



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

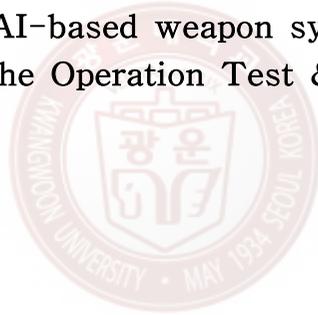
이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

박사학위 청구논문
2022학년도

인공지능(AI) 기반 무기체계
시험평가 개선방안 연구
(운용 시험 평가를 중심으로)

A study on the improvement plan for T&E
of AI-based weapon system
(Focus on the Operation Test & Evaluation)



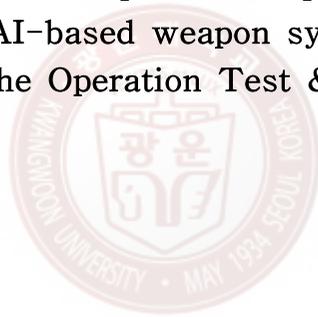
광운대학교 대학원

방위사업학과

김백중

인공지능(AI) 기반 무기체계
시험평가 개선방안 연구
(운용시험평가를 중심으로)

A study on the improvement plan for T&E
of AI-based weapon system
(Focus on the Operation Test & Evaluation)



광운대학교 대학원

방위사업학과

김백중

인공지능(AI) 기반 무기체계
시험평가 개선방안 연구
(운용 시험 평가를 중심으로)

지도교수 정석재
공동지도교수 박찬봉

이 논문을 공학 박사학위 청구논문으로 제출함.

2023년 6월

광운대학교 대학원

방위사업학과

김백중

김백중의 공학 박사학위논문을 인준함.

심사 위원장 손 채 봉 (인)

심 사 위 원 박 종 재 (인)

심 사 위 원 김 장 엽 (인)

심 사 위 원 박 찬 봉 (인)

심 사 위 원 정 석 재 (인)

광운대학교 대학원

2023년 6월

감사의 글

학위 논문을 마무리하며 감사의 글을 작성하고 있는 이 시간이 감회가 새롭습니다. 군 생활을 하면서 배움의 길에 입문하고 중도에 포기하지 않고 이 자리까지 올 수 있도록 격려하고 도움을 주신 원우회 동기, 선·후배분들과 힘든 코로나 상황에도 더 많은 것을 가르쳐 주시고 도와주신 교수님들 덕분에 지금까지 올 수 있었습니다.

우선, 박사과정 동안 학문에 대한 열정과 올바른 방향으로 나아갈 수 있도록 지도해 주시고, 응원해 주신 지도교수님께 진심으로 감사를 드립니다. 또한, 바쁘신 가운데에도 학위 논문이 완성될 수 있도록 꼼꼼하게 심혈을 기울여 심사와 지도를 해주신 교수님께 머리 숙여 감사를 드립니다.

논문을 진행하면서 도움을 주신 모든 분들과 설문에 적극적으로 응해주신 국방 전력업무 관계자분들께 깊은 감사의 인사를 드립니다.

사랑하는 제 가족들에게 감사를 전합니다. 부족한 아들을 오늘날 제가 있도록 키워주시고 항상 응원해 주신 양가 부모님께 감사를 드립니다. 일과 공부를 하면서 함께 해주지 못했음에도 건강하게 잘 성장해준 나의 최고의 보물이자 우리 가족의 희망인 희원·희건 두 아들에게 아빠의 아들로 태어나줘서 고맙다고 지면을 빌어 전합니다. 뒤늦은 학업에도 끝까지 학위과정을 끝낼 수 있도록 응원해 준 내 인생 최고의 선물이자 영원한 나의 동반자 이도윤 여사! 감사하고 사랑합니다. 당신에게 이 박사학위 취득의 영광을 돌리고 싶습니다.

지면으로 미처 언급하지 못했지만, 저에게 아낌없는 사랑과 격려를 해주셨던 모든 분들께 진심으로 깊은 감사의 인사를 드립니다. 앞으로 국방 전력분야에 꼭 필요한 존재가 되도록 더욱더 정진하며 노력하겠습니다. 감사합니다.

2023년 6월

김 백 중 올림

국문요약

인공지능(AI) 기반 무기체계 시험평가 개선방안 연구 (운용시험평가를 중심으로)

우리 군은 인공지능·무인·로봇 등 첨단 과학기술 기반 ‘AI 과학기술 강군’ 육성을 목표로 인공지능(AI) 기반 무기체계 전력화를 위해 노력하고 있다.

본 연구는 인공지능 기술이 적용된 무기체계 시험평가 개선 수요를 도출하고 우선순위를 평가하여 개선방안을 제시하는데 목적이 있다. 무기체계 개발 및 시험평가 환경변화와 시험평가 사례분석을 통해 시험평가 개선 수요를 도출하였다. 국방 전력분야 전문가를 대상으로 기초조사 및 AHP 설문을 통해 시험평가 제도·구조·기술 분야에서 11개의 세부 평가항목을 도출하고 우선순위를 평가하였다. AI 기반 무기체계 시험평가는 데이터 기반 성능평가와 실물에 의한 시험평가를 병행해야 하며, 별도의 데이터셋을 활용한 성능측정이 필요하다. AHP 분석을 통해 무기체계 시험평가 개선방안에 대한 상대적 중요도를 분석한 결과, 시험평가 제도 - 기술 - 구조 순으로 시험평가 제도 개선이 높게 나타났다. 세부 평가항목의 우선순위는 시험평가 결과판정 개선, 시험평가 조직 및 전문인력 양성, 과학적 시험평가체계 정립 순으로 평가되었다. 우선순위가 높은 평가항목에 대하여 AI 기반 무기체계 시험평가 개선방안을 제시하였다.

주제어 : 첨단 무기체계, 시험평가(T&E) 제도·구조·기술, 인공지능(AI), AHP(계층적 분석기법), 운용시험평가(OT&E)

ABSTRACT

**A study on the improvement plan for T&E
of AI-based weapon system
(Focus on the Operation Test & Evaluation)**

KIM, Baekjung

Dept. of Defense Acquisition Program

The Graduate School Kwangwoon University

The ROK military is making efforts to deploy AI-based weapon systems with the goal of fostering an ‘AI science and technology strong’ based on advanced science and technology such as AI, Unmanned, Robots, and so on.

The purpose of this study is to derive the need for improvement for the test and evaluation(T&E) of a weapon system using AI technology, evaluate the priority, and present an improvement plan. T&E improvement requirements were derived through weapon system development, T&E environment changes, and T&E case analysis. 11 detailed evaluation items were derived and priorities were evaluated in the field of T&E system, structure, and technology fields through basic surveys and AHP surveys targeting experts in the defense domain. AI-based weapon system T&E should be performed in parallel with data-based performance evaluation and actual T&E, and performance measurement using a separate T&E data set

is required for AI models performance evaluation. As a result of analyzing the relative importance of the weapon system T&E improvement through AHP analysis, it came out in the order of T&E system - technology - structure. The priority of detailed evaluation items were evaluated in the order of improving the judgment of T&E result, training organization and education of experts, and establishment of scientific T&E system. For evaluation items with high priority, measures to improve the AI-based weapon system T&E were presented.



Key words : Advanced weapons systems, T&E system·structure·technology, AI, AHP, Operational Test and Evaluation (OT&E)

차 례

감사의 글	i
국문 요약	ii
영문 요약	iii
차 례	v
그림 차례	vii
표 차례	ix
제1장 서 론	1
제1절 연구배경 및 목적	1
1. 연구배경	1
2. 연구목적	3
제2절 연구의 범위 및 방법	5
1. 연구의 범위	5
2. 연구의 방법	6
제2장 무기체계 시험평가 이론적 배경	9
제1절 국방획득체계 변천 과정	9
제2절 시험평가(T&E) 이론적 고찰	14
1. 시험평가의 정의 및 역할	14
2. 무기체계 시험평가 수행절차	16
제3절 인공지능(AI)의 개념 및 기술발전 동향	21
1. 인공지능(AI)의 개념	21
2. 인공지능(AI)의 기술발전 동향	26
제4절 선행연구 고찰	32
1. 무기체계 시험평가 관련 선행연구	32
2. 인공지능(AI)의 군사용 활용 관련 선행연구	35
3. 선행연구 분석 및 시사점	38
제3장 연구의 방법론	40
제1절 연구수행 절차	40
제2절 사례연구	43
제3절 계층적 분석기법	47

제4장 첨단 무기체계 시험평가 개선 필요성	51
제1절 무기체계 개발환경의 변화	51
1. 국방 과학기술의 발전과 미래전 양상	51
2. 첨단 무기체계 발전방향	54
제2절 첨단 무기체계 시험평가 개선 소요	61
1. 일반 무기체계와 AI 기반 무기체계의 시험평가 비교	61
2. 무기체계 시험평가 수행체계 개선 소요	68
제3절 계층적 분석기법(AHP)을 적용한 개선 우선순위 분석	83
1. 계층적 분석기법(AHP) 적용절차	83
2. 계층구조 및 평가항목별 우선순위 분석 결과	86
제4절 소결론	95
제5장 인공지능(AI) 기반 무기체계 시험평가 개선방안	97
제1절 AI 기반 무기체계의 신뢰성 및 시험평가 환경분석	97
1. AI 무기체계 신뢰성 및 운용환경의 제한사항	97
2. AI 기반 무기체계 시험평가 사례	101
제2절 AI 무기체계 시험평가 결과판정 개선	106
1. 시험평가 결과판정 절차	106
2. AI 무기체계 시험평가 결과판정 방안	108
제3절 시험평가 조직 및 전문인력 양성	113
제4절 AI 무기체계의 과학적 시험평가(신뢰성) 정립	117
제5절 AI 무기체계 운용시험평가 절차	126
제6절 소결론	132
제6장 결 론	138
제1절 연구결과 요약 및 의의	138
제2절 연구의 한계 및 향후 연구과제	142
참 고 문 헌	145
부 록	151
#1. 기초조사 설문지	151
#2. 계층적 분석기법(AHP) 설문지	161
#3. 약어표	172

그 립 차 례

그림 1. 국방획득체계의 3요소와 상관관계	9
그림 2. 무기체계 연구개발사업 추진 기간	13
그림 3. SE 기반 무기체계 개발 및 시험평가 검증	15
그림 4. 연구개발 무기체계 수명주기 비용	16
그림 5. 연구개발 무기체계 시험평가 수행절차	17
그림 6. 인공지능, 머신러닝 및 딥러닝과의 관계	22
그림 7. AI 활용 기대효용 정도	24
그림 8. 국방 분야 AI 활용에 따른 문제 발생	25
그림 9. 2016~2021년 우리나라와 주요국의 AI분야 총괄 기술수준 추이 ...	28
그림 10. 국방 AI 기술의 발전 방향(3단계 발전 모델)	31
그림 11. 연구수행 절차 및 연구모형	41
그림 12. 계층적 분석기법(AHP) 기본 공리	48
그림 13. AHP 기법 적용 절차	50
그림 14. 미래 전장 환경과 4차 산업혁명 관점	52
그림 15. 전 세계 상위 15개국 국방비 비율(2020년) 및 주요 무기 수출입국(2016~2020)	53
그림 16. 반자율형 유·무인 복합체계 시범부대 및 주요체계	60
그림 17. AI 영상감시 체계	62
그림 18. AI 데이터 기반 기계학습의 과적합	65
그림 19. AI 데이터 학습 및 평가	66
그림 20. AI 및 실물에 의한 시험평가 운용환경(데이터)	67
그림 21. 미 육군시험평가사령부(ATEC) 조직도	70
그림 22. 기초조사 설문 대상자의 통계학적 특성	74

그림 23. 현행 국방 획득체계를 활용한 첨단 무기체계 전력화 가능성	75
그림 24. 첨단 무기체계 시험평가를 위해 개선이 필요한 분야	76
그림 25. '시험평가 제도' 차원에서 개선해야 할 분야	76
그림 26. '시험평가 구조' 차원에서 개선해야 할 분야	77
그림 27. '시험평가 기술' 차원에서 개선해야 할 분야	78
그림 28. '시험평가 법령/훈령' 차원에서 개선해야 할 분야	79
그림 29. '운용시험평가 항목' 개선 필요성	80
그림 30. '운용시험평가 항목' 중 개선이 필요한 항목	81
그림 31. 대분류(계층 1) 항목의 중요도 및 평가 우선순위	87
그림 32. 시험평가 제도(계층 2)의 상대적 중요도 및 평가 우선순위	89
그림 33. 시험평가 구조(계층 2)의 상대적 중요도 및 평가 우선순위	91
그림 34. 시험평가 기술(계층 2)의 상대적 중요도 및 평가 우선순위	92
그림 35. 계층구조별 중요도 및 평가 우선순위(종합)	94
그림 36. AI 데이터 오용사례 (페이크 뉴스의 위험성 경고)	98
그림 37. 음원 활용 인공지능(AI) 경계시스템	103
그림 38. 국내 구매사업 절차	108
그림 39. AI 무기체계 구매사업 시험평가 개념	109
그림 40. AI 무기체계 연구개발사업 시험평가 개념	110
그림 41. 육군 시험평가단 편성	114
그림 42. 육군 시험평가단 인적 구성	114
그림 43. 미 육군 시험평가사령부(ATEC), 육군평가센터(AEC) 편성	116
그림 44. 육군 AI 무기체계 성능인증위원회 구성(안)	122
그림 45. AI 데이터 증강	124
그림 46. AI 데이터셋 구축	124

표 차 례

표 1. 국방획득체계 변천 과정	11
표 2. 사업추진 단계	13
표 3. 무기체계 시험평가 관리 활동	18
표 4. 무기체계 시험평가 항목	19
표 5. 인공지능(AI)의 종류	23
표 6. 인공지능(AI) 기술수준(ATL) 지표	27
표 7. 국방 인공지능(AI) 기술수준	29
표 8. 시험평가 관련 연구 및 보고서	34
표 9. 인공지능(AI)의 군사용 활용 관련 연구 및 보고서	37
표 10. 연구설계 검증을 위한 네 가지 방법과 진술	45
표 11. JAIC가 수행하는 주요 인공지능 프로젝트	56
표 12. 중국의 차세대 인공지능 발전계획	57
표 13. 국방 전력체계 혁신 분야 핵심기술, 군사능력, 핵심전력	58
표 14. 운용시험평가 항목	63
표 15. 일반 무기체계와 AI 기반 무기체계 시험평가 비교	64
표 16. 신속시험획득사업 추진현황(2020~2022년)	69
표 17. 첨단 무기체계 시험평가 개선 잠정 세부 평가항목	73
표 18. 기초조사 설문 전문가 집단의 분포	74
표 19. 첨단 무기체계 시험평가 개선 소요 세부 평가항목	82
표 20. AHP 전문가 집단의 세부 현황	83

표 21. AHP 계층구조 설계	84
표 22. AHP 계층구조 및 평가요소의 개념	85
표 23. AHP 설문지 회수율 및 유효한 응답 결과	86
표 24. 대분류(계층 1) 항목의 중요도 및 평가 우선순위	87
표 25. 시험평가 제도(계층 2)의 상대적 중요도 및 평가 우선순위	89
표 26. 시험평가 구조(계층 2)의 상대적 중요도 및 평가 우선순위	91
표 27. 시험평가 기술(계층 2)의 상대적 중요도 및 평가 우선순위	92
표 28. 계층구조별 중요도 및 평가 우선순위(종합)	93
표 29. AI 신뢰성의 핵심 속성 및 의미	100
표 30. AI 기반체계 시험평가 현황	102
표 31. 음원 활용 인공지능(AI) 경계시스템 성능확인 결과	104
표 32. 운용시험평가 결과판정 절차	106
표 33. 작전운용성능 외 항목 미충족으로 전력화 중단 및 지연 사례	107
표 34. AI 무기체계 시험평가 판정(안)	110
표 35. 핵심·일반항목으로 구분 분류 기준(예)	111
표 36. 인공지능 신뢰성의 주요 핵심요소 및 의미	119
표 37. 분류모델과 회귀모델 성능평가의 목적 및 척도	120
표 38. AI 유형별 척도와 성능 수준(2022 AI Index Report)	121
표 39. AI 기반 무기체계 시험평가 수행간 시험평가단의 역할	126
표 40. 운용시험평가 단계별 수행절차	129
표 41. 운용시험평가 시험준비 검토회의, 주요 확인사항 및 진입조건	131
표 42. 첨단 무기체계 시험평가 개선방안 우선순위	139

제1장 서론

제1절 연구배경 및 목적

1. 연구배경

2022년 2월 24일 러시아의 우크라이나 침공으로 시작된 우크라이나 전쟁은 수많은 사상자와 수천 대의 장비가 파괴된 전쟁이자, 재래식 무기와 첨단 무기의 각축장으로 미래전의 양상을 예측해볼 수 있다. 4차 산업혁명의 시대, 미래 전장은 첨단 국방과학기술이 적용된 지상·해상·공중에서 우주·사이버전 및 인지·심리 영역으로 확대되어 전쟁의 복잡성이 증가하고, 인공지능(AI, Artificial Intelligence)·드론봇 등 첨단 과학기술 기반의 무기체계 등장으로 미래전 수행 개념이 변화할 것이다.¹⁾ 또한, 인구감소와 4차 산업혁명의 영향으로 우리 군은 ‘병력 위주의 양적 구조’에서 ‘기술 위주의 질적 구조’로 변화하기 위해, ‘국방혁신 4.0’²⁾과 ‘육군 비전 2030’을 넘어 ‘육군 비전 2050’³⁾을 내실 있게 추진하고 있다. 4차 산업혁명 기술은 인공지능(AI), 사물인터넷(IoT), 빅데이터(Big Data), 클라우드(Cloud) 등 기초기술을 기반으로 지능형 로봇, 무인·자율체계 등 스마트 국방을 구현하게 될 것이다. 이에 군은 민·관·산·학·연과 협업을 통

-
- 1) 장상국, OT(기회요인-위협요인) 분석을 통한 차기차주포(K9F) 개발 필요성 연구, 선진국방연구 제3권 3호, 2020, p.27
 - 2) ‘국방개혁 4.0’은 AI·무인·로봇 등 4차 산업혁명 과학기술을 기반으로 북핵·미사일 대응, 군사 전략 및 작전개념, 핵심 전력, 군구조 및 교육훈련, 국방R&D·전력증강체계 분야를 혁신하여 경쟁 우위의 AI 과학기술 강군을 육성하는 것이다. 윤석열 정부 110대 국정과제에 포함되어 제2 창군 수준의 국방혁신 4.0을 추진하고 있으며, 문재인 정부 국방개혁안이었던 ‘국방개혁 2.0’을 대체한다. ‘국방혁신 4.0’의 4.0의 의미는 4차 산업혁명 첨단 과학기술 기반의 국방을 창출하는 상징적인 의미이자, 창군이래 국방의 획기적 변화를 추구하는 4번째 계획이라는 의미를 갖고 있다. (818계획(80년대) → 5개년 국방발전계획(90년대) → 국방개혁(00년대) → 국방혁신(20년대))
 - 3) ‘육군 비전 2030’은 육군의 기본 정책서로 2030년을 전후한 앞으로 10~15년 후의 미래를 대상으로 한다. ‘육군 비전 2050’은 2050년을 내다본 20~30년 후의 미래를 대상으로 하는 기획의 방향을 제시하는 개념서다.

해 혁신적인 기술과 개념을 적용한 첨단과학 기술군으로 변모하기 위해 육군은 ‘Army TIGER 4.0’⁴⁾과 ‘페가수스(PEGASUS) 프로젝트’⁵⁾ 및 ‘레이저’ 정책을 추진하여 한계를 넘어서는 초일류 육군을 구현하기 위해 노력하고 있다.

무기체계 시험평가(T&E, Test and Evaluation)는 무기체계가 연구개발 및 획득과정에서 기술적 또는 운용·관리적 측면에서 소요제기서에 명시된 제반 요구 조건의 충족 여부를 확인 및 검증하는 절차를 말한다.⁶⁾ 즉 사용자의 요구와 일치 여부를 검증하고 운용목적 부합 여부의 적합성을 판단하는 과정이다. 시험평가는 체계개발 및 획득을 위한 과정에서 조기에 무기체계 성능 수준을 확인하고 개발자의 설계결함을 식별하여 개선할 수 있도록 도움을 주는 역할을 한다. 또한 무기체계 획득사업 과정에서 위험(Risk)을 줄이기 위해서는 의사 결정자에게 무기체계의 사용 여부에 대한 정확하고 적절한 정보가 적시에 제공되어야 한다. 이 과정에서 무기체계의 구매, 연구개발 및 생산 단계에서 의사 결정자에게 의사결정에 필요한 정보를 제공하는 것이 시험평가를 하는 주목적이다. 이것은 개발 및 획득단계의 각 의사결정 지점에서 다음 단계로의 전환 여부를 결정하는데 필요한 정보를 제공함으로써 의사결정을 지원하는 역할과 획득 및 개발과정에서 위험관리 수단으로서 역할을 수행한다.⁷⁾

우리 군의 시험평가 제도는 자주국방과 방위산업의 성장이 함께하고 있다.

4) ‘Army TIGER 4.0’은 4차 산업혁명 기술을 기반으로 모든 전투플랫폼을 기동화·네트워크화·지능화를 전투현장에서 구현하는 유·무인 복합전투체계다. Army TIGER 전력은 기반전투체계(첨단 과학기술을 집목한 미래 지상전투체계), 드론봇 전투체계(현용전력과 통합된 드론과 로봇의 유·무인 복합전투체계), 위리위플랫폼(전투원의 치명성과 생존성을 향상시킨 육군의 개인전투체계)의 3대 전투체계로 구성되어 있다.

TIGER: Transformative Innovation of Ground forces Enhanced by the 4th industrial Revolution technology의 약어로 미래 전장을 주도하는 게임체인저로서의 역량을 갖춘 첨단 과학기술군의 결정체를 상징한다.

5) PEGASUS: Pioneer & Explore the Ground And Space for United Space Operations의 약어, 육군 우주력 발전계획으로 통합된 우주작전을 위해 지상 및 우주 개척, 탐험하는 미래 육군을 상징한다.

6) 육군 시험평가단, 시험평가 참고서, 2019, p.17

7) 박원동, 명품 무기체계 탄생의 마지막 진통, 북코리아, 2008, p.34

1970년대부터 본격적으로 시작된 우리나라 방위산업은 초기에는 주로 미군의 군사 부품에 대한 단순 조립 및 면허생산 단계부터 시작하였으나, 1974년부터 율곡사업(국방 8개년 계획, 1974~1981년)을 통해 탄약과 물자, 소총, 박격포 등 무기류의 국산화를 달성하면서 시험평가에 대한 개념이 정립되었다. 2006년 방위사업청(방사청) 개청 이후 방위사업 추진의 투명성, 효율성, 전문성이 제고되고, 국방획득업무의 수행기반이 공고해지는 상황에서 사업관리 절차는 더욱 복잡해지고 폐쇄성으로 인한 획득이 장기화 되고 있다. 또한, 무기체계 시험평가의 전문성·투명성·신뢰성에 중점을 둔 정책추진으로 신속성과 효율성이 결여되었다는 평가와 결과의 완전성 보장을 위한 세밀한 확인·검증과 평가에 대한 지속적인 요구가 상존하고 있다. 4차 산업혁명의 시대, 첨단 과학기술 강군 육성을 위한 획득 프로세스의 전면적인 보완을 위해 국방부 주관 ‘한국형 전력증강 프로세스 정립’을 추진하고 있으며, 소요군에서는 첨단 무기체계 시험평가 수행체계를 개선하기 위해 노력하고 있다.

2. 연구목적

국방 무기체계 획득과정에서 신뢰성 있는 무기체계를 개발하기 위해 무기체계 개발의 최종 단계에서 무기체계의 작전운용성과 군 운용의 적합성, 전력화지원요소의 실용성 등 시험평가를 통해 검증하고 있으며, 소요군의 시험평가 전담 기관에서 이를 선도하고 있다. 최근 국방 무기체계 획득제도와 과학기술의 발전, 장기간 소요되는 무기체계 개발과정의 특성으로 기술의 진부화와 개발 후 품질 및 성능개선 요구가 증가하고 있다. 또한, 4차 산업혁명의 시대에 민간 첨단 과학기술을 적기에 국방 무기체계에 도입될 수 있도록 국방 획득체계 전반에 대한 제도 개선이 요구되고 있다. 최근 전력화 전(前) 무기체계 성능과 안전의 완전성을 보장하기 위한 마지막 단계인 시험평가와 관련하

여 첨단 무기체계를 전력화하는 데 걸림돌로 인식하는 기관들이 많아지고 있다. 이는 전력화 여부를 판단하는 중요한 단계인 시험평가에 대한 역할의 재정립과 합리적인 수행체계의 발전이 요구된다는 것을 입증한다고 할 수 있다.

4차 산업혁명의 시대에는 무기체계의 급속한 첨단화·복합화로 시험평가 업무는 더욱더 전문지식과 높은 난이도를 요구하고 있다. 이에 따른 전문성을 구비한 인력이 더 필요하지만 조직의 보강은 현실적으로 많은 어려움이 있다. 무기체계 시험평가는 설계 및 기술적 완성도(개발시험평가, DT&E)와 작전운용성능 및 군 운용 적합성(운용시험평가, OT&E)에 대한 기준충족 여부를 적법하고 객관적이며, 타당성 있게 판단하는 것이 시험평가의 기본이자 원칙이다. 이와 같이 기본과 원칙을 준수하면서 4차 산업혁명의 첨단 과학기술이 국방 무기체계에 신속하게 적용되어 전력화될 수 있도록 진화적 획득의 활성화와 무기체계 시험평가 수행체계를 근본적으로 변화를 모색할 필요가 있다.

따라서 본 연구에서는 4차 산업혁명의 시대, 첨단 과학기술이 적용되는 AI 기반 무기체계 시험평가를 위해, 국방획득체계와 무기체계 시험평가 관련 이론적 고찰 및 선행연구를 실시하였다. 이를 바탕으로 미래전 양상의 변화와 우리 군이 직면하게 될 국방 무기체계 개발환경 및 시험평가 환경변화를 통해 무기체계 시험평가 개선 소요를 도출하였다. 또한, 계층적 분석기법(AHP, Analytical Hierarchy Process)을 활용하여 첨단 무기체계 시험평가를 위한 시험평가 제도·구조·기술적 분야에서의 상대적 중요도와 하위 세부 평가항목의 우선순위를 결정하였다. AHP 분석 결과 우선순위가 높은 평가항목과 관련하여 인공지능(AI) 기반 무기체계⁸⁾ 시험평가 개선방안을 제시하고자 한다.

8) 인공지능(AI) 기반 무기체계는 인공지능 기술을 포함하여 설계되고 운용되는 무기 시스템을 말한다. 이는 기존의 전통적인 무기체계와는 달리 AI 기술을 사용하여 자동화(Automation), 자기학습(Self-learning), 의사결정(Decision-making) 등의 능력을 갖추고 있다.

제2절 연구의 범위 및 방법

1. 연구의 범위

무기체계 시험평가(T&E)는 특정 무기체계가 기술적, 운용·관리적 측면에서 소요제기서에 명시된 제반 요구조건 충족 여부를 확인·검증하는 절차다. 시험평가는 무기체계 획득과정에서 위험(Risk) 요소를 줄이기 위한 기술적 도구(Tool)로써 사용되며, 각각의 사업추진단계별로 반복적으로 수행한다.⁹⁾ 4차 산업혁명의 시대, 첨단 과학기술을 반영한 무기체계를 현 시험평가 제도와 방법 및 절차를 준수하여 수행해야 한다. 첨단 과학기술은 급속도로 변화하고 있으며, 신기술을 무기체계에 반영하여 획득하는 과정에서 시험평가를 실시해야 한다. 하지만, 현 국방획득제도에서 시험평가 수행체계로는 무기체계의 요구조건을 확인하고 검증하는데 많은 제한사항이 발생하고 있다. 따라서 첨단 무기체계 시험평가 수행 절차를 체계적이고 효율적으로 진행될 수 있도록 시험평가 제도·구조·기술적으로 새롭게 정립하고 개선해야 한다.

본 연구는 국방 무기체계 개발 및 획득과정에서 우리 군이 추진하고 있는 첨단 무기체계에 대한 시험평가 개선 수요를 식별하고, 시험평가 제도·구조·기술적 분야에서 상대적 중요도와 우선순위를 분석하는데 있어 무기체계 분류, 시험평가 방법 등에서 매우 광범위하고 다양하였다. 따라서 첨단 무기체계에 대한 시험평가 개선방안을 제시하는 데 있어 그 연구범위를 다음과 같이 제한하였다.

첫째, 우리 군이 추진하고 있는 첨단 무기체계 전력화는 ‘국방혁신 4.0’과 ‘Army TIGER 4.0’ 등 기획문서 및 정책서에 반영된 한반도의 미래전 양상과 이를 대비하기 위한 무기체계 개발 수요로, 육군에서 추진하고 있는 무기체계에 한정하여 분석하였다.

9) 육군 시험평가단, 전계서, 2019, p.17

둘째, 국방획득체계는 1972년에 최초로 도입된 이후 방사청 개청과 함께 획득업무 추진에 있어 투명성, 효율성, 전문성이 제고되었으나, 신속성과 효율성이 결여되는 점이 있어 급변하는 첨단 과학기술을 국방 무기체계에 적용하는데 많은 제한사항이 있다. 본 연구에서는 방위사업청 개청 이후 국방획득체계에서 소요결정으로부터 체계개발, 전력화된 사업으로 한정하여 사업추진단계를 분석하였다.

셋째, 한반도의 미래 전장은 AI 기반 드론과 로봇, 무인체계 등 유·무인 복합전투체계가 초지능·초연결 네트워크로 연결되어 다영역 동시 통합 작전을 수행할 것이다. 군은 미래 전장에서 운용될 다양한 첨단 무기체계 전력화를 추진하고 있으나, 육군이 추진하고 있는 AI 기반 무기체계로 한정하여 분석하였다. 또한, 현시점은 첨단 무기체계 전력화를 위한 시작 단계로 AI 기반 무기체계 시험평가 사례가 제한되는 점을 고려하여 육군에서 신속시범획득사업으로 추진한 AI 영상감시체계 성능확인 시험으로 한정하고, 기존 무기체계 시험평가 사례와 비교·분석하여 개선방안을 제시하였다.

넷째, 계층적 분석기법(AHP)을 적용하여 무기체계 시험평가 개선을 위한 시험평가 제도·구조·기술 분야의 상대적 중요도와 세부 평가항목의 우선순위를 평가하였으며, AI 기반 무기체계 시험평가 개선방안은 평가항목 중 우선순위가 높게 평가된 항목에 대해 운용시험평가를 중심으로 제시하였다.

2. 연구의 방법

국방획득체계에서 첨단 무기체계 시험평가 중 AI 기반 무기체계에 대한 시험평가 개선방안을 도출하기 위한 연구 방법은 다음과 같다.

4차 산업혁명과 시대적 변화에 적합한 첨단 무기체계 시험평가 개선방안을 도출하기 위한 방법론으로 시험평가에 관한 질적 사례연구 방법을 적용하였

다. 무기체계 개발환경 변화에 따른 무기체계 시험평가 한계와 미래 시험평가 환경변화에 대한 사례분석을 통해 첨단 무기체계 시험평가 개선방안을 도출하고, 첨단 무기체계 시험평가 분야별 개선방안에 대한 정량적 분석은 계층적 분석기법(AHP)을 적용하였다. 첨단 무기체계 시험평가 개선방안을 도출하기 위한 우선순위는 선행연구와 국방 전력분야 전문가를 대상으로 기초조사 설문을 실시하여 시험평가 제도·구조·기술 분야의 평가항목을 도출하고, 질적 연구 방법으로 전문가를 대상으로 AHP 설문을 통해 우선순위를 평가하였다. 세부 평가항목 중 우선순위가 높은 평가항목에 대하여 AI 기반 무기체계 시험평가를 위한 개선방안을 제시하였다. 제시된 방안에 대한 추가적인 AI 기반 무기체계 유형별 세부 시험평가 방법 및 절차, 전문가 인재 Pool 확보 및 조직개편, 네트워크 통합시험 방안 등은 향후 연구과제로 제안하였다.

앞서 제시한 연구의 방법을 바탕으로 본 논문의 목적인 인공지능(AI) 기반 무기체계 시험평가 개선방안(운용시험평가를 중심으로)을 제시하기 위해서 다음과 같이 구성하였다.

본 논문은 6개의 장으로 구성되었다. 먼저 1장에서는 연구의 배경 및 목적, 연구범위 및 방법에 대하여 설명하였다.

제2장에서는 본 연구와 관련된 전반적인 이해를 돕기 위해 국방획득체계 변천 과정과 시험평가 정의 및 역할, 시험평가 수행 절차, 인공지능(AI) 개념 및 기술발전 동향 등 이론적 배경을 기술하고, 선행연구를 통하여 첨단 무기체계 전력화를 위한 시험평가 제도·구조·기술적 분야에서 개선 소요를 도출하였다.

제3장에서는 본 연구의 수행 절차의 분석틀, 주요 방법론인 사례연구와 계층적 분석기법(AHP)에 대한 이론적 배경을 기술하였다.

제4장에서는 첨단 무기체계 시험평가 개선의 필요성을 도출하기 위해 무기체계 개발 및 시험평가 환경변화와 무기체계 시험평가 사례를 분석하여 첨단

무기체계 시험평가 개선 소요를 도출하였다. 그 결과와 국방 전력분야 전문가를 대상으로 기초조사 설문을 통해 첨단 무기체계 시험평가 개선방안에 대한 분야 및 세부 평가항목을 도출하고, 다음 단계로 계층적 분석기법(AHP)을 적용하여 시험평가 제도·구조·기술 분야의 중요도와 세부 평가항목에 대한 상대적 우선순위를 결정하였다.

제5장에서는 AHP 분석 결과 상대적 우선순위가 높은 평가항목에 대하여 인공지능(AI) 기반 무기체계 시험평가 개선방안을 제시하였다.

제6장에서는 인공지능(AI) 기반 무기체계 시험평가 개선방안에 대한 연구 결과를 요약하고 분석 결과 및 개선방안에 대한 시사점을 제시하였으며, 본 연구를 통해 제시하지 못한 연구의 한계와 향후 연구과제에 대해 제시하였다.



제2장 무기체계 시험평가 이론적 배경

제1절 국방획득체계 변천과정

국방획득체계란 우리 군에게 필요한 무기나 물자와 같은 군수품을 시의적절하게 경제적으로 조달하여 운용하도록 하는 제도 및 조직, 절차, 문화를 총칭한다.¹⁰⁾ 우리 군이 현재와 같은 국방획득체제로 발전하는 데에는 미군과 비교했을 때 아주 짧은 역사를 갖고 있다. 6·25전쟁 이후 우리 군은 1970년대까지 미국의 군사원조에 의존하고 베트남전 참전을 통해 군 현대화와 군사력 증강을 시작하였으나 그 당시에는 전력기획이라는 개념이 없었다.

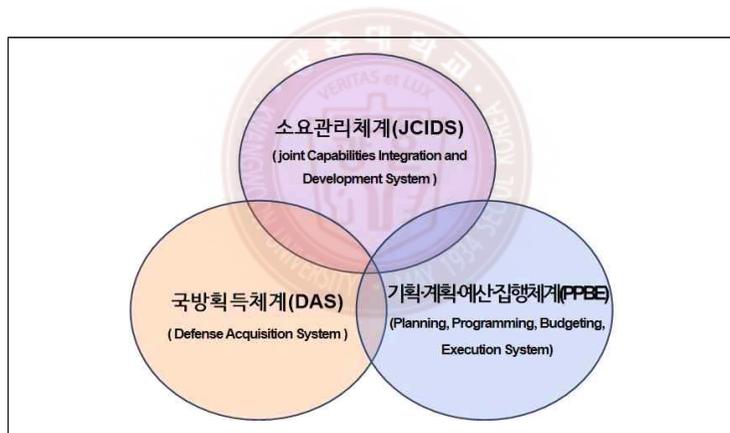


그림 1. 국방획득체계의 3요소와 상관관계¹¹⁾

Figure 1. Correlation with the three elements of the defense acquisition system

박정희 대통령은 주한미군의 철수가 가시화되자 독자적인 방위 능력을 갖추기 위해 자주국방 계획을 입안하고 국방과학연구소(국과연, ADD) 설립과 국방부에 방위산업담당관실, 합동참모본부(합참)에 연구개발과(작전기획국)를 신

10) 송학, 국방획득체계, 투명정보보다 전문성·효율성 강화해야, 녹색경제 신문, 2021. 12. 12, (<http://www.greened.kr>, 검색일: 2022. 4. 5.)

11) 상계서, <http://www.greened.kr>, (검색일: 2022. 4. 5.), 연구자 제작성

설하는 등 방위산업 육성을 통한 자주국방 체계를 확립하였다. 이와 같은 제도 도입을 통해 전력증강 8개년 계획을 시작으로 1차 율곡사업을 통해 노후 장비 교체 및 F-4 전투기를 구매하였다. 이후 5개년 단위로 3차 율곡사업을 시행하여 자주포와 한국형 전차, F-5와 F-16 기술도입 생산, 전투함정과 잠수함 등을 전력화하여 적정 군사력을 유지할 수 있도록 하였다.

국방획득체계는 1972년 최초로 도입되어 방사청 창설(2006. 1월)과 함께 지속적인 제도 개선을 통해 변화되어왔다. 방사청 창설 이전에는 국방부가 기획·계획·예산·집행·평가체계(PPBEES)¹²⁾에 따라 소요, 획득, 운영유지 및 예산의 모든 과정을 직접적으로 통제해 각 군과 함께 획득업무를 수행해 왔다. 2003년 출범한 노무현 정부는 획득체계를 둘러싸고 파생되는 다양한 문제들 - 비용 초과, 성능결함, 전력화 지연, 부패 등 - 을 개선하기 위해 ‘정부조직법’(법률 제 7613호, 2005. 7. 22. 공포)을 개정하고, 2006년 1월 국방부 및 각 군의 획득관련 8개 기관을 통합한 단일 획득조직인 방사청이 개청 되었다.¹³⁾ 방사청 개청 이후부터 국방부, 방사청, 합참, 각 군간 업무조정이 지속적으로 이뤄져 왔고, 현재 소요·중기계획작성·시험평가·운영유지는 국방부·합참·각 군에서, 예산편성·계약·사업관리·양산 및 배치는 방사청에서 수행하고 있다.¹⁴⁾

2006년 방위사업법 제정과 방사청 개청을 통해 국방획득업무의 투명성, 전문성 및 효율성을 높이고자 하였으나 소요결정에 관한 권한은 국방부 장관에게, 전력증강을 위한 예산편성 및 집행과 관리의 책임, 연구개발업무의 통제, 방산물자의 수출지원, 시험평가·분석평가에 대한 권한은 방위사업청에 둬으로써 소요기획 업무를 제외한 획득관련 기능이 방위사업청에 집중되게 되었다.

12) PPBEES: Planning(기획), Programming(계획), Budget(예산), Execution(집행), Evaluation(평가), System(체계)

13) 김종하, 국방획득체계 현황 및 문제점, 그리고 개선방향, 과학기술정책 제27권 제11호, 2017, p.46

14) 상계서, p.46

표 1. 국방획득체계 변천 과정¹⁵⁾

Table 1. The evolution of defense acquisition system

구 분	제도 변천	주요 특징
1972년	◦ 연구개발업무체계 정립 (최초 국방획득체계)	◦ 국방부 방위산업담당관실, 합참 연구개발과 신설 ◦ 병기개발과제 선정위원회(합참) 및 병기개발기술위(국과연)에서 의결
1974 ~ 1978년	◦ 울곡사업 착수 - 연구개발 업무절차 마련 - 무기체계 채택절차 마련	◦ 중기계획 기능 없이 연구개발 절차 위주 제도 ◦ 연구개발 소요기획 → 합참, 예산 → 국방부 ◦ 사용자와 개발자간 업무분담 및 책임소재 명확
1979 ~ 1982년	◦ 미국 PPB제도 도입 - 무기체계 선정절차 마련	◦ 국방기획관리체계 최초 정립 ◦ 계획기능 신설, 목표지향적 전력증강 ◦ 합참에서 소요식별 및 계획업무 담당
1983 ~ 1990년	◦ 미국 PPBEES제도 도입 - 해외구매절차 제도화 - 획득 전단계 품질보증	◦ 울곡계획 업무를 합참에서 국방부로 이관 ◦ 투자비와 운영유지비 통합 ◦ 국방중기계획서 명칭 사용
1991 ~ 1996년	◦ 818 군구조 개편 - 소요제기 기능 합참 이관 - 각 군 소요제기 환원('95)	◦ 국방중기부대기획서 최초 작성 ◦ 소요제기 기능의 시행착오 (합참으로 이관 후, 각 군/기관으로 환원)
1997 ~ 1999년	◦ 문민정부 제도 개선 - 획득절차 단축(9 → 6단계) - RFP에 의한 기종 결정	◦ 합동중기전략목표기획서(JSOP) 폐지, 전력소요 기획 및 계획기능 국방부 통합 ◦ 국방부 획득기능 확대(합참기능 축소), 합참 전력분석 및 시험평가 기능 확대
1999 ~ 2002년	◦ 국민의 정부 제도 개선 - 17개 규정 통·폐합 - 결재 처리방식 의사결정	◦ 국방개혁추진위 구성, 국방업무 전반 개혁 추진 ◦ 회의체 운영에 따른 기간 장기화, 책임 회피 등 문제 해소를 위해 회의체 축소, 결재방식 채택 ◦ 전력소요기획기능 합참 환원, JSOP 재발간 ◦ 소요기획 → 각 군/합참, 획득관리 → 국방부
2003 ~ 2005년	◦ 참여정부 제도 개선 - 방위사업법 개정('06년) (방위사업청 개청)	◦ 국방획득관리 최초의 법 제정 ◦ 소요기획 제외, 기타 획득기능 방사청 집중
2006 ~ 2013년	◦ 방위사업법 제정 - 개방형 회의체 의사결정 ◦ 전문화·계열화제도 폐지('08년)	◦ 획득관리규정 → 국방전력발전업무훈령 ◦ 의사결정 방식 전환(결재방식 → 개방형회의체)
2014년 ~ 현재	◦ 방위사업법 개정 - 정책과 집행 분리	◦ 무기체계 소요결정 주체 : 국방부 → 합참 ◦ 국방중기계획 작성 주체 : 방사청 → 국방부 ◦ 시험평가 주체 : 방사청 → 국방부 ◦ 무기체계 소요의 적절성 검증, 중기계획에 반영 근거 마련

15) 위검·하철수 등 4명, 무기체계 소요결정과 사업추진 방법의 효율적 연계방안, 한국군사문제연구원, 2014, p.6, 연구자 제작성

이후 2014년 4월 16일 『방위사업법 일부개정법률안』 개정에 따라 무기체계 등의 소요결정 주체를 국방부에서 합참으로 이관하고 국방중기계획 작성 주체를 국방부장관으로 조정하였다. 또한, 무기체계 소요의 적절성을 검증할 수 있는 근거를 신설하고, 시험평가 계획수립과 결과판정의 주체를 방사청에서 국방부로 변경하여 평가의 객관성을 제고할 수 있도록 하였다.

현 국방획득 제도는 방위사업청 개청(2006. 1. 1.) 이후 방위사업 추진의 투명성, 효율성, 전문성이 제고되고 획득업무의 수행기반이 공고해진 긍정적인 측면이 있는 반면, 사업관리 절차의 복잡성과 폐쇄성으로 인하여 무기체계 획득이 장기화하고 전력화 지연에 대한 우려가 제기되고 있다. 이는 방위산업의 투명성, 전문성에 중점을 둔 정책추진으로 신속성과 효율성이 결여되었다는 평가와 결과의 완전성 보장을 위한 세밀한 검증과 평가에 대한 지속적인 요구가 상존하고 있다. 또한, 4차 산업혁명의 민간 과학기술을 국방 무기체계에 신속하게 반영될 수 있도록 현 국방획득체계의 프로세스를 전면적으로 보완할 필요가 있다.

방위력개선사업 추진은 획득 방법에 따라 연구개발사업과 구매사업으로 구분된다. 표 2와 같이 연구개발사업은 기획-계획-체계개발-최초 양산-후속양산 순으로 진행되며, 총 19단계로 체계개발 단계가 7개 단계로 가장 많다. 구매사업은 기획-계획-기종 결정-최초 구매-후속 구매순으로 진행되며, 총 16단계로 기종 결정 단계가 6개 단계로 가장 높은 비율을 차지하고 있다. 사업추진 기간 측면에서 무기체계 연구개발사업 총 49개 사업을 대상으로 분석한 결과 그림 2와 같이 소요결정으로부터 최초 전력화까지는 16.8년, 전력화 완료까지는 23.7년이 소요되는 것으로 분석되었다. 이는 기획·계획단계에서 소요결정으로부터 사업타당성조사까지 9.5년으로 과다하게 소요되고, 전력화 완료까지는 약 24년이 소요되어 기술의 진부화 문제가 제기되는 등 제도적 개선이 필요하다.

