

EAI 신년기획 특별논평 시리즈

한국외교 2025 전망과 전략 ⑨

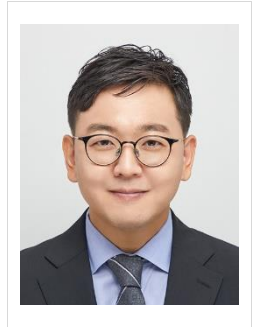
2025 세계안보 현실 대 인식: 인공지능의 군사적 이용과 공격송배 현상

김양규 (동아시아연구원; 서울대학교)

2025 세계안보 현실 대 인식: 인공지능의 군사적 이용과 공격송배 현상

김양규

동아시아연구원 수석연구원, 서울대 정치외교학부 강사



I. 트럼프 2.0 시대 세계군사 환경 전망

5일 후면 트럼프 대통령이 백악관으로 복귀한다. 캐나다, 아르헨티나, 헝가리, 이탈리아 총리 또는 대통령에서부터 애플, 오픈AI, 메타, 아마존, 일본 소프트뱅크그룹 기업 최고경영자에 이르기까지 대통령 인수위 팀이 꾸려진 플로리다 마러라고를 앞다투어 찾고 있다. 다양한 국가 및 기업의 리더십들이 10% 이상의 보편관세 부과를 예고한 트럼프의 분노를 피하기 위해 안간힘을 쓰고 있다는 분석이다(이재립 2025). 이 가운데 한국은 윤석열 대통령의 탄핵소추안이 지난 12월 14일 가결된 데 이어, 12월 27일에는 한덕수 대통령 권한대행 국무총리까지 직무가 정지되면서 최상목 경제부총리 겸 기획재정부 장관이 권한대행의 대행을 맡는 초유의 사태를 맞았다. 미 주요 싱크탱크의 전망처럼, 트럼프 취임 이후 100시간 내로 한국 관련 주요 사안인 주한미군, 관세, 반도체법 문제에 대한 미국의 정책들이 결정된다면(박성민 2024), 대표성을 띠고 마러라고를 방문할 수 있는 국가 리더십이 부재한 한국은 다른 국가들보다 불리한 상황에서 2025년을 시작하게 되었다고 보아야 할 것이다.

그러나 한편으로는 트럼프 2기나 한국 탄핵 국면과 같은 리더십 변수와 무관하게 세계안보 질서 저변에서 작동하는 구조적인 변수가 실질적으로는 더 중요하다고 지적할 수 있다(Waltz 1979). “혁명이 지형을 바꾸지 않고, 혁명은 지리적 필요도 바꾸지 않는다(Revolutions do not change geography, and revolutions do not change geographical needs)”는 애틀리(Clement Attlee) 전 영국 총리의 말처럼, 트럼프 개인의 특이성에도 불구하고 미국 전략적 이익의 근본적인 속성은 바뀌지 않고, 한국의 국익도 일정 수준에서는 국내정치적 혼란과 구별되어 존재하는 측면이 있다. 특히 규칙이나 가치 기반 외교보다 선명한 거래주의적 접근(transactional approach)을 취하는 트럼프주의의 외교적 속성을 고려할 때(전재성 2025), 리더십 개인 변수보다 더 중요한 것은 실제 현실의 구

조직 변화이며, 이를 맥락에 깔고 형성되는 각국의 전략적 이해관계일 수 있다. 한국이 국내정치적 혼란을 극복하고 외교안보 정책을 정상적으로 다시 가동하는 시기를 대비하여, 이런 세계안보 환경을 결정할 구조적 변수들에 대한 분석이 필요하다.

본고는 2025년 세계군사질서의 변화 방향을 국가 대전략이나 리더십의 특성, 국내정치적 요인보다 군사기술의 변화 차원에서 살펴본다. 특히 인공지능(artificial intelligence: AI)의 군사적 이용이 현재 공격-방어 균형(offense-defense balance) 측면에서 어떤 변화를 가져오고 있는지, 자율 무기체계(autonomous weapon systems), 사이버 안보, 핵무기-AI 넥서스 차원에서 검토하고, 최근 21세기 버전의 ‘공격숭배(cult of the offensive) 현상’이 미국과 중국의 정책 커뮤니티 내에서 부상하고 있는 이유에 대해 분석한다. 아울러, 궁극적으로는 군사기술의 변화보다 그것을 어떻게 인식하는지가 향후 세계안보 환경의 속성을 결정하는 중요한 요인이 된다는 점을 지적하고, 한국의 안보정책 방향을 제시한다.

