Atomic Scientists*. 4월 5일. <https://thebulletin.org/2021/04/meet-the-future-weapon-of-mass-destruction-the-drone-swarm/>.
- Lamberth, Catherine A. 2023. "Decoding the Defense Department's Updated Directive on Autonomous Weapons." *Lawfare*. 8월 17일. <https://www.lawfaremedia.org/article/decoding-defense-departments-updated-directive-autonomous-weapons>.

- Linney, Charlie, and Xiaomin Tang. 2024. "AI and Autonomous Weapons Systems: The Time for Action Is Now." *Saferworld*. 5월 15일.
<https://www.saferworld.org.uk/resources/news-and-analysis/post/1094-ai-and-autonomous-weapons-systems-the-time-for-action-is-now>.
- Miller, Seumas. 2023. "Cognitive Warfare: An Ethical Analysis." *Ethics and Information Technology* 25: 46. <https://doi.org/10.1007/s10676-023-09717-7>.
- Morelle, Marie, Julien Cegarra, Damien Marion, and Jean-Marc André. 2023. "Towards a Definition of Cognitive Warfare." Paper presented at the Conference on Artificial Intelligence for Defense, DGA Maîtrise de l'Information. <https://hal.science/hal-04328461>.
- Pujol Chica, Irene, and Quynh Dinh Da Xuan. 2024. "The Battle for the Mind: Understanding and Addressing Cognitive Warfare and Its Enabling Technologies." IE Center for the Governance of Change. <https://coilink.org/20.500.12592/zs7hb9h>.
- Rickli, Jean-Marc, Federico Mantellassi, and Gwyn Glasser. 2023. Peace of Mind: Cognitive Warfare and the Governance of Subversion in the 21st Century. GCSP Policy Brief No. 9, Geneva Centre for Security Policy. <https://www.gcsp.ch/publications/peace-mind-cognitive-warfare-and-governance-subversion-21st-century>.
- Sârbu, Annamaria, and Anca Gavrilăş. 2023. "Using Artificial Intelligence Tools for Obtaining Cognitive Warfare Advantages." *The Defence Horizon Journal*. 10월 23일. <https://tdhj.org/blog/post/artificial-intelligence-cognitive-warfare-twitter/>.
- Tripathi, Prateek. 2024. "When AI Crosses the Line: The Impending Threat of Autonomous Weapons Systems." *Observer Research Foundation*. 5월 6일.
<https://www.orfonline.org/expert-speak/when-ai-crosses-the-line-the-impending-threat-of-autonomous-weapons-systems>.
- U.S. Department of Defense. 2023. "Deputy Secretary of Defense Kathleen Hicks Keynote Address: 'The Urgency to Innovate' (As Delivered)." 8월 28일.
<https://www.defense.gov/News/Speeches/Speech/Article/3507156/deputy-secretary-of-defense-kathleen-hicks-keynote-address-the-urgency-to-innov/>.

Vasist, Pramukh Nanjundaswamy, Debashis Chatterjee, and Satish Krishnan. 2023. "The Polarizing Impact of Political Disinformation and Hate Speech: A Cross-country Configural Narrative." *Information Systems Frontiers*, 1–26.

Zhang, Guangshen, Yongli Li, Haoxian Wang. 2022. "A Brief Analysis of the Basics of Cognitive Domain Operations [浅析认知域作战的基本要义]." *PLA Daily*. 9월 8일.

http://www.81.cn/jfjbmap/content/2022-09/08/content_323692.htm.

■ **저자:** 전재성_EAI 국가안보연구센터 소장, 서울대학교 정치외교학부 교수.

■ **담당 및 편집:** 박지수_EAI 연구원

문의: 02-2277-1683 (ext. 208) jspark@eai.or.kr

본 대담록을 인용할 때에는 반드시 출처를 밝혀주시기 바랍니다.

EAI는 어떠한 정파적 이해와도 무관한 독립 연구기관입니다.

EAI가 발행하는 보고서와 저널 및 단행본에 실린 주장과 의견은 EAI와는 무관하며 오로지 저자 개인의 견해를 밝힙니다.

발행일 2024년 9월 6일

“[AI와 신문명 표준 스페셜리포트] 군사도전② AI 기반 자율무기체계, 인공지능의 발전과 군사안보질서의 변화”
979-11-6617-798-9 95340

재단법인 동아시아연구원

03028 서울특별시 종로구 사직로7길 1
Tel. 82 2 2277 1683 Fax 82 2 2277 1684

Email eai@eai.or.kr

Website www.eai.or.kr