표 2. 사업추진 단계¹⁶⁾

Table 2. Business promotion stage

연구개발(19단계)		구매(16단계)	
기획 (2)	1. 소요제기 2. 소요결정	기획 (2)	1. 소요제기 2. 소요결정
계획 (4)	3. 선행연구 4. 소요검증 5. 사업추진기본전략 6. 사업타당성조사	계획 (4)	3. 선행연구 4. 소요검증 5. 사업추진기본전략 6. 사업타당성조사
체계개발 (7)	7. 체계개발기본계획 8. 제안요청서 9. 계약체결 10. 시제제작 11. 시험평가 착수 12. 시험평가 종료 13. 규격화	기종결정 (6)	7. 구매계획 8. 제안요청서 9. 대상장비 선정 10. 시험평가 착수 11. 시험평가 종료 12. 기종결정
최초양산 (4)	14. 제조성속도평가(MRA) 15. 방산물자지정 16. 양산계획 17. 최초 양산 완료	최초구매 (2)	13. 최초 구매계약 14. 최초 구매 완료
후속구매 (2)	18. 후속 양산계획 19. 후속 양산 완료	후속구매 (2)	15. 후속 구매계약 16. 후속 구매 완료



그림 2. 무기체계 연구개발사업 추진 기간¹⁷⁾

Figure 2. Weapon system R&D project implementation period

16) 육군, 획득제도 개선방향, 2022, p.5

17) 상계서, p.5

제2절 시험평가(T&E) 이론적 고찰

1. 시험평가의 정의 및 역할

시험평가(T&E, Test and Evaluation)는 무기체계 개발 및 획득과정의 한 분야로 해당 무기체계에 대해 성능, 기술, 품질, 운용관리 측면에서 제반 요구 조건 및 개발목표의 충족 여부와 운용 적합성, 효율성 등을 확인 검증하는 절차로 시험과 평가의 합성어로 사용하고 있다.¹⁸⁾ 시험(Test)은 개발 및 운용 측면에서 대상 체계의 객관적 성능을 검증하고 평가하기 위한 기초 자료를 획득하는 과정을 의미한다.¹⁹⁾ 평가(Evaluation)는 시험을 통해 수집된 자료 또는 시뮬레이션 자료, 연구자료 등 기타 수단을 활용하여 획득한 자료를 근거로 평가 기준과 비교 분석하는 과정으로 포괄적으로는 시험을 통해 수집된 자료를 활용하여 체계의 가치를 판단하거나 의미 있는 정보를 도출하는 과정을 말한다.²⁰⁾ 즉, 사용자의 요구와 일치 여부를 검증하고 운용목적 부합 여부의 적합성을 판단하는 관리 도구(Tool)라 할 수 있다.

시험평가는 일반적으로 개발시험평가(DT&E, Development Test & Evaluation)와 운용시험평가(OT&E, Operation Test & Evaluation)로 구분되어 수행된다. 이중 개발시험평가는 연구개발주관기관이 표준화된 시험환경에서 요구성능을 측정하고 설계상의 중요한 문제점을 식별하고 보완하여 기술적으로 개발목표가 충족되었는지 검증하는 것이다. 운용시험평가는 소요군의 운용자에 의해 무기체계가 운용되는 작전환경 또는 유사 환경에서 작전운용성능(ROC)의 충족과 군 운용의 적합성, 효율성, 안전성 등을 확인하고 전력화지원요소 등에 대한 적합성을 시험평가 하는 것으로 무기체계 획득과정에서 매우 중요한 단계이다. 이러한 시험평가는 그림 3과 같이 체계개발 및 획득을 위한 체계공학

18) 합참, 무기체계 시험평가 업무 규정, 합참 규정-541-01('21. 2. 26. 제정), p.110

19) 육군 시험평가단, 전계서, 2019, p.17

20) 상계서, p.17

(SE, System Engineering) 과정의 한 집약된 부분으로 체계개발 시 초기에 성능 수준을 확인하고 개발자가 설계결함 등을 식별하여 보완할 수 있도록 도움을 주는 역할을 수행한다.

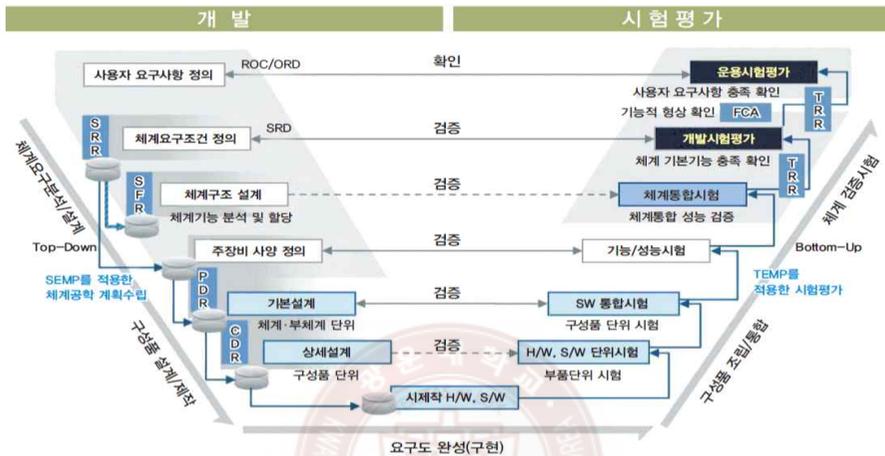


그림 3. SE 기반 무기체계 개발 및 시험평가 검증²¹⁾

Figure 3. SE-based weapon system development and T&E verification

또한, 시험평가는 개발 및 획득된 무기체계의 수명주기 비용을 결정하는 역할을 한다. 과거 무기체계의 획득과정에서 연구개발, 생산, 운용 및 유지단계의 비용 중에서 시험평가 비용이 차지하는 비율이 낮았으나, 최근에는 무기체계의 고도화됨에 따라 시험 비용이 급격하게 증가하고 있다. 일반적으로 그림 4와 같이 연구개발 무기체계의 전제 수명주기비용 중 연구개발 및 시험평가 비용이 전체 비용의 약 10%, 양산단계 비용이 약 30%, 운용유지단계 비용이 약 60%를 차지한다. 비록 연구개발 및 시험평가 비용이 총수명주기 비용에 10%에 불과하지만, 양산 및 운용유지 단계에서 무기체계 운용의 결정적 요소가 되는 활동으로 무기체계 획득과정에서 적정 비용과 일정을 통제하는 의사

21) 류중수, 업체주관 연구개발사업의 시험평가 발전방안 연구, 국방과 기술, 2019, p.101

결정에 결정적인 자료를 제공한다.²²⁾ 미 국방성 자료에 의하면 무기체계의 결함을 수정하는데 소요되는 비용은 개발비의 10~30%로 추정되고, 재설계 및 수정 비용은 개발 및 획득단계에 초기 시험평가를 통하여 결함을 식별하고 수정한다면, 감소시킬 수 있음을 제시하고 있다.²³⁾

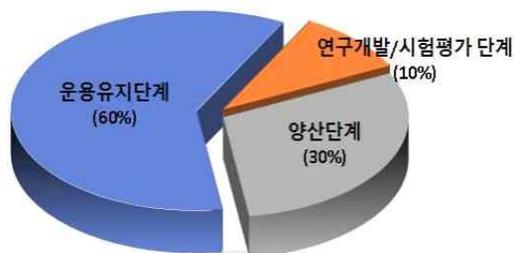


그림 4. 연구개발 무기체계 수명주기 비용²⁴⁾
 Figure 4. R&D weapon system life cycle cost

이와 같이 시험평가는 무기체계의 유용성과 비용 대비 효과 측면에서 유리성을 평가하기 위하여 적시에 정확하고 적절한 정보를 의사결정권자에게 제공한다. 무기체계 개발 및 획득과정에서 위험관리 도구로서 위험 요소를 최소화할 수 있도록 조치함으로써 획득 관련 기관으로 하여금 운용 측면에서 효과성과 적합성이 높은 무기체계가 획득되도록 보증하고, 원활한 사업관리가 진행될 수 있도록 하고 있다.

2. 무기체계 시험평가 수행절차

시험평가는 무기체계 개발 및 획득과정으로 합참에서 무기체계의 소요를 결

22) 방사청, 정밀유도무기 시험평가 기준 정립 및 절차 연구, 2014, p.92

23) 합참, 무기체계 시험평가 실무 가이드북, 2016, p.6

24) 상계서, p.6

정하면 방사청은 선행연구를 거친 후 방위력개선사업에 대한 추진 방법을 결정한다. 일반적인 무기체계 연구개발은 기술 수준과 획득 기간을 고려하여 결정된 사업추진 방법에 따라 그림 5와 같이 탐색개발과 체계개발 과정에서 시험평가를 수행한다.

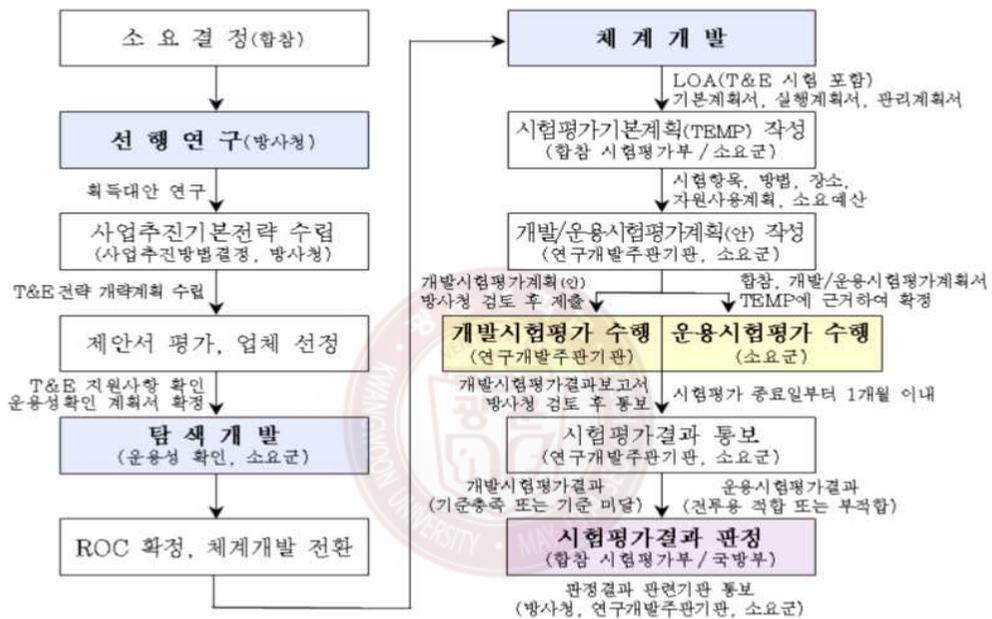


그림 5. 연구개발 무기체계 시험평가 수행절차²⁵⁾

Figure 5. R&D weapon system T&E procedure

선행연구 이전 소요제기로부터 소요결정 단계까지 소요군의 통합개념팀(ICT)에 참여하여 시험평가 수행과 관련 사항을 검토하고 합동 및 각 군의 무기체계 운영개념을 구현할 수 있는 작전요구성능(ROC)을 확인한다. 선행연구 단계에서는 연구개발 무기체계에 대한 개념을 형성하고 사업추진기본전략을 수립하는 단계로 시험평가 개념 및 추진계획을 수립한다.

25) 국방부, 국방전력발전업무훈령, 2021, pp.187~189, 연구자 재정리

탐색개발 단계는 무기체계의 핵심부분에 대한 기술을 개발하고 핵심기술의 완성도 및 무기체계에 대한 적용 가능성을 확인하여 체계개발단계로 진행 여부를 판단하는 단계로써 무기체계 기술개발과 시물레이션 및 모형제작, 시험 또는 시제품 제작과 시험을 통해 개발기술을 입증한다.²⁶⁾

표 3. 무기체계 시험평가 관리 활동²⁷⁾
Table 3 Weapon system T&E management activities

구 분	소요제기 /소요결정	선행연구	탐색개발	체계개발
요구도 관리	작전운용능력 및 기부성능 개발 및 관리			
	운용요구서(ORD) 개발 및 발전			체계요구서(SRD) 개발 및 발전
시험평가 계획관리	시험평가 개념 개발/발전			
	시험평가전략 수립/적용		시험평가기본계획서 개발/적용	
	통합시험계획 수립/관리			
	운용성확인			
	개발시험평가 계획/적용			
	운용시험평가 계획/적용			

또한 작전운용성능을 포함한 구체적인 운용요구서(ORD) 및 시스템을 설계하기 위한 체계요구서(SRD) 초안을 작성한다. 이를 바탕으로 예비 시험평가기본계획서를 작성하고 완성된 시제품 또는 실제 체계개발 시험 이전 단계의 해석, M&S, 시범 등을 통해 적용시스템의 잠재적인 운용 효과와 운용 적합성에 관한 의사결정 자료를 조기에 제공하기 위한 평가로 운용성확인을 수행한다.

체계개발 단계는 무기체계를 설계하고 이에 따른 시제품을 생산하여 시험평가를 통해 양산에 필요한 규격을 완성하는 단계로 시험평가기본계획서(TEMP)

26) 합참, 전계서, p.34

27) 방사청, 정밀유도무기 시험평가 기준 정립 및 절차 연구, 2014, p.92

를 기준으로 상세한 시험평가 계획을 수립하고 시험평가를 수행한다. 현재 무기체계 획득과정에서 체계개발 단계의 시험평가는 개발시험평가와 운용시험평가로 구분된다.

개발시험평가는 연구개발주관기관에서 시제품에 대한 기술상 성능(신뢰도, 유지성, 호환성, 적합성, 내환경성, 안전성 등)을 측정하고, 설계상의 중요한 문제점이 해결되었는지를 확인하고 평가한다. 운용시험평가는 소요군에서 개발장비의 시제품에 대하여 각종 작전환경 또는 이와 동등한 조건에서 작전운용 성능 충족 여부 및 군 운용 적합 여부를 확인하기 위한 시험평가로서 사업의 특성 및 효율적인 예산 집행을 위해 필요하다고 판단할 경우 최초 시험평가 단계와 후속 양산 결정을 위한 후속 시험평가 단계로 구분하여 수행할 수 있다.²⁸⁾ 이를 위하여 무기체계 시험평가는 표 4와 같이 국방전력발전업무 훈령에 정의된 개발시험평가 및 운용시험평가의 주요 시험평가 항목에 대하여 연구개발주관기관과 소요군의 시험평가 주관기관에 의해 수행한다.

표 4. 무기체계 시험평가 항목²⁹⁾

Table 4. Weapon system T&E items

개발시험평가 항목	운용시험평가 항목
1. 무기체계 기능·성능시험 가. 주요 작전운용성능 시험 나. 기술적·부수적성능 시험 다. 설계검토를 통하여 확정된 기능 및 성능시험 라. 소프트웨어시험 마. 환경시험 바. 전자파 적합성 시험 등	1. 작전운용성능의 충족성 시험 가. 주요 작전운용성능 나. 기술적/부수적 성능
2. 핵심부품·구성품 시험 가. 성능시험 나. 환경시험 다. 신뢰성시험	2. 합동성 및 상호운용성 가. 운용개념 및 체계특성 나. 연동성 및 정보교환 다. 표준 및 아키텍처 라. 정보보호 마. 주파수

28) 상계서, pp.47~48

29) 국방부, 전계서, pp.51~68, 연구자 재정리

3. 소프트웨어 신뢰성시험	3. 군 운용 적합성
4. 전력화지원요소의 기술적 입증시험	가. 운용 및 조작 적합성, 안정성 나. 전술적 운용의 적합성
5. 합동성 및 상호운용성 시험 가. 운용개념 및 체계특성	다. 기존 무기체계와의 상호운용 적합성 라. 환경 적응성
나. 연동성 및 정보교환 다. 표준 및 아키텍처 라. 정보보호 마. 주과수	4. 전력화지원요소의 실용성 확증시험



제3절 인공지능(AI)의 개념 및 기술발전 동향

1. 인공지능(AI)의 개념

오늘날 우리는 부지불식간에 인공지능과 함께 살아가고 있다. ‘인공지능(AI, Artificial Intelligence)’이란 인간이 가진 지적 능력을 컴퓨터를 통해 구현하는 기술로 인공지능은 ‘소프트웨어(SW, Software)’로 존재할 수도 있고 로봇처럼 ‘몸(body)’을 가질 수도 있다.³⁰⁾

인공지능이라는 용어는 1956년 ‘다트머스 회의(Dartmouth Conference)’로 알려진 모임에서 처음으로 사용하였다. 다트머스대학교의 수학자이자 컴퓨터 과학자인 매카시(J. McCarthy)는 ‘AI 하계 연구 프로젝트’(J. McCarthy et al., 1955)를 기획하면서 10여 명의 과학자가 두 달 동안 공동으로 AI 연구를 수행할 것을 제안하였다. 이때 참석자 A. 뉴얼(Newell)과 사이먼(H. Simon)³¹⁾은 “향후 20년 안에 기계는 사람이 할 수 있는 일은 무엇이든지 할 수 있게 될 것”이라고 주장했다.³²⁾

인공지능(AI)의 사전적 의미는 기억·지각·이해·학습·연상·추론 등 인간의 지성을 필요로 하는 행위를 기계·컴퓨터를 통해 실현하는 기술이다.³³⁾ 인공지능은 ‘머신러닝(Machine Learning)’이라는 기술을 기반으로 하고 있다. 머신러닝은 컴퓨터가 학습 모형을 기반으로 스스로 학습하고 AI의 성능을 향상시키는 것을 말한다. 빅데이터를 분석 가공해서 새로운 정보를 얻어 내고 미래를 예측할 수 있는 기술로, 집약된 데이터 기반으로 상관관계 및 특성을 식별하고

30) 소형근, 포스트 휴머니즘 시대에 인간과 인공지능 이해하기: 강한 인공지능의 등장에 어떻게 대처할 것인가?, 구약논단 제28권 4호(통권 86집), 2022, pp.250~253

31) 1978년 노벨 경제학상을 수상하고, 참석 당시에는 컴퓨터 언어 정보 처리와 컴퓨터 프로그램인 Logic Theorist와 General Problem Solver에 관해 연구함.

32) 오승환·이주량·최병삼 등 11명, 인공지능 기술 활용 강국을 향한 과학기술정책 제고 전략, 과학기술정책연구원(STEPI), 2020, p.7

33) <https://terms.naver.com/entry.naver?docId=4356939&cid=43665&categoryId=43665>, 네이버 시사경제용어사전, (검색일: 2023. 4. 9.)

결론을 도출한다.

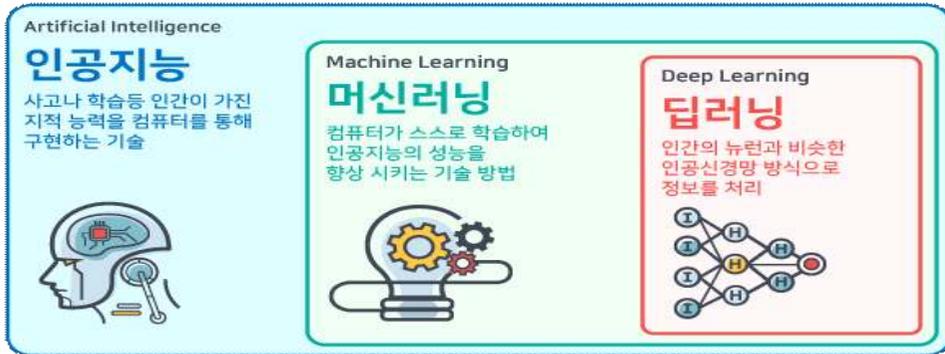


그림 6. 인공지능, 머신러닝 및 딥러닝과의 관계³⁴⁾

Figure 6. Relationship between Artificial Intelligence, Machine Learning and Deep Learning

이와 같은 ‘머신러닝’ 중 각광 받는 기술은 바로 ‘딥러닝(Deep Learning 또는 representation learning)’이다. 딥러닝은 컴퓨터가 수많은 데이터를 이용하여 마치 인간의 뉴런과 같이 스스로 학습할 수 있는 인공신경망 방식으로 정보를 처리하고 학습하는 기술이다. 딥러닝은 알고리즘을 계층으로 구성하고, 축적된 데이터를 분석만 하지 않고 스스로 데이터를 학습하는 능력을 활용하여 결론을 도출한다. 인공지능(AI)의 종류에는 개발 방법 및 인간의 지능을 모방한 정도 등에 따라 표 5와 같이 약한 AI(Weak AI)과 강한 AI(Strong AI), 슈퍼 AI(Super AI)로 구분할 수 있다.

일반적으로 분류하고 있는 인공지능의 분류기준과 관련 내용을 정리해 보면 약한 AI는 특정한 분야의 주어진 과업을 인간의 지시에 따라 수행하는 것으로 구글의 알파고 등이 해당된다. 강한 AI는 어떤 문제를 스스로 ‘학습·판단·해결’할 수 있는 인간 수준의 인공지능으로 아이언맨의 자비스 등이 해당된다. 만약 AI 기술이 더 발전한다면, 지금의 ChatGPT는 다음 단계로 진화하여 강한

34) <https://m.blog.naver.com/pwj6971/221614497987>, (검색일: 2023. 4. 9.)

AI를 넘어 슈퍼 AI로 나아갈 수 있을 것으로 본다.

표 5. 인공지능(AI)의 종류³⁵⁾
Table 5. Types of Artificial Intelligence

유형	설명	예시
약한 AI (Weak AI)	<ul style="list-style-type: none"> 주어진 조건 아래서만 작동 가능 사람의 지능으로만 가능하던 작업을 수행하는 AI 	<ul style="list-style-type: none"> 애플의 시리 아마존의 알렉사 구글의 알파고 IBM의 왓슨 카네기멜론대 딥블루
강한 AI (Strong AI)	<ul style="list-style-type: none"> 스스로 사고하고 해결하는 AI 인간과 같은 사고가 가능한 AI 	<ul style="list-style-type: none"> 아이언맨 자비스, HER 사만다 비서 로봇
슈퍼 AI (Super AI)	<ul style="list-style-type: none"> 모든 영역에서 인간을 뛰어넘는 AI 	<ul style="list-style-type: none"> 어벤저스의 비전 터미네이터 스카이넷

군사적으로 인공지능(AI)은 사전적 의미의 AI 개념을 바탕으로 지휘관이 군사 업무를 수행할 때 의사결정(지휘결심)을 지원하는 기술을 말한다. 현재 군에서 분류하고 있는 AI 기반의 무기체계는 감시·정찰, 지휘통제, 유·무인 복합, 사이버 및 기타 체계 등에 지능화 능력을 구비하여 운용하는 체계를 말하며, AI 기반 전력지원체계는 전투지원장비, 전투지원물자, 의무, 교육훈련, 국방정보시스템, 그 밖의 체계(시설 등)에 지능화 능력을 구비하여 운용하는 체계로 구분하고 있다.³⁶⁾ 하지만 AI 기반 무기체계와 전력지원체계를 명확히 구분하는 것은 어렵다. 예를 들어 무기체계의 지휘통제 AI와 전력지원체계의 의무 AI 기능이 결합된 경우 어떤 체계로 분류할 것인지는 관련 기관의 협의를 통해 구분해야 한다.

국방 분야 AI 기술의 활용 수준에 대한 과학기술정책연구원(STEPI)에서 전

35) 오승환·이주량·최병삼 등 11명, 전게서, p.8 연구자 재정리

36) 육군, 인공지능 종합발전 계획, 2022, p.12

문가를 대상으로 한 인식조사 설문 결과, 전문가의 74%가 ‘매우 미흡’하거나 ‘미흡’하다고 응답하여 대체로 부정적인 평가를 보였으며, 국내 국방 분야 인공지능 기술 도입의 중요성은 전체 응답자의 94%가 ‘많은 분야에서 중요’하거나 ‘대부분의 분야에서 중요’하다고 대답하였다.³⁷⁾ 이는 국방 AI 기술의 도입이 매우 중요하나 현재 활용 정도는 미흡하다고 생각하고 있다. 또한 국방 분야에서 AI를 활용하여 얻을 수 있는 기대효용에 대한 응답은 그림 7과 같이 인공지능 기술의 효용이 현재까지는 전력 자원 관리나 위험 임무 대체와 같은 보조적인 임무 수행에 가장 많은 초점을 두고 있으나, 전투력 강화 및 지원 역량 측면에 대해서도 기대를 하고 있음을 암시하고 있다.³⁸⁾

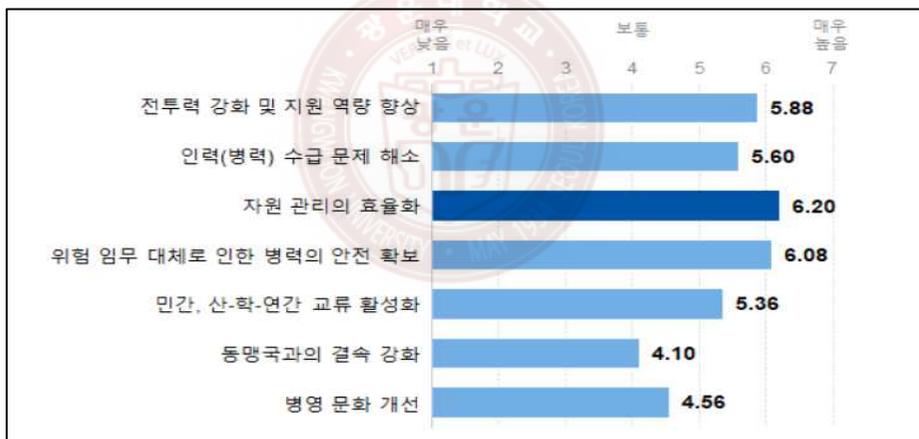


그림 7. AI 활용 기대효용 정도³⁹⁾

Figure 7. Expected utility of using AI

반면, AI 기술 활용이 높아짐에 따라 예상되는 문제점으로 그림 8과 같이 ‘데이터 확보에 따른 보안 문제’를 가장 많이 제기했다. 이어서 ‘관련 법제 미

37) 오승환·이주량·최병삼 등 11명, 전게서, pp.236~242

38) 상게서, pp.236~242

39) 상게서, p.241

비 및 충돌'과 '윤리적 문제' 순으로 응답하였다. 이것은 향후 국방 AI 전력 향상에 필수적인 데이터의 확보를 위해 까다로운 보안 절차에 대한 규제 개선이 필요한 것임을 시사하고 있어 국방예산 급증의 문제는 사회적 합의의 문제로 상대적으로 큰 이슈로 보지 않고 있다.⁴⁰⁾

국방 분야의 AI 기술 도입을 위해서는 'AI 전문인력의 부족'을 최우선적으로 해결해야 할 과제이며, 'AI 활용을 저해하는 각종 규제' 등에 대해서도 상대적으로 해결해야 할 우선 과제로 보고 있었다.⁴¹⁾

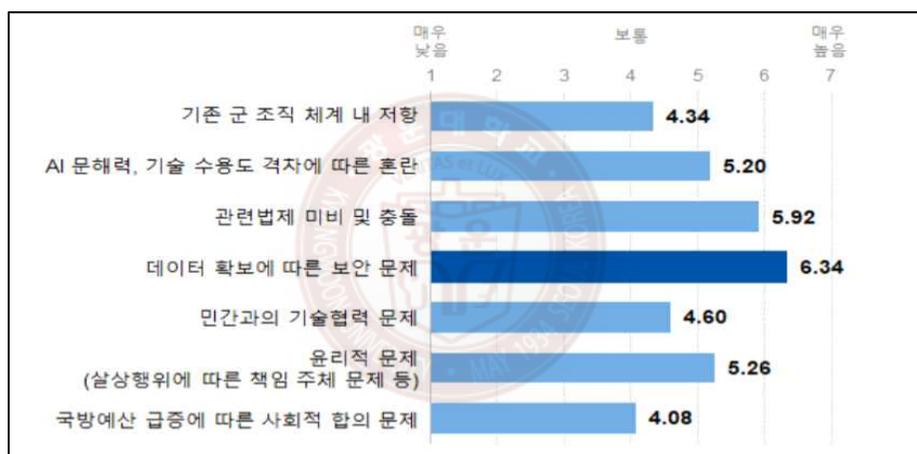


그림 8. 국방 분야 AI 활용에 따른 문제 발생⁴²⁾

Figure 8. Problems arising from the use of AI in the defense sector

이상과 같이 AI를 군사적으로 활용하기 위해서는 인프라 구축이 필요하다. 또한, AI를 활용하기 위해서는 데이터의 활용과 관련 문제를 해결해야 할 필요가 있다. 보안 문제나 제도적 규제가 여기에 해당 된다고 할 수 있다.

40) 윤정현, 국방 분야의 인공지능 활용성 제고 방안과 시사점, 과학기술정책 포커스, 과학기술정책연구원, 2021, pp.46~49

41) 상계서, p.49

42) 상계서, p.47

2. 인공지능(AI)의 기술발전 동향

“미래는 이미 와 있다. 단지 널리 퍼져 있지 않을 뿐이다(The future is already here. It's just unevenly distributed).” 1982년 소설 “뉴 로맨서”를 통해 과학소설의 한 장르인 사이버 펑크의 거장이 된 윌리엄 깁슨(William F. Gibson)이 이코노미스트와 인터뷰(2003년)에서 한 말이다.⁴³⁾ 오늘날 우리는 실생활속에서 사이버공간(Cyberspace)이 없으면 살아가기 힘들 정도로 인터넷과 인공지능(AI)에 의존하고 있다. 코로나-19(COVID-19) 이전부터 우리의 삶 속에 깊숙이 자리를 잡고 있던 각종 인터넷 강의와 인터넷 뱅킹, 카카오톡(Kakao Talk), 페이스북(Facebook), 인스타그램(Instagram) 등은 코로나-19로 우리의 삶을 완전히 바꿔놓고 있다. 또한, 최근에는 관공서를 비롯하여 민간기업에서 상담원 대신 챗봇(chatbot)⁴⁴⁾이 업무를 담당하고 있다. OpenAI가 개발한 ChatGPT-4⁴⁵⁾는 현재 가장 진화한 자연어 기반 인공지능 챗봇으로 평가되고 있다. ChatGPT는 스마트폰 이후, 전 세계에 기술 쇼크를 불러일으키고 있다.

인공지능의 기술 수준은 인공지능 기술을 개발하고, 응용이나 서비스를 사업화하기 위해 표준화된 지표로서 표 6과 같이 한국전자통신연구원(ETRI)에서 제시한 기술수준(ATL, Artificial Intelligence Technology Level)의 지표를 사용하고 있다.

인공지능 기술은 현재 다양한 산업 분야에서 사용되고 있다. 대표적으로 앞에서 제시한 OpenAI가 개발한 ChatGPT-4가 있으며, Google Home⁴⁶⁾, Tesla

43) <https://www.sedaily.com/NewsView/1VQVVVLXHAV>, 서울경제(2019. 11. 21.), (검색일: 2022. 11. 20.)

44) 챗봇: 기업용 메신저에 채팅하듯 질문을 입력하면 인공지능(AI)이 빅데이터 분석을 바탕으로 일상언어로 사람과 대화를 하며 해답을 주는 대화형 메신저를 말한다. [네이버 지식백과] (한경 경제용어사전)

45) ChatGPT: 2023년 3월 14일 챗GPT를 개발한 오픈AI가 공개한 최신 인공지능(AI) 언어 모델로서, 기능 자체는 이전에 나온 챗봇과 다를 바 없으나 차이점은 결과물이다. 전문가가 썼다고 해도 손색없는 수준의 보고서를 작성하는가 하면 시를 짓기도 한다. [네이버 지식백과] (한경 경제용어사전)

Autopilot⁴⁷⁾, 챗봇 등 이외에도 인공지능 관련 제품은 매년 증가하고 있다. 인공지능 기술의 발전에 따라 다양한 제품이 출시되고 있다.

표 6. 인공지능(AI) 기술수준(ATL) 지표⁴⁸⁾
Table 6. AI technology level indicator

Level	1)지식성장	2)지능범위	3)도메인	4)인간-AI 협업	5)모달리티	6)통합인지
ASI (Artificial Super Intelligence)	6 인간의 개입 없이 지식 성장	인간의 6대 지능 이상을 지원하고 빠르고 정확함	오픈 영역	AI 스스로 미션을 생성하고 수행	5개 이상	AI 자기인지와 인간보다 빠른 직관
AGI (Artificial General Intelligence)	5 인간의 부분개입 지식성장	인간의 4대 지능 이상을 통합	이중 분야로 확장 가능	AI가 판단근거 설득, 인간-AI 상호협의	4개 이상	인간/상황/이력/지식 종합판단
	4 인간 개입으로 지식성장	인간의 2대 지능 이상을 통합	유사 분야로 확장 가능	AI가 판단근거를 설명하여 추천	2~3개	인간/상황/이력/지식 부분종합
ANI (Artificial Narrow Intelligence)	3 비지도 학습 (고정지식)	1개 분야 전문가 수준으로 지식처리 가능	동일 분야 내 확장	AI 추천으로 AI-인간 협업 미션 수행	ND	인간과 주변 상황 연결판단
	2 지도 학습 (고정지식)	단일 분야에서 단순한 업무처리	ND	AI-인간 협업으로 사전 정의된 미션 수행	ND	언어, 이미지, 감정, 현상 등 동종 인지
	1 규칙 기반 (고정지식)	정해진 조건 비교로 대응처리	고정 분야 적용	사전에 정해진 순서에 따라 수행	1개	HW 인지 수치 의존 (온도 센서 등)
똑똑한 AI(IQ)			인간과 소통하는 AI(EQ)			

46) Google Home: 사용자의 음성을 인식해 질문을 파악하고 음악 재생, 예약, 스케줄 조회, 메시지 전송 등을 수행하는 AI 비서 시스템인 구글 어시스턴트를 집안에서 사용할 수 있도록 한 단말기로, 구글이 2016년 5월 18일 구글 어시스턴트와 함께 공개되었다. 집안의 디지털 기기를 연결하는 사물인터넷의 허브 역할을 하며 와이파이, 마이크, 스피커, 조명 등이 탑재되어 있다. [네이버 지식백과] (시사상식사진)

47) Tesla Autopilot: 테슬라에서 출시한 자율주행 기능으로, 카메라, 레이더 및 초음파 센서 등을 사용하여 차량의 주행을 자동화한다. 차선 유지, 차간 거리 유지, 자동차 감지 등 다양한 기능을 수행할 수 있다. [네이버 지식백과] (ICT 시사상식)

48) 한국전자통신연구원. ATL 1.0 정의, 2020

- 1) 지식성장: 새로운 사실을 인지하고, 지식을 쌓는 과정의 자율성
- 2) 지능범위: 인간의 9대 지능⁴⁹⁾에 대한 커버리지(적용 범위)
- 3) 도메인: AI 모델이 적용될 수 있는 도메인의 크기, 광범위성
- 4) 인간-AI 협업: 주어진 미션, 또는 새로운 문제를 해결하기 위한 인간-AI 간의 협력 방법
- 5) 모달리티: 시각, 청각, 테스트, 센서 등 입력 데이터 형식의 통합 정도를 정의
- 6) 통합인지 수준: AI가 대상자와 연결된 정보, 상황, 이력 등을 종합하여 인지하는 정도

우리나라 및 주요국의 인공지능(AI) 기술수준과 관련하여 최근 변화 추이에 대해 정보통신기획평가원(IITP)의 'ICT기술수준조사'에 따르면 그림 9와 같이 AI 분야의 세계 최고기술 보유국은 미국으로, 미국 대비 AI 기술격차는 중국(0.8년), 유럽(1.0년), 한국(1.3년), 일본(1.5년) 순으로 나타났다.⁵⁰⁾

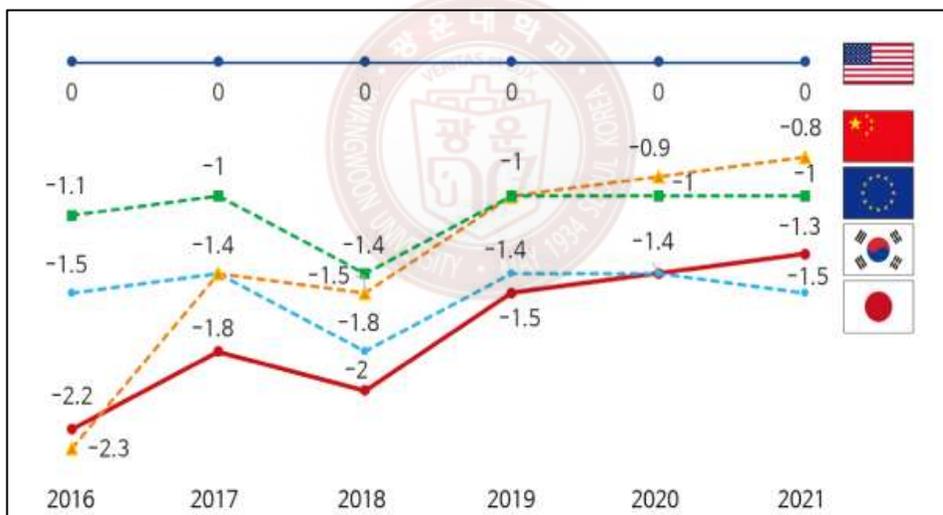


그림 9. 2016~2021년 우리나라와 주요국의 AI분야 총괄 기술수준 추이⁵¹⁾

Figure 9. 2016-2021 Trends in overall technology level in the AI field in Korea and major countries

49) 인간의 9대 지능요소(하워드 가드너, 1983): 언어지능, 논리-수학적 지능, 공간지능(시각 포함), 신체·운동적 지능, 음악지능, 개인 내 지능, 자연주의적 지능, 대인관계 지능, 실존지능

50) https://spri.kr/posts/view/23555?code=data_all&study_type=industry_trend, SPRI소프트웨어 정책연구소(2023. 3. 2.), (검색일: 2023. 4. 1.)

51) 정보통신기획평가원, ICT 기술수준 조사 및 기술경쟁력보고서

반면, 국방 AI 분야 기술 수준은 표 7과 같이 국방기술진흥연구소(KRIT)에서 2021년 기준 미국 대비 77%의 기술 수준으로 약 2.5년의 기술격차가 존재하고 있어 국가적인 인공지능(AI) 분야 총괄 기술 수준의 차이가 더 크게 나타나고 있다.

표 7. 국방 인공지능(AI) 기술수준⁵²⁾
Table 7. Defense AI technology level

구 분		기술수준(%)		기술격차(년)	
		미국	한국	미국	한국
학습지능	머신러닝	100	69	0	2.6
신뢰지능	설명 가능한 AI	100	82	0	1.3
	견고한 AI	100	67.5	0	4.3
	공정한 AI	100	82.7	0	1.9
고성능 AI H/W	모델 경량화	100	84	0	1.5
	서버용 H/W	100	60	0	5
	모바일 / Edge H/W	100	85	0	1.5
전장인식	영상인식	100	95	0	0.5
	음성인식	100	100	0	0
	언어인식	100	86	0	1.3
	감정인식	100	60	0	7
자율판단	전장환경 및 상황인지	100	70	0	3
	국방추론 및 지식표현	100	80	0	2.5
지휘결심	임무계획	100	70	0	3
	임무할당	100	70	0	3
임무수행	행동지능	100	80	0	2
	유·무인 협업제어	100	75	0	2.5
전 체		100	77.4	0	2.5

이에 국내 방위산업의 기술격차를 줄이기 위한 R&D 사업은 핵심기술, 미래 도전기술 등 169건에 약 1.1조 원, AI 서비스 구축사업은 체계개발, 부처협력

52) 국방기술진흥연구소, 미래 국방 2030 기술전략(AI), 2021

정보통신기술(ICT) R&D 등 53건에 0.7조 원 등 다양한 유형의 사업을 진행하고 있다.⁵³⁾ AI, 무인·자율 분야 등 첨단기술 분야에 국방 R&D 예산 규모를 확대하여 선진국과의 기술격차를 극복하기 위해 R&D 사업 및 전력화 사업을 추진하고 있다.

미국을 비롯한 주요 선진국들은 AI 기반 군사력을 강화하고 있다. 미국은 AI를 미래전에 대비한 ‘3차 상쇄전략’ 구현의 핵심 수단으로 보고 국방 AI 거버넌스를 구축하여 기술개발에 투자를 확대하고 있다. 중국과 러시아에 대비하여 2030년까지 250~300명의 전투병과 수천 대의 로봇으로 구성된 새로운 전투단을 만들 계획이다. 중국은 미국에 대한 상대적인 열세를 극복하기 위해 강군 건설의 핵심 요소로 AI를 선정하고 대규모 예산 투자로 미래 지능화전을 선제적으로 준비하고 있다. 2027년까지 ‘AI 기반 인민해방군 현대화’를 공식화하고 ‘군민융합’을 강조하면서 민간업체들과 무인 드론·지능형 전투체계를 구축하기 위해 노력 중이다. 러시아는 2030년까지 군대의 30%를 원격통제 또는 자율화된 로봇으로 대체하고, 2025년 AI 기반으로 독자적 전투단위로 로봇군 창설을 예고하고 있다. 일본은 오래전부터 저출산, 고령화 등의 사회문제를 겪고 있다. 이에 AI를 활용하여 저성장·고령화 극복을 위한 국가 경제·사회 혁신의 수단으로 AI 기술경쟁력 확보의 중요성을 강조하고, AI 기술혁신을 가속화하고 있다.⁵⁴⁾

이와 같이 AI 기술을 적용한 선진국의 무기체계는 미국의 자율무인잠수함과 4족 보행 로봇 등이 있으며, 중국은 무인지상장비 ‘침병’과 수중 로봇 ‘인조 상어’ 등이 있다. 우리 군도 최근 5년간(2018~2022년) AI를 군에 적용하는 사업에 약 1조 8,903억을 투자하여 국방과학연구소(ADD)를 중심으로 무인 자율체계를 개발 중이며, 최근 무인수색차량⁵⁵⁾ 탐색개발을 완료하고 위성영상 식별

53) 육군, 전계서, 2022, p.17

54) 오승환·이주량·최병삼 등 11명, 전계서, pp.32~42

기술, AI 공중교전기술 등 44개의 연구를 진행하는 등 상당 부분 기술을 축적하고 있다. 또한, 국방부는 지난 2022년 6월 AI 과학기술 강군 개념 및 추진 방향을 발표하면서 국방 AI의 전면적 확산을 위한 핵심 요소 및 접근 전략으로 그림 10과 같이 초기 단계의 현 국방 AI 수준을 높이기 위해 추진력을 확보하고 역량을 결집시키겠다고 했다. 중요 분야에 대해서는 민·관·군 협업을 긴밀하게 추진하고 있다.



그림 10. 국방 AI 기술의 발전 방향(3단계 발전 모델)⁵⁶⁾

Figure 10. Development direction of defense AI technology(3-stage development model)

55) 전술도로 자율주행, 충돌 방지·회피, 지능형 다중표적 탐지 등 AI 기술을 적용하고 있다.

56) 국방부, AI 과학기술 강군 개념 및 추진 방향, 국방부 연두보고, 2022

제4절 선행연구 고찰

1. 무기체계 시험평가 관련 선행연구

무기체계 시험평가와 관련된 연구사례는 시험평가 정책 및 제도, 기술적인 측면에서 기존 무기체계에 대한 시험평가의 신뢰성을 향상시킬 수 있는 방안 등을 중심으로 연구가 이루어지고 있었다. 다만, 4차 산업혁명과 연계한 첨단 무기체계 변화에 관한 시험평가 연구는 부족하여 추가 연구가 필요하다.

임영봉(2014)은 “육군 무기체계 운용시험평가에 관한 연구”에서 시험평가 제도와 절차를 재정비하여 시험평가를 획득 전(全) 순기에 적극적으로 참여하여 개발 초기부터 개발위험을 적극적으로 방지하고, 해외 시험시설 활용을 위한 국제협력 강화 등을 제시하였다.

박종완(2015)은 “무기체계 시험평가의 신뢰성 향상방안”에서 시험평가 신뢰성에 영향을 미치는 내외부 요인을 분석하여 시험평가 전문화, 개발시험평가의 운용시험평가 전환을 위한 검증체계 정립과 내구성 환경시험에 대한 인식을 제고하고 시험평가 적정 시제 산출 및 시험평가 업무의 표준화 정립 등을 제시하였다.

강응섭(2016)은 “국방 기술 환경변화에 부응한 무기체계 시험평가 신뢰도 제고에 관한 연구”에서 안정성(정책·조직) 측면에서 시험평가계획을 종합적으로 작성 검토하고 조정과 통제할 수 있는 조직의 신설과 과학적(투명·객관성) 측면에서 입증된 시험은 통합시험을 시행하고, 시스템엔지니어링 등 다양한 기법 적용 및 전문성(인력·기술) 측면에서 시험평가 전문 특기를 부여하는 방안을 제시하였다.

정월균(2017)은 “이해관계자 분석을 통한 무기체계 평가제도에 관한 연구”에서 시험평가사령부 창설, 시험장 구축, 전문 시험평가관 양성 및 예산 반영, 운용시험평가 시설 및 장비편성 등 제도적 보완을 제시하였다.

이용학(2017)은 “국방 무기체계 시험평가의 투명성에 관한 제언”에서 독립된 전문 시험기관으로 ‘국방 무기체계 시험평가원’ 설립을 제시하고, 시험평가의 신뢰성과 전문성, 투명성을 높이는 방안을 제시하였다.

김선영·최기일(2019)은 “방위산업 발전을 위한 무기체계 시험평가 개선방안 연구”에서 시험평가 항목 및 기준의 적법성을 강화하고 합참에 의한 시험평가 및 행정 소요 기간 단축과 제안서평가와 시험평가 연계성 강화 등을 제시하였다.

최규옥·송재은·유영민(2019)은 “무기체계에 적용 가능한 적응형 시험평가에 관한 연구”에서 시제품에 대한 평가체계를 개발시험평가 및 운용시험평가 위주에서 연구개발 전(全) 기간에 걸쳐 단계별 평가 및 수시 평가를 통해 진행 수준에 맞게 적시적으로 수정·보완하여 연구개발하는 적응형 시험평가 체계 적용방안을 제시하였다.

이강경(2020)은 “국방 획득환경 변화에 적합한 무기체계 운용시험평가 개선 전략 도출”에서 사회적으로 관심이 증대되고 있는 전투원 안전보장, 국방개혁 2.0 추진과 연계한 수명주기 비용 최적화 등을 통해 국방예산을 절감하고 미래 전장 환경의 변화에 대비한 시대적 요구를 반영할 것을 제시하였다.

이용복·전일국·김상홍(2022)은 “인공지능(AI) 무기체계 시험평가방법 발전방안 연구(분류모델 성능평가를 중심으로)”에서 인공지능 모델의 성능을 평가할 수 있는 척도의 필요성과 분류모델 인공지능 체계의 성능평가 분야에서 혼동 행렬을 중심으로 설명하였다. 시험평가 이해당사자 간 해석의 차이를 줄이고, 시험평가 대상 체계의 본질적인 성능을 정확히 평가할 수 있는 인공지능 성능평가 척도 적용의 필요성을 제안하였다.