II. 군사기술 변화: AI의 군사적 이용 현실 vs. 인식

리더십 변수나 국내정치 변수와 구별되는 구조적 층위에서 작동하는 변수 가운데 군사안보질서 변화에 영향을 미치는 핵심적인 변수는 ‘공격-방어 균형’이다. 이는 동일한 자원을 투입했을 때 공격과 방어 중 어느 쪽이 더 효과적인지의 문제에 관한 것이다. 저비스(Robert Jervis)는 국제 협력의 용이성에 핵심적인 영향을 미치는 안보 딜레마(security dilemma)가 얼마나 심각한지, 그 수준이 바로 이 공격-방어 균형에 의해 결정된다고 강조한 바 있다(Jervis 1978, 187-199). 다양한 국제위기 국면에서 공격 우위는 확전 위험을 증대시키고, 방어 우위는 전략적 안정성을 높이는 경향을 보였다. 이 때문에 공격-방어 균형은 전쟁 발발 가능성, 동맹 정치 동학, 군비 경쟁 등 안보 정세 차원에서 폭넓은 함의를 가진다(김양규 2024a).

중장기적으로 향후 공격-방어 균형 변화에 영향을 미치는 요인들 가운데 가장 주목해야 할 기술은 AI와 그 군사적 활용이다. AI는 다른 기술들의 변화를 이끄는 ‘가능케 해주는 기술(enabler)’이자, 기존 기술을 효과를 강화하는 ‘능력 증폭기(force multiplier)’로서 기능하며, 기술 발전의 기반이자 메타 기술이라는 특성을 보인다(Horowitz 2018). “작은 마당, 높은 담장(small yard, high fence)”으로 표현되는 미국의 디리스크링(de-risking) 전략과 미중 첨단기술 경쟁의 핵심에 AI가 있는 것도 바로 이러한 이유 때문이다(손열 외 2023; 김양규 2024a; 배영자 2025). 최근 발행되는 미국

의 전략문건들에는 AI 기술이 미국의 패권 유지 및 중국에 대한 미국의 초격차 우위를 지속하기 위한 핵심적인 수단이 될 것이라는 기대가 깔려 있다(김양규 2024b; Jacobsen and Liebetrau 2023).

AI는 방대한 데이터를 분석해 신속하고 정확한 결정을 내리며, 군사 작전에서 효율성을 높이는 ‘전력 증강자’ 역할을 하고, 특히 전장 정보 처리, 표적 탐지, 적 전술 대응 속도 향상 등의 분야에서 혁신적인 변화를 일으킬 것으로 예측된다. 또한 미국의 “통합억지(integrated deterrence)”와 중국의 “지능화전(intelligentized warfare)” 개념처럼, AI는 다영역 작전 능력을 강화하며, 육·해·공·우주·사이버 영역을 통합적으로 운영하는 데 중요한 역할을 한다(손열 외 2023; 김양규 2024b; 배영자 2025).

1. 세계안보 환경의 객관적 현실: AI-사이버, AI-자율무기체계, 핵-AI 넥서스

AI의 군사적 이용은 공격-방어 균형 차원에서 어떠한 변화를 가져오게 될까? 이에 답하기 위해서는 기존 전력에 AI가 통합되어 일어나고 있는 최근의 변화들을 사이버 공격-방어 역량 및 자율무기체계, 핵-AI 넥서스 차원에서 살펴볼 필요가 있다(김양규 2024b; 전재성 2024).

첫째, 사이버 안보 영역에서 AI는 공격과 방어 능력을 모두 강화하고 있다(Jacobsen and Liebetrau 2023). 그런데 현재 동 문제에 대한 많은 논의들, 특히 정부 측 문건들은 사이버 안보 균형에서 AI가 공격 우위를 가져온다고 주장하는 경우를 훨씬 쉽게 발견할 수 있다. 물론 이는 기술 자체의 특성상 사이버 공격이 평시와 전시의 경계를 무너뜨리고, 그 공격이 언제 발생하는지 감지하고 방어하는 능력을 구축하는 데 현실적인 어려움이 있기 때문이기도 하다(전재성 2024). 예를 들어, 악성코드에 대한 대응의 경우 그 목적과 목표가 모호하면 이를 탐지하기 매우 어렵다는 기술적 문제가 있는데, 이에 착안하여 딥로커(Deeplocker)와 같이 AI로 구동되는 정교한 맬웨어 공격 프로그램이 개발되기도 하였다. 동 프로그램은 특정 피해자에게 도달할 때까지 의도를 숨기다가 AI 모델이 얼굴 인식이나 음성 인식과 같은 지표를 통해 대상을 식별하는 즉시 악성코드를 심는 방식으로 작동하기 때문에 탐지 및 대응이 상당히 어렵다.

AI로 강화된 사이버 공격 프로그램을 대응하기 위해서는 마찬가지로 AI 역량이 연동된 방어 프로그램이 필요하다. 현재 사이버 보안 회사들은 AI 기술이 연동된 공격 프로그램과 방어 프로그램을 함께 개발하고 있다. 적국이 사이버 공격 프로그램을 어떻게 활용할지 알지 못하면 방어 프로그램 또한 개발할 수 없기 때문이다. 공격-방어 균형 문제를 놓고 사이버 보안 프로그램 개발자라는

단일 행위자 내부에서 창과 방패의 대결이 계속 진행되는 셈이다. 특히, 사이버 방어 역량을 키우기 위해 AI 기술의 통합을 강화할수록 시스템 간 상호연결성이 커지고 ‘적의 공격 대상으로 삼을 수 있는 대상(attack surface)’도 함께 확대되어 방어 체계 전체의 취약성이 심화되는 역설은 사이버 안보 차원에서 공격과 방어 중 무엇이 더 우위에 있는지 확정적으로 논의하기 매우 어렵게 만든다.

둘째, 자율무기체계의 경우도 사이버 안보 영역과 마찬가지로 AI의 활용으로 인해 공격이나 방어 중 하나의 우위를 일방적으로 야기하지는 않을 것으로 보인다. 자율무기체계가 현재 많은 주목을 받고 있는 핵심적인 이유는 전쟁비용을 획기적으로 줄이기 때문이다. 드론이나 킬러 로봇 기술이 보여 주듯, 자율무기체계는 대량 생산이 가능하고, 확장성과 확산성이 뛰어나며, 한 명의 조작자가 여러 무기 체계를 동시에 운용할 수 있기에 적에게 가할 수 있는 피해의 규모가 무기 운용자의 수가 아니라 무기 자체의 수량에 의해 결정된다. 각 로봇이 독립적으로 작동하면서도 효율성을 유지할 수 있어, 부수적 피해(collateral damage)를 최소화하면서 원하는 표적만을 제거할 수 있다(전재성 2024). 뿐만 아니라, 군사작전의 성공을 위해 필요한 인간 병사의 개입을 최소화할 수 있어 전쟁의 장기화에 따른 전사자 증가나 그로 인한 국내 여론 악화를 우려할 필요가 없어진다. AI로 강화된 자율무기체계가 가져오는 전쟁 비용의 감소야말로 진정한 의미의 군사 혁신이라고 주장하는 연구도 있다(Schneider and Macdonald 2024). 전쟁 수행 비용의 감소는 공격-방어 균형에서 공격의 이점을 강화한다고 볼 수 있다.

그런데 AI로 인해 강화된 자율무기체계는 방어 우위를 가져오는 양상도 보인다. 예를 들어, 대만해협을 중심으로 미중 간 충돌이 일어날 경우를 한정해 생각해 보면, 최근의 기술 변화는 방어 우위의 환경을 조성하고 있다. 중국이 대만에 대해 상륙작전이나 봉쇄작전을 시도하더라도, 초정밀 장거리 공격 능력과 정보·감시·정찰(Intelligence, Surveillance and Reconnaissance: ISR) 능력, 자율무기체계를 활용하는 미국의 첨단 방어 체계로 인해 중국은 자신의 군사적 목표를 달성하기 어렵고, 부분적으로 달성하더라도 엄청난 전쟁 비용을 치러야 할 것으로 예상된다. 반대로 중국 또한 장거리 정밀 타격 능력을 통해 제1도련선 내 미국의 주요 지휘부와 공군 기지에 심각한 피해를 입힐 수 있어 유사시 중국 본토를 대상으로 하는 미군의 작전 수행 능력을 상당 부분 제한할 수 있다. 다시 말해, 대만해협에서 미중 간 분쟁 발생시 양국 중 어느 쪽이든 상대방에게 공격을 감행하면 매우 높은 비용을 지불해야 하는 상황인 것이다(Fravel and Heginbotham 2024). 아울러, 생성형 AI가 제대로 작동하기 위해서는 엄청난 양의 데이터가 필요하다는 점에 주목하여, 방어국이 적국보다 자

국 지형과 군사기지에 대해 더 상세한 정보를 보유하고 있기 때문에 AI로 강화된 무인 자율무기체계는 방어에 더 유리한 기술이라는 주장도 제기되고 있다(King 2024).

셋째, 핵무기-AI 넥서스는 ‘1차 공격능력(first strike capability 또는 counterforce capability)’과 ‘2차 공격능력(second strike capability 또는 countervalue capability)’을 모두 강화하는 특징이 있다. ISR 능력 향상, 적국 핵 지휘통제(Nuclear Command and Control: NC2) 시설에 대한 사이버 공격 및 정보 교란, 재래식 대군사타격(counterforce capability)의 정밀도 향상 등은 상대방의 핵 역량을 파괴하여 1차 공격능력을 강화한다. 반대로, 조기경보 및 상황 인식 능력 강화, 사이버 방어, 드론 스웜의 방어적 활용, 핵무기의 자동 보복 시스템 등은 적의 핵 공격에 대한 대응과 방어 측면에서 2차 공격능력을 제고하는 효과가 있다(김양규 2024b). 1차 공격능력과 2차 공격능력 중 무엇이 공격이나 방어에 기여하는 것인지 올바르게 이해하기 위해서는 쌍방이 서로에 대해 신뢰성 있는 2차 공격능력을 구비하였을 때 상호확증파괴(Mutual Assured Destruction: MAD)가 구축되고, 이것이 해당 국가 간 전략적 안정성을 높인다는 점을 고려해야 한다. 즉, 2차 공격능력은 MAD의 신뢰성을 높여 전략적 안정성을 높이고, 1차 공격능력은 이를 파괴하기에 불안정성을 심화시킨다. 이런 맥락에서 AI가 결합된 핵무기 능력이 2차 공격능력뿐 아니라 이를 무력화시키는 1차 공격능력을 동시에 강화하고 있다는 점은 핵무기 균형 차원에서도 AI 기술의 도입이 공격이나 방어 중 어느 하나의 우위를 가져온다고 단정적으로 말하기 어렵게 만든다.