표 8. 시험평가 관련 연구 및 보고서

Table 8. T&E related research and reports

구 분	연구 제목 및 연구 결과
임영봉 (2014)	<ul style="list-style-type: none"> ◦ 육군 무기체계 운용시험평가에 관한 연구 <ul style="list-style-type: none"> - 시험평가 제도와 절차 재정비하여 시험평가를 획득 전(全) 순기에 적극적으로 참여하여 개발 초기부터 개발위험을 적극적으로 방지하고 - 시험평가 판정 절차를 개선과 시험평가 전문 특기 부여 - 개발시험과 운용시험을 통합하여 통합시험평가체계 정착 - 해외 시험시설 활용을 위한 국제 협력 강화 등을 제시
박종완 (2015)	<ul style="list-style-type: none"> ◦ 무기체계 시험평가의 신뢰성 향상방안 <ul style="list-style-type: none"> - 시험평가 신뢰성에 영향을 미치는 내외부 요인을 분석하여 시험평가 전문화, 개발시험평가의 운용시험평가 전환을 위한 검증체계 정립 - 내구성 환경시험에 대한 인식 제고, 적정 시제 산출, 시험평가 업무의 표준화를 정립 등을 제시
강응섭 (2016)	<ul style="list-style-type: none"> ◦ 국방기술 환경변화에 부응한 무기체계 시험평가 신뢰도 제고에 관한 연구 <ul style="list-style-type: none"> - 안정성(정책·조직) 측면에서 시험평가계획을 종합적으로 작성 검토하고 조정과 통제할 수 있는 조직 신설 - 과학적(투명·객관성) 측면에서 입증된 시험은 통합시험 수행 - 시스템엔지니어링 등 다양한 기법 적용, 전문성(인력·기술) 측면에서 시험평가 전문특기 부여하는 방안 제시
정월균 (2017)	<ul style="list-style-type: none"> ◦ 이해관계자 분석을 통한 무기체계 평가제도에 관한 연구 <ul style="list-style-type: none"> - 시험평가사령부 창설, 시험장 구축, 전문 시험평가관 양성 및 예산 반영, 운용시험평가 시설 및 장비편성 등 제도적 보완을 제시
이용학 (2017)	<ul style="list-style-type: none"> ◦ 국방 무기체계 시험평가의 투명성에 관한 제언 <ul style="list-style-type: none"> - 독립된 전문 시험기관으로 '국방무기체계시험평가원' 설립과 시험평가의 신뢰성과 전문성, 투명성을 높이는 방안을 제시
김선영·최기일 (2019)	<ul style="list-style-type: none"> ◦ 방위산업 발전을 위한 무기체계 시험평가 개선방안 연구 <ul style="list-style-type: none"> - 시험평가 항목 및 기준의 적법성을 강화하고 합참에 의한 시험평가 및 행정소요 기간 단축과 제안서평가와 시험평가 연계성 강화 등을 제시
최규옥·송재은· 유영민 (2019)	<ul style="list-style-type: none"> ◦ 무기체계에 적용 가능한 적응형 시험평가에 관한 연구 <ul style="list-style-type: none"> - 시제품에 대한 평가체계를 개발시험평가 및 운용시험평가 위주에서 연구개발 전(全) 기간에 걸쳐 단계별 평가 및 수시평가 - 진행 수준에 맞게 적시적으로 수정·보완하여 연구개발하는 적응형 시험평가 체계 적용방안을 제시

이강경 (2020)	<ul style="list-style-type: none"> ◦ 국방 획득환경 변화에 적합한 무기체계 운용시험평가 개선전략 도출 - 사회적으로 관심이 증대되고 있는 전투원 안전보장, 국방개혁 2.0 추진과 연계한 수명주기비용 최적화를 통해 국방예산을 절감하고 미래 전장환경의 변화에 대비한 시대적 요구를 반영할 것을 제시
이용복·전일국· 김상홍 (2022)	<ul style="list-style-type: none"> ◦ 인공지능(AI) 무기체계 시험평가방법 발전방안 연구 - 인공지능 모델의 성능을 평가할 수 있는 척도의 필요성과 분류모델 인공지능 체계의 성능평가 분야에서 혼동행렬을 중심으로 설명 - 시험평가 이해당사자 간 해석의 차이를 줄이고, 시험평가 대상 체계의 본질적인 성능을 정확히 평가할 수 있는 인공지능 성능평가 척도 적용의 필요성을 제안

2. 인공지능(AI)의 군사용 활용 관련 선행연구

4차 산업혁명의 시대, 인공지능(AI) 기반 무기체계 시험평가 개선방안에 관한 연구를 위해 선행된 무기체계 시험평가 관련 연구와 더불어 미래전에 대비하기 위한 인공지능(AI)의 군사용 활용과 연관된 연구논문 등을 확인하였다. AI의 군사용 활용방안은 AI 기반 무기체계의 소요를 도출하는 데 있어 중요한 과정으로 소요군의 소요제기로부터 전력화되는 과정에서 반드시 실시해야 하는 시험평가를 위한 중요한 자료로 활용할 수 있다.

이종관·한창희(2019)는 “미래전과 국방 인공지능 체계”에서 군사 선진국에서 추진하고 있는 군사용 인공지능 체계의 개발 동향과 인공지능의 한계를 극복하여 미래 전장에서 인공지능이 활용될 수 있도록 체계적인 데이터 관리 및 활용과 군 주도의 인공지능을 연구하고 군 인공지능 플랫폼 구축 및 민관군 연구 협력 체계를 구축할 것을 제시하였다.

안진우·노상우·김태환·윤일웅(2020)은 “인공지능 분야 국방 미래기술에 관한 실증연구”에서 인공지능 분야 미래기술에 관한 특성 평가지표를 분석하고 국방 인공지능 분야 유망도를 특성 평가지표 간 상관관계와 다중회귀분석을 통

해 무기체계 적용성과 경제적 파급효과가 높을수록 유망도가 높아진다는 결과를 제시하였다.

윤정현(2021)은 “국방 분야의 인공지능 활용성 제고 방안과 시사점”에서 전문가 인식조사를 통해 본 국방 인공지능 활용 분야로 전투지원, 의무지원, 교육훈련, 국방정보시스템의 활용도가 높을 것으로 전망했다. 더불어 국방에 특화된 의사결정 지원을 위한 데이터 플랫폼 기술개발과 데이터 활용을 제약하는 보안 규제 개선, 인공지능 혁신을 장려하는 규제 샌드박스의 단계별 도입 등을 제안하였다.

장상국·최기일(2021)은 “미래 국방을 대비한 인공지능 기반의 방위산업 발전 방향 연구”에서 국방 분야에 있어 AI 기술이 적용된 대상과 도입 시기를 단기·중기·장기로 구분하고 대상사업별 요구성능을 제시하였다. 또한, 국방 분야 AI 발전방안으로 AI 전문인력의 육성과 AI 적용체계의 신뢰성과 활용에 있어 윤리적 문제가 발생하지 않도록 관련 정책 및 제도 보완을 제시하였다.

권세일·강지훈·서석호 등 5명(2022)은 “인공지능 기술 및 인증제도 분석을 통한 국방 인공지능 품질관리 방안 연구”에서 AI 적용 무기체계에 대한 품질 확보의 필요성과 AI 기술적용 시 야기될 수 있는 위험 요소 및 국내외 AI 품질인증제도를 분석하여 국방 인공지능 품질관리 방안을 제시하였다.

조현석·고봉준(2023)은 “인공지능의 군사적 활용과 국제 핵질서”에서 인공지능 기술이 활용되는 두 가지 범주(자율무기와 지휘통제)와 다섯 가지 기능영역으로 구분하여 ‘인공지능-핵 넥서스’ 개념의 틀을 바탕으로 오픈 소스로 공개되고 있는 인공지능의 기술에 대한 위협성에 대비한 핵 안전성과 인공지능의 군사적 활용방안을 제시하였다.

표 9. 인공지능(AI)의 군사용 활용 관련 연구 및 보고서

Table 9. Research and reports on military use of AI

구 분	연구 제목 및 연구 결과
이종관·한창희 (2019)	<ul style="list-style-type: none"> ◦ 미래전과 국방 인공지능 체계 <ul style="list-style-type: none"> - 군사 선진국에서 추진하고 있는 군사용 인공지능 체계의 개발 동향과 인공지능의 한계를 극복하여 미래 전장에서 인공지능이 활용될 수 있도록 체계적인 데이터 관리 및 활용과 군 주도의 인공지능을 연구하고 - 군 인공지능 플랫폼 구축 및 민관군 연구 협력 체계를 구축할 것을 제시
안진우·노상우· 김태환·윤일웅 (2020)	<ul style="list-style-type: none"> ◦ 인공지능 분야 국방 미래기술에 관한 실증연구 <ul style="list-style-type: none"> - AI 분야 미래기술에 관한 특성 평가지표 분석 - 국방 AI 분야 유망도를 특성 평가지표 간 상관관계와 다중회귀분석을 통해 무기체계 적용성과 경제적 파급효과가 높을수록 유망도가 높아진다는 결과 제시
윤정현 (2021)	<ul style="list-style-type: none"> ◦ 국방 분야의 인공지능 활용성 제고 방안과 시사점 <ul style="list-style-type: none"> - 전문가 인식조사를 통해 본 국방 인공지능 활용 분야로 전투지원, 의무지원, 교육훈련, 국방정보시스템의 활용도가 높을 것으로 전망 - 국방에 특화된 의사결정 지원을 위한 데이터 플랫폼 기술개발과 데이터 활용을 제약하는 보안 규제 개선, 인공지능 혁신을 장려하는 규제 샌드박스의 단계별 도입 등을 제안
장상국·최기일 (2021)	<ul style="list-style-type: none"> ◦ 미래 국방을 대비한 인공지능 기반의 방위산업 발전방향 연구 <ul style="list-style-type: none"> - 국방 분야에 있어 AI 기술이 적용된 대상과 도입 시기를 단기·중기·장기로 구분하고 대상사업별 요구성능 제시 - 국방분야 AI 발전방안으로 AI 전문인력의 육성과 AI 적용체계의 신뢰성과 활용에 있어 윤리적 문제가 발생하지 않도록 관련 정책 및 제도 보완 제시
권세일·강지훈· 서석호 등 5명 (2022)	<ul style="list-style-type: none"> ◦ 인공지능 기술 및 인증제도 분석을 통한 국방 인공지능 품질관리 방안 연구 <ul style="list-style-type: none"> - AI 적용 무기체계에 대한 품질확보의 필요성과 AI 기술적용 시 야기될 수 있는 위험 요소 및 국내외 AI 품질인증제도를 분석하여 국방 인공지능 품질관리 방안 제시
조현석·고봉준 (2023)	<ul style="list-style-type: none"> ◦ 인공지능의 군사적 활용과 국제 핵질서 <ul style="list-style-type: none"> - 인공지능 기술이 활용되는 두가지 범주(자율무기와 지휘통제)와 다섯가지 기능영역으로 구분하여 '인공지능-핵 넥서스' 개념의 틀을 바탕으로 오픈 소스로 공개되고 있는 인공지능의 기술에 대한 위협성에 대비한 핵 안전성과 인공지능의 군사적 활용방안 제시

3. 선행연구 분석 및 시사점

무기체계 시험평가와 인공지능(AI)의 군사용 활용과 관련한 선행연구 자료를 다양하게 확인한 결과 기존 무기체계 시험평가에 대한 신뢰성 향상과 인공지능(AI)을 국방에 도입하기 위한 정책적·제도적 방안 등을 제시하고 있었다. 반면, AI 기반 무기체계 시험평가를 위한 분야는 아직 활발한 연구가 진행되고 있지 않다는 것을 알 수 있었다.

본 연구와 관련하여 선행연구 분석은 무기체계 시험평가 분야와 인공지능(AI)의 군사용 활용 등에 관련된 분야로 구분하였다.

먼저, 무기체계 시험평가 선행연구 분야에서는 “무기체계 시험평가의 신뢰성 향상방안”에서 개발시험평가와 운용시험평가 전환을 위한 검증체계 개선 및 시험평가 업무의 표준화 방안과 “국방 무기체계 시험평가의 투명성에 관한 제언”에서 독립된 전문 시험기관을 설립하여 시험평가의 신뢰성을 높이는 방안을 제시하였다. 또한, 시험평가 제도 및 구조와 관련하여 “국방기술 환경변화에 부응한 무기체계 시험평가 신뢰도 제고에 관한 연구”와 “이해관계자 분석을 통한 무기체계 평가제도에 관한 연구”에서 시험평가 전문특기 부여 및 전문 시험평가관 양성을 위한 예산 반영 등을 제안하였다. 첨단 무기체계 시험평가와 관련해서 “국방 획득환경 변화에 적합한 무기체계 운용시험평가 개선 전략 도출”과 “인공지능 무기체계 시험평가방법 발전방안 연구”에서 국방개혁 추진과 미래 전장환경 변화에 대비하고, AI 무기체계의 시험평가시 AI 성능평가 척도 적용의 필요성을 제안하였다.

두 번째, 인공지능(AI)의 군사용 활용 관련 선행연구 분야에서는 “미래전과 국방 인공지능 체계”에서 미래 전장에서 AI가 활용될 수 있도록 체계적인 데이터 관리, AI 플랫폼 및 민·관·군 연구협력 체계 구축을 제기하였다. 또한, AI 도입을 제약하는 보안 및 규제개혁과 관련하여 “국방 분야의 인공지능 활

용성 제고 방안과 시사점”에서 데이터 활용을 제약하는 보안규제 개선과 규제 샌드박스의 도입을 제안하였다. AI 기반 방위산업 및 품질관리와 관련하여 “미래 국방을 대비한 인공지능 기반의 방위산업 발전방향 연구”와 “인공지능 기술 및 인증제도 분석을 통한 국방 인공지능 품질관리 방안 연구”에서 국방 분야 AI 기술 적용대상과 시기를 구분하여 요구성능을 제시하고, AI 기술적용 시 야기될 수 있는 위험 요소를 분석하고 대비하기 위한 전문인력 육성과 품질관리 방안을 제시하였다.

선행연구 결과를 종합해 보면, 무기체계 시험평가와 관련하여 시험평가 제도와 구조, 기술적 분야의 개선을 통해 시험평가의 신뢰성 향상과 전문성을 강화하고, 인공지능(AI)의 군사적 활용을 위한 실증연구를 통해 체계적인 데이터 관리와 AI 플랫폼 구축, 군의 AI 도입을 제약하는 보안 및 규제를 개선하고 AI 적용체계의 신뢰성을 보장할 수 있도록 정책 및 제도 보완 등을 제시하였다.

본 연구자는 이러한 선행연구 결과를 고려하여 국방 무기체계 개발 및 시험평가 환경변화와 무기체계 시험평가 사례를 분석하고, 다음 단계로 첨단 무기체계 시험평가 개선을 위하여 시험평가 제도·구조·기술 분야의 상대적 중요도와 세부 평가항목의 우선순위를 분석하고, AI 기반 무기체계의 시험평가 개선에 중점을 두고 구체적인 개선방안을 연구하여 제시하고자 한다.

제3장 연구의 방법론

제1절 연구수행 절차

본 연구는 무기체계 개발환경 변화에 따른 첨단 무기체계 시험평가 개선방안을 도출하는 데 있어 인공지능(AI) 기반 무기체계로 한정하였다. 우리 군이 추진하고 있는 AI 기반 무기체계의 전력화는 초기 단계로 AI 체계 시험평가 사례가 제한되는 점을 고려하여 육군(지상군)에서 수행하고 있는 AI 경계시스템(영상 감시체계)으로 국한하였다.

시험평가는 일반적으로 개발시험평가(DT&E)와 운용시험평가(OT&E)로 구분되어 수행하는데 무기체계 시험평가 개선방안은 운용시험평가를 수행하고 있는 소요군을 중심으로 그림 11과 같이 연구를 단계화하여 수행하였다.

먼저, 연구의 배경 및 필요성을 제기하기 위해 국방 과학기술의 발전과 미래 전장환경의 변화를 고찰하였다. 이를 위해 우리 군이 추진하고 있는 ‘국방 혁신 4.0’과 ‘육군비전 2030(육군 기본정책서)’, ‘Army TIGER 4.0’ 등 기획문서 및 정책서를 검토하고, 한반도의 미래전장 환경의 변화와 이를 대비하기 위한 무기체계 개발 소요와 육군에서 추진하고 있는 첨단 무기체계 소요를 살펴보았다. 이를 바탕으로 기존 연구자들의 무기체계 시험평가와 인공지능 관련 분야별 주요 논점과 문제 제기를 통해 연구 문제를 설정하였다.

두 번째로 이론적 배경은 국방획득체계의 변천 과정과 사업추진단계를 분석하고, 국방전력발전업무 훈령과 합참 무기체계 시험평가 업무 규정, 육군 시험평가 참고서 등 현 무기체계 시험평가 역할과 수행절차를 중심으로 살펴보았다. 더불어 민간과 군에 적용하고 있는 인공지능(AI) 개념과 기술 동향을 확인하고 무기체계 시험평가 개선 소요를 식별하였다.

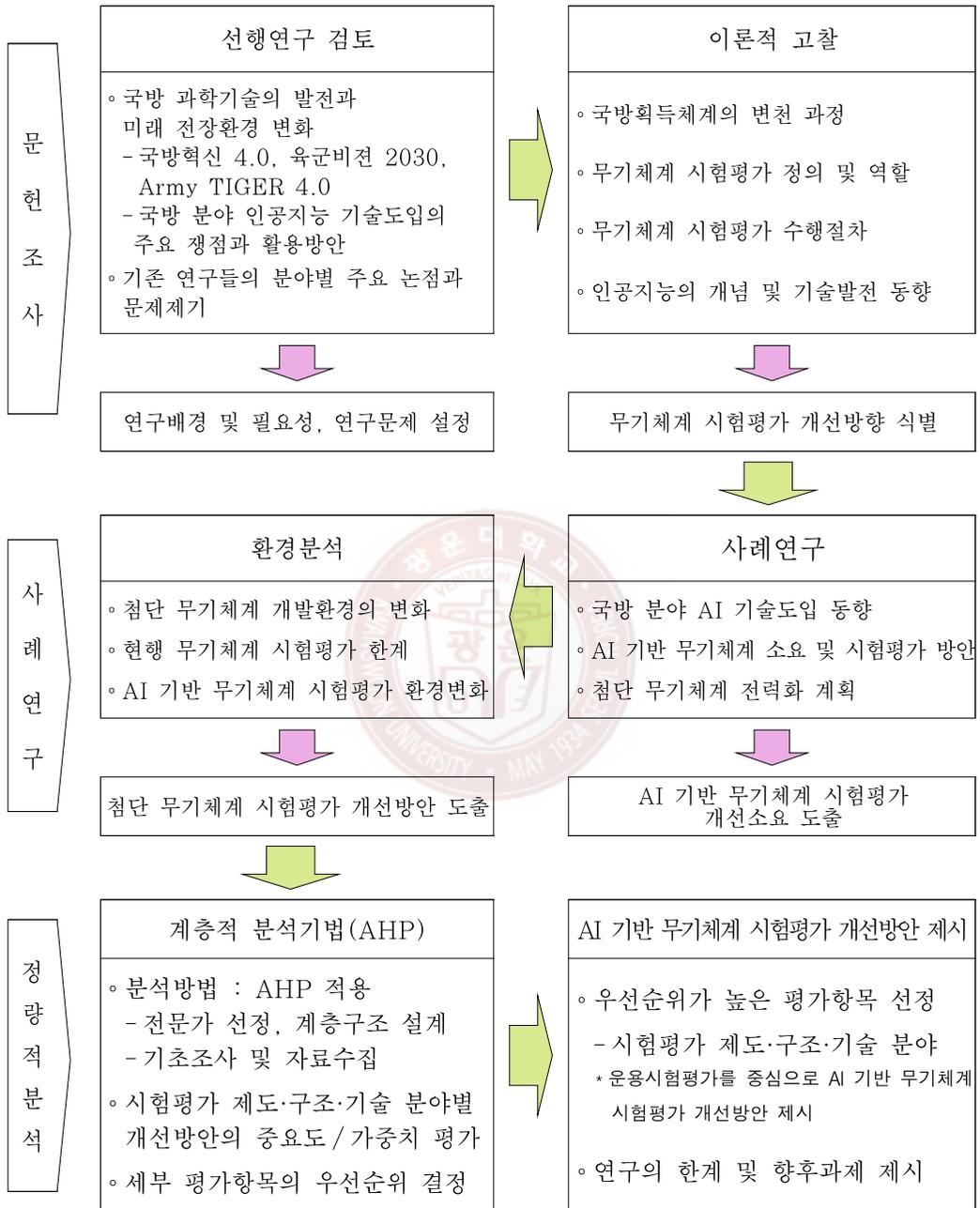


그림 11. 연구수행 절차 및 연구모형
Figure 11. Research process and model

세 번째는 국방 분야 AI 기술도입 동향과 AI 기반 무기체계 소요 및 시험 평가 방안, 첨단 무기체계 전력화 계획을 중점으로 살펴보고, 현행 무기체계 시험평가와 AI 경계시스템(영상 감시체계) 시험평가 사례를 비교 분석하여 AI 기반 무기체계 시험평가의 개선방안을 도출하였다.

네 번째로 정성적 분석의 한계를 보완하기 위해 정량적 분석으로 계층적 분석기법(AHP)을 적용하여 첨단 무기체계 시험평가 개선방안의 시험평가 제도·구조·기술 분야의 상대적 중요도와 세부 평가항목의 우선순위를 결정하였다.

마지막으로 AHP를 통해 분석된 세부 평가항목 중 우선순위가 높은 항목에 대하여 운용시험평가를 중심으로 인공지능(AI) 기반 무기체계 시험평가를 위한 개선방안을 제시하였다.



제2절 사례연구

사례연구(CSR, Case Study Research)는 미국 COMOS corporation의 회장인 Robert K. Yin이 제시한 방법으로 주로 사회과학 분야에서 특정 현상이나 집단을 깊이 있게 자료를 수집하고 분석하여 심층적으로 기술하고 해석하는 질적·정성적으로 연구하는 방법이다.⁵⁷⁾ 사례연구는 특정 사례에 대한 종합적인 이해를 시도하는 방법으로 실증적이면서 해석적인 연구설계가 가능하고 다양한 현상들에 대해 이론을 검증하고 구축하는 것이 가능하다.⁵⁸⁾

사례연구의 정의에 앞서 ‘사례’의 개념을 전통적인 사례연구에서 엔(Robert K, Yin, 2021)은 ‘사례’는 ‘개인’을 의미한다고 했으며, 일반적인 사례에 대해서는 학자마다 정의하는 것이 다양하고 개인 외에도 조직의 변화, 사건이나 개념, 프로그램, 의사결정 등 무엇이든 사례가 될 수 있다고 보고 있다. 게링(John Germing, 2006)에 의하면, 사례는 어느 특정 순간 또는 일정 기간에 걸쳐 관찰되는 공간적으로 제한된 현상이라고 하였고, 크레스웰(John W.Creswell, 2007)은 시간의 경과에 따라 하나의 경계를 가진 체계를 탐색하고 다양한 정보 원천들을 포함한 심층적인 자료를 수집하여 사례에 대해 기술을 하는 것으로 정의하였다.

사례연구의 목적은 연구자가 연구하고자 하는 주제, 사례에 대해 상세하고 정확하게 ‘기술’하고, 기술을 토대로 ‘설명’하고 현상에 대한 ‘평가’를 통해 개선점이나 대안을 제시하는 것이다. 기술이란 현상을 객관적으로 있는 그대로 서술하는 것이며, 설명은 연구자가 알게 된 것을 독자와 공유하는 것, 평가는 가치, 기준 등에 적합 여부를 판정하는 것이다.⁵⁹⁾

57) 신경식, 송민채, 신현섭 외 5인, 사례연구 방법(6판)(서울: 한경사, 2021), pp.25~56

58) 유기용, 정종원, 김영석, 김한별, 질적 연구방법의 이해(제2판)(서울: 박영사, 2018), pp.26~36

59) 최하영, 교육 과정적 통합을 위한 초등 특수교사와 통합학급 담임교사 간 협력에 관한 사례 연구, 한국교원대학교 대학원, 2006, pp.46~48

Yin은 완전하고 좋은 사례연구 조건으로 “대안적인 관점을 고려해야 한다는 점과 충분한 증거를 제시할 수 있어야 하며, 흥미로워야 한다는 점”을 강조하고 있다.⁶⁰⁾ 또한 사례연구의 필요성은 복잡한 사회현상을 이해하려는 동기에서 출발하고 그 과정에서 맥락을 통해 그 당시의 현상을 연구하는 것으로 설명했으며 연구목적에 따라 사례연구의 유형을 탐색적, 기술적, 설명적 사례연구로 분류하였다.⁶¹⁾ 탐색적 사례연구는 연구주제가 새로운 것이거나 특이한 것이어서 많이 알려지지 않은 상황일 경우에 이루어지며, 종속변수에 대한 명확한 구분이 없고 무엇이 중요한지 누구를 인터뷰하는지 연구자가 알기 힘들어 연구의 성격이 개방적인 경향이 있다.⁶²⁾ 기술적 사례연구는 현상을 기술하는 연구로써 탐색적 연구보다는 훨씬 더 구조화되어 있어 기본적으로 정의된 대상에 대해 상대적으로 적은 변수에 초점을 맞추고 체계적이고 정확하게 측정하는 인구조사나 여론 조사가 예가 될 수 있다.⁶³⁾ 설명적 사례연구는 현상에 대한 인과적 관계를 검증하는 데 목적이 있으며, 기술적 연구는 단순히 고립된 변수에 대한 정보를 찾는 반면, 설명적 연구는 변수들 사이에 존재하는 관계를 설명하는 것이다.⁶⁴⁾ 사례연구는 사례가 되는 연구대상의 독특한 성질을 구체적이고 상세하게 연구할 수 있고, 대상의 특성이나 문제의 원인을 포괄적으로 파악하여 인과 관계를 규명할 수 있다. 또한, 소수의 대상을 시간적인 변화에 따라 현상의 동태 파악 및 문제의 원인 등을 규명할 수 있다.⁶⁵⁾

Yin은 좋은 사례연구로 연구설계 검증을 위한 네 가지의 방법과 기술을 표 10과 같이 제시하고 있다.⁶⁶⁾

60) 신경식, 송민채, 신현섭 외 5인, 전게서, pp.370~379

61) 상계서, pp.29~33

62) 상계서, pp.29~31

63) 상계서, pp.31~33

64) 상계서, p.32

65) 이영철, 사회과학에서 사례연구의 이론적 지휘: 비판적 실재론을 바탕으로, 한국행정학보 40(1), 2006, pp.71~90

표 10. 연구설계 검증을 위한 네 가지 방법과 전술⁶⁷⁾

Table 10. Four methods and tactics for study design validation

검 증	사례연구 전술	전술적용 단계
구성 타당성 (Construct Validity)	<ul style="list-style-type: none"> ◦ 다양한 자료원을 사용하라 ◦ 사례연구 보고서 초안을 검토할 중요 정보 제공자를 확보하라 	자료수집단계 보고서 작성 단계
내적 타당성 (Internal Validity)	<ul style="list-style-type: none"> ◦ 패턴 매칭 기법을 사용하라 ◦ 인과 관계를 설명하라 ◦ 경쟁 가설을 설정하라 ◦ 논리모델을 사용하라 	자료수집단계 자료수집단계 자료수집단계 자료수집단계
외적 타당성 (External Validity)	<ul style="list-style-type: none"> ◦ 단일사례연구에서는 이론을 사용하라 ◦ 다중사례연구에서는 반복 연구 논리를 사용하라 	연구설계단계 연구설계단계
신뢰성 (Reliability)	<ul style="list-style-type: none"> ◦ 사례연구 프로토콜을 사용하라 ◦ 사례연구 데이터베이스를 개발하라 ◦ 증거사슬(Chanin of evidence)을 유지하라 	자료수집단계 자료수집단계 자료수집단계

네 가지 검증 방법중 먼저, 구성 타당성은 연구하고자 하는 개념들을 정확하게 측정하는지를 검증하는 것으로서 적절한 개념을 도출하고 도출한 개념을 측정하기 위한 도구는 적절한지 입증해야 확보할 수 있다. 두 번째, 내적 타당성은 설명적 사례연구에만 해당되는 개념으로 독립변수와 종속변수 사이에 타당한 인과 관계를 설정해야 확보할 수 있다. 세 번째, 외적 타당성은 사례연구의 발견이 어떻게 일반화될 수 있는지의 가능성 정도를 의미하는 것으로 연구 설계 단계에서 타당성의 기초를 만드는데 가장 적절한 시점이다. 네 번째, 신뢰성은 선행된 연구를 다른 연구자가 연구수행 절차를 똑같이 따라 했을 때 결과를 동일하게 얻을 수 있을 때 확보되는 것으로, 신뢰성을 검증하는 목적

66) 신경식, 송민채, 신현섭 외 5인, 전게서, pp.85~92

67) 상계서, p.86

은 동일하게 수행된 연구에 있어서 오류와 편견을 최소화하는 데 있다.

본 연구에서는 연구목적 달성을 위해 Yin의 완전하고 좋은 사례연구 조건에서 제시한 것과 같이 충분한 증거에 의한 대안적 관점을 제시하기 위해 국방 과학기술의 발전과 미래 전장환경, 첨단 무기체계 획득 및 인공지능(AI)의 군 활용성, 무기체계 시험평가 관련 참고문헌을 수집하고 분석자료로 활용하였다. 첨단 무기체계 시험평가 사례연구는 인공지능(AI) 기반 시험평가 사례를 주로 설명적 사례연구와 기술적 사례연구를 병행하였다.

사례연구를 수행하기 위한 참고문헌은 국방 과학기술의 발전과 미래 전장환경, 첨단 무기체계 획득 및 인공지능(AI)의 군 활용성, 무기체계 시험평가와 관련된 국내 박사학위 논문과 학술지 논문, 연구보고서, 육군의 기획문서 등을 수집하여 분석하였다. 자료수집은 광운대학교 전자도서관과 구글 및 네이버 학술검색 등 인터넷 사이트를 활용하여 검색하고 수집하였으며, 일부 민감한 군사 자료는 관련 부서를 방문하여 열람하였다.

제3절 계층적 분석기법

계층적 분석기법(AHP, Analytic Hierarchy Process)은 다수의 대안에 대하여 다면적인 평가 기준과 다수주체에 의한 의사결정을 위해 설계된 의사결정 방법의 하나로 미국 Pennsylvania 대학교의 Thomas L. Saaty⁶⁸⁾ 교수가 1970년대 초에 개발한 기법으로 계층 분석적 의사결정 기법이라고 불린다.⁶⁹⁾ AHP는 의사결정자의 직관적이고, 합리적인 또는 비합리적인 판단을 근거로 가중치, 일관성 지수와 같은 정량적인 요소와 쌍대비교라는 정성적인 요소를 동시에 고려하면서 의사결정을 내릴 수 있도록 한 방법이다. AHP 기법은 이론의 명확성과 단순성, 편리성이라는 특성으로 인해 넓게 응용되어 왔으며,⁷⁰⁾ 한국에서는 타당성을 분석하는 도구로 활용되고 있다.⁷¹⁾ AHP 기법은 주관적인 판단에 의존하는 상황에 적합하고, 정성적·정량적 요소를 동시에 반영하는 것이 가능하다. 의사결정 과정에서 수평적 요소 간 정량적 비교가 불가능한 사항에도 주관적으로 판단한 내용을 객관화하여 1:1 쌍대비교를 실시한다. 이를 통하여 가중치를 결정한 다음 하위계층의 평가 결과가 상위계층에 전달되는 효과를 추정할 수 있도록 하는 합리적인 시스템적 접근방법이다.⁷²⁾

AHP 기법은 인간의 사고체계와 비슷한 접근방법으로 문제를 분석 후 구조화할 수 있고, 상대적 중요도 혹은 선호도를 체계적으로 비율 척도화하여 정량적 형태의 결과를 얻을 수 있다는 점에서 유용하다.⁷³⁾ 또한 AHP의 기본 공

68) T. L. Saaty, Decision Making for Leaders, II, AHP series, RWS Publication, 1995.

69) 권오정, 다기준 의사결정 방법론 이론과 실제(북스힐, 2018), p.182

70) 조근태·조용곤·강현수, 계층분석적 의사결정(동현출판사, 2003), p.3

71) 임성훈·조기홍·박승, 무기체계의 효과분석과 의사결정을 위한 다기준 분석방법론 연구, 한국군사과학기술학회지, 제12권 제5호, 2009, p.560

72) Howard, Ronald A., Decision Analysis: Practice and Promise, Management Science, Vol.34, No.6, 1988, pp.681~685

73) 김홍빈, 무기체계 소요기획에 관한 영향요인의 우선순위 결정 방안 연구, 광운대학교 박사학위 논문, 2015, p.72

리는 그림 12와 같이 이론적 전제로 역수성(reciprocal), 동질성(homogeneity), 독립성(independency), 기대성(expectation) 4가지에 대한 배경을 지니고 있다.



그림 12. 계층적 분석기법(AHP) 기본 공리⁷⁴⁾

Figure 12. AHP fundamental axiom

역수성이란 두 요인에 대한 선호의 강도는 상호 역수 조건을 만족해야 한다는 것으로 의사결정자는 동일한 계층 내에 있는 2개의 요인을 짝지어 비교할 수 있어야만 하고, 그 선호의 강도를 표현할 수 있어야 한다.⁷⁵⁾ 동질성이란 특정 항목에 대한 선호도는 범위가 제한된 척도로 표현 가능하다는 것이다.⁷⁶⁾ AHP에서 주로 사용되는 평가척도는 9점 척도이다. 독립성은 한 계층의 특정한 요소는 관련된 상위계층의 요소에는 종속되고, 동일 계층의 다른 요소에 대해서는 독립적이어야 한다는 것이다. 의사결정 계층은 독립성의 공리에 따라 MECE (Mutually Exclusive, Collectively Exhaustive)가 되도록 설계해야 한다. 기대성은 의사결정의 목적을 계층이 완전하게 포함하고 있어야 한다는 것이다. 의사결정 계층은 기대성의 공리에 따라 Collectively Exhaustive 하게 설계되어야 한다.⁷⁷⁾

74) 권오정, 전계서, p.184

75) 상계서, p.184

76) 상계서, p.184

77) 육군교육사령부, 전계서, p.8장-34-3, 연구자 재정리

AHP의 요소에는 목적, 대안, 평가 기준 등 세 가지가 있다. 먼저 목적은 해결하고자 하는 의사결정이고, 대안은 의사결정을 위해 도출한 것이다. 마지막으로 평가 기준은 대안을 평가하기 위해 개발한 것을 말한다. 목적, 대안, 평가 기준의 요소를 계층의 형태로 만들고 나서, 그 계층을 구성하고 있는 요소 간에 상호비교를 통해 가중치를 구하고 이를 토대로 의사결정을 내린다.

AHP 기법의 기본적 형태는 상위단계로서 의사결정자의 최종목표가 있고 중간단계로서 하부단계의 비교 및 평가의 요소, 대안들로 구성되는 계층구조를 이루는 것이 일반적이다. 문제 해결을 위한 단계를 구성한 후에는 최종 목표를 위해 각 평가 기준의 가중치를 산출하고, 이때 평가 기준 가중치는 그대로 하위단계에 전달되고, 최종 목표에 최적 대안을 선택하게 된다.⁷⁸⁾

AHP는 다기준 의사결정 방법으로 상당히 유용한 방법이다. 어떠한 이론적 근거를 적용하기 쉽고, 의사결정이 요구되는 어떠한 분야의 문제도 적용하기가 수월한 범용적 모델이다. AHP는 정량적 효과, 계층적 구성, 측정, 일관성, 그룹의사결정 등 5가지의 장점을 갖고 있다. AHP 적용시 유의할 점은 첫째, 설문지의 결함으로 인해 부정확한 결과가 도출될 가능성이 내재되어 있다. 계층구성이 MECE(Mutually Exclusive·Collectively Exhaustive)를 충족하지 못하는 경우 왜곡이 발생할 수 있다. 둘째, AHP는 전문가를 대상으로 설문을 시행하는 데 있어 대상자 선정시 경험과 능력을 보유하고 해당 문제를 객관적으로 판단할 수 있도록 해야 한다. 셋째, 평가자의 비합리적 판단 가능성이 내재되어 있다. 설문조사를 할 때 대안별 상대적 평가점수를 부여하게 되는데 의사결정 요소와 대안이 많을 경우, 평가자의 비합리적 판단 가능성이 내재되어 있으며, 이러한 불합리성과 비일관성을 일관성 검사를 통해 검증해야 한다.⁷⁹⁾

78) Wind, Yoram and T. L. Saaty, Marketing Application of the AHP, Management Science, Vol.1, No.27, 1980, pp.645~668

일반적으로 AHP에서 대안을 평가하는 절차는 그림 13과 같이, 먼저 의사결정 요소 결정 단계로 문제 해결의 목표와 평가 기준, 대안을 설정한다. 2단계는 의사결정 문제의 계층적(Decision Hierarchy) 구조화로 각 계층간 항목별로 중복되지 않도록 MECE 원리가 그대로 적용되어야 한다. 3단계는 계층내 요소간에 쌍대비교 실시, 4단계 설문지 응답의 일관성($CR \leq 0.1$)을 측정하여 검증한다. 5단계는 의사결정 요소의 상대적 가중치(중요도) 산출하고, 마지막 단계에는 계층별 획득한 가중치(중요도)를 종합하여 최적의 대안(우선순위)을 선택하는 절차로 진행한다.

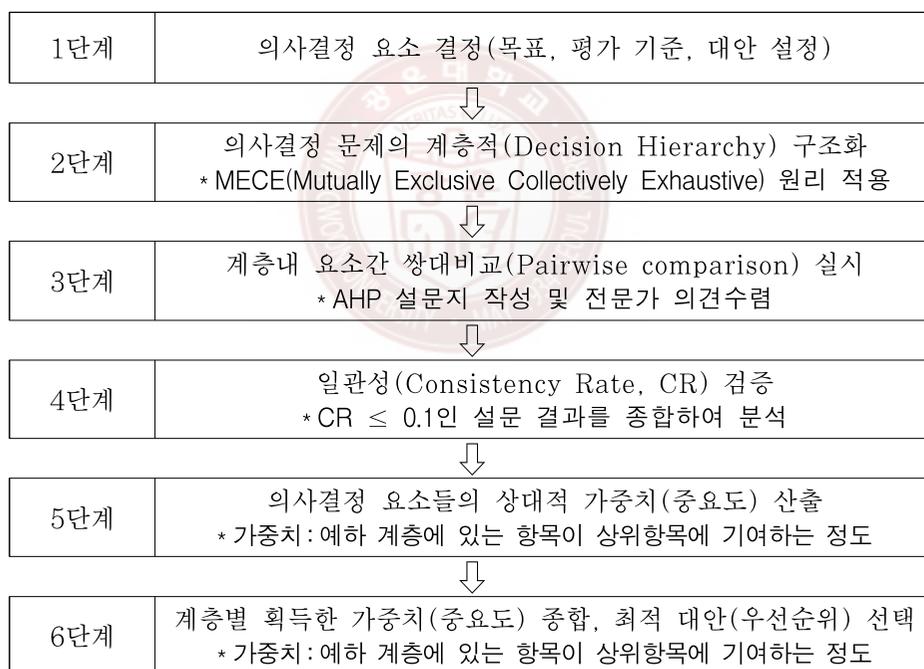


그림 13. AHP 기법 적용 절차⁸⁰⁾

Figure 13. AHP technique application procedure

79) 육군교육사령부, 국방운영분석 이론과 실제(제8장 계층적 분석기법), 2021, pp.8장-34-1~2, 연구자 재정리

80) 권오정, 전계서, pp.182~253, 연구자 재정리

제4장 첨단 무기체계 시험평가 개선 필요성

제1절 무기체계 개발환경의 변화

1. 국방 과학기술의 발전과 미래전 양상

현대전을 흔히 과학기술의 전쟁이라고 부른다. 미래전의 양상을 진단하고 예측하는 것은 여러 가지가 있지만, 미래전은 신개념의 무기체계와 국방 과학기술이 주도할 것으로 예측하고 있다. 민간의 상용과 민군 과학기술의 수준이 급속도로 발전하고 있으며, 미래전에 대비하기 위해 우리 군은 군사전략과 작전술을 구현하기 위한 무기체계 전력화 소요를 발굴하고 있다. 하나의 예로서 미래전은 인공지능(AI)을 기반으로 로봇(Robot)이 전장을 주도해 갈 것이다. 이러한 로봇은 전장에서 인간과 함께 유·무인 복합체계로 전투를 수행하고 전쟁에 대한 정책적 결정권자나 전투현장 지휘관들의 개입할 수 있는 여지가 없어질 수도 있다는 우려가 제기되고 있다.⁸¹⁾

미래전의 양상은 4차 산업혁명과 함께 우리가 생각하고 있는 것보다 더 획기적인 변화를 가져올 것이다. 4차 산업혁명 도래에 따라 다양한 정보통신기술(ICT)들이 상호 융합·촉진·진화하면서 과학기술은 비약적으로 발전하게 될 것이다. 인공지능(AI), 사물인터넷(IoT), 클라우드(Cloud), 빅데이터(Big Data), 로봇(Robot) 등의 과학기술은 미래 국가경쟁력을 이루는 핵심기술로써, 경제와 산업은 물론 군사 분야에도 커다란 변화를 불러일으킬 것이다. 또한, 미국의 경제·산업·기술 전반에 걸친 중국 견제와 중국의 ‘중국식 현대화’ 전략으로 미·중 간 기술 패권 경쟁은 더욱 심화될 것이다. 이러한 미·중 간의 기술 패권 경쟁 배경 중 하나는 4차 산업혁명 과학기술이 군사 분야에 접목될 경우 기존

81) The Economist, special report: The future of War, Jan 2018, pp.3~16
<https://www.economist.com/special-report/2018-01-27>, (검색일: 2022. 4. 12.)

의 전쟁수행 개념과 체계를 획기적으로 변화시킬 수 있는 ‘와해적 기술 (disruptive technology)’⁸²⁾ 이라는 사실에 있다.⁸³⁾



그림 14. 미래 전장 환경과 4차 산업혁명 관점⁸⁴⁾

Figure 14. Future battlefield environment and the perspective of the 4th industrial revolution

국방 과학기술의 발전과 관련하여 일찍이 미국 국방장관을 역임한 페리 박사는 국방력을 높이는 핵심은 ‘이용 가능한 최신 과학기술을 최소한의 비용으로 최단 시간에 전력화하는 것’이라고 강조했다. 국방력의 핵심이 과학기술이라는 말이다. 군사 대국인 미국을 제외하고라도 과학기술을 군에 적절하게 활용해 자주국방 능력을 갖춘 나라들은 많이 있다. 이스라엘이 대표적인 국가다. 작은 나라임에도 불구하고 아랍과의 전쟁에서 연전연승하고 있는 이스라엘은 첨단 무기체계에 필요한 기술을 외국에 의존하지 않고 스스로 개발한다는 국방 과학기술정책을 유지해 왔다. 그 결과 무인항공기 시스템과 미사일 등 첨단 무기를 비롯해 각종 재래식 무기의 성능개량 기술은 세계적으로 가장 앞선 나라가 됐다.⁸⁵⁾

82) AI, IoT, Big Data, Robot 등 완전히 새로운 기능이나 속성으로 기존 기술 및 시장 진입장벽을 무력화시키는 급진적 혁신의 기술체계를 말한다.

83) 류태규, 4차 산업혁명과 국방분야 과학기술 적용, 한국국가전략 제5권 3호, 2020, pp.228~230

84) 상계서, p.230

85) <https://www.joongang.co.kr/article/4370801#home> [과학기술과 국방] 첨단기술도 중요한 전쟁 억제수단, (검색일: 2022. 4. 18.)

국가별 첨단기술 수준은 국방비 증가와 세계 방산시장 점유율 간의 상관관계를 통해서도 확인할 수 있다. 전 세계 국방비 비율과 주요 무기 수출국으로 세계 방산시장 점유율에서 그림 15와 같이 부동의 1위를 고수하고 있는 미국이 항공·우주, 융합기술 분야 등 대부분의 첨단 국방 과학기술 분야에서 세계 최고 수준의 위치를 점하고 있다.



그림 15. 전 세계 상위 15개국 국방비 비율(2020년) 및 주요 무기 수출입국(2016~2020)⁸⁶⁾
 Figure 15. Global defense spending percentage of top 15 Countries (2020) and major arms import/export countries (2016~2020)

우리나라는 인구감소와 생명 중시 사상의 확산 등 사회변동 요인과 4차 산업혁명의 시대에 국방 과학기술의 발전으로 격동의 시대적 변화상황이 펼쳐지고 있다. 한반도를 둘러싼 미·중 갈등과 경쟁 속에서 과학기술의 비약적 발전과 전쟁 양상의 변화로 우리에게 혁신적이고 도약적인 변화를 요구하고 있다. 4차 산업혁명의 과학기술을 군에 적시적으로 도입하고 무기체계에 적용하여 상대적 우위를 달성할 수 있어야 한다. 러시아·우크라이나 전쟁을 통해서 4차 산업혁명의 과학기술이 전쟁 양상에 미치는 파급력이 얼마나 큰지 보여주고 있다고 할 수 있다. 따라서 우리 군은 국방 과학기술에 기반한 혁신적이고 도약적인 변화를 추구해야 할 중요한 변곡점에 와 있다고 할 수 있다.

86) 스톡홀름 국제평화 연구소(SIPRI), SIPRI 연감 2021(요약), p.15

2. 첨단 무기체계 발전방향

최근 군사 분야에서 ‘4차 산업혁명’만큼이나 자주 등장하는 단어가 ‘게임 체인저(Game Changer)’다. 게임 체인저는 특정 무기체계가 전쟁이나 분쟁의 판도를 바꿀 수 있다며 사용하는 표현이다. 사전적인 의미로는 “새롭게 소개된 요소나 요인으로 현재의 상황이나 활동을 상당히 뚜렷하게 변화시키는 것”, 또는 “어떤 일에서 결과나 흐름의 판도를 뒤바꿔 놓을 만한 중요한 역할을 한 인물이나 사건, 제품” 등을 일컫는 말이다.⁸⁷⁾

미래 무기체계에서 게임 체인저의 정의는 군사적인 작전 상황과 전쟁 수행에 있어서 지대한 영향을 미쳐서, 전쟁의 승패를 결정할 정도로 새롭고 혁신적인 군사 과학기술이 적용된 무기체계라고 할 수 있다.⁸⁸⁾ 미래 전장의 판도를 뒤바꿀 게임 체인저는 기존의 일반적인 무기체계가 구비하고 있는 기동성, 화력, 생존성 등 특성 이외 4차 산업혁명 시대의 핵심기술인 인공지능(AI)을 구비한 무인체계로 발전될 것이다. 일반적으로 4차 산업혁명의 핵심 요소로써 인공지능(AI), 사물인터넷(IoT), 빅데이터, 로봇, 3D 프린팅 등 5가지의 ICT 기반 기술과 무기체계가 기능을 발휘하기 위해 필수적으로 요구되는 센서, 에너지, 새로운 소재 등이 적용된 체계로 발전될 것으로 전망하고 있다. 이러한 첨단 무기체계는 AI, 빅데이터 및 VR/AR/MR⁸⁹⁾ 기술의 적용이 확대되고 언

87) 네이버 지식백과, <https://terms.naver.com/entry.naver?docId=1398167&cid=43667&categoryId=43667>, (검색일: 2022. 4. 18.)

88) 김종열, 미래 무기체계와 군사 과학기술 발전추세 분석, 전략연구 통권 제76호, 2018, p.86

89) VR(Virtual Reality, 가상현실): 우리가 살고 있는 물리적인 공간이 아닌 컴퓨터로 구현한 가상 환경 또는 그 기술 자체.

AR(Augmented Reality, 증강현실): VR과 달리 위치, 지리정보를 송수신하는 GPS 장치 및 중력 그리고 자이로스코프에 따른 위치정보 시스템을 기반으로 우리가 경험하는 현실 세계에 가상의 물체나 정보가 합성되어 실제 현실과 가상현실의 상호작용하는 공간으로 만들어 주는 기술.

MR(Mixed Reality, 혼합현실): VR이나 AR보다 한 단계 더 나아간 기술로 VR과 AR 두 기술의 장점만을 합친 기술, MR은 현실과 가상의 정보를 융합해 조금 더 진화된 가상세계를 구현하고 냄새 정보와 소리 정보를 융합해 사용자가 상호 작용할 수 있는 기술. VR과 AR은 시각에 전적으로 의존하지만 MR은 시각 외에 청각 촉각 등 인간의 오감을 접목시킬 수 있

어·음성·시각·감성의 인지능력과 학습, 추론 등 지능을 구현하여 수집한 방대한 자료를 실시간 분석하여 처리하고 종합된 정보를 자율적인 의사결정 능력이 확충됨으로써 자율성이 크게 향상될 것이다. 이는 다수의 무인체계 또는 유인체계 간 실시간 소통이 보편화되고 인간의 간섭이 최소화될 것이다. 또한, 드론과 로봇 등 무인 전투체계가 미래 전장에서 활용이 급속도로 확대됨으로써 유·무인복합 전투수행체계로 발전되고, 전장에서 인간과 로봇이 모듈화된 부대구조로 다영역 동시통합작전을 수행하는 초연결(hyper connectivity) 시대로 첨단 무기체계는 발전할 것이다.⁹⁰⁾

미국을 비롯한 주요 선진국들은 4차 산업혁명의 근간인 AI를 기반으로 한 군사력을 강화하기 위해 다음과 같이 추진하고 있다. 미국은 미래전에 대비한 첨단 과학기술 확보를 위해 국방 인공지능 기술개발에 투자를 확대하고 있다. 미 국방성은 2017년 4월에 Algorithmic Warfare Cross-Function Team으로도 알려진 ‘Project Maven’을 출범하면서 인공지능(AI)을 이용한 무인기 기반의 영상 정보를 수집·분석하는 프로젝트로 당시 Google 직원들의 많은 항의와 이직을 불러일으켰던 프로젝트다. 당시 미 국방성은 ‘대반군과 대테러 작전을 지원하기 위해 매일 수집하는 엄청난 양의 풀 모션 비디오 데이터로 인해 어려움을 겪고 있는 군사 및 민간 분석가를 돕는데 필요한 컴퓨터 버전 알고리즘을 개발하고 통합하는 것을 목표로 한다.’라고 했다. 현재는 고등방위연구계획국(DARPA)⁹¹⁾과 국방혁신단(DIU)⁹²⁾, 합동 인공지능센터(JAIC)⁹³⁾가 협업하여

다는 점이 다름. <https://www.ahnlab.com/kr/site/securityinfo/secunews/secuNewsView.do?seq=29885>, 상상이 현실로, VR, AR, MR, XR, SR 뭐가 다를까, (검색일: 2023. 4. 9.)

90) 함참, 4차 산업혁명 기반 첨단 무기체계 시험평가 발전방안 연구, SMI, 2018, pp.73~75

91) 고등방위연구계획국(DARPA)는 미 국방성에 소속된 미군 관련 기술 연구개발기관으로, 1957년 소련이 쏘아 올린 스푸트니크 1호에 위기의식을 체감한 드와이트 D. 아이젠하워 대통령의 주도로 1958년에 창설된 ARPA에 뿌리를 두고 있다. 우리나라의 국방과학연구소와 비슷해 보이지만, 실제로는 美 국방성 산하 R&D 기획·평가·관리 전담 기관으로 즉, 무언가를 직접 개발하기보다는 어떠한 사업에 여러 방산업체나 연구기관들을 경합시켜 그중에서 가장 적합한 결과물을 주로 선정하는 기구다. <https://www.darpa.mil/>, (검색일: 2022. 4. 12.)

표 11과 같이 인공지능 프로젝트를 수행하고, 민간의 기술적 성과를 국방에 적극적으로 도입하고 있다. JAIC은 美 국방성 내 인공지능 프로젝트의 실행을 전담하고, 수행 중인 주요 국방 인공지능 프로젝트는 전투수행 뿐만 아니라 전력자원의 상태 점검, 정보 능력의 확대에 이르기까지 국방체계 전반을 망라하고 있다.⁹⁴⁾

표 11. JAIC가 수행하는 주요 인공지능 프로젝트⁹⁵⁾

Table 11. Major AI projects carried out by JAIC

1. 합동전투수행 분야	2. 전투원 건강 상태	3. 업무체계 혁신
전쟁양상 변화를 위해 시스템, 센서, 표적화 등에 AI 활용 *밀폐공간에 초소형무인기 20개 군집비행	질병치료, 작전태세 향상, 비용 절감, 돌파구 마련을 위해 군 데이터 활용 *전투원 데이터에 기초, 전국적 데이터 수집 전자인프라 개발	AI기술 활용, 업무절차 개선을 통해 국방정보능력 향상 비용절감, 효율성 증대 *비표준 양식 식별, 4만개 양식 표준화 및 1,300인시 절감
4. 인명보호 및 위험감소	5. 합동군수	6. 합동정보전
재난현장 긴급대응 조직에 인명구조 및 시설보호에 적절한 정보제공에 AI 활용 *캘리포니아산불 예측 모델 개발 및 고도화 추진	상태진단, 훈련, 절차 개선, 수요예측, 공급망 최적화를 통해 전투태세 향상 *MH-60M GE YT706 엔진시동 실패 예측 모델개발 적용	국방부와 합동군 정보 우수 달성을 위해 AI 기반의 정보능력 제공 *MADAT 개발모금

92) DIU(Defence Innovation Unit)은 서비스 및 컴포넌트에서 전투 명령 및 국방 기관에 이르기까지 국방성의 모든 조직과 협력하여 국가 보안 과제를 해결하는 첨단 상용 솔루션을 신속하게 프로토타이핑하고 현장에 배치하는 역할을 담당한다. 실리콘 벨리, 보스턴, 오스틴, 펜타곤에 지사를 둔 DIU는 국방성 파트너를 미국 전역의 선도적인 기술 기업과 연결해 주는 역할을 한다. DIU는 미군 전역에서 중대한 문제를 해결하는 데 도움이 되는 상업 기술을 집중적으로 다루며, 특히 민간 분야에서 주도적으로 개발되고 있는 5가지 최첨단 기술 분야(AI, 자율화, 사이버, 휴먼 시스템, 우주) 등에 중점을 두고 있다. 정보통신기획평가원, (GT주간브리프-실리콘벨리)미국, AIU(국방혁신단) 주요 현황 및 추진동향, 2020, pp.2~11

93) JAIC(Joint AI Cente): 미 국방성은 2018년 7월 27일 JAIC을 창설하여 AI를 국방성 업무에 적용하고 있다. 이는 4차산업 혁신에 대한 포괄적 기준과 원칙을 갖고 합참, 각 군, 통합군사령부 그리고 동맹국까지 적용할 예정이다. <https://www.ai.mil/> (검색일: 2022. 4. 12.)

94) 윤정현, 국방 분야 인공지능 기술 도입의 주요 쟁점과 활용 제고 방안, 과학기술정책연구원(STEPI), 2021, pp.10~11

95) 상계서, p.11

중국 국무원은 2017년 7월 ‘차세대 인공지능 발전계획’을 발표하면서 ‘인공지능의 선두주자, 글로벌 혁신 중심지’로 만들겠다는 목표와 구체적인 전략을 표12와 같이 수립하였다.⁹⁶⁾ 2018년 3월에는 ‘군민융합전략요강’을 통해 광범위한 빅데이터 분석과 학습 능력을 갖춘 첨단군의 육성계획을 발표하고, 2027년까지 ‘AI 기반 인민해방군 현대화’를 공식화하고 민간업체들과 무인 드론·지능형 전투체계를 구축하기 위해 노력 중이다.⁹⁷⁾

표 12. 중국의 차세대 인공지능 발전계획⁹⁸⁾
Table 12. China's next-generation AI development plan

구 분	추진 내용
1단계 (2017 ~ 2020)	<ul style="list-style-type: none"> 중국의 AI 산업을 글로벌 경쟁업체와 비슷한 수준으로 발전 이를 위해 인공지능 산업을 새로운 경제 성장축으로 육성
2단계 (2021 ~ 2025)	<ul style="list-style-type: none"> 일부 AI 분야에서 세계 최고에 도달하는 것을 목표로 함 이를 위해 중국의 산업고도화와 경제발전을 선도하는 주요 원동력으로 지능사회 건설을 추진
3단계 (2026 ~ 2030)	<ul style="list-style-type: none"> 글로벌 AI 혁신의 중심센터로 도약 AI의 글로벌 거버넌스에 적극적으로 참여·추도하려는 의지를 포함

유럽연합은 제조 분야의 강점과 우수한 연구 및 협업 생태계를 기초로 국방을 포함한 공공부문의 인공지능 도입을 촉진하고 있다. 2018년 4월 ‘AI for Europe’를 채택하고 2020년 2월에 EU 인공지능 백서를 발간하여 EU가 가진 기술 및 산업 분야 강점을 바탕으로 인공지능 정책 및 규제 프레임워크를 제시하였다. 최근에는 국방 분야에서 EU 차원의 공동안보방위정책(CSDP,

96) 상계서, pp.12~13

97) 이창영, 중국의 민군융합을 통한 지능화군 건설 전략, KIMS Periscope(한국해양전략연구소) 제166호, 2019, pp.1~3

98) 윤정현, 전계서, p.13

Common Security and Defense Policy)을 발표하고 광범위한 CSDP 임무를 수행하기 위한 AI 기술을 데이터 관리, 상황인식, 훈련 및 시뮬레이션, 공급 및 유지보수, 추적 감시 등에 적극적으로 도입하는 중이다.⁹⁹⁾

우리나라는 2018년 국방부와 과기정통부 간 ‘과학기술 기반 미래 국방 발전 전략’을 공동 수립하고 미래 국방력 확보 및 미래전장 변화를 선도하는 가교 연구와 혁신적인 기초·원천기술 개발을 위한 R&D 로드맵을 제시하였다. 2019년에는 ‘국방 인공지능 발전계획’을 수립하고 국방운영, 기술 기반, 전력체계의 3개 분야 혁신에 초점을 두고 표 13과 같이 미래 8대 국방 핵심기술, 10대 군사능력, 30개 핵심 전력을 구분하여 국방혁신 과제를 추진하고 있다.¹⁰⁰⁾

표 13. 국방 전력체계 혁신 분야 핵심기술, 군사능력, 핵심전력¹⁰¹⁾

Table 13. Defense power system innovation field Core technology, Military capability, Core force

구 분	세 부 내 용
미래 8대 국방 핵심전략	① 첨단센터 ② 신소재 ③ 신추진 ④ 고출력 / 신재생에너지 ⑤ AI ⑥ 가상현실 ⑦ 무인로봇 ⑧ 사이버
10대 군사능력	① 고위력 ② (극)초음속 ③ 초정밀 ④ 소형·경량화 ⑤ 무인/유·무인 복합 ⑥ 스텔스 ⑦ 장사정·신추진 ⑧ M&S사이버 ⑨ 초연결·네트워크 ⑩ 비살상·전자전체계
30개 핵심전력	AI 기반 초연결 전투체계, 휴머노이드 로봇, 초소형 곤충형 정찰로봇, AI 지휘통제체계, 초고속, 지능형 어뢰체계 등 30개

현재 우리 군은 AI 과학기술 강군 육성을 목표로 ‘국가 인공지능 전략’과 연계해 AI 기반 핵심 첨단전력을 확보하기 위해 유·무인 복합전투체계 구축과 우주, 사이버, 전자기스펙트럼 영역 작전수행능력을 강화하고, 합동 전(全) 영

99) 윤정현, 국방 분야 인공지능 활용성 제고 방안과 시사점, 과학기술정책 포커스, 2021, p.45

100) 윤정현, 전계서, pp.18~20

101) 국방부, 2020년 국방부 업무계획, 2022

역 지휘통제(JADC2) 체계를 구축하기 위해 ‘국방혁신 4.0’을 추진하고 있다.

국방부는 2023년 연두 업무보고를 통해 전쟁양상 변화에 따라 전투 효율성을 제고하고 인명 손실을 최소화하기 위해 AI 등 첨단 과학기술 기반의 유·무인 복합체계 구축을 가속화 하겠다고 했다.¹⁰²⁾ AI 기반 유·무인 복합전투체계는 AI를 기반으로 인간 전투원과 무인전투체계를 통합하여 전투임무를 수행하는 체계를 말한다. 국방부는 AI 기반 유·무인 복합전투체계로 단계적으로 전환¹⁰³⁾하고 운용성을 보장할 수 있도록 인프라를 구축중에 있다.

군은 2024년 ‘국방 AI센터’를 창설하고, 그림 16과 같이 ‘반자율형 체계’로의 효율적인 전환을 위해 각 군별로 운용 중인 시범부대 전력을 조기에 확보하고, 운용성과 검증은 통해 각 군별 특성에 부합하는 유·무인 복합체계를 확대해 나가겠다고 하였다.¹⁰⁴⁾ 반자율형 유·무인 복합체계는 인간과 로봇이 함께 작전을 수행하는 미래형 군사 시스템으로 무인항공기체계(UAVs)¹⁰⁵⁾, 무인지상차량체계(UGVs)¹⁰⁶⁾, 무인수중·수상기체계(UUVs)¹⁰⁷⁾, 인공지능(AI)¹⁰⁸⁾, 네트워크 체계¹⁰⁹⁾ 등으로 구성된다. 이와 같은 주요 체계들이 융합되어 반자율형

102) 국방부, 2023년 국방부 연두보고, 2023 1월, <https://www.fnnews.com/news/202207221828281450>, 파이낸셜 뉴스, (검색일: 2023. 4. 11.)

103) 1단계 (원격통제형 체계): 사람이 무인체계를 조종하는 방식으로, 작전 효율성 향상을 위해 무인체계를 보조적인 수단으로 운용

2단계 (반자율형 유·무인 복합전투체계): AI 기술이 적용되어 무인체계가 사람의 판단 기능을 제한적으로 대체(실시간 대응, 실시간 다중 기체 협력 등)

3단계 (자율형 유·무인 복합전투체계): AI 기술이 적용되어 사람 개입이 최소화된 무인체계가 작전임무를 수행(전장의 모든 것을 인식하고 분석하며 임무수행을 위한 최소한의 지침만 필요)

104) 국방부, 전계서, <https://www.fnnews.com/news/202207221828281450>, (검색일: 2023. 4. 11.)

105) UAVs는 고고도 탐사, 정찰, 지능 수집 등 다양한 임무에 사용되며, 대부분 지상에서 운영되고 항공기에서 무선으로 제어된다.

106) UGVs는 로봇 기술을 활용하여 만들어진 차량으로, 인간의 개입 없이 임무를 수행할 수 있다. 예를 들어, 탐색, 폭발물 제거 등이 있다.

107) UUVs는 해양 위와 아래에서 수행되는 임무에 사용되며, 지능 수집, 해양 탐사, 수중 폭발물 제거 등이 대표적이다.

108) AI는 인간의 뇌를 모델링하여 구현된 인공적인 지능으로, 병력 감축 및 작전성과 향상 등 많은 기여를 할 수 있을 것이다.

109) 미래형 군사작전에서 수많은 체계가 함께 동작하여 하나의 큰 시스템으로 작동한다. 이를

유·무인복합체계가 구현된다면 미래의 전투는 더욱 효율적이면서 안전하고 정확하게 작전을 수행할 수 있는 단계로 발전할 것이다.



그림 16. 반자율형 유·무인 복합체계 시범부대 및 주요체계¹¹⁰⁾

Figure 16. Semi-autonomous manned and unmanned complex system demonstration unit and main system

위해 초지능·초연결 네트워크 체계가 필요하다.
 110) 국방부, 전게서, 2023년 국방부 연두보고

제2절 첨단 무기체계 시험평가 개선 소요

1. 일반 무기체계와 AI 기반 무기체계의 시험평가 비교

한반도의 미래 전장은 지상·해상·공중을 넘어 우주와 사이버 영역 등 5차원의 전장 공간의 확장과 융합기술이 접목된 첨단 무기체계의 운용으로 무기체계간 상호연동성과 복잡성이 증가할 것이다. AI 기반 드론과 로봇, 무인 차량 등 유·무인 복합전투체계로 초지능·초연결 네트워크로 연결된 다영역 동시통합작전을 수행하는 작전환경에서 AI를 기반으로 하는 무기체계에 대한 시험평가를 수행해야 하는 시험환경으로 변화되고 있다. 하지만 지금까지 AI를 기반으로 하는 무기체계 전력화 소요는 다양하고 많으나, 현재 시험평가를 진행한 사례는 제한되어 하나의 특정 사례를 기준으로 일반적인 무기체계와 AI 기반 무기체계 시험평가를 비교하고자 한다.

중요시설 경계시스템은 구매사업으로 복수업체를 대상으로 약 1년의 사업기간 중 시험평가는 통상 4~6개월을 수행했다. 이 사업은 병력 위주의 경계시스템을 병력과 과학화 경계시스템이 통합된 경계 체계로 전환하기 위한 중요시설 경계시스템으로 시작하여 지금까지 2차 사업을 종료하고 3차 사업이 진행 중이다. 이와 병행하여 신속시범획득사업¹¹¹⁾으로 추진하고 있는 AI 영상감시체계가 있다. 운영개념은 병력과 과학화 경계시스템을 통합 운영하여 침투하는 적을 최대한 조기에 발견하여 경고하기 위해 감시, 감지, 통제시스템을 구축하여 운영한다. 감시시스템은 적 예상 침투로 및 취약지역에 중점적으로 운용하되, 사각 지역이 최대한 발생하지 않도록 운용한다. 감지 시스템은 울타리에 설치하여 통과 및 월책시 감지하여 책임지역 상황실로 경보가 전파되도

111) 신속시범획득사업: 신기술이 적용된 민간제품 또는 국방기술개발사업 성과물을 군이 시범운용하여 군사적 활용성을 확인하고 군이 필요로 하는 무기체계를 전력화하는 사업이다. [신속시범획득사업 업무관리 지침(2021)]

록 운용하는 것이다. 통제시스템은 주둔지 및 시설 단위로 상황실을 구축하여 감시 및 감지 시스템을 중앙통제하고 필요한 정보를 관련 부대로 전송할 수 있는 체계로 그림 17과 같이 구성되어 있다.

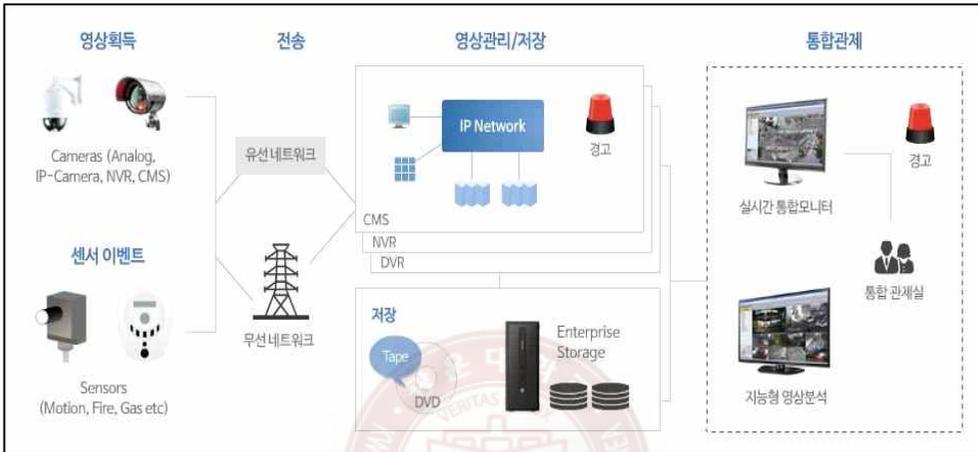


그림 17. AI 영상감시 체계¹¹²⁾
 Figure 17. AI video surveillance system

일반적인 무기체계 시험평가는 개발 장비의 시제품에 대하여 표 14와 같이 작전운용성능의 충족성, 합동성·상호운용성, 군 운용의 적합성, 전력화지원요소의 실용성 확인을 통해 ‘전투용 적합’ 또는 ‘전투용 부적합’ 판정 자료를 제공하기 위해 시험평가를 소요군(시험단)이 주관하여 수행한다.

일반 무기체계의 시험평가 수행은 무기체계를 운용하고 통제하는 주체가 사람(장병)이기 때문에 실제 운용 환경하에서 대상 무기체계를 야전에서 직접 탑승해서 조작과 기동을 하면서 작전지역의 극한의 조건에서 무기체계의 운용능력을 비롯하여 각종 센서의 능력, 최대속도, 전투중량, 기동능력 등을 직접 확인하고 있다.

112) http://www.koits.net/HOME/kits/sub.htm?nav_code=kit1529408669, (검색일: 222. 4. 13.)

표 14. 운용시험평가 항목¹¹³⁾
 Table 14. Operational T&E items

구 분	시험평가 항목
작전운용성능의 충족성	◦ 주요 작전운용성능 ◦ 기술적/부수적 성능
합동성 및 상호운용성	◦ 운용개념 및 체계특성 ◦ 연동성 및 정보교환 ◦ 표준 및 아키텍처 ◦ 정보보호 ◦ 주파수
군 운용 적합성	◦ 운용 및 조작 적합성, 안정성 ◦ 전술적 운용의 적합성 ◦ 기존 무기체계의와의 상호운용 적합성 ◦ 환경 적응성
전력화지원요소의 실용성	◦ 전투발전지원요소 ◦ 통합체계지원(IPS) 요소

화력 무기체계의 경우에는 해당 무기체계의 특성을 고려하여 최대·최소사거리, 최대발사속도, 정확도(CEP, Circular Error Probability) 등 무기체계의 HW와 SW의 성능을 확인하고 검증하는 시험을 진행하고 있다. 해당 무기체계를 운용하는 사람(장병)들에 의해 발생할 수 있는 오차를 줄여 신뢰성을 높이기 위해 제한된 과학적인 시험평가 방법을 적용하고 있다. 시험평가 계획수립시부터 시험평가 실시, 결과분석까지 시험평가분석 인원과 통계 SW (Minitab, G*Power) 등을 활용하여 지원하고 있다. 다만, 사람에 의한 오차는 시험 항목의 이해관계자의 다양한 관점(Framing Effect)에 따라 동일하게 해석되지 않을 경우도 발생할 수 있다. 이러한 오차를 줄이기 위해 시험평가 방법 및 절차를 수행하는 데 있어 무기체계를 조작하는 인원에 대한 무작위 선발과 운전자 교육 등을 통해 최소화하고 있다. 따라서, 무기체계 운용간 발생할 수 있는 오차는 무시할 수 있는 우연의 오차로 가정하고 있다.

AI는 학습을 통한 성능향상이 시험평가 전·중·후 지속적으로 이루어지고, 동작 특성이 기존의 SW와는 다르다. 일반적인 SW는 정해진 입출력을 산출하기 위해 알려진 절차와 방법을 구현하기 때문에 프로그램의 코드를 수정하기 전

113) 국방부, 전계서, p.55

까지는 항상 동일한 결과를 산출한다. 반면, AI 기능은 자료에 기반하여 원하는 출력을 얻기 위한 모형을 기계적으로 생성하기 때문에 모형은 동적으로 변화되고 확률적 결과를 제시한다. 즉 기존의 시험평가 프레임에 의한 이분법적 합·불 판정 적용이 제한된다. 또한, AI 기반 무기체계의 1차 통제 주체는 AI, 2차 통제 주체는 사람이 수행한다. 따라서 개발 장비의 시제품인 해당 무기체계에 대해서 실물에 의한 시험평가를 수행하기 전에 무기체계의 기반이 되는 핵심 SW인 AI에 대한 시험평가를 위한 별도의 절차가 선행되어야 한다.

표 15. 일반 무기체계와 AI 기반 무기체계 시험평가 비교¹¹⁴⁾

Table 15. General weapon system and AI-based weapon system T&E comparison

구 분	일반 무기체계	AI 기반 무기체계
시험평가 대상	<ul style="list-style-type: none"> 통제 주체가 사람 실제 운용 환경하에서 무기체계 시험평가 야전에서 직접 탑승해서 조작과 기동을 통해 극한의 조건에서 운용 능력 확인 실물에 의한 기동능력, 최대속도, 사거리, 정확도 등 직접 확인 다양한 환경에서 요구되는 성능 발휘 여부 등 확인 /검증 	<ul style="list-style-type: none"> 통제 주체: 1차 인공지능(AI), 2차 사람 AI 성능에 대한 시험평가 무기체계 장착된 각종 센서에서 수집된 영상 데이터를 AI가 먼저 처리 AI가 처리한 결과를 사람이 추가 확인 AI 시험평가 대상: 무기체계 성능을 결정하는 알고리즘의 적절성, 데이터 품질, 컴퓨터 성능 확인
시험평가 절차	<ul style="list-style-type: none"> HW/SW 특성은 연구개발주관기관 자체시험을 통해 성능 범위 예측 /확인 혹한·혹서기 시험 등 극한 환경(특정 조건)에서 실물에 의한 현장 시험평가 실시 	<ul style="list-style-type: none"> 개발과정에서 실제 운용될 환경에 대한 충분한 데이터 확보 제한 시험평가 착수 전까지 AI 성능 예측 곤란 운용환경에서 실물에 의한 현장 시험평가 진행 시 AI 운용환경에 대한 성능측정 불가능
SW 시스템	<ul style="list-style-type: none"> 주어진 입력 조건에 따라 정해진 결과를 출력할 수 있도록 사전에 프로그래밍 시스템의 정확성은 개발자의 기술력에 의존 (설계서와 일치된 결과물 출력) 	<ul style="list-style-type: none"> AI 알고리즘이 동작할 때마다 다른 결과 출력 시스템의 정확성은 입력된 데이터와 학습된 데이터에 의존

114) 합참, 인공지능 무기체계 시험평가 기준, 방법 절차 연구, 국방대학교, 2022, pp, 17~21, 연구자 재정리

AI 기반 경계시스템의 사례를 보면 기존의 경계시스템은 카메라 등 센서가 수집한 영상을 AI 기능 없이 영상감시를 담당하고 있는 사람(장병)이 직접 보고 판단을 했었다. 하지만 AI 기반 경계시스템에서는 카메라 등 각종 센서가 수집한 영상을 AI가 먼저 처리를 해서 결과 값을 영상감시를 담당하는 사람이 추가 확인하는 절차로 수행하고 있다. 따라서 일반 경계 시스템에서의 감시체계와 AI 기반 경계시스템의 감시체계에 대한 시험평가 방법 및 절차가 바뀌어야 한다.

예를 들어 일반적인 경계시스템의 영상 감시체계는 안개·우천·황사 등 자연 현상에 의해 바뀌는 기상변화에 대한 시험환경을 구축하고, 다양한 자연환경에서 감시 센서의 정상 작동 여부 및 요구성능을 충족하는지를 확인 검증하는 시험을 실시한다. 하지만 AI 기반 경계시스템은 기존 경계시스템의 감시체계의 시험 결과에 추가하여 AI 성능이 정상적으로 발휘되는지를 검증하기 위한 AI 성능평가를 실시해야 한다. AI 성능평가는 그림 18과 같이 데이터 기반으로 AI 분류 능력의 과소적합 및 과대적합 여부를 평가해서 실제 적용 가능성을 확인한다.

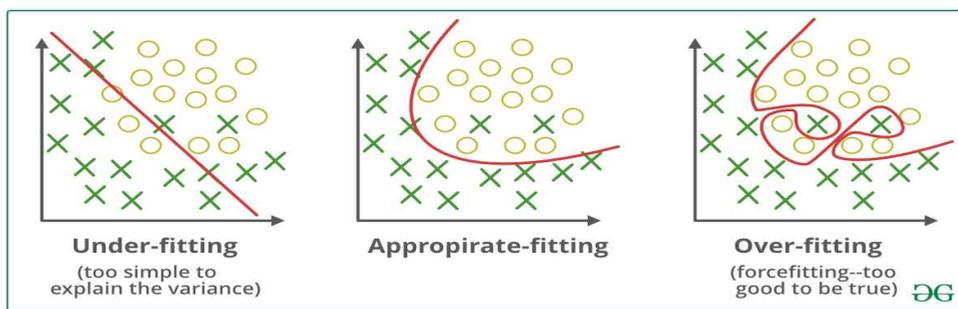


그림 18. AI 데이터 기반 기계학습의 과적합¹¹⁵⁾

Figure 18. Over-fitting in AI data-driven machine learning

115) https://colab.research.google.com/drive/1zx9k16VII8MZ3_n6VcnsBeiOoQQxavBF?usp=sharing, (검색일: 22. 4. 13.)

AI 기반체계의 시험평가에 있어 AI 성능평가에 영향을 주는 요소가 있다는 것을 인식할 필요가 있다. 첫째, AI 성능은 학습과 성능평가 정도에 따라 성능 발휘 수준을 결정한다. 양질의 데이터로 충분히 학습하고 평가되어야 원하는 기대성능을 발휘할 수 있는 것이다.

둘째, AI의 성능은 데이터의 품질과 성능평가의 타당성에 의존한다. 하지만 아직까지 군은 AI를 위한 학습과 시험평가용 데이터를 미보유하고 있을 뿐만 아니라, 그동안 군사용 AI 성능을 지원해주었던 한국인터넷진흥원(KISA)¹¹⁶⁾이 2021년부터는 국방분야 인증을 지원하지 않고 있어 군사용 AI 성능을 인증할 수 있는 공신력 있는 인증기관이 없다는 것이다.

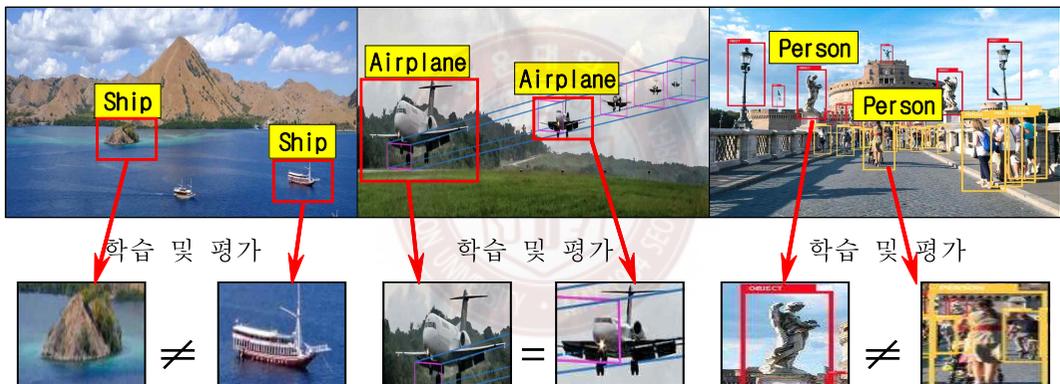


그림 19. AI 데이터 학습 및 평가

Figure 19. AI data learning and evaluation

마지막으로, 이러한 AI 체계의 특성을 이해하지 못하고 AI 기반 무기체계의 시험평가를 현장에서 실물에 의한 시험평가만 진행한다는 것은 AI 기반 무기체계의 성능을 충분히 검증하는데 제한이 된다. 예를 들어 AI 데이터의 성능

116) 한국인터넷진흥원(KISA, Korea Internet & Security Agency): 인터넷·정보보호 전문기관 2009년 창립되어 사회 전 분야로의 사이버 안전망 구축 확대 및 데이터 경제 활성화 기반 조성에 힘쓰고, 블록체인·5G 등의 신기술을 기반으로 한 기반시설 구축 및 혁신 서비스 발굴에도 앞장서고 있다. <https://www.kisa.or.kr/>, (검색일: 2022. 5. 2.)

평가를 미실시한 상태에서, 시간과 상황조성 등 여건이 제한되는 시험평가 현장에서 실물에 의한 시험평가를 수행한 결과 영상 감시체계의 성능이 기준을 미충족했다면, 기준 미충족에 대한 원인을 어떻게 분석해야 하는 것일까? 기준 미충족 원인이 영상 감시 센서의 문제인지 아니면 AI의 문제인지 구분할 수가 없다. 따라서 AI 기반 무기체계는 실물에 의한 현장평가 이전에 AI의 성능을 평가하는 데이터 기반 성능평가가 선행되어야 한다는 것이다.

군사용 AI는 군 운용환경의 특수성으로 인해 민간 AI와는 구별된다. 먼저, AI를 학습시키기 위해서는 분야별로 데이터가 일정량 이상으로 필요하고 데이터의 품질(Quality)에 따라 AI의 성능이 결정된다. 군사용 데이터의 경우 군사보안을 이유로 데이터의 수집과 관리가 어려워 개발자에게 데이터를 제공할 때 많은 제약이 따른다. 또한, 일부 특수분야는 센서의 다양성과 연동 문제로 데이터 확보가 제한되는 점이 있다.



그림 20. AI 및 실물에 의한 시험평가 운용환경(데이터)

Figure 20. T&E operational environment (data) by AI and prototype equipment

둘째, 문제를 해결하기 위해 정의된 규칙 및 절차의 모임인 프로그램 언어로 구현되기 때문에 국방환경의 특수성과 군의 요구사항에 부합한 군사용 인공지능을 개발하기 위해서는 민간의 기술력과 군사 전문가의 전문적 식견이

결합된 군사용 알고리즘 개발이 필요하다.

셋째, 데이터를 수집하고 관리하기 위한 하드웨어와 인공지능 개발을 위한 다양한 애플리케이션의 작동기반인 소프트웨어로 구성된다. 군사용 AI 연구개발 확산과 전·평시 실시간 사용을 위해 알고리즘의 개발, 저장, 공유를 위한 오픈형·폐쇄형 군사용 AI 플랫폼 구축이 필요하다. 또한, 부족한 컴퓨팅 환경과 전투현장에서도 작동이 가능하도록 Edge Computing¹¹⁷⁾, 모바일 기반의 체계 등 소형화, 경량화된 플랫폼도 필요하다.

국방 분야에서 인공지능을 활용하기 위해서는 체계적이고 실질적인 연구·개발이 필요하지만, 아직 성과를 거두고 있지는 못하고 있다. 중장기적으로 AI를 국방 분야에 활용할 수 있도록 체계적인 추진이 필요하다.

2. 무기체계 시험평가 수행체계 개선 소요

가. 전문가 설문을 위한 잠정 평가요소 선정

앞서 살펴본 바와 같이 무기체계 개발환경의 변화, 일반 무기체계와 AI 기반 무기체계의 시험평가 차이를 통해 본 연구에서는 4차 산업혁명과 시대적 변화에 적합한 첨단 무기체계 시험평가 개선방안을 도출하였다. 이를 위한 방법론으로 문헌연구 및 사례연구 등 선행연구와 국방 전력분야 전문가를 대상으로 기초조사 설문 및 인터뷰를 통해 시험평가 제도·구조·기술·법령 및 훈령 측면에서 첨단 무기체계 시험평가 개선을 위한 세부 평가항목을 도출하였다.

먼저, 시험평가 제도적 측면에서 제2장의 국방획득체계 변천 과정에서 제시한 바와 같이 현 무기체계 획득제도는 투명성, 전문성에 중점을 둔 정책추진으로 신속성과 효율성이 결여되어 있다. 빠르게 변화하는 첨단 과학기술을 신

117) Edge Computing: 데이터 처리를 중앙 서버가 아닌 데이터가 발생하는 주변(Edge)에서 데이터를 처리하는 기술

속하게 반영하지 못하고 있어 무기체계 소요결정으로부터 최초 전력화까지 평균 16.8년이 소요되고 있다. K2 전차의 경우에는 1992년 소요결정 이후 2015년 전력화까지 23년이 소요되었으며, TMMR 무전기의 경우 1996년 소요결정 이후 전력화까지 19년이 소요되는 등 4차 산업혁명의 신기술 기반, 첨단 무기체계를 전력화하는 데 많은 제한사항이 있다. 이러한 국방획득체계에 무기체계 시험평가 제도 분야가 포함되어 있다. 최근 이를 개선하기 위해 방위사업청에서는 2020년 신속시범획득사업을 신설하고 2022년까지 총 34개 사업 중 육군 사업 26개를 선정(종료 16개, 진행 10개)하여 추진한 결과, 표 16과 같이 종료사업(16개) 중 전력화와 연계한 사업은 8개 사업으로 소요와 연계가 낮거나 성능이 일부 미흡하여 전력화 연계율이 50%로 낮은 수준을 보이고 있다.

표 16. 신속시범획득사업 추진현황(2020~2022년)¹¹⁸⁾

Table 16. Current status of rapid demonstration acquisition project(2020-2022)

구 분	총계	Army Tiger	드론	대드론	현존전력 강화
사업(수)	16개	4개	5개	1개	6개
전력화 연계	8개	2개	1개	1개	4개

신속시범획득사업이 기존 무기체계 획득 소요 기간과 비교할 때 상대적으로 매우 짧은 기간이다. 하지만 신속시범획득사업 선정시 업체가 제시한 과제 위주의 시범사업이 아닌 군이 원하는 시범사업이 선정되어 군사적 활용성이 검증되고 소요가 결정될 수 있도록 제도적 보완이 필요하다. 또한, 선행연구 고찰에서 임영봉(2014) “육군 무기체계 운용시험평가에 관한 연구”에서 제시한 바와 같이 국제시험평가 교류 및 협력을 강화하고 국제 시험평가 프로그램

118) 육군, 신속시범획득사업 추진현황, 2023, 연구자 재정리

(ITEP, International T&E Program)¹¹⁹⁾ 참여 및 국방 연구개발 협력을 통해 해외 시험시설 활용을 비롯하여 시험평가 기술 노하우와 시험기준 표준화 등 K방산의 수출에 기여할 수 있도록 제도적 보완이 필요하다.

시험평가 구조적 측면에서는 그림 21과 같이 미국의 시험평가 조직을 벤치마킹할 필요가 있다. 미국은 역사적으로 수많은 전쟁을 통해 군사 대국이 되었다. 국방 획득제도가 1972년 미국으로부터 도입되었듯이 시험평가 체계 또한 미국의 무기 원조로부터 시작하여 면허생산 이후 자체 무기 개발까지 시험평가의 중요성을 깨닫게 되었다. 시험평가는 무기체계 연구개발에 있어 성과와 군 운용 적합성을 검증하는데 핵심 과정이다.



그림 21. 미 육군시험평가사령부(ATEC) 조직도¹²⁰⁾

Figure 21. U.S. Army test and evaluation command(ATEC) organization

119) ITEP는 시험평가 절차, 기법의 공동발전 및 국가별 시험장 활용 등 협력적 시험평가 활동을 제공하기 위해 운영중인 국제협력 프로그램으로서, 미국은 국제 파트너 11개국(영국, 프랑스, 독일, 이탈리아, 덴마크, 필란드, 네덜란드, 노르웨이, 스웨덴, 호주, 캐나다)과 양자 협정을 맺고 참여하고 있다. 현재 한국과 일본, 이스라엘은 가입 협상을 진행하고 있다.

120) U.S Army Test and Evaluation Command Website, <https://www.atec.army.mil/organization.html>, (검색일: 2022. 5. 8.)

선행연구 고찰에서 강응섭(2016) “국방기술 환경변화에 부응한 무기체계 시험평가 신뢰도 제고에 관한 연구”, 정월균(2017) “이해관계자 분석을 통한 무기체계 평가제도에 관한 연구”, 이용학(2017) “국방 무기체계 시험평가의 투명성에 관한 연구”에서 제시한 바와 같이 미국의 시험평가 관련 조직과 시험평가 인프라 등을 참고하여 한국군의 특성에 맞는 시험평가 조직과 인프라를 구축할 수 있도록 발전시킬 필요가 있다. 또한, 4차 산업혁명의 신기술을 무기체계에 적용하는 데 있어 무기체계 시험평가의 신뢰성을 높이기 위해서는 소요 결정으로부터 시험평가 준비·실시·종료 이후까지 시험평가를 지원하고 분석할 수 있도록 민·군 기술협력 및 첨단기술 전문 시험평가관 양성 및 예산을 반영할 필요가 있다.

시험평가 기술적 측면에서는 첨단 과학기술의 발전에 따라 다양한 무인 전투체계들이 전투원을 대신하여 인공지능을 갖춘 무인 전투체계들이 전투를 수행하는 무인·로봇전이 될 것이다. 또한, 모든 전투플랫폼과 전투지휘 체계들이 고속기동화, 지능화, 네트워크화되어 초지능·초연결될 것이다. 이와 같은 인공지능(AI) 기반 무기체계는 전통적인 무기체계와 다른 시험평가가 필요하다. 인공지능은 학습을 통한 성능향상이 시험평가 전·중·후 지속적으로 이루어지고, AI 모델의 성능평가를 위해서는 학습에 사용되지 않은 별도의 평가의 데이터셋이 필요하다. AI는 기존 무기체계와 같이 현장에서 실물에 의한 시험평가만으로 시험기준을 충족할 수가 없다. 실물에 의한 시험평가 이전에 AI 기반의 데이터 성능평가를 우선적으로 실시해야 한다.

이와 같은 첨단 무기체계 시험평가를 위해서는 선행연구 고찰의 박종완(2015) “무기체계 시험평가의 신뢰성 향상방안”, 최규옥 등 3명(2019) “무기체계에 적용 가능한 적응형 시험평가에 관한 연구”, 권세일·강지훈 등 5명(2022) “인공지능 기술 및 인증제도 분석을 통한 국방 인공지능 품질관리 방안 연구”,

이용복 등 3명(2022) “인공지능(AI) 무기체계 시험평가 방법 발전방안 연구”에서 제시한 바와 같이 과학적 시험평가 및 시험평가 기술지원과 Army TIGER 기반체계와 드론봇 전투체계의 초연결망을 기반으로 하는 네트워크 통합시험이 필요하다. 또한 AI 기반 무기체계 시험평가를 위한 한국형 시험표준을 정립할 필요가 있다.

마지막으로 시험평가 법령 및 훈령은 시험평가 제도를 뒷받침하는 것으로 현 국방전력발전업무훈령과 합참 시험평가 업무 규정에 개발시험평가(DT&E)의 경우 “기준충족 또는 기준미달”, 운용시험평가(OT&E)는 “전투용 적합 또는 전투용 부적합”으로 시험평가 결과를 판정한다. 시험 결과가 평가 기준과 비교하여 1개 항목이라도 기준에 미달하는 경우 “개발 실패 또는 전투용 부적합”으로 분류하고 있다. AI 기반 무기체계는 데이터 기반으로 지속적인 학습이 이루어지는 과정으로 이분법적으로 시험평가 결과를 판정할 수가 없다. 또한 AI와 무인·자율주행과 관련하여 아직까지 윤리적 기준과 법률적으로 정립이 되지 않은 점이 있어 무기체계에 적용하기 위해서는 관련 법령 제정이 필요하다.

첨단 무기체계 시험평가를 위해서는 선행연구 고찰의 강웅섭(2016) “국방기술 환경변화에 부응한 무기체계 시험평가 신뢰도 제고에 관한 연구”, 김선영·최기일(2019) “방위산업 발전을 위한 무기체계 시험평가 개선 연구”, 장상국·최기일(2021) “미래 국방을 대비한 인공지능 기반의 방위산업 발전방향 연구”에서 제시한 바와 같이 통합시험 확대와 AI·무인·자율주행 등 관련 법령을 제정하고, 환경시험 관련 국방표준 정립, 안전기준 인증 및 품질보증 등이 필요하다.

이상과 같이 전문가 기초조사 설문을 위하여 첨단 무기체계 시험평가 개선을 위한 잠정 평가항목을 표 17과 같이 분야별로 그룹화하였다.

표 17. 첨단 무기체계 시험평가 개선 잠정 세부 평가항목

Table 17. Provisional detailed evaluation items required for advanced weapon system T&E improvement

구 분	세부 평가항목	
시험평가 제도	<ul style="list-style-type: none"> ◦ 신속획득방안 개선/ 확대 ◦ 국방연구개발 협력 	<ul style="list-style-type: none"> ◦ 국제시험평가 교류/협력
시험평가 구조 (조직/편성)	<ul style="list-style-type: none"> ◦ 시험평가 관련 조직 ◦ 시험평가 인프라 구축 	<ul style="list-style-type: none"> ◦ 민·군 기술협력 / 교류 ◦ 시험평가 지원
시험평가 기술 (방법/절차)	<ul style="list-style-type: none"> ◦ 과학적 시험평가체계 정립 ◦ 시험평가 기술지원 	<ul style="list-style-type: none"> ◦ 네트워크 통합시험체계 구축 ◦ 한국형 시험표준 정립
시험평가 법령 / 훈령	<ul style="list-style-type: none"> ◦ 통합시험 확대 ◦ AI·무인·자율주행 등 관련 법령 제정 ◦ 환경시험관련 국방표준 정립 	<ul style="list-style-type: none"> ◦ 시험평가 결과관정 ◦ 안전기준 인증 및 품질보증

나. 전문가 기초조사 설문을 통한 평가항목 확정

기초조사 설문은 선행연구를 통해 도출된 첨단 무기체계 시험평가 개선 소요에 대한 잠정 평가항목을 확정하기 위한 과정으로 도출된 평가항목은 첨단 무기체계 시험평가 개선방안에 대한 정량적 분석인 계층적 분석기법(AHP)의 계층구조와 세부 평가항목으로 반영하였다.