상기 내용을 토대로 AI의 군사적 이용에 대한 최근 연구들이 공통적으로 지적하는 사항들을 다음 세 가지로 정리할 수 있다. 첫째, AI 기술은 어떤 측면에서 살펴보아도 그 자체만으로 공격이나 방어 우위를 가져온다고 보기 어렵다. 이는 해당 기술이 가진 범용성 및 능력 증폭기적인 속성을 고려한다면 매우 자연스러운 귀결이다. 애초에 군사기술 혁신을 연구하는 학자들은 신기술의 군사적 이용과 이에 대응하는 상대국의 적응 사이의 작용-반작용(action-reaction)을 고려할 때, 기술 변수 하나로 공격-방어 균형 문제에 대한 답을 확정적으로 내릴 수 없다는 점을 강조한다(Biddle 2023). AI의 범용성은 이러한 작용-반작용 문제를 더욱 선명하게 드러낸다고 볼 수 있다.

둘째, AI 기술의 효율성-안정성 긴장관계이다. 기술 변화에 따른 공격 우위 양상의 부각에 대응하여 AI 기반 방어 역량을 제고하기 위해서는 인간의 개입 및 통제 영역을 제한하고(human out of the loop) 기계의 자율성(autonomy)을 높여야 한다. 그런데 이렇게 되면 전쟁 비용은 축소되지만 ‘의도하지 않은 확전’ 가능성은 커진다. 반대로, 책임 있고 안전한 방식으로 AI를 활용할수록 AI 역량을 군사력에 접목시키는 효과는 줄어들고 이로 인한 효율성 감소는 애초에 AI를 군사적으로

활용하고자 하는 목적에 반하는 결과를 가져온다. 특히, 적국은 안전한 방식의 AI 기술 적용을 포기하고 기계의 자율성에 맡기는 효율성에 치중한 전략을 선택하였는데 아군은 책임 있는 AI 사용을 강조하여 인간 지휘관의 개입과 통제를 강화할 경우, 적국의 AI-사이버, AI-자율무기체계, AI-핵무기 시스템의 효율성이 아군의 AI 역량을 압도해 버리는 문제에 봉착할 수 있다.

셋째, AI의 군사적 이용의 결과가 공격이나 방어 중 일방의 우위로 귀결되는지 단정하기 어렵고, 효율성-안정성 길항 관계에 따른 역설이 존재하는 것이 객관적인 현실이지만, AI로 강화된 사이버 공격 능력이나 자율무기체계에 AI가 결합된 형태의 군사역량에 대한 연구들을 보면 AI의 군사적 사용 이후 ‘공격 우위 시대’가 올 것이라고 전망하는 경우가 많다. 특히 미국과 중국 정치 엘리트들의 인식을 보면, AI가 기존 군사 역량에 통합될 때 공격 우위를 가져온다는 기대가 만연해 있어, 마치 1차 세계대전 이전의 “공격숭배(cult of the offensive)” 현상과 비슷한 양상을 보인다(Selden 2024).

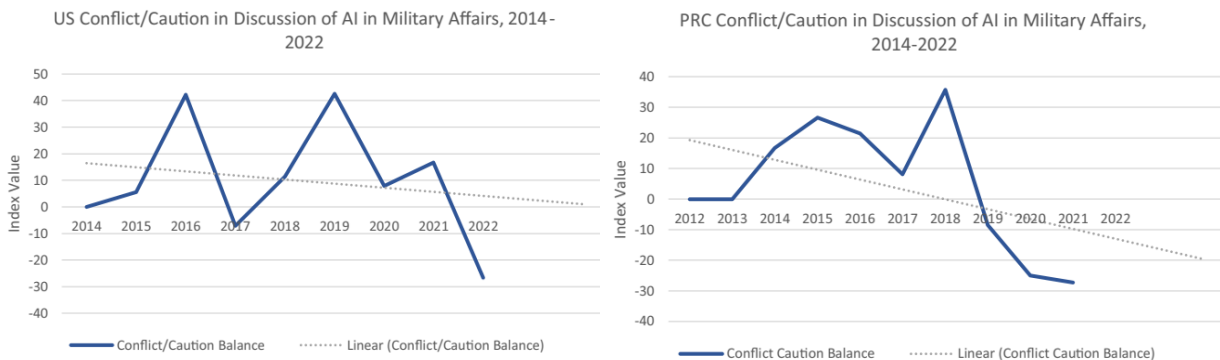
2. 세계안보 현실에 대한 해석: 분쟁 불가피성에 대한 인식과 공격숭배 현상