선행연구를 통해 도출된 첨단 무기체계 시험평가 개선 소요에 대한 기초조사 설문을 표 18과 같이 국방 전력분야 전문가 45명을 대상으로 실시하였다. 기초조사 설문은 현재 국방 전력분야에 근무하고 있거나 과거에 전력업무를 수행한 경험이 있는 관계자를 대상으로 첨단 무기체계 시험평가 개선 필요성에 대해 객관식 질문과 단일 및 복수 답변이 가능하도록 하여 응답자들의 설

문조사가 수월할 수 있도록 했다. 설문서는 직접 접촉과 E-mail을 통해 99% 회수하였다. 설문서를 검토한 결과 신뢰성이 모두 있는 것으로 총 45부를 분석에 활용하였다. 이러한 설문조사 응답자의 전력분야 근무경력과 근무기간, 직업(신분) 등 통계학적 특성은 그림 22와 같다.

표 18. 기초조사 설문 전문가 집단의 분포

Table 18. Distribution of expert groups in the basic survey questionnaire

구분	계	실무경력				신분			
		5년 미만	5~10년 미만	10~20년 미만	20년 이상	군인	공무원	연구원	방산업체
기간									
인원(명)	45	13	17	11	4	22	8	10	5
비고	<ul style="list-style-type: none"> ◦ 국방부/합참(6명) ◦ 방사청(5명) ◦ 육.기참부(4명) ◦ 시평단(11명) ◦ 분평단(4명) ◦ ADD·KIDA·기품원(10명) ◦ 방산업체(5명) ◦ 군인(22명): 장군(2명), 영관장교(12명), 준사관(8명) 								



그림 22. 기초조사 설문 대상자의 통계학적 특성

Figure 22. Statistical characteristics of the subjects of the basic survey

기초조사 설문은 객관식 문항에 대한 복수 응답과 답변의 이유를 기술하고

있어 엑셀을 활용하여 기초조사 설문을 종합하였다.

기초조사 전문가 설문을 분석한 결과, 4차 산업혁명의 시대에 미래 첨단 무기체계를 전력화하는 데 있어, 현행 국방획득 절차 적용이 가능한가에 대해서 그림 23과 같이 응답자의 82.2%가 일부 가능 또는 가능하다고 하였다. 하지만, 현 획득제도에 대해서 신속성과 효율성이 결여되어 급변하는 첨단 신기술을 무기체계에 신속하게 적용하기 위해서는 제도적 개선이 필요하다는 의견을 제시하였다.

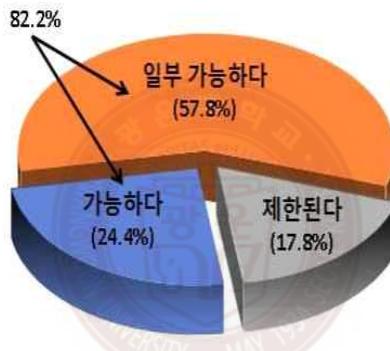


그림 23. 현행 국방 획득체계를 활용한 첨단 무기체계 전력화 가능성
Figure 23. Possibility of deploying advanced weapon systems using the current defense acquisition system

첨단 무기체계 시험평가를 위해서 개선해야 할 분야로는 그림 24와 같이 시험평가 기술(방법 / 절차)(37%) > 시험평가 구조(조직 / 편성)(29.3%) > 시험평가 법령 및 훈령(17.4%) > 시험평가 제도(15.2%) 순으로 개선이 필요하다는 의견을 제시하였다. 추가의견으로 ‘시험평가 제도’에 ‘법령 및 훈령’을 포함하고 있어 두 가지 항목으로 구분하는 것보다 통합해야 한다는 의견을 다수가 제시하였다. 국립어학원의 표준국어대사전에 ‘제도’의 사전적 의미로 관습, 도덕, 법률 따위의 규범적이나 사회구조의 체계로 정의하고 있다. ‘제도’의 사전

적 의미에 ‘법령’을 포함하고 있는 것으로 분석되어 ‘시험평가 제도’에 ‘법령 및 훈령’을 통합하는 것으로 검토하였다.

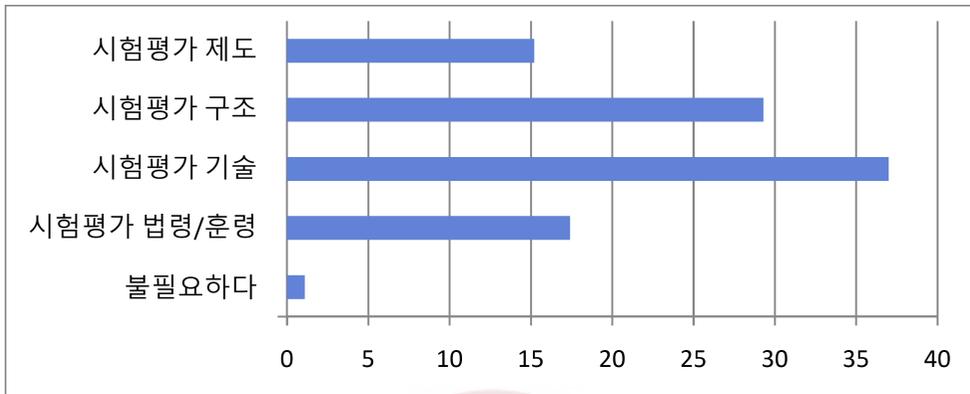


그림 24. 첨단 무기체계 시험평가를 위해 개선이 필요한 분야

Figure 24. Areas that need improvement for T&E of advanced weapon systems

시험평가 제도적 차원에서 개선해야 할 분야는 그림 25와 같이 신속획득 방안(41.9%) > 국방 연구개발 협력(핵심기술 공동연구, 시험평가 기술공유 등)(41.9%) > 국제 시험평가 교류 및 협력(16.1%) 순으로 개선이 필요하다는 의견을 제시하였다.

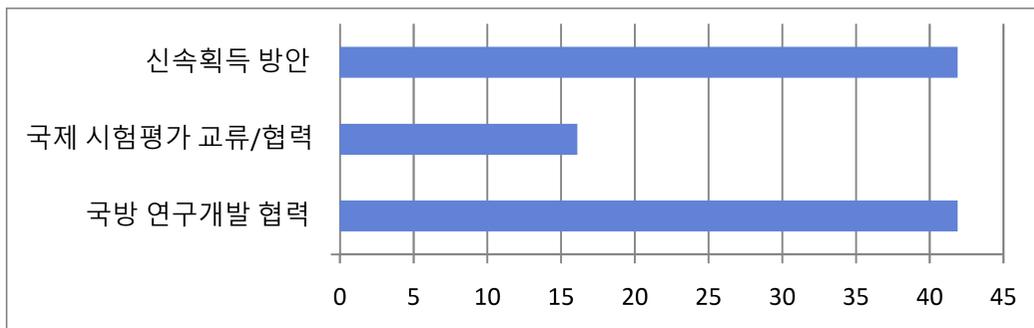


그림 25. ‘시험평가 제도’ 차원에서 개선해야 할 분야

Figure 25. Areas to be improved in terms of the ‘T&E system’

이는 앞서 제시한 현 국방획득 제도하에서 첨단 과학기술을 적용한 무기체계의 신속한 획득을 위해서는 제도적 개선이 필요하다는 의견과 같은 사항이다. 현 신속시범획득사업은 시범사업의 과제선정 과정에서 업체에서 제시된 과제 위주로 시범사업이 선정되고 소요군(시험평가단)에서 군사적 활용성이 검증되어도 군에서 필요로 하는 기술이 아닌 경우가 많아 군이 참여해 원하는 무기체계가 선정될 수 있도록 시범사업의 제도적 개선이 필요하다는 의견을 제시하였다.

시험평가 구조¹²¹⁾적 차원에서 개선해야 할 분야는 그림 26과 같이 시험평가 인프라(전문 시험시설 및 장비, 시험장 등)(29.7%) > 시험평가 지원(시험기술 데이터 관리/계측, 시험평가 기술연구 등)(25.7%) > 시험평가 관련 조직(23.8%) > 민·군 기술협력(20.8%) 순으로 개선이 필요하다는 의견을 제시하면서, 시험평가 구조적 측면에서는 어느 특정 분야가 아닌 전체적인 분야를 균등하게 개선이 필요하다는 의견이었다.

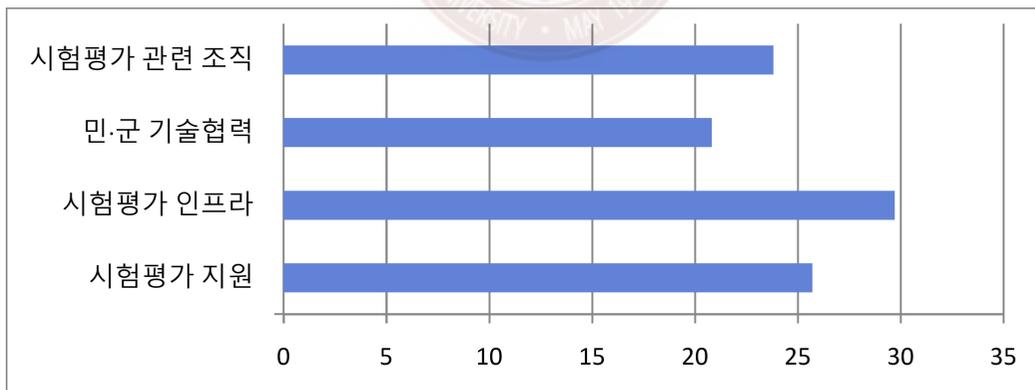


그림 26. '시험평가 구조' 차원에서 개선해야 할 분야

Figure 26. Areas to be improved in terms of the 'T&E structure'

121) 사전적 의미 : 부분이나 요소가 어떤 전체를 짜 이룸 또는 그렇게 이루어진 일개, 각 부분이나 요소를 모아 어떤 전체를 만들. 또는 그렇게 만들어진 전체의 뼈대로 정의하고 있다. [국립어학원 표준국어대사전·고려대 한국어대사전]

이는 첨단 무기체계 시험평가를 위해서는 인공지능(AI)을 비롯하여 드론과 로봇 등 4차 산업혁명의 신기술을 적용하는 데 있어 소요군의 시험평가 기관만의 자체적인 노력으로 할 수 없는 사항으로 정부 차원에서 관련 예산과 인력을 편성하고 지원을 해야 한다는 의견을 제시하였다.

시험평가 기술¹²²⁾(방법/절차) 차원에서 개선해야 할 분야는 그림 27과 같이 과학적 시험평가(39.1%) > 시험평가 기술지원(지원프로그램)(24.1%) > 한국형 시험표준 정립(20.7%) > 네트워크 통합시험(16.1%)순으로 개선이 필요하다는 의견을 제시하였다.

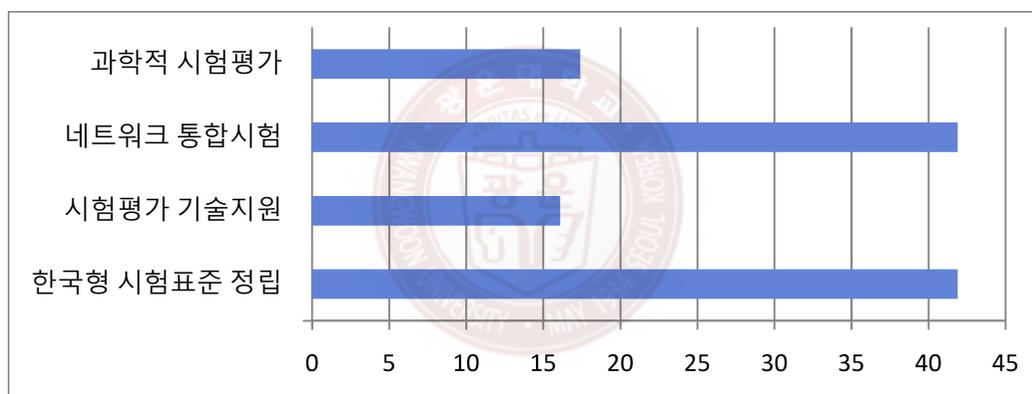


그림 27. '시험평가 기술' 차원에서 개선해야 할 분야

Figure 27. Areas to be improved in terms of the 'T&E technology'

이는 현 시험평가를 수행하면서 신뢰성에 많은 의문을 제기하고 있다는 점을 들면서 첨단 과학기술이 접목된 신기술에 대한 전문성과 시험평가 절차가 아직 정립되지 않은 시점에 있다. 시험평가 전문기관으로 시험평가를 수행하는 데 있어 과정과 결과에 대해 신뢰성을 확보하기 위해서는 시험평가 기관의

122) 사전적 의미 : 과학 이론을 실제로 적용하여 사물을 인간 생활에 유용하도록 가공하는 수단, 어떤 원리나 지식을 자연적 대상에 적용하여 인간 생활에 유용하도록 만드는 구체적이고 실제적인 수단으로 정의하고 있다. [국립어학원 표준국어대사전·고려대 한국어대사전]

자구노력과 함께 제도적 지원이 필요하다는 의견을 제시하였다.

시험평가 관련 법령 및 훈령 차원에서 개선해야 할 분야로 그림 28과 같이 AI·무인·자율주행 등 관련 법령 제정(37.5%) > 시험평가 결과의 판정(조건부 적합 추가 등 유연성 확대)(21.9%) > 안전기준, 인증 및 품질보증 기준(15.6%) > 통합시험 확대 적용(14.6%) > 환경시험 관련 국방표준 제정(10.4) 순으로 개선이 필요하다는 의견을 제시하였다.

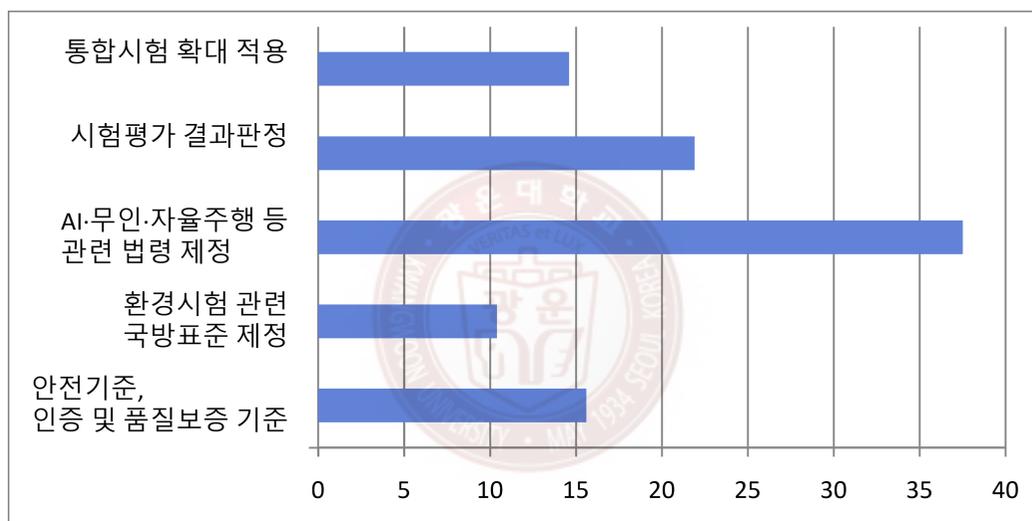


그림 28. '시험평가 법령/훈령' 차원에서 개선해야 할 분야

Figure 28. Areas to be improved at the level of the 'T&E laws/directives'

시험평가 관련 법령 및 훈령은 앞에서 제시했듯이 시험평가 제도에 통합하는 것으로 검토되었다. 시험평가 관련 법령 및 훈령으로 검토되는 사항은 민간차원에서 인공지능의 윤리나 킬러로봇에 대한 윤리에 대해 많은 검토가 이루어지고 있으며, 美 국방혁신단(DIU)의 '책임감 있는 AI 활용을 위한 실무 지침서'에 개발 중인 AI 프로그램이 모든 개발단계에서 국방부 AI 윤리 원칙에 부합하도록 하고, 공정성, 책임소재 및 투명성을 준수하도록 지침을 제공하고

있다. 이러한 점을 고려하여 AI 기반 유·무인 복합무기체계를 개발하는 데 있어 법률적으로 제한이 없도록 관련 법령을 제정해야 한다는 것이다. 또한 현 시험평가 제도에서도 문제를 제기하고 있는 시험평가 결과판정은 과학기술의 발전추세에 맞도록 진화적으로 무기체계가 개발되어야 한다는 점과 인공지능의 학습과 평가가 단계화되어 이루어지는 것을 고려해야 한다. 또한, 이분법적으로 전투용 적합·부적합으로 판정하는 것이 아니라 융통성 있게 조건부 적합 등 진화적으로 시험평가 결과를 판정할 수 있도록 검토되어야 한다는 의견을 제시하였다.

첨단 무기체계 운용시험평가를 수행하는 데 있어 그림 29와 같이 현 운용시험평가 항목에 대해 개선은 불필요하다는 의견이 다수였다. 지금까지 현 운용시험평가 항목으로 드론이나 무인수색차량, AI 경계시스템 등 운용시험평가를 수행하는데 제한사항이 없었다는 점을 제시하고 있다. 다만, 신개념 첨단 무기체계 시험평가를 수행하는 데 있어 중요한 것은 시험평가 방법 및 절차에 있을 뿐 시험평가 항목과는 관련이 없다는 의견이 많아 운용시험평가 항목에 대한 계층적 분석(AHP) 및 개선방안 연구에서 제외하는 것으로 하였다.

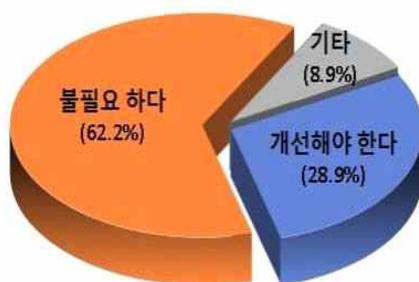


그림 29. ‘운용시험평가 항목’ 개선 필요성
Figure 29. Need for improve ‘OT&E items’

운용시험평가 항목과 관련하여 일부 의견으로 필요한 항목은 그림 30과 같이 군 운용 적합성(42.1%) > 작전운용성능의 충족성(30.3%) > 전력화지원요소의 실용성(24.4) 순으로 개선이 필요하다는 의견을 제시하였다.



그림 30. '운용시험평가 항목' 중 개선이 필요한 항목
 Figure 30. Items in need of improvement among 'OT&E items'

이는 운용시험평가를 수행하면서 군 운용 적합성 항목과 관련하여 사업관리 부서나 연구개발주관 기관에서 많은 의문을 제기하고 있는 사항이다. 세부 항목 선정 시 운용개념과 사업관리 회의를 통해 운용부대에서 제시한 요구사항을 시험평가 항목으로 반영될 수 있도록 하는 체계를 구축한다면 사업관리 부서 및 연구개발주관 기관의 의문을 해소하고 운용시험평가에 대한 신뢰성을 갖게 할 수 있을 것이다.

운용시험평가 항목 가운데 작전운용성능의 충족성 중 개선이 필요한 항목은 주요 작전운용성능(55.6%) > 기술적 / 부수적 성능(37.8%) > 기타(6.7%) 순이며, 군 운용 적합성 중 개선이 필요한 항목은 전술적 운용의 적합성(30.1%) >

기존 무기체계와의 상호운용 적합성(27.4%) > 운용 및 조작 적합성, 안정성(27.4%) > 환경 적응성(15.1%) 순으로 제시하였다. 전력화지원요소의 실용성은 통합체계지원(IPS) 요소(57.6%) > 전투발전지원요소(36.4%) > 기타(6.1%) 순으로 의견을 제시하였다.

이상과 같이 무기체계 개발환경의 변화에 따른 첨단 무기체계 시험평가 개선 수요를 도출하기 위해 선행연구와 AI 기반 시험평가 사례, 전문가 기초조사 설문문을 분석하여 시험평가 개선 수요에 대한 세부 평가항목을 시험평가 제도·구조·기술적 분야로 정리한 결과, 표 19와 같다.

표 19. 첨단 무기체계 시험평가 개선 수요 세부 평가항목

Table 19. Detailed evaluation items required for advanced weapon system T&E improvement

구 분	세부 평가항목
시험평가 제도 (법령/훈령)	<ul style="list-style-type: none"> ◦ 신속획득방안 개선/ 확대 ◦ 시험평가 결과관정 개선 ◦ AI·무인·자율주행 등 관련 법령 제정 ◦ 안전기준 정립
시험평가 구조 (조직/편성)	<ul style="list-style-type: none"> ◦ 시험평가 조직 및 전문인력 양성 ◦ 민·군·국제 기술협력 / 교류 <ul style="list-style-type: none"> * 산·학·연·군 협업체계, 국제 시험평가 프로그램(TEP) 참여 ◦ 시험평가 인프라 구축
시험평가 기술 (방법/절차)	<ul style="list-style-type: none"> ◦ 과학적 시험평가체계 정립 ◦ 네트워크 통합시험체계 구축(Army TIGER 기반체계) <ul style="list-style-type: none"> * AI·빅데이터, 클라우드 기반 초연결 전장 네트워크, 유·무인 복합, 자율지능화 무기체계 등 ◦ 시험평가 기술지원(평가지표, 방법·절차) ◦ 한국형 시험표준 정립

제3절 계층적 분석기법(AHP)을 적용한 개선 우선순위 분석

1. 계층적 분석기법(AHP) 적용절차

본 연구는 4차 산업혁명의 시대에 첨단 과학기술이 적용된 무기체계 시험평가 개선방안을 도출하기 위해 국방정책 및 소요기획, 사업관리, 시험평가, 연구개발주관기관 등 국방 전력업무 관련 전문가를 대상으로 계층적 분석기법(AHP) 설문조사를 실시하고 분석하였다.

무기체계 시험평가 개선방안에 대한 기초조사 설문 대상자 중 실무경력 2년 미만은 AHP 설문에서 제외하고 실무경력 3년 이상인 전문가를 추가 반영하여 표 20과 같이 46명을 대상으로 첨단 무기체계 시험평가 개선방안을 도출할 수 있도록 AHP 설문조사(전문가 인터뷰)를 실시하였다.

표 20. AHP 전문가 집단의 세부 현황
Table 20. Details of the AHP expert group

전력분야 경력	3~5년 미만	13명	소속기관
	5~10년 미만	17명	◦ 국방부/합참(6명) ◦ 방사청(5명)
	10~20년 미만	11명	◦ 육.기참부(5명) ◦ 시평단(11명)
	20년 이상	5명	◦ 분평단(4명) ◦ 방산업체(5명)
계급/신분	군 인	23명	* 군인(23명) : 장군(2), 영관장교(13), 준사관(8)
	공(군)무원	8명	
	연구원	10명	
	방산업체	5명	
계		46명	

계층적 분석기법(AHP)의 계층구조 및 평가항목에 대한 MECE(Mutually Exclusive, Collectively Exhaustive) 충족 여부를 검증하기 위해 합참 시평부 및 육군 시평단, 방사청 및 국과연 등 국방 전력분야 전문가 그룹과 토의를 통해 보완하였다. 선행연구와 기초조사 설문 및 인터뷰를 통해 도출된 계층구조 및 평가항목 간 MECE 충족이 필요하다는 의견이 있어 AHP 계층구조 및 평가항목을 수정하고, 표 21과 같이 첨단 무기체계 시험평가 개선방안을 위한 AHP 평가항목의 계층구조를 설계한 후 표 22와 같이 평가 요소에 대한 개념을 구체화하였다.

표 21. AHP 계층구조 설계
Table 21. AHP hierarchical structure design

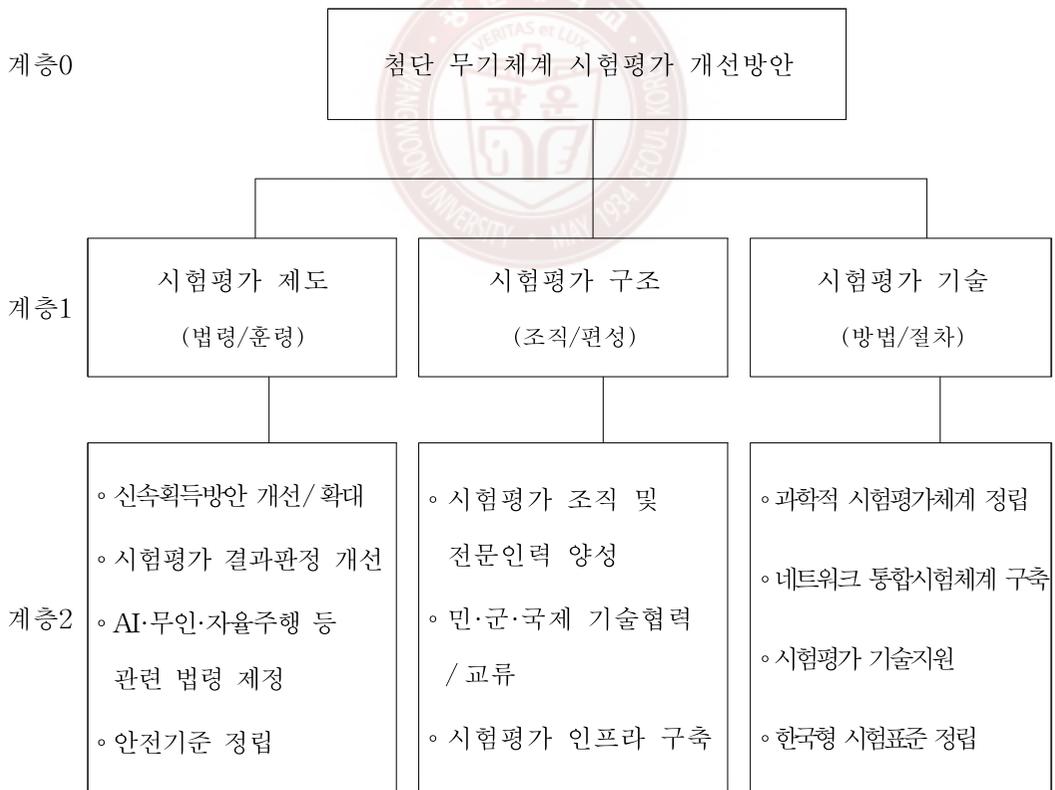


표 22. AHP 계층구조 및 평가요소의 개념

Table 22. Concept of AHP hierarchy and evaluation factors

계층0	계층1	계층2	평가요소의 개념
첨단 무기 체계 시험 평가 개선 방안	시험 평가 제도 (법령 /훈령)	신속획득방안 개선/ 확대	미 국방혁신단(DIU) COS(Commercial Solution Open)를 벤치 마킹하여 한국형 국방혁신단(K-DIU) 신설 획득체계 제도개선 및 범위 확대, 제도적 법적 기반 구축
		시험평가 결과관정 개선	소요군이 요구하는 무기체계의 성능 수준을 최적화하면서 전력화 시기를 충족할 수 있는 시험평가 결과관정의 유연성과 제도 적인 보완(결과관정의 안정적인 검증체계 구축)
		AI·무인자율주행 등 관련 법령 제정	AI·무인차량·로봇·자율주행차량 주행을 위한 관련 규제 및 법령 체계 정립, 국방 데이터 활용을 제약하는 보안규정 개선 등
		안전기준 정립	무기체계 연구개발 및 획득과정에 안전성 관련 제도 및 관리 체계 구축, 안전기준 인증 및 품질보증 기준 제정
	시험 평가 구조 (조직 /편성)	시험평가 조직 및 전문인력 양성	AI·로봇·자율주행·무인체계 등 미래 첨단 무기체계 시험평가를 위한 전문인력 양성, 시험평가 조직개편 및 보강
		민·군·국제 기술협력 / 교류	민·군 기술협력체계 구축, 시험평가와 관련하여 산·학·연·군 협업체 계 구축, 국제시험평가 프로그램(TEP) 참여 등 * 첨단 무기체계 시험평가 인프라 활용 및 제도 벤치마킹
		시험평가 인프라 구축	AI, 무인로봇, 자율주행 등 첨단과학기술이 적용된 무기체계 전문 시험평가 시설 구축
	시험 평가 기술 (방법 /절차)	과학적 시험평가체계 정립	통계적으로 신뢰도와 타당도를 높일 수 있는 방법적용 * 시험 대상 체계 특성을 고려 시험항목 및 시험방법 적용, 객관적으로 신뢰할 수 있는 시험결과 제시 * 시뮬레이션에 의한 시험평가 프로세스(STEP) 및 HILS 시스템, M&S 등 적용
		네트워크 통합시험체계 구축	AI·빅데이터·클라우드 기반 초연결 전장 네트워크, 유·무인 복합 자율지능화 무기체계 등 네트워크 통합시험평가 체계 구축 * Army TIGER 기반체계 네트워크 통합시험
		시험평가 기술지원	시험평가 DB 구축 및 관리, AI 등 미래 무기체계 시험평가기술 (평가지표, 방법·절차) 연구 및 지원 * 시험계획 수립시부터 결과분석 전 과정에 신뢰성 있는 시험계획·시험환경 구축·시험 데이터 분석/ 변환 등 기술지원
		한국형 시험표준 정립	대한민국 작전환경 특성에 적합한 시험평가 기준 및 절차 제정 * 군사 표준규격(MIL-STD), 시험운용절차(TOP)

AHP 설문조사는 미래 첨단 무기체계 시험평가 개선방안에 대한 우선순위를 체계적으로 평가할 수 있는 방법으로, 리커드 9점¹²³⁾ 평가척도를 사용하여 설문 문항을 구성하였다. 설문 조사간 응답자의 판단 오차를 최소화하고 신뢰성을 높이기 위해 직접 방문 및 화상통화를 통해 설문 배경과 목적, 설문 방법 및 절차에 대해 자세하게 설명을 실시한 후 진행하였다. 이후 전문가 인터뷰를 통해 설문 결과에 대한 추가의견을 받았다.

AHP 분석 방법에 활용한 도구는 육군 분석평가단에서 분석용으로 보유 중인 전문분석 소프트웨어 ‘메이크잇(Make It)’으로서, AHP 분석을 위해 활용하는 유용한 분석프로그램이다. 설문의 일관성 비율(CR, Consistency Ratio)을 확인하기 위해 설문조사 결과값을 입력 후 분석한 결과, 설문 대상자 46명의 설문 중 일관성 비율이 0.1 이상인 7부는 일관성 값이 유효하지 않아 제외하고, 가중치는 39부의 평가 결과에 대해 보간법을 적용하여 산출하였다.

2. 계층구조 및 평가항목별 우선순위 분석 결과

국방획득 관련 소요제기, 사업관리, 시험평가, 국과연을 비롯한 지원기관과 방위산업 업체 등 국방전력 분야 전문가를 대상으로 AHP 설문을 실시한 결과, 설문지 배부에 대한 회수 및 유효한 응답은 표 23과 같이 분석되었다.

표 23. AHP 설문지 회수율 및 유효한 응답 결과

Table 23. AHP questionnaire return rates and valid response results

설문 대상	설문지 회수 결과		응답 결과	
	회수	회수율	유효응답	응답율
53명	46부	86.8%	39부	84.8%

123) 9점 리커드 척도: Saaty 교수는 쌍대비교를 할 때 요인 A에 대한 요인 B의 수준을 나타내는 척도를 7±2개로 제안하면서 AHP 쌍대비교에서 일반적으로 사용하고 있다.

본 연구는 첨단 무기체계 시험평가 개선방안에 대한 계층구조별 상대적 중요도 및 세부 평가항목에 대한 우선순위를 계층적 분석기법(AHP)을 활용한 것으로 제3장 연구의 방법론에서 분석 절차를 세부적으로 살펴보았다.

가. 대분류(계층 1) 항목에 대한 평가결과

선행연구 및 사례연구 결과 도출된 첨단 무기체계 시험평가 개선방안에 대한 대분류 항목은 시험평가 제도, 시험평가 구조, 시험평가 기술 등 3개 요소이며 대분류에 대한 상대적 중요도를 평가한 결과, 시험평가 제도(0.4700) > 시험평가 기술(0.2790) > 시험평가 구조(0.2510) 순으로 우선순위가 확인되었으며, 세부 결과는 표 24 및 그림 31과 같다.

표 24. 대분류(계층 1) 항목의 중요도 및 평가 우선순위

Table 24. The importance and evaluation priority of the major classification (tier 1) items

계층 0(목표)		계층 1(대분류)	중요도	우선순위
첨단 무기체계 시험평가 개선방안	1.0000	시험평가 제도	0.4700	1
		시험평가 구조	0.2510	3
		시험평가 기술	0.2790	2

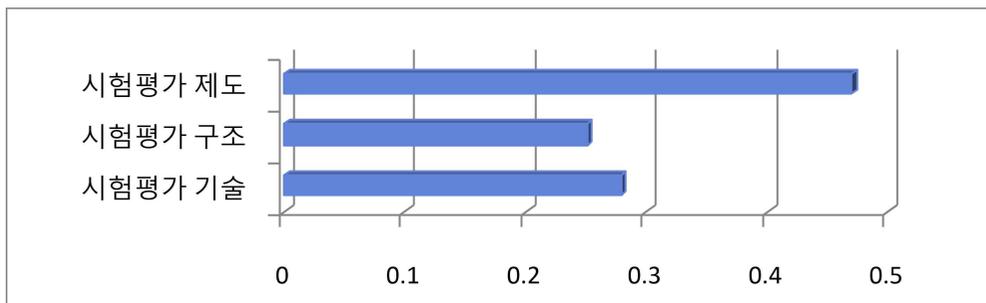


그림 31. 대분류(계층 1) 항목의 중요도 및 평가 우선순위

Figure 31. The importance and evaluation priority of the major classification (tier 1) items

첨단 무기체계 시험평가 개선방안의 상대적 중요도 및 평가 우선순위 결과를 분석해 보면, 선행연구와 전력분야 전문가의 기초조사 설문에서도 제시된 바와 같이 첨단 무기체계를 현행 국방획득 절차를 적용하는 것은 가능하나 신속성과 효율성이 결여되어 관련 법령 및 훈령을 포함한 제도적 개선이 필요하다. AI 기반 무기체계를 비롯하여 로봇·무인체계 등 유·무인 복합전투체계에 대한 시험평가를 수행하기 위해서는 현 시험평가 제도하에서 제한적으로 시험평가를 수행하는 것은 가능하다. 하지만, 제4장에서 제시한 바와 같이 AI를 기반으로 하는 무기체계는 일반 무기체계와 시험평가 절차를 동일하게 수행하는데 제한이 있어 가장 우선적으로 제도적인 기반을 구축해야 한다는 것으로 분석되었다. 다음으로 시험평가 기술 분야가 상대적으로 우선순위가 높게 평가된 것은 미래전은 5차원의 전장 공간으로 확장되고 신개념 첨단 과학기술이 적용된 드론과 로봇, 무인차량 등 유·무인 복합전투체계가 초지능·초연결 네트워크로 연결된 무기체계 시험평가 방법 및 절차 등 시험평가 기술을 개선해야 하는 점을 고려하여 중요성이 높게 평가된 것으로 해석된다. 시험평가 구조적 분야의 중요도가 상대적으로 낮게 평가된 것은 하위 평가항목인 시험평가 조직 및 전문인력 양성이 높게 평가되었음에도 불구하고 국제적 기술협력 및 교류 측면에서 중요도를 상대적으로 가장 낮게 평가되었다. 이는 시험평가 제도와 시험평가 기술 분야보다 각 세부평가 항목과 비교할 때 무기체계 시험평가 개선에 있어 시험평가 조직을 보강하고 인프라를 구축하는 것은 군 구조 개편과 국방재원 확보 등 제한사항들이 예상되어 상대적으로 낮게 평가하였다.

나. 중분류(계층 2) 항목에 대한 평가결과

첨단 무기체계 시험평가 개선을 위한 시험평가 제도 분야의 상대적 중요도를 평가한 결과, 시험평가 결과판정 개선(0.3466) > AI·무인·자율주행 등 관련

법령 제정(0.2869) > 신속획득방안 개선/확대(0.2357) > 안전기준 정립(0.1308) 순으로 우선순위가 평가되었으며, 세부 결과는 표 25 및 그림 32와 같다.

표 25. 시험평가 제도(계층 2)의 상대적 중요도 및 평가 우선순위
Table 25. Relative importance and evaluation priority of the T&E system (Tier 2)

계층 1(대분류)		계층 2(중분류)	중요도	우선순위
시험평가 제도	1.0000	신속획득방안 개선 / 확대	0.2357	3
		시험평가 결과판정 개선	0.3466	1
		AI·무인·자율주행 등 관련 법령 제정	0.2869	2
		안전기준 정립	0.1308	4

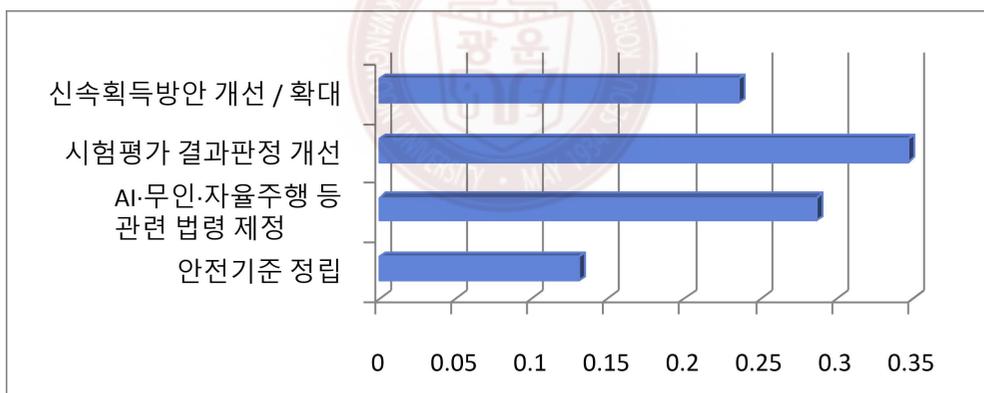


그림 32. 시험평가 제도(계층 2)의 상대적 중요도 및 평가 우선순위
Figure 32. Relative importance and evaluation priority of the T&E system (Tier 2)

무기체계 시험평가 제도 분야의 상대적 중요도를 분석해 보면, 다수의 전문가들은 현 국방획득 제도하에서 무기체계의 적기 전력화를 저해하는 요인으로 시험평가를 들고 있다. 일부 연구개발주관기관의 관계자들은 운용시험평가 항

목 중 ‘군 운용 적합성’의 세부 시험 항목에 대해 ‘작전운용성능(ROC)’ 보다 기준을 충족하기 힘들어 신뢰할 수 없다는 의견도 있다. 하지만, 이와 같은 사례는 시험평가 계획수립 간에 합참 통합시험평가팀(통시팀) 회의를 통해 관련 기관의 협의에 의해 합의된 사항으로 일부 기술 수준이 부족하여 기준을 충족하지 못한 경우라 할 수 있다. 따라서, ROC 이외의 기타 시험 항목에 대해 기준을 미충족했을 경우 조건부 충족 등 ‘시험평가 결과판정’과 관련하여 제도적으로 융통성이 부여되어야 한다는 것이다. 첨단 무기체계를 전력화하기 위해서 반드시 실시해야 하는 시험평가의 결과판정이 4차 산업혁명의 빠른 기술변화를 반영하지 못하고 전력화가 지연되거나 중단되는 사례를 막기 위해서 제도적으로 가장 중요하다고 판단한 것으로 해석된다.

다음은 AI·무인·자율주행 등 관련 법령 제정의 중요도가 높게 평가된 것은 AI와 살상 로봇, 무인 시스템에 대해 인간이 통제하지 않는 무기체계 혹은 기체에 자율성을 부여하거나 인정할 수 있는지, 윤리적으로나 법률적으로 제도적 정립이 선행되어야 한다는 점이 반영된 것으로 해석된다. 신속획득 방안 개선은 도입 취지에 맞게 민간의 첨단 과학기술이 적용된 제품이 신속하게 도입될 수 있도록 하고, 안전기준은 무기체계 개발에 있어 전력화 이후 사용자의 입장에서 안전이 최우선적으로 반영되어야 하는 점에서 중요성이 높게 평가된 것으로 판단된다.

첨단 무기체계 시험평가 개선을 위한 시험평가 구조 분야의 상대적 중요도를 평가한 결과, 시험평가 조직 및 전문인력 양성(0.5920) > 시험평가 인프라 구축(0.2687) > 민·군·국제 기술협력 및 교류(0.1394) 순으로 우선순위가 평가되었으며, 세부 결과는 표 26 및 그림 33과 같다.

표 26. 시험평가 구조(계층 2)의 상대적 중요도 및 평가 우선순위

Table 26. Relative importance and evaluation priority of T&E structure (Tier 2)

계층 1(대분류)		계층 2(중분류)	중요도	우선순위
시험평가 구조	1.0000	시험평가 조직 및 전문인력 양성	0.5920	1
		민·군·국제 기술협력/교류 강화	0.1394	3
		시험평가 인프라 구축	0.2687	2

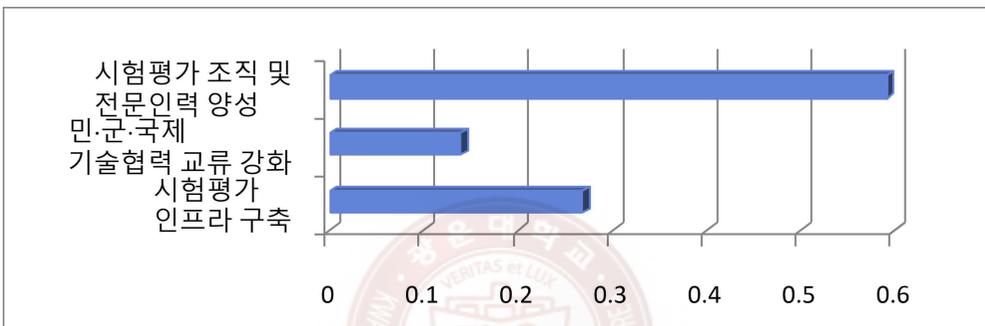


그림 33. 시험평가 구조(계층 2)의 상대적 중요도 및 평가 우선순위

Figure 33. Relative importance and evaluation priority of T&E structure (Tier 2)

무기체계 시험평가 구조 분야의 상대적 중요도를 분석해 보면, 전문가들은 4차 산업혁명의 신기술이 적용된 AI를 기반으로 하는 무기체계 즉, 로봇과 드론, 무인체계에 대해 기본적인 전문지식을 갖고 있는 전문인력의 양성이 필요하고, 육군 시험단의 무기체계 분류에 의해 편성된 기존 조직을 미래전 대비 첨단 무기체계를 시험 평가할 수 있는 조직편성과 전문가를 보강하는 것으로 국방개혁 4.0과 Army TIGER 4.0을 원활하게 추진할 수 있다는 점이 상대적으로 중요도를 높게 평가하였다. 시험평가 인프라 구축과 민·관·국제 기술협력 교류 강화는 앞서 살펴보았듯이 군사용 인공지능(AI)을 위한 데이터, 알고리즘, 플랫폼, 인재 등 인프라 구축과 신기술이 적용된 무기체계 시험평가를 위해서는 민간·국제적으로 기술교류 협력 강화가 필요하나, 전문인력양성이나 인

프라 구축에 비해 상대적으로 중요도를 낮게 평가하였다. 이는 전문인력과 인 프라가 구축되면 자연스럽게 민·군·국제기술 협력 교류가 활성화된다는 전문가들의 의견이 반영되었다.

첨단 무기체계 시험평가 개선을 위한 시험평가 기술 분야의 상대적 중요도를 평가한 결과, 과학적 시험평가체계 정립(0.5125) > 네트워크 통합시험체계 구축(0.1854) > 시험평가 기술지원(0.1651) > 한국형 시험표준 정립(0.1370) 순으로 우선순위가 평가되었으며, 세부 결과는 표 27 및 그림 34와 같다.

표 27. 시험평가 기술(계층 2)의 상대적 중요도 및 평가 우선순위
Table 27. Relative importance and evaluation priority of T&E techniques (Tier 2)

계층 1(대분류)		계층 2(중분류)	중요도	우선순위
시험평가 기술	1.0000	과학적 시험평가체계 정립	0.5125	1
		네트워크 통합시험체계 구축	0.1854	2
		시험평가 기술지원	0.1651	3
		한국형 시험표준 정립	0.1370	4

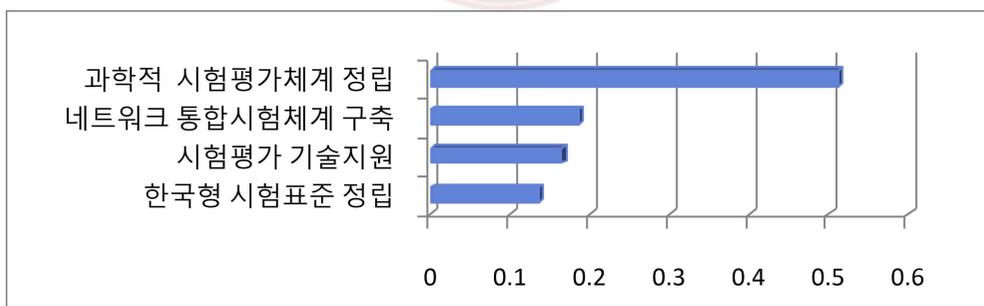


그림 34. 시험평가 기술(계층 2)의 상대적 중요도 및 평가 우선순위
Figure 34. Relative importance and evaluation priority of T&E techniques (Tier 2)

무기체계 시험평가 기술 분야의 상대적 중요도를 분석해 보면, 과학적 시험평가체계 정립은 현행 시험평가 수행체계의 제한사항과 시험평가 역량에 대한

신뢰의 문제와 더불어 인공지능(AI)을 기반으로 로봇, 드론, 자율주행 무기체계에 대한 시험평가 수행의 효율성과 신뢰성 차원에서 상대적으로 중요도를 높게 평가받았다. 네트워크 통합시험 체계는 앞에서 서술하였듯이 육군이 Army TIGER 4.0 추진과 AI를 기반으로 초연결·초지능화된 네트워크를 구축하기 위해서는 반드시 네트워크 통합시험체계 구축이 필요하여 상대적 중요도를 높게 평가받았다. 시험평가 기술지원과 한국형 시험표준 정립은 과학적 시험평가를 활성화할 수 있는 기반으로 첨단 무기체계 시험평가를 위한 기술지원 등 시스템 구축이 필요한 것으로 평가된다.

이상으로 국방 과학기술의 발전과 무기체계 발전에 따른 첨단 무기체계 시험평가 개선을 위한 시험평가 제도·구조·기술 분야에 대한 중요도와 세부 평가 요소에 대한 우선순위를 살펴보았다. 전력분야 전문가 그룹을 대상으로 AHP를 적용한 중요도와 우선순위를 종합하여 평가한 결과, 표 28 및 그림 35와 같다.

표 28. 계층구조별 중요도 및 평가 우선순위(종합)

Table 28. Importance and evaluation priority by hierarchy (Comprehensive)

계층 1(대분류)	계 2 (중분류)		(가중치)	중요도(종합)	우선순위	
시험평가 제도 (0.4700)	1	신속획득방안 개선 / 확대	(0.2357)	3	0.1108	5
		시험평가 결과판정 개선	(0.3466)	1	0.1629	1
		AI·무인·자율주행 등 관련 법령 제정	(0.2869)	2	0.1348	4
		안전기준 정립	(0.1308)	4	0.0615	7
시험평가 구조 (0.2510)	3	시험평가 조직 및 전문인력 양성	(0.5920)	1	0.1486	2
		민·군·국제 기술협력 / 교류 강화	(0.1394)	3	0.0350	11
		시험평가 인프라 구축	(0.2686)	2	0.0674	6
시험평가 기술 (0.2790)	2	과학적 시험평가체계 정립	(0.5125)	1	0.1430	3
		네트워크 통합시험체계 구축	(0.1854)	2	0.0517	8
		시험평가 기술지원	(0.1651)	3	0.0461	9
		한국형 시험표준 정립	(0.1370)	4	0.0382	10

* 일관성 비율(CR) < 0.1

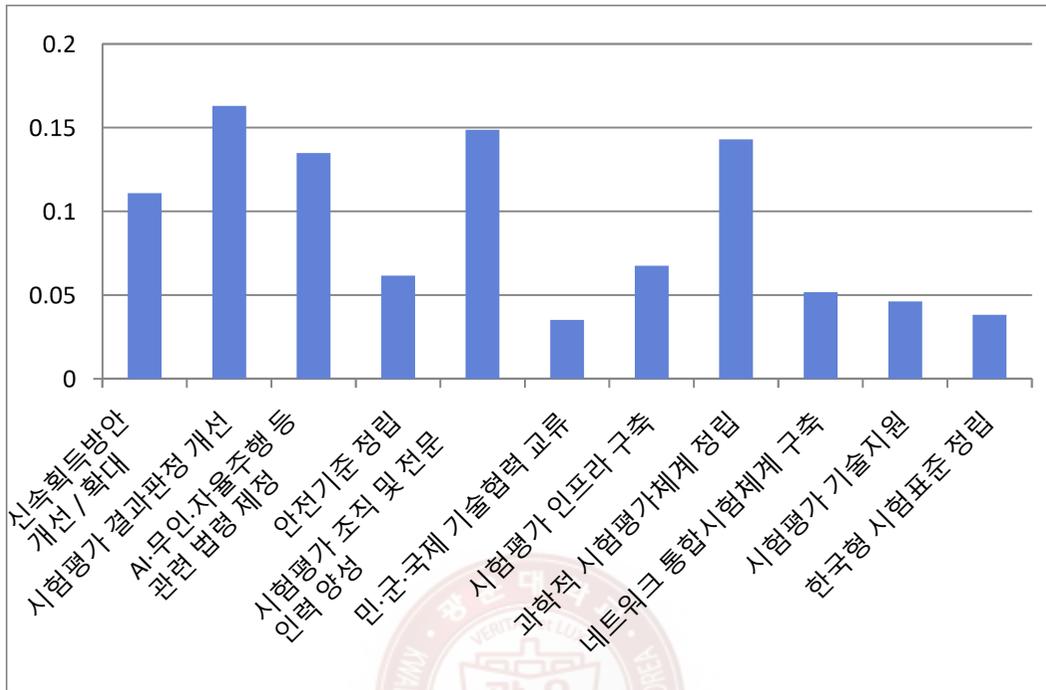


그림 35. 계층구조별 중요도 및 평가 우선순위(종합)

Figure 35. Importance and evaluation priority by hierarchy (Comprehensive)

제4절 소결론

본 장에서는 인공지능(AI) 기반 무기체계에 대한 시험평가 수행체계 개선방안을 도출하기 위해, 무기체계 개발 및 시험평가 환경변화와 무기체계 시험평가 사례분석을 통해 시험평가 개선 소요를 도출하였다. 이 결과와 국방 전력 분야 전문가를 대상으로 기초조사 설문을 실시하여 첨단 무기체계 시험평가 개선방안에 대한 시험평가 제도·구조·기술 분야로 분류하고 평가항목을 조정 및 통합하는 과정을 통해 첨단 무기체계 시험평가 개선방안에 대한 계층구조 및 평가항목을 3개 분야 11개 세부 평가항목으로 그룹화하였다. 또한 확정된 세부 평가항목에 대한 상대적 중요도 및 우선순위를 분석하기 위해 국방 전력 분야 전문가 46명을 대상으로 AHP 분석기법에 의한 설문조사를 실시하였다.

설문 결과에 대해서는 AHP 기법 전문분석 소프트웨어인 ‘메이크잇(Make It)’을 이용하여 일관성 비율을 검증하고, 상대적 중요도와 우선순위를 평가하고 분석하였다.

첨단 무기체계 시험평가 개선방안에 대한 시험평가 제도·구조·기술 분야(계층 1)에 대한 상대적 중요도를 분석한 결과, 시험평가 분야별 중요도 측면에서 시험평가 제도 > 시험평가 기술 > 시험평가 구조 순으로 시험평가 제도 개선이 높게 나타났다. 이러한 결과는 AI 기반 무기체계를 비롯하여 로봇·무인체계 등 유·무인 복합전투체계에 대한 시험평가를 수행하기 위해서는 현 시험평가 제도하에서 제한적으로 시험평가를 수행하는 것은 가능하나 AI를 기반으로 하는 무기체계는 일반 무기체계와 동일하게 시험평가 수행 절차를 준수하는데 제한사항이 있어 가장 먼저 제도적 개선이 필요한 분야라는 것으로 해석된다.

첨단 무기체계 시험평가 개선방안에 대한 11개 세부 평가항목의 우선순위를 종합 분석한 결과, 시험평가 결과판정 개선 > 시험평가 조직 및 전문인력 양

성 > 과학적 시험평가체계 정립 > AI·무인·자율주행 등 관련 법령 제정 > 신속획득방안 개선 및 확대 > 시험평가 인프라 구축 순으로 높게 평가되었다. 반면, 상대적으로 낮은 우선순위는 민·군·국제 기술협력 및 교류, 한국형 시험 표준 정립, 시험평가 기술지원, 안전기준 정립 순이었다. 이와 같이 세부 평가 항목의 우선순위가 높은 평가항목과 낮은 평가항목을 비교했을 때 약 3.5~4.7 배 정도의 큰 차이를 보였다. 이는 중요도가 상대적으로 높게 평가된 시험평가 결과판정이나 시험평가 조직 및 전문인력 양성에 대해 우선적으로 개선함으로써 4차 산업혁명의 시대 첨단 무기체계 시험평가를 원활하게 수행할 수 있는 기반을 구축할 수 있는 것으로 분석되었다.



제5장 인공지능(AI) 기반 무기체계 시험평가 개선방안

제1절 AI 기반 무기체계의 신뢰성 및 시험평가 환경분석

1. AI 무기체계 신뢰성 및 운용환경의 제한사항

인공지능(AI) 기술은 우리의 삶과 밀접하게 적용되고 지속적으로 발전하고 있다. 또한, AI 기술은 국방 무기체계에 적용하여 각종 위기 상황을 인간보다 빠르게 인지하고 대응할 수 있는 인간의 조력자이자 동반자가 될 기술이다. 하지만 AI를 활용하는 과정에서 AI의 역기능과 충분한 시험평가 부족으로 사회적·윤리적으로 위험(Risk)과 오류로 인해 많은 부작용이 발생하고 있다.

AI 기술은 소프트웨어(SW)로 구현되는 의사결정 시스템으로 대량의 데이터를 학습해서 분류나 예측형 알고리즘으로 구현된다. 알고리즘은 학습된 데이터를 바탕으로 형성된 수리적인 모형으로, 다양한 요인에 의해 영향을 받는다. 예를 들어, 편향된 데이터로 학습한 알고리즘은 불공정한 의사결정을 하거나 권리를 침해할 수 있는 오류가 발생할 수도 있다.¹²⁴⁾

대표적인 사례를 보면, 2020년 10월, 한 영국 매체에서 프랑스 헬스케어 기업인 Nabla가 개발한 인공지능 정신과 상담 챗봇을 시범 운영한 결과, 모의 환자에게 자살 충동을 실행에 옮길 것을 권유하는 사례가 있었다.¹²⁵⁾ 2018년 미국의 버락 오바마 전 대통령이 당시 미국 트럼프 대통령을 욕하는 것처럼 딥러닝 방식으로 얼굴과 목소리, 손짓 등 행동을 합성한 가짜뉴스 및 영상을 제작해 물의를 일으킨 사례도 있었다.¹²⁶⁾ 또한, 테슬라 자동차의 자율주행 기능이 역광·강한 빛·우천 시에는 사람과 도로를 구분하는 능력이 저하되는 문

124) https://blog.naver.com/kpfjra_/222456508992, 알고리즘 편향을 해결하기 위한 노력들, 신문과 방송, (검색일: 2023. 4. 8.)

125) 신예진, 2022 신뢰할 수 있는 인공지능 개발 안내서, 한국정보통신기술협회(TTA), 2022, p.22

126) <https://ditoday.com/> 페이크 뉴스의 위험성을 경고하는 페이크 뉴스, 2018, (검색일: 2023. 4. 8.)

제에 의해 주행 방향 맞은편 역광 속에 있는 흰색 트레일러를 하늘로 오인해서 충돌하는 사고가 있었다. 운전자가 완전자율주행(FSD, Full Self Driving)을 켜고 주행하다가 고속도로에서 갑자기 급제동 되면서 8중 추돌사고가 발생하기도 했다. 미국에서 벌어지는 자율주행 사고의 약 70%가 테슬라 차로 인해 발생했다는 통계도 있다.¹²⁷⁾



그림 36. AI 데이터 오용사례 (페이크 뉴스의 위험성 경고)¹²⁸⁾

Figure 36. AI data misuse (Warning about the dangerousness of Fake News)

이와 같은 문제를 해결하기 위해 주요 선진국과 국제기구에서는 AI 관련 가이드라인(미 DIB AI 원칙, UNESCO 윤리권고, OECD AI Principle, EU GDPR 등)을 내놓고 있다. 국제표준화기구 ISO/IEC의 인공지능 위원회 JTC1/SC42에서는 AI 관련 기술 영역의 표준화 작업을 주도하고 있다. 이러한

127) <https://www.joongang.co.kr/article/22521196#home> 테슬라사고 태양 역광 탓, 2018, (검색일: 2023. 4. 8.)

128) 중앙일보, 페이크 뉴스, 전게서, (검색일: 2023. 4. 8.)

흐름에서 과학기술정보통신부에서도 2021년 5월에 「신뢰할 수 있는 인공지능 실현전략」을 발표했다.¹²⁹⁾ 앞으로 인공지능(AI)을 기반으로 무기체계를 개발하고 시험평가를 수행할 때 신뢰할 수 있는 AI 기반 무기체계가 될 수 있도록 해야 된다.

ISO/IEC TR 24028 Overview of Trustworthiness in artificial intelligence에서 신뢰성(Trustworthiness)을 ‘검증 가능한 방식으로 이해관계자의 기대에 부응하는 능력’이라고 정의하고 있다. EC의 Ethics Guidelines for Trustworthy AI에서는 신뢰할 수 있는 인공지능(TAI, Trustworthy AI)은 관련 법규를 준수하고, 윤리적 원칙과 가치를 준수하며, 좋은 의도에도 불구하고 의도치 않게 발생할 수 있는 피해에 대해 기술적으로나 사회적으로 강건하다는 것을 의미한다고 서술했다. 이렇듯 AI 신뢰성(Trustworthiness)이란 데이터 및 모델의 편향, AI 기술에 내재된 블랙박스 성격의 위험과 한계를 해결하여 AI를 운용하는 과정에서 발생할 수 있는 부작용 방지를 위해 준수해야 하는 가치 기준이다.¹³⁰⁾ AI 신뢰성의 핵심 속성은 AI가 사람과 환경에 미칠 악영향을 완화하고 제거하는 것으로 표 29와 같이 안전성, 설명 가능성, 견고성, 공정성, 투명성과 같은 핵심 속성을 갖는다.¹³¹⁾

우리 군이 적용하는 AI 기반 무기체계가 AI의 신뢰성이 확보되지 않는다면, 심각한 물적·인적 피해를 일으킬 수 있으며, 전쟁의 승패를 좌우할 수도 있다. 하지만 군에서 적용하고자 하는 AI 기반 무기체계의 신뢰성을 확보하기에는 많은 어려움이 있다. AI가 적정 수준의 품질을 확보하기 어려운 이유는 데이터와 알고리즘, 시스템적인 측면에서 AI가 갖는 특성과 함께 운용환경이나 개발환경 측면에서 제한사항이 있다.

129) 신예진, 전게서, p.22

130) 상게서, p.22

131) 합참, 전게서, pp.9~30

표 29. AI 신뢰성의 핵심 속성 및 의미¹³²⁾

Table 29. Core attributes and implications of AI trustworthiness

핵심 속성	의 미
안전성 (Safety)	AI가 판단·예측한 결과로 체계가 동작하거나 기능할 때 사람과 환경에 악영향을 미칠 가능성이 완화 또는 제거된 정도
설명 가능성 (Explainability)	AI가 판단·예측하는 과정을 사람이 이해할 수 있는 방식으로 제시 또는 추적할 수 있는 정도
견고성 (Robustness)	AI가 외부 간섭이나 극한적인 운영 환경에서 사용자가 의도한 수준의 성능과 기능을 유지하는 정도
공평성 (Fairness)	AI가 특정 그룹에 대한 차별이나 편향을 포함한 결론을 도출하지 않는 정도
투명성 (Transparency)	AI의 동작 과정이 보편적으로 합리적인 정도

군이 AI를 운용하는 작전환경은 민간기업이나 사회에서 운용하는 환경에 있어 다른 점이 있다. 민간은 특수한 목적이나 한정된 조건에서 운용하기 위해 개발하고 운용되지만, 군의 작전환경은 변화 요소가 무수히 많으며, 불확실하고 예측할 수 없는 환경에서 100% 완벽한 AI를 기반으로 하는 무기체계 개발을 요구하고 있다. 예를 들어, 자율주행 차량을 일반 포장도로에서 운행할 경우 도로에 설치된 각종 교통 표지판과 차선, 신호등, 안전장치 등은 양호한 도로 환경에서 자율주행 차량에 입력된 데이터에 의해 제어가 가능하지만, 군 작전지역에서 자율주행을 하는 차량은 차선도 없는 비포장도로와 야지에서 각종 나무와 풀들 사이에서 움직이는 동물과 병력, 각종 장비가 혼재된 환경과 입력된 데이터를 통해 차량을 제어한다는 것은 신뢰성을 기대하기 어렵게 만든다. 따라서 AI 기반 무기체계는 운용환경 측면에서 AI가 성능을 발휘하는데 아주 난해한 약조건이라고 할 수 있다. 또한, 군이 작전하는 전투 현장은 전쟁

132) 상계서, p.9

영화의 전투처럼 예측불허의 환경으로 AI가 학습 데이터로 경험하지 못한 특이값(Outlier)¹³³⁾이 발생하고 포탄과 공용화기의 파열음, 각종 장비의 굉음으로 AI가 경험하지 못한 잡음(Noise)들은 AI 운용의 신뢰성을 보장하는데 제한을 줄 것이다.

개발환경 측면에서 데이터 보유 및 활용에 있어 민간 AI는 Google, 공공데이터 포털, 과학기술정보통신부의 AI hub 등 데이터의 생산이 다양하고 지속적으로 축적되고 있다. 또한 민감한 개인정보를 제외한 데이터의 공유가 가능하여 무제한으로 데이터를 활용할 수 있다. 하지만, 군사용 AI 데이터는 군 작전환경의 특수성으로 해당 AI 기반 무기체계가 학습하고 성능을 발휘할 수 있는 데이터가 충분하지 못하다. 특히, 경계지역(GP 및 GOP, 해안) 등 군사 보호구역 내에서 AI 기반 무기체계가 학습할 수 있는 데이터를 확보하는 것이 제한된다. 또한, 보안적인 측면에서도 기보유한 데이터를 활용하고 공유하기가 제한되는 문제점이 있다. 이러한 제한사항이 있음에도 불구하고, 군사용 AI 기반 무기체계에 대한 성능은 완벽해야 한다고 요구하고 있다.

2. AI 기반 무기체계 시험평가 사례

국방 AI 기반 무기체계 시험평가 사례는 아직까지는 진행형으로 최근 탐색개발을 통해 운용성 확인시험을 종료한 무인수색차량¹³⁴⁾이나 지뢰탐지robot¹³⁵⁾ 등이 있으나, 기존 일반 무기체계와 동일하게 시험평가를 진행하였다. 무인수색차량과 지뢰탐지robot 모두 AI와 관련된 분야는 센서를 통해 자신의 위치 및

133) 특이값(outlier): 본체에서 분리되거나 따로 분류되어있는 물건, 표본 중 다른 대상들과 확연히 구분되는 통계적 관측치, 각 분야에서 큰 성공을 거둔 탁월한 사람. [네이버, 교양영어사전]

134) 무인수색차량은 기계화부대의 효과적인 수색정찰 및 경계작전을 위한 연구개발사업으로 통제차량 1대와 무인차량 2대로 편성되어 운용된다.

135) 지뢰탐지robot은 장병을 대신해 위험한 지역의 지뢰를 제거하거나 급조폭발물 등을 탐지할 수 있다. 원격으로 운용되는 케도형 로봇으로 임무에 따라 지뢰탐지기, X-ray 투시기, 케이블 절단기 등 다양한 임무 장비를 조작 팔에 자동으로 탈부착할 수 있다.

주변 장애물을 검출하고 인식하는 부분과 임무와 목적을 달성하기 위한 의사 결정 및 실행하는 분야가 있지만 모두 실물에 의한 시험평가만 진행하였다. 또한, 구매시험평가로 진행하고 있는 중요시설 경계시스템 구축사업은 1·2차 사업을 종료하고 해당 부대에 설치되어 운용중에 있다. 3차 중요시설 경계시스템 사업은 2차 사업을 진행하면서 도출된 문제점을 보완하고 AI 기반 영상 감지 경보에 대한 개념과 시험기준을 설정하고 시험평가를 진행하고 있다.

2021년과 2022년 신속시범획득사업으로 진행한 AI 영상감시체계 시범운용으로 성능확인 사례는 표 30과 같이 2건이 있다. 이는 민간의 창의적인 신기술이 적용된 제품을 구매 및 군 시범운용 후 소요결정과 연계하여 후속 물량을 신속히 전력화하는 사업이다.

표 30. AI 기반체계 시험평가 현황
Table 30. AI-based system T&E status

구 분	운영개념	AI 관련 기술 요구사항(시험기준)
음원활용 AI 경계시스템	<ul style="list-style-type: none"> 카메라는 철책과 평행하는 방향으로 감시 음원감시기는 철책 전방의 소리를 집음(集音) AI기술로 영상·음성을 실시간 분석하여 문제 상황을 식별, 관제인원에게 경보 	<ul style="list-style-type: none"> 지능형 영상분석 SW는 딥러닝 기반으로 영상을 분석, 침입·배회 탐지 기능을 가져야 하며, 평가점수는 95점 이상 딥러닝 기반으로 음성을 분석하여, 이상 음원을 감지, 검출률은 90% 이상
레이더연동 AI 경계시스템	<ul style="list-style-type: none"> 원거리에서 침투하는 적 및 거동수상자를 식별 레이더가와 카메라간 장비연동 *레이더가 탐지한 객체를 카메라가 추적 / 확대 수집된 영상으로부터 학습데이터 생성, 객체 탐지 알고리즘을 개선할 수 있도록 실증랩 구축 	<ul style="list-style-type: none"> 수집된 영상정보를 분석하여 딥러닝 기술을 이용하여 객체 검출 검출된 객체를 카메라와 레이더로 추적하여 전방 000m 이내 객체(사람) 검출률은 95% 이상 전방000m 이내 객체(사람) 검출률은 98% 이상

AI 영상감시체계 사업은 AI 경계시스템을 구축하는 방안으로 음원을 활용한 AI 경계시스템과 레이더를 활용한 AI 경계시스템 시범운영으로 경계지역 (GOP, 해안 등)이나 경계가 필요한 중요경계지역에 설치하여 원거리에서 침투하는 적 및 거동수상자를 식별하는 체계다. 운용개념은 그림 37과 같이 경계지역에 일정한 간격으로 설치된 레이더와 카메라, 음원감시기를 고정 설치하고 장비간 연동을 통해 수집된 정보를 딥러닝 알고리즘을 이용하여 물체의 움직임과 패턴, 소리 데이터를 분석하여 이상 소리나 소음, 움직임을 탐지하는 시스템이다.

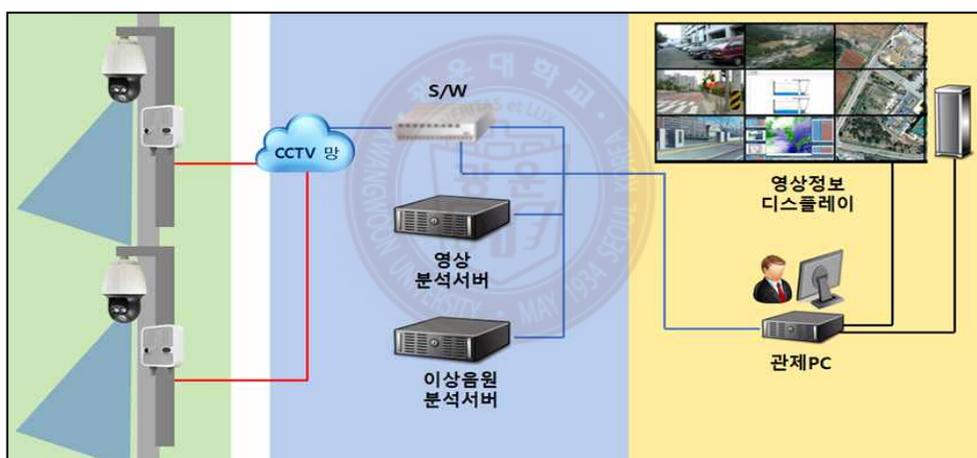


그림 37. 음원 활용 인공지능(AI) 경계시스템136)

Figure 37. AI alert system using sound source

음원 활용 AI 영상감시체계 성능 확인은 3차에 걸쳐 실시하였다. 시범운영 1차 성능 확인은 개발업체에서 납품 전에 자체 데이터를 활용하여 학습을 시킨 후 현지 지형에서 침입자를 투입하여 식별능력을 확인하였다. 이후 2차, 3차 성능 확인은 현지 지형에서 직접 획득한 데이터를 가지고 심화학습을 수행

136) <https://www.aitimes.kr/news> 인공지능신문, (검색일: 2023. 4. 7.)

한 후 침입자를 투입하여 식별 능력을 확인하였다. 학습과 성능 확인을 단계화하여 실시한 결과, 표 31과 같이 평가점수(F1 Score)와 정탐지율(%)이 비약적으로 향상되었다. 평가점수(F1 Score)는 식(1)과 같이 검출률(r)과 정밀도(p)의 조화평균이다.

표 31. 음원 활용 인공지능(AI) 경계시스템 성능확인 결과

Table 31. Result of performance validation of AI alert system using sound

구분	상황조성 (회)	정탐지137) (회)	미탐지138) (회)	오탐지139) (회)	검출률 (r)	정밀도 (p)	평가점수 (F1 Score)	정탐지율 (%)
1차	1,742	702	1,020	20	0.408	0.972	57.5	40.2
2차	2,393	1,352	967	75	0.583	0.947	72.2	56.4
3차	2,371	1,670	626	75	0.727	0.957	82.6	70.4

$$F1\ Score = 2 \times \left(\frac{r^* \times p^{**}}{r + p} \right) \dots\dots\dots(1)$$

$$*검출률(r) = \frac{\text{정탐지}}{(\text{정탐지} + \text{미탐지})} \quad **정밀도(p) = \frac{\text{정탐지}}{(\text{정탐지} + \text{오탐지})}$$

레이더 연동 AI 경계시스템은 레이더를 이용하여 움직이는 객체를 우선 식별한 후 식별된 객체를 영상 카메라와 열상 카메라로 감시 및 추적하는 체계다. 즉, 레이더에 의해 포착된 객체의 영상을 AI 시스템으로 인식하여 관제 인원에게 알려주는 시스템이다. 성능 확인은 식(2)와 같이 검출률(r)로 나타낼 수 있으며, 실제 True인 것 중에서 모델이 True라고 예측한 것을 나타낸다. 성능 확인은 2차에 걸쳐 실시되었으며, 1차 성능확인 결과 50m~150m 거리의

137) 정탐지: 실제로 존재하는 객체를 탐지하고 종류까지 올바르게 식별하는 경우
 138) 미탐지: 실제로 존재하는 객체를 탐지하지 못하는 경우
 139) 오탐지: 실제로 존재하지 않는 객체를 존재하는 것으로 탐지하는 경우

객체는 73%였다. 약 1개월 후 2차 성능 확인 결과 98%로 군에서 요구하는 95% 이상의 검출률을 충족하는 수준으로 도달하였다.

$$\text{검출률}(\text{recall, detection rate}) = \frac{\text{사람 침입 검출 횟수}}{\text{실제 사람 침입 횟수}} \dots\dots\dots(2)$$

이상과 같이 AI 영상감시체계에 대한 군사적 활용성 확인 결과, 학습과 성능 확인을 진행할수록 평가점수(F1 Score)와 검출률은 향상되는 것을 확인할 수 있었다. 다만, 학습된 시스템은 해당 작전지역 내에서만 학습 결과에 대한 성능을 낼 수 있는 것으로 다른 지역에 배치하여 사용할 경우 다시 학습과 성능 확인을 해야 한다는 것과 현 경계작전지역 내에서 고정적으로 운용하면서 학습과 성능을 확인한다면 지속적으로 향상될 수 있다.



제2절 AI 무기체계 시험평가 결과판정 개선

1. 시험평가 결과판정 절차

시험평가 결과판정은 국방전력발전업무훈령 제61조(시험평가 결과의 판정 및 보고)와 제69조(운용시험평가 결과 조치), 합참 무기체계 시험평가 업무 규정 제20조(시험평가 결과판정 및 보고)에 표 32와 같은 절차가 반영되어 있다.

표 32. 운용시험평가 결과판정 절차¹⁴⁰⁾

Table 32. Operational T&E result judgment procedure

구분	결과판정 절차 / 주관부서		비 고
1	운용시험평가 결과보고서 작성, 검토 / 결재, 통보	소요군(시평단) → 합참	OT 종료 후 1개월 이내
2	운용시험평가 결과보고서 접수 / 검토	합참 시험평가부	
3	관련기관 검토 의뢰 / 통합시험 평가팀 검토회의		
4	운용시험평가 결과판정(안) 작성 / 위원회(실무위) 검토	합참, 시험평가 위원회	시험평가위원회 심의: 총사업비 3,000억 이상
5	운용시험평가 결과판정	국방부(전력정책관)	
6	판정 결과 통보	합참 → 방사청, 소요군, 연구개발주관기관	전투용 부적합 판정시: 사업중단 또는 계속 추진 위원회(분과위) 심의

소요군(시평단) 에서는 운용시험평가 완료 후 1개월 이내에 합참(시평부)에 시험평가 결과를 제출해야 한다. 이후 합참은 시험평가 결과를 검토한 후 관련 기관에 통보하여 검토의견을 수렴하고 통시팀 검토회의를 통해 시험평가 결과에 대한 의견을 확인한다. 시험평가 결과검토 시 시험평가계획서를 기준으로 결과의 적법성, 객관성, 타당성을 검토하고 시험평가위원회(실무위원회)

140) 상계서, p.73

에서 시험평가 결과 판정(안)을 검토해서 작성한 결과를 국방부(전력정책관실)에 제출하면, 제출한 판정(안)을 근거로 시험평가 결과를 판정하여 합참 및 방사청에 통보한다. 시험평가 결과 시험평가 전(全) 항목이 기준충족일 경우에는 전력화를 위한 양산계획을 수립하고 규격화를 통해 전력화를 진행된다. 하지만 시험평가 결과 작전운용성능(ROC) 이외의 일부 항목이 기준을 미충족할 경우, 합참은 소요군의 기준미달 항목에 대한 수용 여부와 필요시 연구개발주관기관의 보완계획 등을 각종 회의체를 통해 종합적으로 검토하여 전투용 적합으로 판정(안)작성이 가능하다. 즉, ‘획득추천’을 통해 사업추진이 가능하고, ROC 이외의 시험 항목이 기준 미달된 경우에는 소요군의 기준미달 항목에 대한 수용 여부로 ‘전투용 적합’ 판정이 가능하다. 하지만 시험평가 관련 기관들의 인식의 문제와 개발과정의 완벽성을 추구하는 제도의 경직된 적용으로 표 33과 같이 전력화가 중단되거나 지연되는 사례가 발생하고 있다.¹⁴¹⁾

표 33. 작전운용성능 외 항목 미충족으로 전력화 중단 및 지연 사례
Table 33. Cases of force deployment suspension and delaying due to non-fulfillment of items other than ROC

구 분	내 용
OO형 장갑차	<ul style="list-style-type: none"> ◦ ROC / 기술적·부수적 성능에 미포함된 ‘소음기준’ (업체 추가개발 제안)이 시험기준 미충족 ◦ 후속 조치로 전력화 시기 0년 이상 지연
OOO 무전기-II	<ul style="list-style-type: none"> ◦ ROC 전(全) 항목 기준충족 ◦ 군 운용 적합성 항목(0개) 기준미달 <ul style="list-style-type: none"> ☞ 전투용 부적합 판정, 전력화 불가 ◦ 개발업체 소송제기, 대체전력이 없어 P-00K 운용
000mm사거리연장탄	<ul style="list-style-type: none"> ◦ ROC 전(全) 항목 기준충족 ◦ 기타 성능 / 군 운용 적합성 항목(0개) 기준미달 <ul style="list-style-type: none"> ☞ 전투용 부적합 판정

141) 국방부, 전게서, pp.58~59, 합참, 전게서, pp.22~33

2. AI 무기체계 시험평가 결과판정 방안

일반 무기체계의 SW와 HW 특성은 작전운용성능(ROC)과 기술적·부수적 성능(기부성능)에 반영되어 업체주관으로 개발시험평가를 통해 성능의 구현 및 범위가 예측된다. 소요군의 운용시험평가를 통해 야전 운용의 극한의 시험 환경에서 특정 조건 위주로 현장에서 실물에 의한 시험평가를 실시하고, 합참과 국방부를 통해 시험평가 결과판정을 받는다. 반면, AI 기반 무기체계는 개발과정에서 실제 운용될 환경에 대한 충분한 데이터를 확보하는데 제한이 많아 시험평가 착수 전까지 AI 성능평가를 수행하는 것은 곤란하다. 앞 절에서 설명한 AI 영상감시체계에서 보았듯이 AI 체계는 업체 자체적으로 축적한 데이터에 의한 학습을 하고 성능확인 결과, 성능인증을 받았음에도 불구하고 군의 작전환경에서 1차 성능확인 결과는 매우 낮은 평가점수를 받았다. 2·3차 학습과 성능확인 과정을 통해 1차 평가 대비 약 40% 수준의 성능이 향상되는 것을 알 수 있었다.

현행 획득 절차를 적용한 구매사업의 경우 그림 38과 같이 사업공고로부터 시험평가 완료 시까지 약 1년 이내의 기간이 적용되며, 시험평가 기간은 통상 5~6개월 정도의 기간을 부여한다. 또한 현행 획득 프로세스는 인공지능 소요 데이터 확보, 딥러닝, 성능 강화, 성능 검증 등을 시험평가 기간 내에 모두 수행하는 것은 불가능하다.

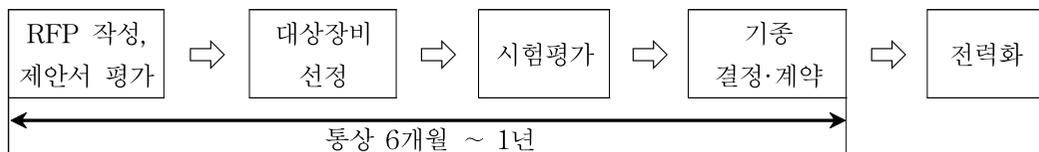


그림 38. 국내 구매사업 절차¹⁴²⁾

Figure 38. Domestic purchasing procedure

142) 국방부, 전계서, pp.62~63, 육군 시험평가단, 전계서, pp.90~112, 연구자 재정리

AI 기반 무기체계 구매사업의 시험평가는 그림 39와 같이 수행체계를 개선할 필요가 있다. 특히 구매시험평가 이전에 업체에서는 AI 모델의 기본성능 확인을 하고 구매시험평가 기간에 AI 모델 학습 능력을 검증하며, AI 모델 실환경에서 구매시험평가를 수행해야 한다. 이때 현장 데이터 기반 AI 성능평가를 실시하여 성능 확인이 되었을 때 실물에 의한 시험평가를 실시해야 한다.

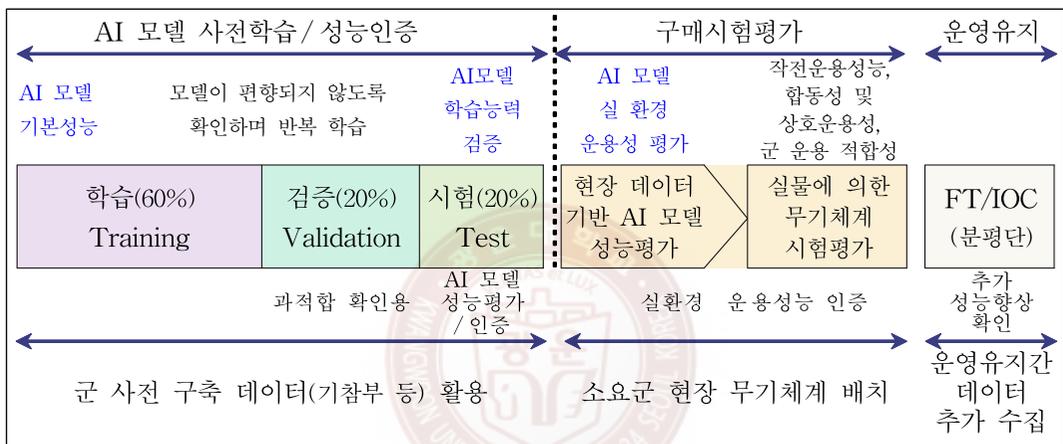


그림 39. AI 무기체계 구매사업 시험평가 개념
Figure 39. AI weapon system purchase test evaluation concept

AI 기반 무기체계 연구개발사업의 시험평가는 그림 40과 같이 개발시험평가 단계에서 AI 모델 성능평가 전문기관에 의해 성능인증을 받은 모델에 대해 현장 배치 후 구매시험평가와 동일하게 현장 데이터 기반 AI 성능평가를 실시하고 실물에 의한 무기체계 시험평가를 진행하면 된다.

AI 기반 무기체계의 재정립과 합리적인 시험평가 수행체계를 개선하기 위해서는 먼저, 시험평가 결과관정을 데이터 기반 성능평가와 실물에 의한 시험평가 결과를 종합적으로 평가하여 표 34와 같이 전투용 ‘적합’, ‘조건부 적합’, ‘부적합’이나 전투용 ‘적합’, ‘가능’, ‘개발’과 같은 시험평가 등급제¹⁴³⁾를 적용하여

결과를 판정해야 한다. 또한 반복 개발학습이 필요한 AI 적용사업 등은 양산 간 지속 개발 및 확인을 위한 별도의 절차 마련이 필요하다.

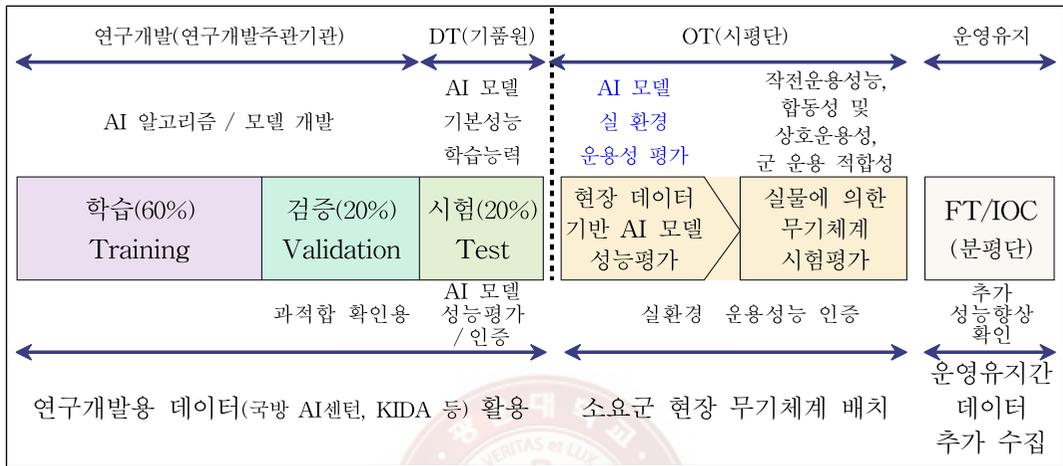


그림 40. AI 무기체계 연구개발사업 시험평가 개념
Figure 40. AI weapon system R&D project T&E concept

표 34. AI 무기체계 시험평가 판정(안)
Table 34. T&E Judgment of AI weapon system (draft)

데이터 기반 성능평가			실물에 의한 시험평가	판 정
기본성능	학습능력	최종성능		
×	×	×	×	전투용 부적합 (전투용 개발)
○	×	×	×	
○	○	×	○ / ×	조건부 전투용 적합 (전투용 가능)
○	○	○	○	전투용 적합

143) 이진호·정석재, 무기체계 연구개발 시험평가 판정 개선방안에 대한 연구, 군사발전연구 제15권 제2호, 2021, p.134

둘째, 시험평가 항목을 표 35와 같이 ROC와 같은 핵심항목과 기부성능, 업체 제안 등 일반항목으로 구분¹⁴⁴⁾하고 일반항목 중 미충족 항목은 양산과정에서 보완할 수 있도록 하여, 일정 수준 이상 충족 시 ‘전투용 적합’으로 판정할 수 있는 제도적 보완이 필요하다.

표 35. 핵심·일반항목으로 구분 분류 기준(예)

Table 35. Classification criteria for core-general items (Example)

구분	주요내용
핵심항목 (필수항목)	<ul style="list-style-type: none"> ◦ 작전수행 및 전투력 발휘에 영향을 미치는 기본성능 항목 * AI 영상감시체계: 영상분석 SW 딥러닝 기술 이용 객체 검출(95% 이상) * AI 지휘통제체계: 표적정보 자동으로 통합관리 / 항목별 분류, 복합체계 분석, 상황에 예 전시(평가점수 95점 이상) ◦ 전투원의 생존성 및 안전성에 영향을 미치는 항목 ◦ 타 무기체계의 성능 발휘에 영향을 주는 상호운용성 등 항목 ◦ 전력화지원요소에 중대한 영향을 주는 상호운용성 등 항목
일반항목 (선택항목)	<ul style="list-style-type: none"> ◦ 작전에 미치는 영향이 미미하고, 운용적인 방법으로 해결(대안제시)이 가능한 항목 ◦ 전투원의 편의성과 관련된 항목 ◦ 단순히 경제적 운영에 기여하는 항목

셋째, 군의 필요성을 고려하여 일부 성능이 미충족된 전력의 적기 전력화를 위해 ‘획득추천’ 제도를 활성화해야 한다. ‘획득추천’은 국방전력발전업무 훈령과 합참 무기체계 시험평가 업무 규정에 기 반영되어 있는 제도로 앞에서 제시한 AI 기반 무기체계 시험평가 결과판정과 같이 현 제도의 이분법적(○, ×) 판정에 유연성을 갖으면서, 소요군에서는 기준은 일부 미흡하지만 단기간 내에 보완할 수 있다고 판단될 경우 적극적으로 활용하여 전력화를 지연하거나

144) 육군 시험평가단, 시험평가 제도개선 심층 토론회 안건 검토의견, 2023, pp.2~3

중단되는 것을 막아야 한다. 이와 관련하여 국방전력발전업무 훈령과 합참 무기체계 시험평가 업무 규정에 이미 반영된 사항으로 소요군에서 적극적으로 활용할 수 있도록 해야 한다.

넷째, 시험평가도 무기체계 개발의 일부 과정으로 인식하고, 양산과정에서 발생하는 일부 문제점에 대해서는 부실 시험평가로 판단하는 인식을 탈피할 수 있는 제도적용의 유연성이 필요하다. 더불어 각종 협의체의 판단 권한을 강화하고 협의체 판단의 면책 부여제도 등을 도입하여 시험평가 간소화에 대한 부담을 경감해야 한다. 이러한 제도가 정립되지 않는다면 시험평가간 쟁점 사항이 발생할 경우 각 기관별 책임에 대한 우려 등으로 경직된 결과관정을 내릴 수밖에 없다.



제3절 시험평가 조직 및 전문인력 양성

우리 군의 시험평가는 미국 등 선진 군의 시험평가 방법 등을 벤치마킹하여 많은 발전을 해왔으나, 아직까지 시험평가 인프라(시험환경 및 인력 등)가 부족하여 일부 제한적으로 적용하고 있다. 우리 군에서는 아직까지 무기체계 시험평가(T&E)에서 시험(Test)과 평가(Evaluation)의 개념을 혼용하여 사용하고, 시험설계와 시험 결과에 대한 분석 및 평가 방법이 체계적으로 정립되지 못하고 있다.

미래의 첨단 무기체계는 기존의 일반적인 무기체계가 구비하고 있는 기동성, 화력, 생존성 등 특성 이외 4차 산업혁명 시대의 핵심기술을 적용한 AI 기술을 구비한 무인체계로 발전하고 있다. 제4장에서 제시한 바와 같이 미국을 비롯한 주요 선진국들은 AI 기반 군사력을 대폭 강화하고 있다. 신개념 과학기술이 적용된 AI 기반 첨단 무기체계를 시험평가하기 위한 전문인력의 필요성은 증가하고 있지만, 시험평가를 수행할 수 있는 전문인력 양성과 조직편성 등은 국방혁신을 위한 우선순위에 반영되지 못하고 있다. 육군 시험평가단¹⁴⁵⁾의 현재 편성은 8개과 1개실로 그림 41과 같이 일반 무기체계를 시험평가하기 위한 조직으로 AI, 드론, 자율주행 등 첨단 과학기술이 적용된 시험과 평가를 분석할 수 있는 조직으로 미흡하다. 또한, 현 시험평가단은 그림 42와 같이 인원 구성은 장교(47.1%) > 공(군)무원(30.6%) > 준사관(14%) 순으로 현역(장교, 준사관)이 61.5%로 편성되어 있다. 시험평가 근무경력 측면에서 보면 1~3년 미만(64.9%) > 3~5년 미만(28.8%) > 5~10년 미만(3.6%) > 10년 이상(2.7%)로 다수의 인원이 3년 미만의 근무경력을 갖고 있는 것에 비해 5년 이상의 시험평가 경력자는 6.3%에 불과하다. 학위취득 측면에서는 학사(47.5%)

145) 육군 시험평가단은 1981년 5월 1일부로 교육사 전투발전부 시험평가처로 창설되었으며, 1995년 5월에 시험평가단으로 독립된 이후, 교육사에서 육군본부로 예속변경(2006. 4. 1.)되어 현 계통대에 위치하고 있다.

> 석사(40.7%) > 준학사(9.3%) > 박사(2.5%)로 석·박사 학위를 취득한 인원은 43.2%를 보이고 있다.

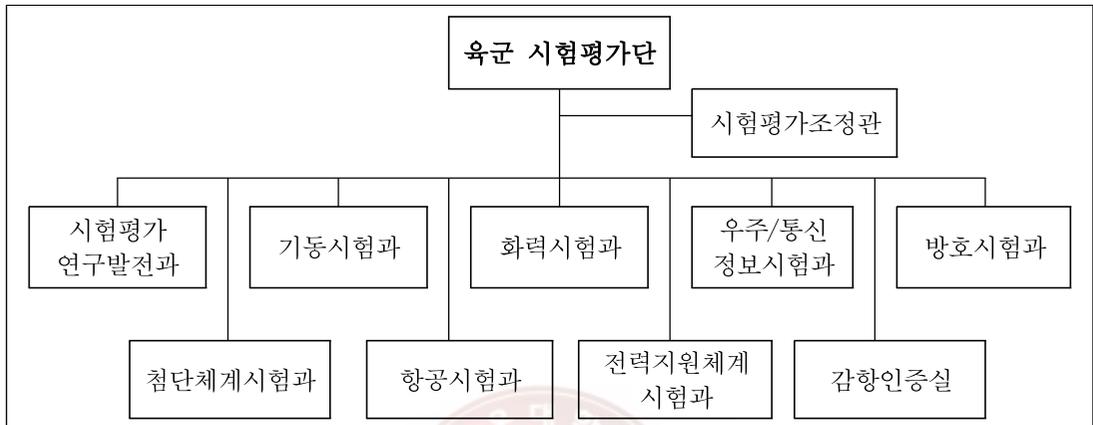


그림 41. 육군 시험평가단 편성

Figure 41. Organized ROK army T&E command



그림 42. 육군 시험평가단 인적 구성

Figure 42. Personnel organization of ROK army T&E command

이와 같은 인적 구성을 분석해 보면, 시험평가단은 영관장교 위주로 편성되어 있어 장교 인사 운영방침에 따라 시험평가단에 최장 근무할 수 있는 기간은 3년

미만이다. 시험평가 업무를 체계적으로 준비하고 수행하는데 제한이 많아 사업관리 측면에서 시험평가를 준비하는 과정에서 장교 전·출입에 따라 업무담당자가 수시로 변경될 수밖에 없는 상황이다. 또한, 다수에 인원들이 석사 및 학사학위를 취득하고 있으나, 전력업무(시험평가)와는 관련이 없는 비전공자다. 특히, 현역 장교의 경우 석사학위를 58%나 취득하고 있으나, 시험평가단에 근무하기 전까지 전력업무를 경험한 인원은 극소수로 야전에서 지휘관 및 참모직을 수행한 인원이 전입 즉시 무기체계 시험평가를 수행한다는 것은 매우 어려운 일이다.