왜 정책결정권자들은 군사기술 변화의 객관적 현실을 정확하게 인지하지 못하고 이처럼 공격 우위의 시대가 올 것이라고 예상하는가? 공격-방어 균형의 현실과 인식 사이 뚜렷한 간극이 발생하는 현상은 동 변수의 중요성을 강조한 저비스의 초창기 연구에서부터 지적된 문제이다. 저비스는 ‘공격-방어 균형’과 ‘공격-방어 구분(offense-defense differentiation)’을 중심으로 이론적으로 가능한 4가지 세계를 제시하는데, ‘방어 우위’ 상황에 ‘공격과 방어 기술의 구분이 가능’할 경우 안보딜레마가 존재하지 않는 “매우 안정적인(doubly stable)” 안보 환경을 형성한다고 설명한다. 저비스는 인류 역사상 이렇게 안정적인 여건이 형성되었던 거의 유일한 시기로 1차 세계대전 이전 20세기의 첫 10년을 꼽는다. 그리고 기관총 무기체계와 참호전의 의미를 제대로 이해했다면, 유럽 열강들은 모두 당시 다른 안보 정책들을 선택했을 것이고, 1차 대전은 발발하지 않았을 것이라고 주장한다. 그럼에도 불구하고 당시 주요국의 거의 모든 전략가들은 공격이 우위에 있다고 믿었고, 공격과 방어 기술은 따로 구분되지 않는다고 생각했다. 이러한 ‘인식’이 물리적 ‘현실’을 압도했던 것이다(Jervis 1978, 211-214).

최근 연구는 정책결정권자들이 군사기술 차원에서 일어나는 현실의 변화를 제대로 읽어내지 못하는 이유에 대해 “모호한 기술 발전과 불가피한 분쟁에 대한 인식의 증가(mix of ambiguous technological advances and growing sense of inevitable conflict)가 서로를 강화하는 인식의 악순환을 형성”하기 때문이라고 설명한다(Selden 2024, 6). 해당 연구는 미국 대통령, 국방부 장관의 인터

뷰, 연설, 국방부 공식 전략문건, 각종 정부 위원회 보고서, 그리고 의회 성명 등의 문건과 중국 정부의 안보 전략, 정부 보고서, 시진핑 주석의 인터뷰, 고위급 중국공산당 간부들 연설, 그리고 신화통신과 같은 중국 국영 언론의 논평 등의 문건 가운데 AI의 군사적 이용에 관한 논의를 2014년부터 2022년까지 샘플링하여 분석한 결과 이러한 결과를 얻었다고 보고한다. 아울러, 미국과 중국의 정책결정권자들의 양국 간 분쟁 가능성에 대한 인식이 [그림 1]과 같은 양상을 보였다고 설명한다. 점선 아래로 파란색 실선이 내려가는 부분이 양국 간 분쟁이 불가피하다는 인식이 압도적인 국면이다.

[그림 1] 미국 및 중국 정책결정권자들의 분쟁/자제 논의



미국의 경우 트럼프 1기 초반에 중국에 대한 적대적 인식이 강세를 보이다 2018년부터 긍정적으로 돌아서지만, 2021년 바이든 행정부 출범 이후 중국과 분쟁이 불가피하다는 인식이 급부상하는 양상을 보인다. 중국은 2019년부터 나빠진 대미 인식이 그 이후로 회복되지 않고, 양국 간 충돌과 분쟁을 피하기 어렵다는 인식이 주류를 이루는 패턴을 보인다. 미중 정치 엘리트들 사이에 이러한 인식이 고착화되고, 기술이 야기하는 공격-방어 우위 효과가 모호한 현상이 지속되면, 본 연구의 관찰이 정확하다는 가정하에 미중 리더십은 AI의 군사적 활용이 공격 우위의 세계를 가져올 것이라고 인식하게 될 가능성이 크다. 이러한 주장은 동기기반 추론(motivated reasoning)에 대한 기존 연구들이 지적하는 바와 같이 정책결정권자들은 “행동이 불가피하다고 느낄 경우(feel compelled to act),” 부정(denial), 선택적 해석(selection) 및 기타 심리적 기제를 적용하여 적의 의지와 의도에 대한 모든 신호를 의도적으로 무시하는 양상을 보인다는 지적(Lebow 1981)과도 같은 맥락에 있기 때문에 상당히 설득력이 있다. 적과 반드시 싸워야만 하기 때문에, 기술발전이 공격에 유리한 여건이 구비되었다고 믿고 싶은 것이다.

III. 2025 미중 군비경쟁과 한국의 국방정책

지도자들이 현재 자국이 공격 우위의 군사기술적 환경에 놓여있다고 믿을 때, 전쟁으로 얻을 수 있는 이득이 크다고 기대하게 되고, 군비경쟁은 심화되며, 동맹 질서는 고착화된다. 안보딜레마는 악화되고, 현상유지 세력들 간에도 분쟁이 일어날 가능성이 높아진다(Jervis 1978, 189-190). 1차 세계대전 이전 유럽 열강들의 지도자들이 그러했듯, 인식의 힘은 물리적 현실을 거스르기도 한다. 그렇다면 결국 AI의 군사적 활용에 따라 변화하는 현실의 실제적 진실보다, 그것을 해석하는 지도자들의 인식이 더 중요할 수 있다는 결론에 이른다. 그렇다면 트럼프의 귀환과 한국 국내정치 혼란 및 정치 리더십의 위기는 심각한 문제로 받아들여야만 한다.

물론 2025년 내에 미중 간 직접적인 군사충돌이 발생할 가능성은 상당히 낮다. 2024년 11월 페루 리마에서 열린 아시아-태평양경제협력체(Asia-Pacific Economic Cooperation: APEC) 정상회의를 계기로 가진 바이든-시진핑 정상회담의 결론도 “미중 경쟁이 분쟁과 대립으로 빠지는 것을 막고 관리(manage competition responsibly and prevent it from veering into conflict or confrontation)”하기 위한 양국 소통 채널을 지속적으로 가동해야 한다는 것이었고(The White House 2024), 왕이 외교부장이 12월 17일 개최된 ‘2024년 국제형세와 중국 외교’ 토론회에서 강조한 5대 외교 원칙 가운데에서도 제1원칙은 “평화”였다(Ministry of Foreign Affairs, The People’s Republic of China 2024). 그러나 앞서 논의된 군사기술 차원의 변화와 이를 바라보는 미중 양국 리더십의 인식을 염두에 둘 때, 이것은 “우리 시대의 평화(peace for our time)”를 예고하는 것이기 보다는 ‘폭풍 전의 고요’인 면이 더 크다. 리마 미중 정상회담에서도 미국은 대만 문제, 항행의 자유, 러시아 방위산업에 대한 중국의 지원, 미국 민간 인프라에 대한 중국의 사이버 공격에 대한 불쾌감을 뚜렷하게 드러냈고, 왕이 부장의 연설의 제2원칙 “단결”과 제4원칙 “정의”에서 강조되는 것은 글로벌 사우스와 브릭스와 중국의 연대 강화 및 “공평한(equitable)” 국제질서였다. 미국의 주요 싱크 탱크들은 대만과 남중국해에서 중단기적으로 미중 간 군사적 충돌이 발생할 가능성과 이에 대한 미국의 대응 방안을 제시하는 보고서를 발행하고 있다(Cancian, Cancian, and Heginbotham 2023; O’Hanlon 2024; Sisson and Patt 2024).

현재는 미중이 핵무기 사용 결정에 대한 인간의 통제력을 유지하는 것으로 합의하였기 때문에 핵-AI 넥서스 경쟁과 이로 인한 의도하지 않은 핵 전쟁 발발 가능성은 상대적으로 낮은 편이다. 그러나 2035년까지 핵탄두 생산을 700~1,500개까지 확대하려는 중국(Kristensen, Korda, Johns, and

Knight 2024)과 잠수함발사탄도미사일(submarine-launched ballistic missile: SLBM)용 저위력 전술 핵탄두 W76-2 생산 확대, F-35에 탑재할 수 있는 개량형 저위력 전술핵폭탄 B61-12 생산 등 전술핵 능력을 강화하려는 미국의 의 정책 방향으로 인해 고전적 의미의 미중 수직적 핵확산 및 군비경쟁은 지속될 것이다. 아울러, ‘잠재적 위험을 신중하게 고려하고 책임감 있게 개발한다’는 일반적인 원칙 이외에 AI 기술 개발 및 그 군사적 이용에 대한 구체적인 합의가 결여되어 있는 가운데, AI-자율무기체계와 AI-사이버 공격·방어 능력 문제를 놓고 미중 간 경쟁은 더욱 심화될 것으로 보인다.

지난주 젠슨 황(Jensen Huang) 엔비디아 최고경영자가 양자컴퓨터의 상용화에 앞으로 15-20년은 더 걸릴 것이라고 전망함에 따라 주식 시장에서 양자컴퓨터 관련주가 폭락하였다. 주가 폭락보다 더 중요한 사실은, 만약 양자컴퓨터의 시대가 예상보다 더디 오게 될 경우 AI 기술의 중단기적 중요성은 더욱 증대된다는 점이다. 계엄과 탄핵 국면이 끝나고 한국 리더십이 안정기에 들어설 때, 장기적 관점에서 엄중한 도전과제들을 고려한 국방정책 방향의 수립이 필요하다. ‘우리 시대의 평화’가 아닌 ‘폭풍 전야’의 상황이라 해도, “결정적 10년(decisive decade)”을 지나는 동안 아직 한국에게 기회의 창은 열려 있기 때문이다.