4차 산업혁명의 시대에 인공지능, 빅데이터, 사물인터넷, 로봇, 3D 프린팅 등 5가지의 ICT 기반 기술과 인공지능을 구비한 무인체계, 드론봇 전투체계 등 첨단 무기체계 성능에 대해 시험평가를 수행하기 위해서는 전문성이 확보되어야 한다. AI를 기반으로 하는 무기체계를 포함하여 드론, 로봇, 자율주행 등 유·무인 복합무기체계를 시험평가할 수 있도록 전문인력을 양성하고 첨단 기술 학위소지자 및 전력업무 유경험자를 우선적으로 보직시켜야 한다. 시험평가관으로 일정 자격심사나 자격시험에 합격한 시험평가관이나 시험평가 유경험자들은 시험평가 전문 특기를 부여하여 전문형 장교와 같은 보직 관리를 해야 한다. 그렇지 않다면 매년 영관장교들의 전·출입으로 시험평가 준비 및 실시간 시험평가 Know-How가 부족하여 전력화의 최종 Gate-Keeper로서 역할을 수행하는데 많은 어려움을 겪을 수 있다. 또한, 현 시평단은 시험 위주의 조직편성으로 시험 데이터에 대한 과학적이고 전문적인 평가분석이 제한된다. 미군의 경우 그림 43과 같이 육군시험평가사령부(ATEC)에 7개 평가부서를 두고 100여 명의 공학, 통계, 수학 전문가로 구성된 육군평가센터(AEC)를 편성하여 시험부서와 평가부서를 분리 운용함으로써 평가의 독립성과 전문성을 보장하고 있다.

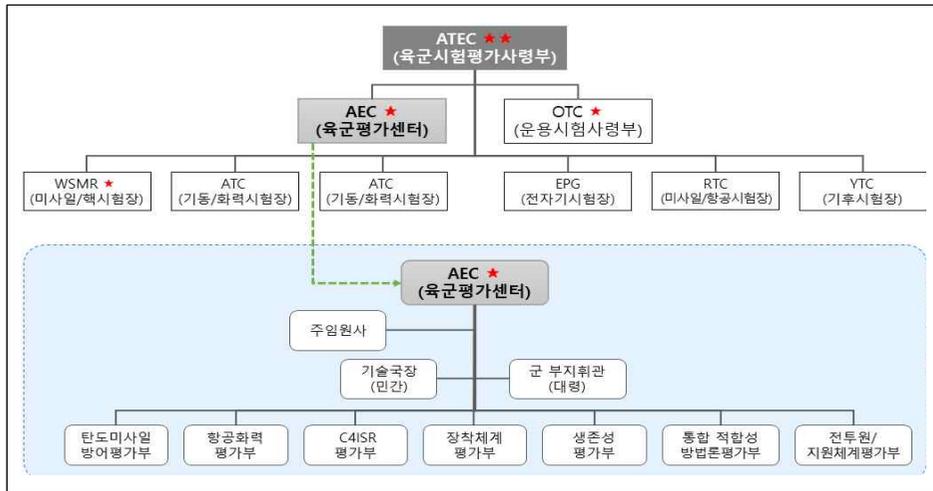


그림 43. 미 육군 시험평가사령부(ATEC), 육군평가센터(AEC) 편성¹⁴⁶⁾

Figure 43. Organization of the U.S. army T&E command (ATEC) and the army evaluation center (AEC)

이처럼 소요군의 시험평가 조직편성을 ‘시험’과 ‘평가’ 기능을 분리하고, 평가 조직을 보강해야 한다. 또한, 시험계획 수립시 과학적인 실험계획법을 수립할 수 있는 전문인력과 시험평가 관련 직무연수 및 관계기관(산업기술시험원, 항공우주연구소 등) 실무 위탁 교육과 견학을 확대하여 전문성을 강화해야 한다.

146) <https://www.army.mil/atec>, (검색일: 2022. 6. 5.)

제4절 AI 무기체계의 과학적 시험평가(신뢰성) 정립

과학적 시험평가는 시험평가의 본질인 연구개발을 통해 제작된 시제품(Prototype)에 대한 시험과 평가를 거쳐 양산된 무기체계 및 전력지원체계의 전력발휘 안정성을 보장하는 활동을 합리적이고 효율적으로 수행하기 위한 활동이다. 시제품을 대상으로 양산품의 품질을 보증해야 하는 시험평가의 현실적인 제약에도 불구하고 다양한 과학적 시험평가 방법을 적용하여 한정된 자원을 효율적으로 사용하면서 통계적인 신뢰성을 향상하기 위한 노력이 필요하다. ‘과학적’은 사전적으로 ‘과학의 바탕에서 본 정확성이나 타당성이 있는 것’으로 정의하고 있다.¹⁴⁷⁾ ‘과학적’에 대한 사전적 정의에 시험평가의 신뢰성과 효율성의 의미를 복합적으로 고찰하면 다음과 같다. ‘과학적 시험평가’를 ‘시험평가 자원을 효율적으로 사용하면서 시험평가 결과의 신뢰성을 향상하기 위해 보편적인 진리 또는 법칙을 적용하여 시험평가를 체계적으로 계획·시행·분석하는 방법’으로 정의할 수 있다. 즉, 시험평가를 과학적으로 수행했다면 시험당시와 동일한 조건 또는 유사한 조건에서 같은 시험을 반복할 때 통계적으로 동일하거나 유사한 결과를 획득할 수 있어야 한다.¹⁴⁸⁾

미군은 시험평가를 ‘위험 식별 및 감소에 활용할 수 있는 데이터, 시험을 통해 획득한 데이터 등 성능과 관련된 정보 제공을 위해 체계를 시험 및 분석하는 과정’으로 정의하고 있다.¹⁴⁹⁾ 미 육군 시험평가사령부는 시험평가 규모와 범위에 대해 논리적으로 설명이 가능하도록 지침을 제공하고 있다. 실험 또는 시험을 계획·수행·분석할 수 있도록 과학적 시험 및 분석기법(STAT) 지침을 시험평가기본계획서(TEMP)에 포함하여 하달하고 있다. 우리 군도 과학적 시

147) 국립어학원, 표준국어대사전, 2022년

148) 육군 시험평가단, 과학적 시험평가를 위한 Guide Book, 2022, p.10~16

149) The Defence Acquisition University, Test & Evaluation, 2012년, US Marine corps, Operational Test & Evaluation Manual, 2013

험 및 분석을 통해 위험을 식별하고 감소에 활용할 수 있도록 과학적 시험 및 분석기법을 TEMP에 반영하여 개발·운용시험평가계획서가 작성되도록 해야 한다. 시험평가를 수행하는 데 있어 사업관리기관, 소요군, 연구개발주관기관 등 관련 기관들이 시험평가 결과에 대해 신뢰할 수 있어야 한다. 따라서 과학적 시험설계를 통해 시험평가 항목과 항목별 시험 횟수 등 통계적으로 신뢰할 수 있도록 시험방법 및 절차를 적용해야 한다.

과학적 시험평가는 통계적으로 신뢰도(reliability)¹⁵⁰⁾와 타당도(validity)¹⁵¹⁾를 높일 수 있는 방법을 통해 구현할 수 있다. 시험 대상 체계 특성을 고려하여 시험평가 항목과 시험방법이 시험 대상 체계의 본질적인 특성으로 타당하게 시험하면서 시험 결과가 객관적으로 신뢰할 수 있으면 과학적이라고 할 수 있기 때문이다. 시험평가는 시간·예산·인원·장비 등 제한된 시험평가 자원을 효율적으로 사용해야 한다. 이를 위해서 시험계획 단계에서 시험방법 최적화를 통한 시험평가 일정과 통계적 합리성을 고려한 시제품 수량 및 적정 시험 횟수 등을 결정해야 한다. 신뢰성(Reliability)을 높이기 위해서는 시험수행 단계에서 계통오차·우연오차 등을 고려한 시험환경 조성 및 통제, 시험 오차 최소화를 위한 측정시스템을 분석해야 한다. 또한, 결과분석 단계에서 시험 데이터를 통계적인 관점에서 합리적으로 분석하여 확보하고 제공해야 한다.

과기정통부는 「신뢰할 수 있는 인공지능 실현전략」 발표(2021년)에서 ‘인공지능(AI)의 신뢰성은 인공지능이 내포한 위험과 기술적 한계를 해결하고, 활용·확산과정에서의 부작용을 방지하기 위한 가치 기준’으로 정의하고 있다. 인공지능 신뢰성(Trustworthiness)의 주요 핵심요소 및 의미는 표 36과 같이 안

150) 신뢰도(reliability)란 통계적으로는 측정 도구가 측정하고자 하는 현상을 일관성 있게 측정하는 능력으로 0~1 사이의 값을 갖는다. [이종익, 사회조사분석사 2급, 시대고시기획, 2018년]

151) 타당도(validity)란 어떤 측정으로 얻어진 데이터가 가리키는 대상이 조사자가 알고자 하던 것과 일치하는 정도, 평가하고자 하는 시험 항목의 본질을 타당한 측정 수단을 활용하여 정확히 확인하는 정도를 의미한다. [네이버 나무위키] <https://namu.wiki/w/>, (검색일: 2023. 4. 9.)

전, 설명 가능, 투명, 견고, 공정 등의 속성을 포함하고 글로벌 주요국에서도 인공지능의 신뢰성에 대한 개념을 설정하고 정책적으로 추진하고 있다.¹⁵²⁾

표 36. 인공지능 신뢰성의 주요 핵심요소 및 의미¹⁵³⁾
Table 36. Key elements and meaning of AI trustworthiness

속 성	주요 의미
안전 (Safety)	◦ AI의 판단·예측 결과로 인한 시스템 동작과 기능 수행이 사람과 환경에 악영향을 미치지 않도록 예방할 수 있는 상태
설명 가능 (Explain ability)	◦ AI의 판단·예측의 근거와 결과에 이르는 과정이 사람이 이해 가능한 방식으로 제시되거나, 문제 발생 시 결과 도출과정의 분석이 가능한 상태
투명 (Transparency)	◦ AI의 판단·예측 등 작동과정과 이를 구현하기 위한 구성 요소에 있어 이용자가 인지하고 확인·검사가 가능한 상태
견고 (Robustness)	◦ AI가 외부의 간섭 및 극한적인 운영 환경에서도 사용자가 의도한 수준의 성능 및 기능을 유지하는 상태
공정 (Fairness)	◦ AI가 데이터를 처리하는 과정에서 특정 그룹에 대한 차별이나 편향을 포함하는 결론을 도출하지 않도록 하는 가능성

AI 기반 무기체계의 과학적 시험평가를 실시하는 목적은 AI 무기체계의 신뢰성을 정립하고 향상시키는 것에 있다. AI는 추론 과정을 인간이 이해할 수 없기 때문에 AI의 역기능에 의한 부정적인 영향을 최소화할 수 있도록 AI 무기체계의 신뢰성¹⁵⁴⁾을 보장해야 한다.

AI 체계의 신뢰성 향상을 위해서는 첫째, AI 모델 개발과정에서 사용하지 않은 시험 데이터에 대한 성능을 평가하여 과소적합(Under-fitting), 적절한 적

152) 과기정통부, 신뢰할 수 있는 인공지능 실현 전략(안), 2021, pp.2~4

153) 상계서, p.3

154) AI 무기체계 신뢰성(AI Weapon System Trustworthiness): 군 요구성을 충족하면서 AI 모델의 편향 최소화, AI 모델이 의도하지 않은 방향으로 동작할 때 해당 체계의 해제 기능 보유 등 AI 기술에 내재된 위험과 한계를 해결할 수 있는 성능 및 품질을 유지하는 특성, [Jared Dunnmon 등 5명, RESPONSIBLE AI GUIDELINES IN PRACTICE, Defense Innovation Unit, 2021]

합(Appropriate-fitting), 과대적합(Over-fitting) 여부를 확인하고 실제 가능성(Generalization)을 평가해야 한다. AI 모델의 성능평가는 표 37과 같이 분류모델과 회귀모델로 구분하여 수행할 수 있다. 분류모델은 특정 숫자의 이미지가 주어졌을 때 어떤 숫자인지 판별, CCTV에 포착된 물체의 종류 판별 등과 같이 데이터를 종류별로 분류할 때 사용하는 모델이며 AI 데이터를 잘 분류했는지를 평가할 수 있는 혼동행렬(Confusion Matrix) 기반의 척도를 사용한다. 회귀모델은 교육 수준에 따른 예상 연봉, 풍속별 미사일의 예상 궤적 등 연속된 값을 예측할 때 사용하는 인공지능 모델로 AI가 예측한 값과 실제값과의 오차(Error)를 평가할 수 있는 척도로 사용한다.¹⁵⁵⁾

표 37. 분류모델과 회귀모델 성능평가의 목적 및 척도¹⁵⁶⁾

Table 37. Objectives and scales of classification and regression model performance evaluation

구분	성능평가 목적 및 척도
분류 모델	<ul style="list-style-type: none"> ◦ 모델이 분류한 결과가 제대로 분류되었는지 평가 ◦ 혼동행렬, ROC Curve, 이익도표 등 척도 활용
회귀 모델	<ul style="list-style-type: none"> ◦ 모델이 예측한 결과의 정확도 평가 ◦ 평균제곱 오차, 평균 제곱근 오차, 이익도표 등 척도 활용

AI 성능평가 척도는 미 Stanford 대학에서 발간하는 ‘2022 AI Index Report’ 표 38을 참조하여 무기체계의 요구성능 및 시험기준으로 적용이 가능하다. 국내 AI 무기체계 관련 기술발전 수준을 고려하여 적정 수준을 결정하는데 있어 국방 분야는 민간에 비해 상대적으로 데이터 사용이 제한되어 충분한 기술조사가 선행되어야 한다.

155) 이용복·전일국·김상홍, 전계서, pp.3~4

156) 합참, 전계서, pp.34~39

표 38. AI 유형별 척도와 성능 수준(2022 AI Index Report)¹⁵⁷⁾

Table 38. Scales and performance levels by AI type (2022 AI Index Report)

구 분		목 적	지표	기준	Low	High
이미지/ 영상	Computer Vision	Denoising	PSNR	2019	30	38
		Image classification	Accuracy	2022	-	90.88%
		Image Generation	FID Score	2022	7.71	-
		Object Detection	mAP	2022	-	79.5%
		Semantic Segmentation	MIou	2022	-	86.2%
소리/ 음성	Audio 인식	Audio Classification	mPA	2021	0.35	0.5
		Language Identification	Accuracy	2011		97
		Sound Event Detection	F ₁ Score	2021	38%	48%
		Voice	Similarity	2019	3.5	3.8
	Speech Recognition	Dialogue Generation	AVGF-1	2020	19	20
		Speaker Verification	EER	2021	10.4	10.9
		Speech Enhancement	PESQ-NB	2020	2.5	3.4
		Speech Recognition	WER	2022	1.4	-
	Speech Synthesis	mOs	2018	4.25	4.5	
언어/ 텍스트	자연어 처리	Language Modeling	Perplexity	2021	10	30
		Machine Translation	BLEU Score	2022	31.26	46.4
		Question Answering	F ₁ Score	2022	93.21	95.72
		Text Generation	NLL	2019	-	328
Time Series		Classification	AUC	2020	80%	87%
		Forecasting	MAE	2021	0.3	0.128
		Time Series	Accuracy	2021	95%	97%
Reasoning		Common Sense Reasoning	Accuracy	2019	75%	80%
		Decision Making		-	-	-
		Visual Commonsense Reasoning	Q→AR Score	2022	-	72%
		Visual Reasoning	AUSSES	2020	80%	85%

둘째, AI 무기체계 시험평가를 단계별로 추진해야 한다.

(1단계) AI 모델 성능평가는 개발시험평가 단계에서 AI 모델 성능평가 전문 기관에서 성능인증을 받은 모델에 대해 실 운용환경에서 시험평가를 실시하

157) Standford University Human-Centered AI, Artificial Interlligence Index Report 2022

고, 소요군(시험평가단)에서는 입회하여 성능인증을 확인해야 한다.

(2단계) 실 운용환경 시험평가는 먼저, AI 모델 성능인증을 받은 AI 무기체계를 실 운용 현장에 배치한다. 다음으로 실제 운용되는 환경에서 데이터를 수집하고 학습할 수 있는 기간을 부여한다. 이때 수집된 데이터는 8:2 비율로 학습과 평가용으로 분할하여 AI 모델 고도화 이후 현장 데이터를 활용하여 성능평가를 실시한다. 체계성능 시험평가는 AI 모델 성능평가를 제외한 작전운용성능, 합동성 및 상호운용성, 군 운용 적합성 항목에 대한 시험평가를 실시한다.

이와 같이 AI 모델의 성능인증, 현장 데이터 활용 성능평가, 체계성능 시험평가를 종합하여 AI 무기체계 성능인증을 최종 평가할 수 있는 체계를 구축해야 한다. 또한, AI 무기체계 성능인증 제도를 정립해야 한다. 한국인터넷진흥원 등과 같은 민간 인증제도를 벤치마킹하여 국방 분야 인증제도를 정립하고 소요군의 시험평가 기관에서 수행한 AI 무기체계 성능평가 결과에 대한 성능인증 업무를 수행할 수 있도록 그림 44와 같이 조직을 편성하고 국방전력발전 업무훈령 및 합참 시험평가 업무 규정 등에 반영할 수 있도록 해야 한다.

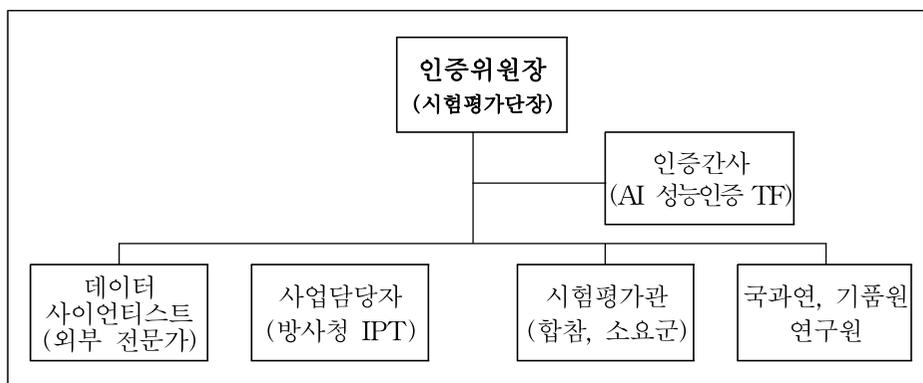


그림 44. 육군 AI 무기체계 성능인증위원회 구성(안)

Figure 44. Composition of army AI weapon system performance certification committee (draft)

셋째, 시험평가용 데이터 확보를 위한 데이터 증강(Data Augmentation)¹⁵⁸⁾이 필요하다. 인공지능 체계의 학습, 성능평가에 필수적인 것이 데이터다. 그러나, 현재 군사용 데이터가 부족한 상태에서 AI 기반체계의 시험평가를 수행하는 것이 제한된다. 따라서 이러한 AI 기반체계의 시험평가를 위하여 데이터를 확보하는 방법으로 그림 45와 같이 데이터 증강을 위한 다양한 기법¹⁵⁹⁾을 활용하여 데이터를 확보해야 한다.

넷째, AI 무기체계는 운영유지단계를 고도화하여야 한다. AI 무기체계를 운용하면서 지속적으로 데이터를 수집하여 추가학습을 통한 고도화를 시행한 후 야전운용시험(FT, Field Test) 및 전력화평가(IOC, Initial Operational Capability)를 통해 성능평가를 할 수 있도록 개선해야 한다. 또한, 그림 46과 같이 시험평가 및 운영유지 간 구축된 데이터셋¹⁶⁰⁾을 지속적으로 축적하여 기존 인공지

158) 데이터 증강(Data Augmentation)은 원본 데이터에 인위적인 변환을 적용하여 새로운 데이터를 대량으로 확보하는 기법을 말한다. 데이터 증강 기법으로 얻을 수 있는 이점으로는 두 가지 정도가 있다. 하나는 부족한 데이터를 충당할 수 있다. 다른 하나는 과적합(over-fitting)을 피할 수 있다. 이에, 왜곡이나 변형에도 인식률이 떨어지지 않도록 강건하게 모델을 학습시킬 수 있게 된다. [네이버 지식백과] 데이터 증강 (AI 용어사전)

159) 데이터 증강(Data Augmentation) 기법(예)

- ① 회전, 이동: 이미지를 회전하거나 이동시켜 새로운 이미지를 생성. 예를 들어, 원본 이미지를 90도 회전하거나 좌우로 이동시킬 수 있다.
- ② 크롭(Crop): 이미지에서 특정 부분을 잘라내어 새로운 이미지를 생성, 이를 통해 이미지의 관심 영역을 강조하거나 원하는 크기로 조정할 수 있다.
- ③ 반전(Flip): 이미지를 수평 또는 수직으로 반전시켜 새로운 이미지를 생성, 이는 이미지의 뒷면이나 반대 방향에서 찍은 사진을 모델이 인식할 수 있도록 한다.
- ④ 노이즈 추가(Noise Addition): 데이터에 노이즈를 추가하여 모델이 노이즈가 있는 환경에서도 정확하게 작동할 수 있도록 한다. 예를 들어, 이미지에 가우시안 노이즈를 추가하거나 텍스트 데이터에 임의의 단어를 삽입할 수 있다.
- ⑤ 확대/축소(Scaling): 이미지를 확대하거나 축소하여 다양한 크기의 이미지를 생성, 이를 통해 모델은 다양한 크기의 물체를 인식하고 처리할 수 있게 된다.
- ⑥ 회색조(Grayscale): 컬러 이미지를 회색조로 변환하여 다양성을 증가시킨다. 이는 색상 정보가 중요하지 않은 경우에 유용하다.
- ⑦ 샘플링(Sampling): 텍스트 데이터의 경우, 단어나 문장을 무작위로 선택하여 새로운 문장을 생성, 이는 언어 모델의 다양한 문체와 문법을 학습하도록 도와준다.

160) 데이터셋은 인공지능 모델을 훈련시키고 평가하기 위해 사용되는 데이터의 모음이다. 이 데이터셋은 다양한 형태와 주제를 가지며, 컴퓨터 비전, 자연어 처리, 음성인식 등 다양한 인공지능 작업을 수행하는 데 사용된다. 데이터셋은 일반적으로 사람이 작업한 레이블, 태그, 주석

능(AI) 무기체계 고도화 및 차후 인공지능(AI) 무기체계 관련 사업에 환류할 수 있도록 해야 한다.

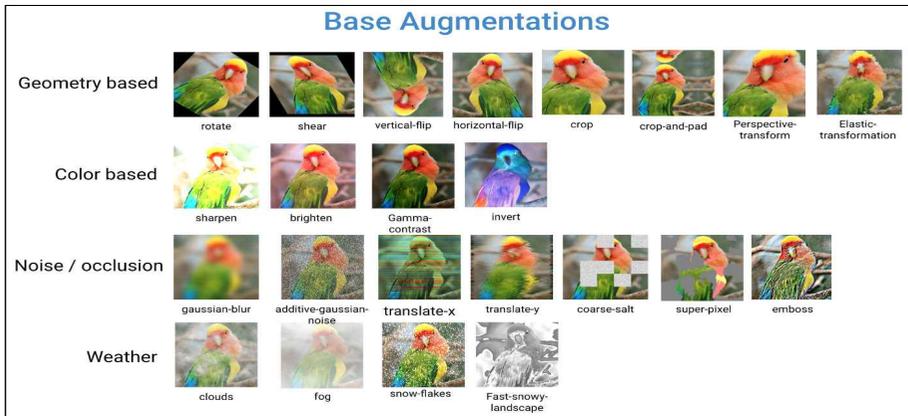


그림 45. AI 데이터 증강¹⁶¹⁾
Figure 45. AI data augmentation

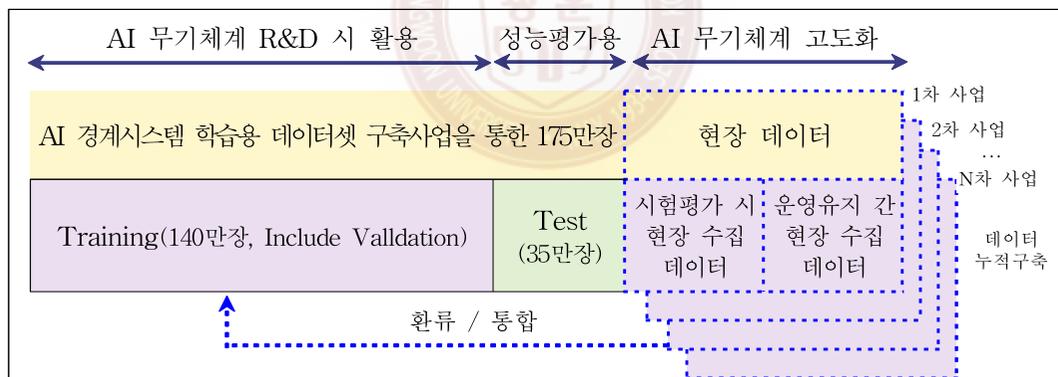


그림 46. AI 데이터셋 구축
Figure 46. Building an AI data set

등의 정보를 포함할 수 있다. 예를 들어, 이미지 분류 작업을 위한 데이터셋은 각 이미지에 대한 레이블 정보를 포함하고 있으며, 자연어 처리 작업을 위한 데이터셋은 텍스트 문장과 해당 문장의 의미나 카테고리에 대한 태그를 포함할 수 있다.

161) <https://blog.insightdatascience.com/automl-for-data-augmentation-e87cf692c366>, (검색일: 2023. 4. 30.)

다섯째, 시험평가용 M&S(Modeling and Simulation) 체계를 구축하고 활용해야 한다. 현재 군에서는 국방 M&S¹⁶²⁾ 체계를 활용하여 소요제기·결정, 획득관리, 분석평가, 교육훈련, 전투실험 등 다양한 업무를 과학적 기법으로 지원하고 있다. 실전적 훈련 보장과 제대별 분석 및 검증 능력 구비를 위한 국방 M&S 체계 전력발전 업무는 적극적으로 추진하고 있는 반면, 시험평가 분야 M&S 체계의 업무 추진은 미흡하다. 국방전력발전업무훈령 제75조(시험평가 수행 등)¹⁶³⁾에 M&S 체계를 활용하여 시험평가가 가능하도록 명시되어 있으나, 관련 정보 부족으로 실제 활용한 실적이 없다.

AI·자율주행·드론·로봇 등 첨단 과학기술을 적용한 무기체계를 비롯하여 장거리·고위력 무기체계 등 실제 운용되는 환경과 조건을 고려한 다양한 시나리오 기반의 시험평가 수행이 가능하다. 또한, 이를 통하여 무기체계 신뢰성 향상과 전력 발휘의 완전성 보장이 가능하다. 시험평가용 M&S 체계는 언제든지 다양한 임의의 환경조건을 묘사하여 추가 비용 소요 없이 반복 시험이 가능하다. AI 기반 신개념 무기체계를 적시에 도입할 수 있도록 시험평가용 M&S 체계를 구축하고 적극적으로 활용하여야 한다.

162) 국방 M&S: 실제 실행이나 체험이 제한되는 현상에 대한 분석 또는 훈련을 통해 현상에 대한 통찰을 얻어 낼 수 있으며, 투입비용 대비 높은 효과를 나타낼 수 있는 과학적 방법이다.

163) 국방전력발전업무훈령 제75조(시험평가 수행 등): ⑤항 “시험시설 등의 미비로 시험평가가 제한되는 항목에 대해서는 M&S를 활용하여 시험평가를 실시할 수 있다.”

제5절 AI 무기체계 운용시험평가 절차

운용시험평가는 개발 장비의 시제품에 대하여 소요군이 각종 작전환경 또는 이와 유사한 조건에서 작전 운용성능의 충족성, 합동성 및 상호운용성, 군 운용의 적합성, 전력화지원요소의 실용성 등을 입증하는 시험평가다. 국방 무기체계 획득 절차에서 소요제기 되는 초기 단계부터 전력화 이후 FT 및 IOC까지 전 단계에 걸쳐 시험평가의 역할이 지속적으로 이루어진다. AI 기반 무기체계 시험평가 역시 표 39와 같이 일반 무기체계 시험평가 절차에 AI 무기체계 시험평가와 관련된 절차를 추가하여 수행해야 한다.

표 39. AI 기반 무기체계 시험평가 수행간 시험평가단의 역할¹⁶⁴⁾

Table 39. The role of the T&E command during the T&E of AI-based weapon systems

국방획득 절차	시험평가의 역할
소요제기	소요군의 ICT에 참여하여 AI 시험평가 관련 사항 검토 및 의견제시
소요결정	합참 ICT에 참여하여 AI 시험평가 관련 사항을 검토하여 의견제시
선행연구	시험평가 관련 시제 수량, 시험평가 기간·예산, AI 데이터 기반 성능평가 및 실물에 의한 현장 시험평가 방안 등 관련 의견제시 * 사업관리카드 ¹⁶⁵⁾ 작성: 소요결정 단계부터 사업종료까지 관리
사업추진기본전략 수립	시험평가 기본전략 ¹⁶⁶⁾ 수립: 시험평가 방법(시험평가 일정, 시험 시제, 시험평가 소요 탄약, AI 데이터 기반 성능평가 및 실물에 의한 현장평가 방안 , 예산 등), 시험평가관 편성 등 구체화하여 반영
연구개발계획서 작성	시험평가 개략계획 ¹⁶⁷⁾ 작성 * AI 기반 무기체계 시험평가 항목 선정 등 시험평가 영향 요소를 식별하여 반영
제안요청서 작성	
제안서 접수/평가	제안서 평가위원 요청 시 참석, 시험평가 준비계획 ¹⁶⁸⁾ 수립 * AI 기반 무기체계 시험평가용 데이터 확보 및 활용계획 등
대상 장비선정	대상 장비 연구 등 운용시험평가 준비 * AI 기반 무기체계 시험평가용 데이터 수집 (데이터 형식 통일, 데이터 분리-학습용, 시험평가용)

164) 육군 시험평가단, 전게서, pp.7~112, 연구자 재정리(추가)

운용시험평가계획(안) 작성 / 통보, 운용시험평가 실시 및 결과 보고, 시험평가 결과판정 * AI 기반 무기체계 시험평가: 현장 데이터 기반 AI 모델 성능평가 + 실물에 의한 무기체계 시험평가 수행 * 시험평가 종료 후 데이터 업데이트 및 AI 모델 성능향상 과정 반복 수행	
야전운용시험 / 전력화평가	시험평가와 관련하여 참여 및 지원, 운용시험평가 후속조치 결과 확인 * AI 데이터 수집 및 성능향상 능력 확인

먼저 소요제기 및 소요결정 단계에서는 시험평가 측면에서 AI 기반 무기체계의 운용개념과 작전운용성능(ROC), 편성 등에 대해 시험평가 측면에서 유사사례 및 현 기술적으로 시험평가간 확인 및 검증을 할 수 있는지 검토하여 소요제기서에 기술된 운용개념 및 ROC 등에 대해 의견을 제시해야 한다. AI 기반 무기체계의 운용개념은 AI의 학습·신뢰 지능과 하드웨어 등 요구 능력을 주요 작전운용성능과 기술적·부수적 성능으로 구분하여 ROC에 제시되어야 한다.

선행연구로부터 사업추진기본전략 수립 단계에서는 시험평가와 직접적인 영향 요소인 시험평가 방법(시험평가 기간, 계절시험 여부, 시험시제, 시험평가 소요 탄약, 예산 등)을 사전에 도출하고, 데이터 기반 성능평가와 실물에 의한 현장 시험평가를 실시하는 방안에 대하여 사업추진기본전략에 반영될 수 있도록 선행연구기관 및 방사청과 수시로 접촉해야 한다. 이와 같은 사항에 대해

-
- 165) 사업관리카드: 소요결정 단계부터 사업종료 시까지 사업 이력을 판단하여 소요결정 시 소요결정 문서를 참고하여 운용개념, 형상 및 주요건, 사업추진 경과, 관련문서, 사업추진 현황, 현안 관리 향후 추진 등을 포함하여 작성한다.
 - 166) 시험평가 기본전략: 사업추진기본전략에 포함될 시험평가 관련 사항을 도출하기 위해 작성하며, 사업추진기본전략수립 전에 시험평가 방법(시험 일정, 시험 시제, 시험평가 소요 탄약 등) 시험평가관 편성 등을 포함하여 작성한다.
 - 167) 시험평가 계약계획: 시험평가항목 선정 등 시험평가 영향 요소를 사전에 판단하여 제안요청서 작성 전에 시험평가 장비 및 수량, 시험평가 항목(계절성 시험항목 확인), 시험평가관 편성, 소요예산 등을 포함하여 작성한다.
 - 168) 시험평가 준비계획: 운용시험평가 계획(안)의 완전성을 보장하기 위해 운용시험평가 착수 1년 전에 시험평가 준비 중점, 통합시험 참여계획, 운용시험평가 계획, 관심 및 착안사항 등을 포함하여 작성한다.

시평단에서는 시험평가 기본전략을 작성하여 보고 후 방사청 IPT에 의견을 제시하여 반영한다.

연구개발계획서 작성 및 제안요청서 작성 단계에서는 AI 기반 무기체계 시험평가 영향 요소인 시험평가 항목(계절성 시험항목 포함)과 소요 예산 등을 사전에 판단하여 제안요청서 작성 전에 반영해야 한다. 이와 같은 사항에 대해 시평단에서는 시험평가 개략계획을 작성하여 보고 후 방사청 IPT에 의견을 제시하여 반영한다.

제안서 접수 및 평가로부터 대상 장비선정 단계에서는 AI 기반 무기체계 시험평가 개략계획을 기초로 시험평가 준비계획을 수립한다. 이때 AI 기반 무기체계 시험평가를 위한 시나리오 등 시험평가 수행계획을 구체화하고, AI 시험평가용 데이터를 확보해야 한다. 데이터 확보는 AI 모델이 해당 무기체계의 목적과 최종상태를 달성할 수 있도록 학습 및 평가할 수 있는 데이터를 수집하고, 데이터 활용을 위한 데이터 형식 통일, 클래스(Class) 분류 및 레이블링(Labeling) 등 전처리(Pre-processing)와 학습용과 시험평가용 데이터로 분리하는 과정을 포함해야 한다.¹⁶⁹⁾

운용시험평가 단계에서는 현장 데이터 기반 AI 모델 성능평가와 실물에 의한 무기체계 시험평가를 일반 무기체계 시험평가와 유사한 절차에 의해 수행한다. 시험평가 종료 후에는 전력화 이전까지 AI 기반 무기체계가 운용하는 과정의 모든 데이터를 업데이트하고 성능을 향상시키기 위한 학습과 성능평가를 반복한다. AI 기반 무기체계 운용시험평가 단계별 수행 절차는 표 40과 같이 일반 무기체계 운용시험평가에 준하여 AI 기반 무기체계 시험평가를 추가하여 1단계 자료수집 및 기초연구로부터 6단계 사후검토 및 후속 조치까지 단계별로 수행해야 한다.

169) 합참, 전계서, pp.82~83

표 40. 운용시험평가 단계별 수행절차¹⁷⁰⁾

Table 40. Operational T&E stepwise procedure

구분	수행 절차	주요 수행 내용
1단계	자료수집 / 기초연구	<ul style="list-style-type: none"> ◦ 소요결정 시 ICT 참여, 사업관리카드 작성 보고 ◦ 사업추진기본전략 수립 전 시험평가 기본전략(안) 작성 보고·통보 * AI 성능평가 데이터 확보 및 활용계획, 시험평가 일정 등 검토 ◦ RFP작성 전 시험평가 계약계획 작성 및 보고 ◦ 소요결정문서, 탐색개발시행계획서 등 자료를 수집하여 체계요구조건 검토회의 참여 등 자료수집 및 기초연구, 준비계획 보고(1년 전) ◦ 운용개념, 환경구축 단계화, 안전성 평가, CDS-MT¹⁷¹⁾ 관련 토의 실시 ⇒ 계획(안) 작성시 반영
2단계	운용시험평가 계획(안) 작성	<ul style="list-style-type: none"> ◦ 개발시험 입회 및 참여 ◦ 운용시험평가 계획(안) 작성 및 보완 ◦ 시험평가계획(안) 작성 및 보완(세미나, 관련기관 검토회의) * AI기반 무기체계의 운용개념 및 시험환경 구축, 시험평가 단계화 방안 등 토의 ◦ 시험평가계획(안) 작성 및 통보: 단장·기참부장 / 합참 시험평가부
3단계	운용시험평가 실시	<ul style="list-style-type: none"> ◦ 시험준비검토(TRR) 회의(운용시험평가 수행 15일 전) * AI 기반 무기체계 준비상태 확인(데이터 등 인프라 구축상태) ◦ 운용시험평가 수행(시험일지 작성, 주간결산, 관련 기관 협조 회의 등)
4단계	운용시험평가 분석 / 평가	<ul style="list-style-type: none"> ◦ 시험평가 결과 종합, 분석 및 평가, 결과 작성 * AI 기반 무기체계 기본성능 충족성, 성능향상 수준 평가 결과 등 ◦ 시험평가 결과 세미나(조정관), 검토회의(과장)
5단계	운용시험평가 결과 보고	<ul style="list-style-type: none"> ◦ 시험평가 결과 보고: 단장, 기참부장 ◦ 시험평가 결과 통보(종료 후 30일 이내): 합참(시험평가부)
6단계	사후검토 / 후속 조치	<ul style="list-style-type: none"> ◦ 시험평가 결과 교훈 작성: 시험평가 종료 2개월 이내 ◦ 후속조치계획 작성 및 보고 ◦ 시험평가가 관련 자료 존안: 단 홈페이지 시험평가 자료관리체계 ◦ 초도양산 전 시험평가 보완 요구사항 조치 등 결과 추적 확인 ◦ 수락 검사, 야전운용시험, 전력화평가 간 참여 * AI 데이터 수집 및 성능향상 능력 확인

1단계 자료수집 및 기초연구는 소요가 결정되면 과업 식별 및 체계개발 단계에서 수행해야 할 운용시험평가를 위해 자료를 수집하고, 운용시험평가 준

170) 육군 시험평가단, 전게서, pp.7~112, 연구자 재정리(추가)

171) CDS-MT(Combat Development Support-Management Team) : 전투발전지원요소 실무조정회의

비를 위해 사업관리카드, 시험평가 기본전략, 시험평가 개략계획, 시험평가 준비계획 등을 통해 시험평가와 관련된 사항을 포함하여 작성하고 방사청 IPT에 반영될 수 있도록 의견을 제시한다. 이때 AI 성능을 평가하기 위한 데이터의 확보 및 활용계획, 시험평가 항목, 시험평가 일정 등 운용시험평가를 준비하는 과정이라 할 수 있다.

2단계 운용시험평가계획(안) 작성은 AI 기반 무기체계의 운용개념 및 시험환경 구축과 시험평가 단계화 방안 등을 토의하여 시험평가계획(안)에 반영하여 작성하고, 관련 기관 검토회의를 통해 시험평가계획(안)을 확정하여 보고 후 합참 시평부에 시험착수 60일 전에 통보한다.

3단계 운용시험평가 실시단계는 합참 시평부로부터 운용시험평가 착수 1개월 전에 접수한 운용시험평가계획서를 근거로 운용시험평가를 수행한다. 운용시험평가 착수 15일 전에 시험준비 검토회의(TRR, Test Readiness Review)를 통해 표 41과 같이 데이터 기반 성능평가 인프라 구축과 실물에 의한 현장 시험평가 준비상태 등을 확인하고 점검해서 운용시험평가 진입 가능 여부를 검토한다. 시험평가 실시는 AI 기반 무기체계의 기본성능의 요구조건충족 여부를 시험하고, AI 학습 능력과 성능향상 수준을 시험한다. 시험평가는 데이터 기반 성능평가와 현장에서 실물에 의한 시험평가를 병행한다.

4·5단계 운용시험평가 분석 및 평가, 결과 보고 단계에는 AI 기반 무기체계의 기본성능 충족 여부와 성능향상 수준 평가 결과, 기타 시험평가 수행 간 수집된 각종 자료를 종합적으로 분석하고, 시험평가 항목별 기준충족 여부를 평가한다. 또한, AI 체계의 특성을 고려 데이터의 부족에 의한 기본성능 및 성능향상체계 구축 여부 등을 분석하고 평가하여 보고 후 합참에 운용시험평가 결과를 통보한다.

표 41. 운용시험평가 시험준비 검토회의, 주요 확인사항 및 진입조건¹⁷²⁾

Table 41. The key confirmations and entry conditions of OT&E TRR

구 분	주요 확인 사항	진입조건
일반 무기체계	<ul style="list-style-type: none"> ◦ 개발(통합)시험평가 결과 ◦ 시험부대 지원 여건 ◦ 시험평가 장비/물자 준비상태 (장비 운용 시설 중 시험평가가 필요한 시설 준공 여부 포함) ◦ 시험평가 예산확보 ◦ 기타 운용시험평가 관련 협조 사항 	<p>실험실 환경에서 무기체계 기술적 요구성능 충족 (개발시험평가 기준충족)</p>
AI 기반 무기체계	<p>일반 무기체계 주요 확인 사항에 추가하여</p> <ul style="list-style-type: none"> ◦ 데이터 기반 성능평가 인프라 구축상태 ◦ 실물에 의한 현장 시험평가 준비상태 ◦ 운용환경 또는 시험평가 환경에서 수집한 데이터를 활용, AI 모델의 성능향상을 수행할 수 있는 상태 	<p>실험실 환경에서 AI 기반 무기체계 기술적 요구성능 충족 (개발시험평가 기준충족)</p>

6단계 사후검토 및 후속 조치 단계로 시험평가 결과 보고 후에는 시험평가 전(全) 과정에 대한 자체 사후검토를 실시하여 교훈을 도출한다. 이는 향후 수행하게 될 다른 무기체계에 대한 동일한 시행착오를 예방하고 새로 보직된 시험평가관에게 Know-How를 제공하는 등 시험평가의 질적인 향상에 기여할 수 있도록 해야 한다. 특히, 시험평가 실시단계에서 보완요구 사항으로 분류된 항목에 대해서는 보완 요구사항 조치 결과를 확인하고, FT에서 AI 무기체계의 특성과 데이터 수집 및 성능향상 능력을 확인해야 한다.

이상과 같이 AI 기반 무기체계를 시험평가 하기 위해서는 국방획득체계 전 단계 시험평가와 관련하여 역할을 수행해야 한다. 또한, 운용시험평가를 위해서 자료수집 및 기초연구로부터 후속조치 단계까지 일반 무기체계 시험평가 수행 절차에 추가하여 AI 무기체계의 특성을 고려하여 데이터 기반 성능평가 및 실물에 의한 현장 시험평가와 관련된 절차를 추가하여 수행해야 한다.

172) 육군 시험평가단, 전게서, pp.7~112, 연구자 재정리(추가)

제6절 소결론

본 장에서는 계층적 분석기법(AHP)을 통해 우선순위가 높게 평가된 시험평가 결과판정 개선과 시험평가 조직 및 전문인력 양성, 과학적 시험평가체계 정립과 운용시험평가 절차에 대하여 인공지능(AI) 기반 무기체계 시험평가 개선방안을 다음과 같이 제시하였다.

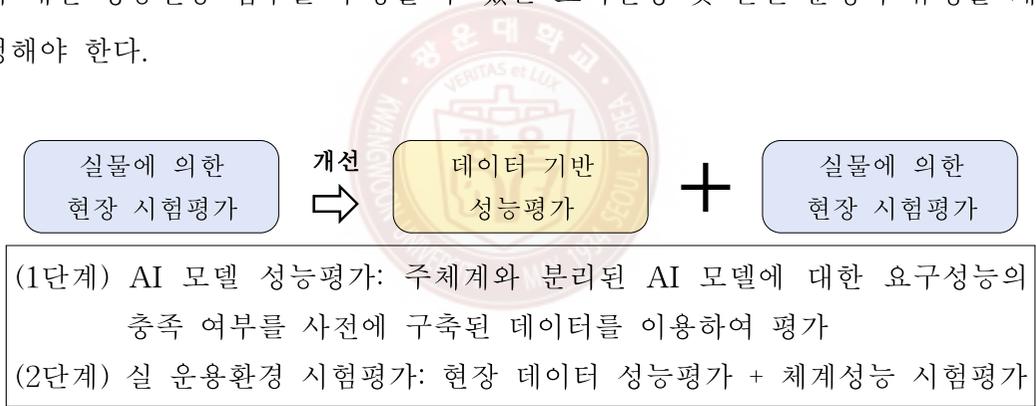
첫째, AI 기반 무기체계 시험평가 결과판정 개선방안으로 먼저, 시험평가 결과판정을 데이터 기반 성능평가와 실물에 의한 시험평가 결과를 종합적으로 평가하여 전투용 ‘적합’, ‘조건부 적합’, ‘부적합’이나 전투용 ‘적합’, ‘가능’, ‘개발’과 같은 시험평가 등급제를 적용하여 결과를 판정해야 한다. 또한 반복 개발학습이 필요한 AI 적용사업 등은 양산 간 지속 개발 및 확인을 위한 별도의 절차 마련이 필요하다. 현 시험평가 항목을 작전운용성능(ROC)과 같은 핵심(필수) 항목과 기술적·부수적 성능, 업체 제안 등 일반(선택)항목으로 구분하고 일반(선택)항목 중 미충족 항목은 양산과정에서 보완할 수 있도록 하여, 일정 수준 이상 충족 시 ‘전투용 적합’으로 판정할 수 있는 제도적 보완이 필요하다. 군의 필요성을 고려하여 일부 성능이 미충족된 전력의 적기 전력화를 위해 ‘획득추천’ 제도를 활성화해야 한다. ‘획득추천’은 국방전력발전업무 훈령과 합참 무기체계 시험평가 업무 규정에 이미 반영되어 있는 제도로 앞에서 제시한 AI 기반 무기체계 시험평가 결과판정과 같이 현 제도의 이분법적(○, ×) 판정에서 유연성을 가지면서, 소요군에서 기준에는 일부 미흡하지만 단기간 내에 보완이 가능하다고 판단될 경우 적극적으로 활용하여 전력화를 지연하거나 중단되는 것을 막아야 한다. 시험평가도 무기체계 개발의 일부 과정으로 인식하고, 양산과정에서 발생하는 일부 문제점에 대해서는 부실 시험평가로 판단하는 인식을 탈피할 수 있는 제도적용의 유연성이 필요하다. 더불어 각종 협의체의 판단 권한을 강화하고 협의체 판단의 면책 부여제도 등을 도입

하여 시험평가 간소화에 대한 부담을 경감해야 한다.