우선 미중 군비경쟁을 전략적으로 활용할 필요가 있다. 공격 우위 환경에 있다는 믿음은 동맹국 간 협력을 강화하고 공격 자산에 대한 투자를 확대하는 강한 유인을 제공한다. 이는 트럼프주의와 별개로 움직이는 미 국방부 내의 독립적인 흐름으로 나타날 가능성이 크다. 따라서 한국은 이를 적극적으로 활용하여 AI 기술 분야를 선도하고 있는 미국과 신기술 기반 안보 협력을 더욱 강화해야 한다. 작년 10월 신설된 전략사령부를 중심으로 한국의 전략 자산을 통합적으로 활용하는 교리를 마련하고, 킬체인 역량 강화, 한미 일체형 확장억제 및 한미 핵·재래식 역량 통합(Conventional-Nuclear Integration: CNI)을 위한 작전 계획 수립 및 연합훈련, 한미일 안보협력 강화, 그리고 무엇보다 AI의 군사적 활용에 관한 미 국방부의 작전 개념 개발과 군 구조 개편, 무기체계 개발 및 조달, 인재 양성 등 광범위한 분야에서 협력을 강화하는 노력을 경주해야 한다. 물론 바이든 행정부와 달리 트럼프 2기 하 한미동맹은 보다 거래적인 형태로 변화될 가능성이 크다. 그렇다면 기술 협력 및 이전의 가격표 변화를 주의 깊게 살피고 이에 선제적으로 대응하는 것이 중요할 것이다.

근본적으로는 단순히 공격 능력 개발에 치중하기보다 군사적으로 활용할 수 있는 한국 AI 기술 역량의 기본 토대를 발전시키는 것이 더욱 중요하다. ‘국방혁신 4.0’이 강조하는 AI 과학기술 강군과 유·무인 복합 전투체계를 구축하기 위해서는 자체적인 AI 기술 인프라 및 반도체 역량뿐 아니라 작전 계획 수립에 필요한 양질의, 그리고 대량의 군사 데이터가 필요하다. 만약 미국이 추진하

고 있는 대규모 언어 모델(Large Language Model: LLM) 방식이 아닌, 한국 데이터나 서버 시장 규모를 고려한 소형 또는 초소형 AI 모델에 특화하는 형태로 AI 국가전략을 설정한다면(배영자 2025), 이것을 군사적으로 활용하는 한국형 AI 군사 모델은 어떤 형태여야 하는지에 대한 본격적인 연구와 적극적인 투자가 반드시 수반되어야 할 것이다. ■

참고 문헌

- 김양규. 2024a. “2024년 세계 군사질서와 한국: 정확성과 투명성 혁명에 따른 공격 우위 시대 한국의 안보정책”. EAI 신년기획 특별논평 시리즈. 1월 5일.
<https://www.eai.or.kr/new/ko/pub/view.asp?intSeq=22294> (검색일: 2025. 1. 14.)
- _____. 2024b. “인공지능-핵무기 넥서스(AI-Nuclear Nexus)와 세계군사질서 전망”. EAI AI와 신
 문명 표준 스페셜리포트. 9월 6일.
https://www.eai.or.kr/new/ko/pub/view.asp?intSeq=22705&board=kor_special (검색
 일: 2025. 1. 14.)
- 박성민. 2024. “빅터 차 “韓혼란 장기화, 트럼프 2기 한미동맹에 최악 시나리오””. 「연합뉴스」. 12월
 13일. <https://www.yna.co.kr/view/AKR20241213002700071> (검색일: 2025. 1. 14.)
- 배영자. 2025. “2025 인공지능 기술 경쟁과 세계정치: 한국의 대응 전략”. EAI 신년기획 특별논평
 시리즈. 1월 13일. <https://www.eai.or.kr/new/ko/pub/view.asp?intSeq=22684> (검색일: 2025.
 1. 14.)
- 손열, 김양규, 이동률, 이승주, 전재성, 하영선. 2023. “‘관리된 경쟁’과 ‘발전권 확보’ 사이에서: 협력
 을 모색하는 2023 APEC 미중 정상회담”. EAI 스페셜리포트. 11월 20일.
<https://www.eai.or.kr/new/ko/pub/view.asp?intSeq=22234> (검색일: 2025. 1. 14.)
- 이재림. 2025. “[트럼프2기 출범] 너도나도 ‘눈도장’...트럼프와의 관계개선 시도 줄이어”. 「연합뉴스」.
 1월 12일. <https://www.yna.co.kr/view/AKR20250111011400087> (검색일: 2025. 1. 14.)
- 전재성. 2024. “AI 기반 자율무기체계, 인지전의 발전과 군사안보질서의 변화”. EAI AI와 신문명 표
 준 스페셜리포트. 9월 6일.
https://www.eai.or.kr/new/ko/pub/view.asp?intSeq=22706&board=kor_special (검색
 일: 2025. 1. 14.)
- _____. 2025. “트럼프주의 외교 전략과 세계질서의 미래, 한미관계”. EAI 신년기획 특별논평 시리
 즈. 1월 3일. <https://www.eai.or.kr/new/ko/pub/view.asp?intSeq=22678> (검색일: 2025.
 1. 14.)

- Biddle, Stephen. 2001. "Rebuilding the Foundations of Offense-Defense Theory." *The Journal of Politics* 63, 3: 741-774.
- Cancian, Mark F., Matthew Cancian, and Eric Heginbotham. 2023. *The First Battle of the Next War: Wargaming a Chinese Invasion of Taiwan*. Washington, DC: Center for Strategic and International Studies.
- Fravel, M. Taylor, and Eric Heginbotham. 2024. "Envisioning a Stable Military Balance in 2034." In *U.S.-China Relations for the 2030s: Toward a Realistic Scenario for Coexistence*, eds. Christopher S. Chivvis, 79-90. Washington, DC: Carnegie Endowment for International Peace.
- Horowitz, Michael C. 2018. "Artificial Intelligence, International Competition, and the Balance of Power." *Texas National Security Review* 1, 3: 36-57.
- Jacobsen, Jeppe T., and Tobias Liebetrau. 2023. "Artificial Intelligence and Military Superiority: How the 'Cyber-AI Offensive-Defensive Arms Race' Affects the US Vision of the Fully Integrated Battlefield." In *Artificial Intelligence and International Conflict in Cyberspace*, eds. Fabio Cristiano, Dennis Broeders, François Delerue, Frédérick Douzet, and Aude Géry, 135-156. London: Routledge.
- Jervis, Robert. 1978. "Cooperation Under the Security Dilemma." *World Politics* 30, 2: 167-214.
- King, Anthony. 2024. "Robot Wars: Autonomous Drone Swarms and the Battlefield of the Future." *Journal of Strategic Studies* 47, 2: 185-213.
- Kristensen, Hans M., Matt Korda, Eliana Johns, and Mackenzie Knight. 2024. "Chinese Nuclear Weapons, 2024." *Bulletin of the Atomic Scientists* 80, 1: 49-72.
- Lebow, Richard Ned. 1981. *Between Peace and War: The Nature of International Crisis*. Baltimore: Johns Hopkins University Press.
- Lieber, Keir A., and Daryl G. Press. 2017. "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence." *International Security* 41, 4: 9-49.

- Ministry of Foreign Affairs, The People's Republic of China. 2024. "Riding the Trend of the Times with a Strong Sense of Responsibility." December 17.
https://www.fmprc.gov.cn/eng/wjbzhd/202412/t20241218_11497818.html (Accessed January 14, 2025)
- O'Hanlon, Michael E. 2024. "Could the United States and China Really Go to War? Who Would Win?" Brookings, August 15, 2024. <https://www.brookings.edu/articles/could-the-united-states-and-china-really-go-to-war-who-would-win/> (Accessed January 14, 2025)
- Schneider, Jacquelyn, and Julia Macdonald. 2024. "Looking Back to Look Forward: Autonomous Systems, Military Revolutions, and the Importance of Cost." *Journal of Strategic Studies* 47, 2: 162–184.
- Selden, Zachary. 2024. "A New 'Cult of the Offensive?' Elite Perceptions of Artificial Intelligence in Military Affairs in the United States and the People's Republic of China." *Foreign Policy Analysis* 20, 4.
- Sisson, Melanie W., and Dan Patt. 2024. "After Primacy: US Military Options in Contemporary East Asia." Brookings, December 20, 2024.
<https://www.brookings.edu/articles/after-primacy-us-military-options-in-contemporary-east-asia/> (Accessed January 14, 2025)
- Waltz, Kenneth Neal. 1979. *Theory of International Politics*. New York: McGraw-Hill.
- The White House. 2024. "Readout of President Joe Biden's Meeting with President Xi Jinping of the People's Republic of China." November 16.
<https://www.whitehouse.gov/briefing-room/statements-releases/2024/11/16/readout-of-president-joe-bidens-meeting-with-president-xi-jinping-of-the-peoples-republic-of-china-3/> (Accessed January 14, 2025)

■ **김양규**_동아시아연구원 수석연구원, 서울대학교 정치외교학부 강사.

■ **담당 및 편집: 박한수**_EAI 연구원

문의: 02-2277-1683 (ext. 204) hspark@eai.or.kr

인용할 때에는 반드시 출처를 밝혀 주시기 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 2025년 1월 15일
"2025 세계안보 현실 대 인식: 인공지능의 군사적 이용과 공격송배 현상" 979-11-6617-854-2 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