둘째, 시험평가 조직 및 전문인력 양성으로, 현 육군 시험평가단은 일반 무기체계 시험평가 조직에서 점차 첨단 무기체를 시험 평가하는 조직으로 변화하고 있다. 하지만, 인적 구성은 장교(47.1%) 및 준사관(14%) 등 현역 위주로 편성되어 있고 시험평가 근무경력은 대다수가 1~3년 미만(64.9%)으로 시험평가 업무를 체계적으로 준비하고 수행하는데 제한이 있다. 사업관리 측면에서 시험평가를 준비하고 실시하는 과정에서 장교의 전·출입으로 업무담당자가 수시로 변경될 수밖에 없는 상황이다. 또한, 다수의 인원들이 전력업무(시험평가)와는 관련이 없는 비전공자로 시험평가단에 근무하기 전까지 전력업무를 경험한 인원은 극소수로 야전에서 지휘관 및 참모직을 수행한 인원이 전입과 동시에 무기체계 시험평가를 수행한다는 것은 많은 어려움이 있다. 이를 개선하기 위해, AI를 기반으로 하는 무기체계를 포함하여 드론, 로봇, 자율주행 등 유·무인 복합무기체계를 시험 평가할 수 있도록 전문인력을 양성하고 첨단기술 학위소지자 및 전력업무 유경험자를 우선적으로 보직시켜야 한다. 시험평가관으로 일정 자격심사나 시험에 합격한 시험평가관이나 시험평가 유경험자들은 시험평가 전문 특기를 부여하여 전문형 장교와 같은 보직 관리를 해야 한다. 그렇지 않다면 매년 영관장교들의 전·출입으로 시험평가를 내실있게 준비하지 못하고, 시험평가 수행간 Know-How가 부족하여 전력화의 최종 Gate-Keeper로서 역할을 수행하는데 많은 어려움을 겪을 수 있다. 또한 시험평가 조직을 ‘시험’과 ‘평가’ 기능으로 분리하여 시험평가를 수행하고 결과에 대해 신뢰성을 확보할 수 있도록 평가조직을 보강해야 한다.

셋째, AI 기반 무기체계 과학적 시험평가 방안으로 과학적 시험 및 분석을 통해 위험을 식별하고 감소에 활용할 수 있도록 과학적 시험 및 분석기법을 TEMP에 반영하여 개발·운용시험평가계획서가 작성되도록 해야 한다. 시험평

가를 수행하는 데 있어 사업관리기관, 소요군, 연구개발주관기관 등 관련 기관들이 시험평가 결과에 대해 신뢰할 수 있어야 한다. 따라서 과학적 시험설계를 통해 시험평가 항목과 항목별 시험 횟수 등 통계적으로 신뢰할 수 있도록 시험방법 및 절차를 적용해야 한다. AI 기반 무기체계 시험평가의 과학적 시험평가는 AI 무기체계의 신뢰성을 향상시키는 것이다. AI 체계의 신뢰성 향상을 위해서는 ① AI 모델의 개발과정에서 사용하지 않은 시험 데이터에 대한 성능을 평가하여 과소적합, 적절한 적합, 과대적합 여부를 확인하여 실제 가능성을 평가해야 한다. ② AI 체계의 특성을 고려하여 데이터 기반 성능평가를 추가하여 아래와 같이 시험평가를 단계별로 추진하고 무기체계 성능평가 결과에 대한 성능인증 업무를 수행할 수 있는 조직편성 및 관련 훈령과 규정을 개정해야 한다.



③ 인공지능 체계의 학습과 성능평가에 필수적인 군사용 데이터를 확보하기 위해 데이터 증강을 통해 시험평가용 데이터의 부족을 극복해야 한다. ④ AI 무기체계는 운영유지단계를 고도화해야 한다. AI 무기체계는 전력화 이후에도 지속적으로 데이터에 의한 학습과 성능향상이 이루어진다. 따라서 전력화 이후 FT에서도 성능평가를 수행할 수 있도록 해야 한다. ⑤ 시험평가용 M&S 체계를 구축하고 활용해야 한다. M&S 체계를 활용하여 실제 운용되는 환경과 조건을 고려한 다양한 시나리오 기반의 시험평가를 수행하여 무기체계 신

뢰성을 향상 및 전력 발휘의 완전성을 보장하고, AI 기반 신개념 무기체계를 적시에 도입할 수 있도록 시험평가용 M&S 체계 구축하여 활용해야 한다.

넷째, AI 기반 무기체계 운용시험평가 절차 개선방안으로 ① 일반 무기체계 시험평가 절차에 AI 무기체계 시험평가와 관련된 절차를 추가하여 수행해야 한다. 먼저 소요제기 및 소요결정 단계에서는 시험평가 측면에서 AI 기반 무기체계의 운용개념과 작전운용성능(ROC), 편성 등에 대해 시험평가 측면에서 유사사례 및 현 기술적으로 시험평가간 확인 및 검증을 할 수 있는지 검토하여 소요제기서에 기술된 운용개념 및 ROC 등에 대해 의견을 제시해야 한다. AI 기반 무기체계의 운용개념은 AI의 학습·신뢰 지능과 하드웨어 등 요구 능력을 주요 작전운용성능과 기술적·부수적 성능으로 구분하여 ROC에 제시되어야 한다. ② 선행연구로부터 사업추진기본전략 수립 단계에서는 시험평가와 직접적인 영향 요소인 시험평가 방법(시험평가 기간, 계절 시험 여부, 시험 시제, 시험평가 소요 탄약, 예산 등)을 사전에 도출하고, 데이터 기반 성능평가와 실물에 의한 현장 시험평가를 실시하는 방안에 대하여 사업추진기본전략에 반영될 수 있도록 선행연구기관 및 방사청과 수시로 접촉해야 한다. 이와 같은 사항에 대해 시험평가단에서는 시험평가 기본전략을 작성하여 보고 후 방사청 IPT에 의견을 제시하여 반영한다.

③ 연구개발계획서 작성 및 제안요청서 작성 단계에서는 AI 기반 무기체계 시험평가 영향 요소인 시험평가 항목(계절성 시험항목 포함)과 소요 예산 등을 사전에 판단하여 제안요청서 작성 전에 반영해야 한다. 이와 같은 사항에 대해 시험평가단에서는 시험평가 개략계획을 작성하여 보고 후 방사청 IPT에 의견을 제시하여 반영한다. ④ 제안서 접수 및 평가로부터 대상 장비선정 단계에서는 AI 기반 무기체계 시험평가 개략계획을 기초로 시험평가 준비계획을 수립한다. 이때 AI 기반 무기체계 시험평가를 위한 시나리오 등 시험평가

수행계획을 구체화하고, AI 시험평가용 데이터를 확보해야 한다. 데이터를 수집하고 활용하기 위해서는 데이터의 형식을 통일시키고, 클래스(Class) 분류 및 레이블링(Labeling) 등 전처리(Pre-processing)와 학습용과 시험평가용 데이터로 분리하는 과정을 포함해야 한다. ⑤ 운용시험평가 단계에서는 현장 데이터 기반 AI 모델 성능평가와 실물에 의한 무기체계 시험평가를 일반 무기체계 시험평가와 유사한 절차에 의해 수행한다. ⑥ 시험평가 종료 후에는 전력화 이전까지 AI 기반 무기체계가 운용하는 과정의 모든 데이터를 업데이트하고 성능을 향상시키기 위한 학습과 성능평가를 반복한다.

AI 기반 무기체계 운용시험평가는 일반 무기체계 시험평가에 준하여 준비하고 AI와 관련된 분야를 추가하여 실시해야 한다. 먼저, 1단계 자료수집 및 기초연구는 소요가 결정되면 과업 식별 및 체계개발 단계에서 수행해야 할 운용 시험평가를 위해 자료를 수집한다. 운용시험평가 준비를 위해 사업관리카드, 시험평가 기본전략, 시험평가 개략계획, 시험평가 준비계획 등을 통해 시험평가와 관련된 사항을 포함하여 작성하고 방사청 IPT에 반영될 수 있도록 의견을 제시한다. 이때 AI 성능을 평가하기 위한 데이터의 확보 및 활용계획, 시험평가 항목, 시험평가 일정 등 운용시험평가를 준비하는 과정이라 할 수 있다. 2단계 운용시험평가계획(안) 작성은 AI 기반 무기체계의 운용개념 및 시험환경 구축과 시험평가 단계화 방안 등을 토의하여 시험평가계획(안)에 반영하여 작성하고, 관련 기관 검토회의를 통해 시험평가계획(안)을 확정하여 보고 후 합참 시평부에 시험착수 60일 전에 통보한다. 3단계 운용시험평가 실시단계는 합참 시험평가부로부터 운용시험평가 착수 1개월 전에 접수한 운용시험평가계획서를 근거로 운용시험평가를 수행한다. 운용시험평가 착수 15일 전에 시험준비 검토회의(TRR)를 통해 데이터 기반 성능평가 인프라 구축과 실물에 의한 현장 시험평가 준비상태 등을 확인하고 점검해서 운용시험평가 진입 가능

여부를 검토한다. 시험평가 실시는 AI 기반 무기체계의 기본성능의 요구조건 충족 여부를 시험하고, AI 학습 능력과 성능향상 수준을 시험한다. 시험평가는 데이터 기반 성능평가와 현장에서 실물에 의한 시험평가를 병행한다. 4·5단계 운용시험평가 분석 및 평가, 결과 보고 단계에는 AI 기반 무기체계의 기본성능 충족 여부와 성능향상 수준 평가 결과, 기타 시험평가 수행간 수집된 각종 자료를 종합적으로 분석하여 시험평가 항목별 기준충족 여부를 평가한다. 또한, AI 체계의 특성을 고려 데이터의 부족에 의한 기본성능 및 성능향상체계가 구축 여부 등을 분석하고 평가하여 보고 후 합참에 운용시험평가 결과를 통보한다. 6단계 사후검토 및 후속 조치 단계로 시험평가 결과 보고 후에는 시험평가 전(全) 과정에 대한 자체 사후검토를 실시하여 교훈을 도출한다. 이는 향후 수행하게 될 다른 무기체계에 대한 동일한 시행착오를 예방하고 새로 보직된 시험평가관에게 Know-How를 제공하는 등 시험평가의 질적인 향상에 기여할 수 있도록 해야 한다. 특히, 시험평가 실시단계에서 보완요구 사항으로 분류된 항목에 대해서는 보완 요구사항 조치 결과를 확인하고, FT에서 AI 무기체계의 특성과 데이터의 수집, 성능향상 능력을 확인해야 한다.

이상과 같이 AI 기반 무기체계를 시험평가 하기 위해서는 국방획득체계 전 단계에 걸쳐 시험평가와 관련하여 역할을 수행해야 한다. 또한, 운용시험평가를 위해서 자료수집 및 기초연구로부터 후속 조치단계까지 일반 무기체계 시험평가 수행 절차에 추가하여 AI 무기체계의 특성을 고려하여 데이터 기반 성능평가 및 실물에 의한 현장 시험평가와 관련된 절차를 추가하여 수행해야 한다.

제6장 결 론

제1절 연구결과 요약 및 의의

4차 산업혁명의 시대, 미래 전장은 첨단 국방과학기술이 적용된 지상·해상·공중에서 우주·사이버전 및 인지·심리 영역으로 확대되어 전쟁의 복잡성이 증가하고, 인공지능(AI)·드론봇 등 첨단 과학기술을 기반으로 하는 무기체계가 등 미래 전장을 주도할 것이다. 이를 증명하고 있는 러시아·우크라이나 전쟁은 민간 첨단과학 기술과 무기체계가 결합 되었을 때 어떠한 시너지를 유발할 수 있는지 보여주고 있다. 한반도에서 발생할 수 있는 미래전은 5차원의 전장 공간으로 확장되고, 인공지능(AI)과 로봇(Robot), 무인 자율시스템 등 첨단 과학기술의 융합으로 유·무인 복합전투체계가 전장을 주도할 것이다. 이러한 전장 환경에서 운용된 AI 기반 무기체계들의 전력화를 위한 최종 Gate-Keeper로서 역할은 무기체계 시험평가가 수행해야 한다.

무기체계 시험평가는 특정 무기체계가 기술적·운용 및 관리적 측면에서 소요제기서에 명시된 제반 요구 조건을 충족하는지 여부를 확인·검증하는 것이다. 시험평가는 무기체계 획득과정에서 위험을 줄이기 위한 기술적 도구로 사용되고, 4차 산업혁명의 시대에 민간 신기술을 반영한 첨단 무기체계의 작전 운용성능(ROC)과 기술적·부수적 성능이 시험평가를 통하여 최종 확인 검증되어야 한다.

이에 따라 본 연구에서는 인공지능(AI) 기반 무기체계에 대한 시험평가 수행체계 개선방안을 도출하기 위해, 무기체계 개발 및 시험평가 환경변화와 무기체계 시험평가 사례분석을 통해 시험평가 개선 소요를 도출하였다. 이 결과와 국방 전력분야 전문가를 대상으로 기초조사 설문을 실시하여 첨단 무기체계 시험평가 개선방안에 대한 시험평가 제도·구조·기술 분야로 분류하고 평가

항목을 조정 및 통합하는 과정을 통해 첨단 무기체계 시험평가 개선방안에 대한 계층구조 및 평가항목을 3개 분야 11개 세부 평가항목으로 그룹화하였다. 이후 확정된 세부 평가항목에 대한 상대적 중요도 및 우선순위를 분석하기 위해 국방 전력분야 전문가 46명을 대상으로 AHP 설문을 실시하고 분석한 결과, 첨단 무기체계 시험평가 개선방안에 대한 우선순위는 표 42와 같다.

표 42. 첨단 무기체계 시험평가 개선방안 우선순위

Table 42. Priorities for improvement plans for advanced weapon system T&E

우선순위	무기체계 시험평가 개선방안
1	시험평가 결과판정 개선
2	시험평가 조직 및 전문인력 양성
3	과학적 시험평가체계 정립
4	AI·무인·자율주행 등 관련 법령 제정
5	신속획득방안 개선 / 확대
6	시험평가 인프라 구축
7	안전기준 정립
8	네트워크 통합시험체계 구축
9	시험평가 기술지원
10	한국형 시험표준 정립
11	민·군·국제 기술협력 / 교류 강화

첨단 무기체계 시험평가 개선방안 중 상대적 우선순위가 높은 평가항목에 대한 인공지능(AI) 기반 무기체계 시험평가 개선방안은 다음과 같다.

첫째, 시험평가 결과판정 개선을 위해 AI 체계의 특성을 고려하여 데이터 기반 성능평가(기본성능, 학습능력, 최종성능)와 실물에 의한 시험평가를 종합적으로 평가하여 전투용 ‘적합’, ‘조건부 적합’, ‘부적합’ 또는 전투용 ‘적합’, ‘가

능’, ‘개발’과 같이 시험평가 등급제 적용하여 결과를 판정해야 한다는 것과 반복 개발학습이 필요한 AI 적용사업은 양산 간 지속 개발 및 확인을 위한 별도의 절차 마련이 필요하다. 또한, 시험평가 항목을 핵심(필수) 와 일반(선택)항목으로 구분하여 일반(선택)항목 중 미충족한 항목에 대해 양산과정에서 보완할 수 있는 융통성과 현 제도상의 ‘획득추천’을 적극적으로 활용하여 전력화 지연 및 중단을 막자는 의견을 제시하였다.

둘째, 시험평가 조직 및 전문인력 양성을 위한 개선방안은 현 육군 시평단의 인적 구성을 분석하여 첨단 무기체계 시험평가를 위해 첨단기술 학위소지자 및 전력업무 유경험자를 우선적으로 보직시켜야 한다. 시험평가관은 일정 자격과 경력을 구비하게 되면 전문 특기를 부여하여 보직 관리를 하고, 미군의 육군 시험평가사령부를 벤치마킹하여 시험과 평가조직을 구분하고 평가조직을 보장해야 한다는 것을 제시하였다.

셋째, AI 기반 무기체계 과학적 시험평가 개선방안은 AI 체계의 신뢰성을 향상시키기 위해 데이터 기반 성능평가와 실물에 의한 현장 시험평가를 실시하고 AI 무기체계 성능평가 결과에 대한 성능인증업무를 수행할 수 있는 조직을 편성해야 한다. AI 무기체계는 운용유지 단계에도 지속적으로 데이터를 수집하고 학습화하여 데이터셋을 구축하고 고도화해야 한다. 또한, AI 기반 신개념 무기체계를 적시에 도입할 수 있도록 시험평가용 M&S 체계를 구축하고 활용해야 한다는 것을 제시하였다.

마지막으로 운용시험평가 절차를 현 체계에서 AI 기반 무기체계와 관련된 분야를 소요제기로부터 운용시험평가 실시, FT 이후까지 전(全) 단계에 걸쳐 시험평가의 역할을 수행해야 한다는 것을 제시하였다.

이와 같은 연구를 통하여 얻은 시사점과 의의는 다음과 같다.

첫째, 4차 산업혁명의 시대, 첨단 과학기술을 국방 무기체계에 적용하기 위

한 추진에도 불구하고 상대적으로 소홀하게 다루고 있는 첨단 무기체계 시험평가의 개선을 위한 현 시험평가체계의 한계와 개선하기 위한 세부 소요를 도출하고 우선순위를 분석해 본 새로운 시도였다.

둘째, 무기체계 시험평가 개선방안을 도출하기 위해 소요기획, 사업관리, 시험평가, 방산업체 등 국방전력 분야에서 근무하고 있거나 근무경력이 있는 전문가들을 대상으로 무기체계 시험평가 개선을 위한 의견을 폭넓게 수렴하여 첨단 무기체계 시험평가 개선을 위한 실질적인 우선순위를 도출하였다. 이로써 전력업무를 담당하고 있는 각 기관 관계관들의 공감대 형성은 물론 협업을 통해 신속하게 개선하고 조치할 수 있는 계기를 마련하였다.

셋째, 국방 전력분야 전문가 집단이 가장 중요하다고 인식하고 있는 첨단 무기체계 시험평가 개선방안의 한 분야는 ‘시험평가 제도’였으며, 세부 평가항목은 ‘시험평가 결과판정의 개선’으로 평가되었다. 이와 같은 분석 결과는 현 시험평가 제도의 한계인 신속성과 효율성이 결여되었다는 것으로 급변하는 기술변화를 무기체계에 신속하게 적용하기 위해서는 시험평가 결과판정을 비롯한 시험평가 제도를 우선적으로 개선해야 한다는 것이다.

이상으로 첨단 무기체계 시험평가 개선방안에 대하여 시험평가 제도·구조·기술 분야의 상대적 중요도와 세부 평가항목에 대한 우선순위를 분석한 결과와 AI 기반 무기체계 시험평가 개선방안 및 시사점을 제시하였다.

본 연구를 통하여 첨단 무기체계 시험평가 개선을 위해, 분석된 시험평가 제도·구조·기술 분야와 11개의 세부 평가항목에 대한 우선순위를 고려하여 국방 전력업무 관련 기관 및 담당자간 상호 공감대를 형성하고, 신속하게 개선할 수 있도록 보다 깊이 있는 연구가 지속되기를 기대한다.

제2절 연구의 한계 및 향후 연구과제

본 연구는 첨단 무기체계 시험평가 개선방안을 도출하기 위해 우리 군이 추진하고 있는 국방혁신 4.0과 Army TIGER 4.0등 기획문서 및 정책서에 등 문헌연구 및 선행연구를 통해 국방 과학기술의 발전과 한반도의 미래전 양상에 대비하기 위한 첨단 무기체계 개발 소요와 무기체계 시험평가 개선 소요를 도출하였다. 국방 전력분야 전문가를 대상으로 기초조사 및 계층적 분석기법(AHP)을 적용하여 첨단 무기체계 시험평가 개선방안에 대한 시험평가 제도·구조·기술 분야의 중요도와 우선순위를 분석하였다. 이를 바탕으로 우선순위가 높은 평가항목에 대해 AI 기반 무기체계 시험평가 개선방안을 제시하는 것으로서, 다음과 같은 현실적인 한계가 있었다.

첫째, 본 연구는 첨단 무기체계 개발 소요와 시험평가 개선 소요를 도출하기 위해 국방혁신 4.0과 Army TIGER 4.0등 기획문서 및 정책서 등을 참고하여 제시하고자 하였으나, 군사보안과 대외 공개가 제한되는 자료들로 선행연구 자료가 빈약하였다. 따라서 미래전 대비 첨단 과학기술군 육성을 위한 국방 과학기술의 수준과 무기체계 소요제기와 관련하여 국방부 연두보고 및 육군의 AI 추진전략과 관련하여 대외 발표 자료 및 연구보고서, 연구논문 등 참고하여 제시하였다.

둘째, 우리 군이 추진하고 있는 인공지능(AI) 기반 무기체계 시험평가 사례가 빈약하였다. 군은 첨단 무기체계 전력화를 추진하고 있으나, 해·공군에서 추진한 인공지능(AI) 기반 무기체계 시험평가 사례에 대한 자료접근이 제한되었다. 이로 인해 AI 체계에 대한 세부 시험평가 절차와 데이터 확인이 불가하여 육군의 신속시범획득사업으로 추진한 육군의 시범 운용부대의 성능확인 결과를 한정하여 제시하였다.

셋째, 국방 전력분야 전문가를 대상으로 기초조사 및 AHP 설문을 실시하는

데 있어 전문가 집단을 각 기관 및 신분별로 균등하게 편성되어야 하나, 현역과 시험평가단이 타 집단에 비해 편중되는 경향이 있다. 이와 같은 점은 연구자와 연구방향 등을 고려하여 타 기관의 전문가 의견을 충분히 반영될 수 있도록 함으로써 분석 결과에 왜곡이 발생하지 않도록 하였다.

넷째, AHP 분석은 의사결정 계층이 각 요소들이 서로 독립적이고 상부 요인은 하부요인을 충분히 함축해야 한다는 가정하에 진행되어야 한다. 따라서 이번 연구에서도 시험평가의 3개 분야, 11개 세부 평가항목이 상호간 독립적이라는 전제하에 진행하였다. 하지만, 세부 평가항목들 사이에 상호 영향성과 종속적인 관점에서 중복적인 의미 등에 대해 충분히 검증하지 못한 점이 있다.

다섯째, AI 기반 무기체계의 운용시험평가 절차와 관련하여 세부적인 시험평가 방법 및 절차를 제시하지는 못했다. 이와 같은 점은 AI 기반 무기체계 시험평가 사례 및 시험평가 유경험자가 제한되어 특정 시험평가 사례 및 시험평가관의 경험이 전반적인 AI 기반 무기체계의 시험평가 절차가 되지 않도록 개괄적인 사항 위주로 제시하였다. 또한 AI 기반 무기체계 유형별 시험평가를 수행하기 위한 시험평가 방법 및 절차를 구체화하여 제시하지 못했다.

따라서, 향후 연구과제는 본 연구 결과를 토대로 다음과 같은 후속 연구가 필요할 것으로 판단된다.

첫째, 국방혁신 4.0 추진과 연계하여 AI·무인·로봇 등 첨단 무기체계 시험평가 시스템을 구축하기 위해서는 시험평가 제도·구조적 기반에서 무기체계 유형별 세부 시험평가 방법 및 절차에 관한 연구가 필요하다. 현재까지는 드론, 폭발물 탐지 및 제거 로봇, 유도무기, 중요시설경계시스템 등 일반 무기체계의 시험평가를 진행했다면, 앞으로는 일반 무기체계에 인공지능(AI)을 기반으로 하는 무기체계로 고도화되어 유·무인 복합전투체계로 발전되고 지능형 지휘통

체체계가 구현될 것이다. 이와 같은 AI 기반 무기체계를 시험평가 하기 위해서는 AI 기반 무기체계의 유형별로 세부 시험평가 방법 및 절차에 대한 연구가 되어야 한다.

둘째, AI·로봇 등 무인체계에 대한 전문가 인재 Pool 확보방안 및 전문성 강화를 위한 전력분야 전문가 인사관리 방안에 관한 연구가 필요하다. 앞장에서 시험평가단의 인적 구성이나 시험평가관의 경력을 분석하여 제시했듯이 현재의 시험평가단은 첨단 무기체계를 시험평가 할 수 있는 조직으로는 미흡한 점이 많다. AI를 기반으로 하는 첨단 무기체계를 시험평가 하기 위해서는 첨단 과학기술에 대한 전문지식과 유형별 무기체계에 대한 시험평가를 수행할 수 있는 경력을 구비해야 한다. 이를 위해서 첨단 무기체계에 대한 인재를 어떻게 확보할 것인지, 전력업무 및 시험평가 유경험자들을 어떻게 활용할 것인지 대해 진지한 연구가 되어야 한다.

셋째, 첨단 무기체계 시험평가의 신뢰성을 확보하기 위해서는 과학적 시험평가 및 네트워크 통합시험에 대한 구체적인 시험평가 방안에 관한 후속 연구가 필요하다. 한반도의 미래 전장은 네트워크 기반의 초지능전(정보융합, 지능정보 제공)이 구현되고 다차원의 통신 네트워크를 구성하여 전 전장에서 드론, 로봇, 지휘통제가 가능한 진술 통신체계가 구현될 것이다. 이를 위해서는 미래 전장에서 운용되는 각각의 무기체계가 초연결되어 지능형 지휘통제체계가 구현되는지 네트워크 통합시험을 통해 확인 검증해서 신뢰성이 확보된 가운데 AI 기반 첨단 무기체계를 운용해야 한다.

이상과 같이 본 연구를 기반으로 후속 연구가 활발하게 이루어진다면, AI 기반 무기체계 시험평가 수행체계가 정립되고, 시험평가에 대한 신뢰성이 확보되어 우리 군이 추진하는 AI 과학기술 강군으로 변화와 혁신을 주도하고 나아갈 수 있을 것이다.

참 고 문 헌

1. 국내 문헌

- [1] 장상국, “OT(기회요인-위협요인) 분석을 통한 차기자주포(K9F) 개발 필요성 연구”, 선진국방연구 제3권 3호, 2020
- [2] 육군 시험평가단, “시험평가 참고서”, 2019
- [3] 박원동, “명품 무기체계 탄생의 마지막 진통”, 북코리아, 2008
- [4] 김종하, “국방획득체계 현황 및 문제점, 그리고 개선방향”, 과학기술정책 제27권 제11호, 2017
- [5] 위검복, 하철수 등 4명, “무기체계 소요결정과 사업추진방법의 효율적 연계 방안”, 한국군사문제연구원, 2014
- [6] 육군, “획득제도 개선방향”, 2022
- [7] 합참, “무기체계 시험평가 업무 규정”, 합참규정-541-01(‘21. 2. 26. 제정), 2021
- [8] 류종수, “업체주관 연구개발사업의 시험평가 발전방안 연구”, 국방과 기술, 2019
- [9] 국방부, “국방전력발전업무훈령”, 제2568호(‘21. 6. 30.), 2021
- [10] 방위사업청, “정밀유도무기 시험평가 기준 정립 및 절차 연구”, 2014
- [11] 합참, “무기체계 시험평가 실무 가이드북”, 2016
- [12] 소형근, “포스트휴머니즘 시대에 인간과 인공지능 이해하기: 강한 인공지능의 등장에 어떻게 대처할 것인가?”, 구약논단 제28권 4호(통권 86집), 2022
- [13] 오승환·이주량·최병삼 등 11명, “인공지능 기술 활용 강국을 향한 과학기술 정책 제고 전략“, 과학기술정책연구원(STEPI), 2020
- [14] 육군, “인공지능 종합발전 계획“, 2022
- [15] 윤정현, “국방 분야의 인공지능 활용성 제고 방안과 시사점“, 과학기술정책 포커스, 과학기술정책연구원, 2021

- [16] 정보통신기획평가원, “ICT기술수준조사 및 기술경쟁력보고서”, 2022
- [17] 국기술진흥연구소, “미래국방 2030 기술전략(AI)”, 2021
- [18] 국방부, “AI 과학기술 강군 개념 및 추진방향”, 2023년 국방부 연두보고, 2022
- [19] 임영봉, “육군 무기체계 운용시험평가에 관한 연구”, 한남대학교 박사학위 논문, 2014
- [20] 박종완, “무기체계 시험평가의 신뢰성 향상방안”, 신뢰성응용연구 15(2), 2015
- [21] 강응섭, “국방기술 환경변화에 부응한 무기체계 시험평가 신뢰도 제고에 관한 연구”, 대전대학교 박사학위 논문, 2016
- [22] 정월균, “이해관계자 분석을 통한 무기체계 평가제도에 관한 연구”, 목원대학교 박사학위 논문, 2017
- [23] 이용학, “국방 무기체계 시험평가의 투명성에 관한 제언”, 김병기 국회의원·한국사회안전범죄정보학회·동계 공공학술심포지, 2017
- [24] 김선영·최기일, “방위산업 발전을 위한 무기체계 시험평가 개선방안 연구”, 한국방위산업학회지 제26권 제1호, 2019
- [25] 최규옥·송재은·유영민, “무기체계에 적용 가능한 적응형 시험평가에 관한 연구”, 예술인문사회 융합 멀티미디어 논문지 제9권 제8호, 2019
- [26] 이강경, “국방 획득환경 변화에 적합한 무기체계 운용시험평가 개선전략 도출”, 충남대학교 박사학위 논문지, 2020
- [27] 이용복·전일국·김상홍, “인공지능(AI) 무기체계 시험평가방법 발전방안 연구”, 신뢰성응용연구 제22권 제1호, 2022
- [28] 이종관·한창희, “미래전과 국방 인공지능 체계”, 한국통신학회논문지 제44권 제4호, 2019
- [29] 안진우·노상우·김태환·윤일웅, “인공지능 분야 국방 미래기술에 관한 실증 연구”, 한국산학기술학회논문지 제21권 제5권, 2020
- [30] 장상국·최기일, “미래 국방을 대비한 인공지능 기반의 방위산업 발전방향

- 연구”, 한국방위산업학회지 제28권 제3호, 2021
- [31] 권세일·강지훈·서석호 등 5명, “인공지능 기술 및 인증제도 분석을 통한 국방 인공지능 품질관리 방안 연구”, 한국품질경영학회 추계학술발표논문지, 2022
- [32] 조현석·고봉준, “인공지능의 군사적 활용과 국제 핵질서”, 정치정보연구 제 26권 제1호, 2023
- [33] 신경식, 송민채, 신현섭 외 5인, “사례연구 방법(6판)”, 서울: 한경사, 2021
- [34] 유기웅, 정종원, 김영석, 김한별, “질적 연구방법의 이해”, 서울: 박영사, 2018
- [35] 최하영, “교육과정적 통합을 위한 초등 특수교사와 통합학급 담임교사 간 협력에 관한 사례연구”, 국교원대학교 대학원, 2006
- [36] 이영철, “사과학에서 사례연구의 이론적 지휘: 비판적 실재론을 바탕으로” 한국행정학보 40(1), 2006
- [37] 권오정, “기준 의사결정 방법론 이론과 실제”, 서울: 북스힐, 2018
- [38] 조근태·조용곤·강현수, “계층분석적 의사결정”, 서울: 동현출판사, 2003
- [39] 임성훈·조기홍·박승, “무기체계의 효과분석과 의사결정을 위한 다기준 분석방법론 연구”, 한국군사과학기술학회지 제12권 제5호, 2009
- [40] 김홍빈, “무기체계 소요기획에 관한 영향요인의 우선순위 결정 방안 연구”, 광운대학교 박사학위 논문, 2015
- [41] 육군교육사령부, “국방운영분석 이론과 실제(제8장 계층적분석기법)”, 2021
- [42] 류태규, “4차 산업혁명과 국방분야 과학기술 적용”, 한국국가전략 제5권 3호, 2020
- [43] 김종열, “미래 무기체계와 군사과학기술 발전추세 분석: 미국을 중심으로”, 전략연구 통권 제76호, 2018
- [44] 함참, “4차 산업혁명 기반 첨단 무기체계 시험평가 발전방안 연구”, SMI안보 경영연구원, 2018
- [45] 정보통신기획평가원, “(GT주간브리프-실리콘벨리)미국, AIU(국방혁신단) 주요 현황 및 추진동향”, 2020

- [46] 윤정현, “국방 분야 인공지능 기술 도입의 주요 쟁점과 활용 제고 방안”, 과학기술정책연구원(STEPI), 2021
- [47] 방위사업청, “신속시범획득사업 업무관리 지침”, 2021
- [48] 이창영, “중국의 민군융합을 통한 지능화군 건설 전략”, KIMS Periscope (한국해양전략연구소) 제166호, 2019
- [49] 윤정현, “국방 분야 인공지능 활용성 제고 방안과 시사점”, 과학기술정책 포커스, 2021
- [50] 합참, “인공지능 무기체계 시험평가 기준, 방법, 절차 연구”, 국방대학교, 2022
- [51] 박홍순·김세용·김용환, “체계적인 방위산업기술보호를 위한 보호체계 우선 순위 분석 연구”, 융합보안논문지 제19권 4호, 2019
- [52] 신예진, “2022 신뢰할 수 있는 인공지능 개발 안내서”, 한국정보통신기술협회(TTA), 2022
- [53] 이진호·정석재, “무기체계 연구개발 시험평가 관정 개선방안에 대한 연구”, 군사발전연구 제15권 제2호, 2021
- [54] 육군 시험평가단, “과학적 시험평가를 위한 Guide Book”, 2022
- [55] 과기정통부, “신뢰할 수 있는 인공지능 실현 전략(안)”, 2021

2. 국외 문헌

- [56] T. L. Saaty, “Decision Making for Leaders, II”, AHP series, RWS Publication, 1995.
- [57] Howard, Ronald A., “Decision Analysis : Practice and Promise”, Management Science, Vol.34, No.6, 1988
- [58] Wind, Yoram and T. L. Saaty, “Marketing Application of the AHP”, Management Science, Vol.1, No.27, 1980
- [59] Howard, Ronald A., “Decision Analysis : Practice and Promise”, Management

Science, Vol.34, No.6, 1988

- [60] The Economist, “Special Report: The future of War”, 2018
- [61] Stockholm International Peace Research Institute (SIPRI), “SIPRI Yearbook 2021 (Summary)”, 2022
- [62] The Defence Acquisition University, “Test & Evaluation”, 2012, US Marine corps, “Operational Test & Evaluation Manual”, 2013
- [63] Jared Dunmon 등 5명, “RESPONSIBLE AI GUIDELINES IN PRACTICE”, Defense Innovation Unit, 2021
- [64] Stanford University Human-Centered AI, “Artificial Interlligence Index Report 2022”

3. 인터넷 및 기타자료

- [65] 송학, “국방 획득체계, 투명성보다 전문성·효율성 강화해야”, 녹색경제신문, <http://www.greened.kr>, (검색일: 2022. 4. 5.)
- [66] <https://terms.naver.com/entry.naver?docId=4356939&cid=43665&categoryId=43665>, 네이버 시사경제용어사전, (검색일: 2023. 4. 9.)
- [67] <https://m.blog.naver.com/pwj6971/221614497987>, (검색일: 2023. 4. 9.)
- [68] https://news.unist.ac.kr/kor/column_202/, UNIST(‘17. 5. 25.), (검색일: 2022. 11. 22.)
- [69] <https://www.sedaily.com/NewsView/1VQVVLXHAV>, 서울경제(‘19. 11. 21.), (검색일: 2022. 11. 20.)
- [70] 한국전자통신연구원(2020), ATL 1.0 정의
- [71] https://spri.kr/posts/view/23555?code=data_all&study_type=industry_trend, SPRI소프트웨어 정책연구소(‘23. 3. 2.), (검색일: 2023. 4. 1.)
- [72] The Economist, special report: “The future of War”, 2018,

- <https://www.economist.com/special-report/>, 2018. 1. 27, (검색일: 2022. 4. 12.)
- [73] <https://www.joongang.co.kr/article/4370801#home>, "첨단기술도 중요한 전쟁 역제수단"(검색일 2022. 4. 18.)
- [74] 네이버 지식백과, <https://terms.naver.com/entry.naver?docId=1398167&cid=43667&categoryId=43667>, (검색일: 2022. 4. 18.)
- [75] <https://www.ai.mil/> 미국방부(DoD), 2018. 7. 27, (검색일: 2022. 4. 12.)
- [76] http://www.koits.net/HOME/kits/sub.htm?nav_code=kit1529408669, (검색일: 2022. 4. 13.)
- [77] https://colab.research.google.com/drive/1zx9k16VII8MZ3_n6VcnsBeiOoQQxavBF?usp=sharing (검색일: 2022. 4. 13.)
- [78] 육군, "신속시범 획득사업 추진현황", 2023
- [79] 미 육군시험평가사령부(ATEC), <https://www.army.mil/atec>, (검색일: 2022. 6. 5.)
- [80] 국방부, "2023년 국방부 연두보고", <https://www.fnnews.com/news/202207221828281450>, 파이낸셜 뉴스, (검색일: 2023. 4. 11.)
- [81] <https://www.aitimes.kr/news> 인공지능신문, (검색일: 2023. 4. 7.)
- [82] https://blog.naver.com/kpfjra_/222456508992, 알고리즘 편향을 해결하기 위한 노력들, 신문과 방송, (검색일: 2023. 4. 8.)
- [83] <https://www.joongang.co.kr/article/22521196#home> 테슬라사고 태양 역광택, 중앙일보(2018. 4. 10.), (검색일: 2023. 4. 8.)
- [84] <https://ditoday.com/> 페이크 뉴스의 위험성을 경고하는 페이크 뉴스(2018. 5. 3), (검색일: 2023. 4. 8.)
- [85] 국방부(2020), "2020년 국방부 업무계획", 2020
- [86] 육군 시험평가단, "시험평가 제도개선 심층 토론회 안건 검토의견", 2023
- [87] <https://www.army.mil/atec>, (검색일: 2022. 6. 5.)

부 록

#1 기초조사 설문지

첨단 무기체계 시험평가 발전방안에 관한 연구

안녕하십니까?

저는 광운대학교 일반대학원 방위사업학과 박사과정을 이수중입니다.

본 설문은 「첨단 무기체계 시험평가 발전방안에 관한 연구」를 수행하기 위한 설문이며, 여러분들께서 바쁘신 시간을 할애하여 답변해주신 내용은 본 연구에 많은 도움이 될 것입니다.

본 설문을 통하여 얻은 자료는 **설문의 목적인 미래 첨단 무기체계 시험평가 개선을 위한 문제점 진단 및 개선방안을 분석하기 위한 자료**로 사용될 것이며, 수집된 자료는 하단에 명시된 법률에 의거 **연구조사 목적**으로만 사용됩니다.

그러므로 각 문항에 정오답은 없으며, 귀하가 느끼신 바를 말씀해주시면 됩니다. 귀하의 소중한 응답이 저의 연구에 많은 도움이 될 것입니다.

설문과 관련하여 **의문사항**이 있으시면 **아래 연락처로 연락**주시면 성실히 답변 드리겠습니다.

설문에 응해 주신 전문가 분들의 관심과 정성어린 답변에 감사의 말씀을 드립니다.

2022년 6월

지 도 교 수 : 광운대학교 대학원 방위사업학과 교수 정 석 재

공동 지도교수 : 광운대학교 대학원 방위사업학과 교수 박 찬 봉

연 구 자 : 광운대학교 대학원 방위사업학과 박사과정 김 백 중

1. 본 연구는 개인정보 보호법 [시행 '20. 8. 5, 법률 제 16930]에 의거 통계조사 목적으로만 사용됩니다.
2. 본 연구의 내용은 통계법 [시행 '20. 6. 9, 법률 제 17339] 제5장 33조(비밀의 보호)에 의거 비밀이 보장되며 통계조사 목적외에는 사용되지 않습니다.

I. 설문 내용

무기체계 시험평가 개선 필요성

육군은 「한계를 넘어서는 초일류 육군」을 구현하기 위해 많은 변화와 혁신을 추진하고 있습니다. 특히 첨단 과학기술군으로 변화하기 위해 기반 전투체계, 드론봇 전투체계, 위리어플랫폼 등 3대 전투체계를 기반으로 기동화, 지능화, 네트워크화가 구현된 군사혁신의 미래 지상군 상징인 Army TIGER 부대의 전력화를 진행 중입니다. 또한 AI(인공지능) 적용과 우주력 발전계획 “폐가수스 프로젝트 및 레이저” 정책을 추진하는 등 AI 과학기술 강군으로 미래 육군건설을 추진하고 있어 신개념 첨단 무기체계 시험평가를 위한 제도적, 방법 및 절차에 대한 연구가 필요합니다.

예를 들어, AI(인공지능)이 탑재된 무기체계를 시험평가할 경우 기존의 시험평가 방법 및 절차로 현장에서 실물(주장비)을 대상으로 AI 기술을 포함한 무기체계 전체의 성능을 시험평가 하는 것이 적절한 것인가? 또한, 인공지능 체계는 수집된 정보를 AI가 판단하기 때문에 인공지능 알고리즘의 성능을 평가할 것인가? 하는 의구심으로 현 무기체계 시험평가 체계를 적용하는데 많은 제한과 한계가 있다.

시험평가는 무기체계 개발하고 획득하는 과정으로 합참에서 소요결정 후 방사청은 선행연구를 거쳐 사업추진기본전략에 따라 체계개발 단계에서 시험평가를 수행하고 있다. 연구개발 및 구매로 획득하는 모든 무기체계는 시험평가 과정을 거쳐야 하며, 일반적으로 아래와 같이 개발시험평가(DT&E)와 운용시험평가(OT&E)로 구분하여 수행하고 있다.

< 무기체계 시험평가 구분 >

구분	개발시험평가(DT&E)	운용시험평가(OT&E)
목적	기술적 개발목표 충족여부 확인	군 운용적합여부 확인
시험환경	전문기술 시험환경	각종 작전환경
시험분야	기능·성능분야 (신뢰도, 유지성, 적합성, 호환성, 내환경성, 안정성)	작전운용성능의 충족성, 군 운용적합성 전력화지원요소의 실용성

* 출처 : 국방전력발전업무훈령 제2568호, 시험평가 참고서(육군 시험평가단)

본 설문에서는 4차 산업혁명 시대에 첨단과학기술이 적용되는 미래 첨단무기체계 연구개발과 연계하여, **현행 무기체계 시험평가에 대한 개선 필요성**을 분석하고자 합니다.

2-2. '불필요하다'를 선택하셨다면 그 이유는 무엇이라고 생각하십니까.

3. 첨단 무기체계 시험평가를 위해 '시험평가 제도' 차원에서 개선해야 할 분야는 어떠한 사항들이 있다고 생각하십니까? (복수선택 가능)

- 가. 신속획득 방안
- 나. 국제 시험평가 교류 및 협력
- 다. 국방 연구개발 협력(핵심기술 공동연구, 시험평가 기술공유 등)
- 라. 기타()

3-1. 개선할 분야(기타 포함)를 선택하셨다면 그 이유는 무엇이라고 생각하십니까.



4. 첨단 무기체계 시험평가를 위해 '시험평가 구조' 차원에서 개선해야 할 분야는 어떠한 사항들이 있다고 생각하십니까? (복수선택 가능)

- 가. 시험평가관련 조직
- 나. 민·군 기술협력
- 다. 시험평가 인프라(전문 시험시설 및 장비, 시험장 등)
- 라. 시험평가 지원(시험기술 데이터 관리 / 계측, 시험평가 기술연구 등)
- 마. 기타()

4-1. 개선할 분야(기타 포함)를 선택하셨다면 그 이유는 무엇이라고 생각하십니까.

5. 첨단 무기체계 시험평가를 위해 ‘시험평가 기술(방법/절차)’ 차원에서 개선해야 할 분야는 어떠한 사항들이 있다고 생각하십니까? (복수선택 가능)

- 가. 과학적 시험평가
- 나. 네트워크 통합시험
- 다. 시험평가 기술지원(지원프로그램)
- 라. 한국형 시험표준 정립
- 마. 기타()

5-1. 개선할 분야(기타 포함)를 선택하셨다면 그 이유는 무엇이라고 생각하십니까.

6. 첨단 무기체계 시험평가를 위해 ‘시험평가 관련 법령/훈령’ 차원에서 개선해야 할 분야는 어떠한 사항들이 있다고 생각하십니까? (복수선택 가능)

- 가. 통합시험 확대 적용
- 나. 시험평가 결과의 판정(조건부 적합 추가 등 유연성 확대)
- 다. AI·무인·자율주행 등 관련 법령 제정
- 라. 환경시험관련 국방표준 제정
- 마. 안전기준, 인증 및 품질보증 기준
- 바. 기타()

6-1. 개선할 분야(기타 포함)를 선택하셨다면 그 이유는 무엇이라고 생각하십니까.

운용시험평가 개선 필요성

운용시험평가는 개발 장비의 시제품에 대하여 각종 작전환경 또는 이와 동등한 조건에서 작전운용성능 충족 여부 및 군 운용 적합 여부를 확인하기 위한 시험평가. 이때, 사업의 특성 및 효율적인 예산 집행을 위해 필요하다고 판단할 경우 최초 시험평가 단계와 후속양산 결정을 위한 후속시험평가 단계로 구분하여 수행할 수 있다.¹⁷³⁾ 운용시험평가는 아래 표와 같이 국방전력발전업무 훈령에 정의된 시험평가 항목에 대하여 수행하고 있습니다.

< 운용시험평가 항목 >

구 분	세부 시험평가 항목
작전운용성능의 충족성	① 주요 작전운용성능 ② 기술적 / 부수적 성능
합동성 및 상호운용성	① 운용개념 및 체계특성 ② 연동성 및 정보교환 ③ 표준 및 아키텍처
군 운용적합성	① 운용 및 조작 적합성, 안정성 ② 전술적 운용의 적합성 ③ 기존 무기체계와의 상호운용 적합성 ④ 환경 적응성
전력화지원요소의 실용성	① 전투발전지원요소 ② 통합체계지원(IPS) ¹⁷⁴⁾ 요소

* 출처 : 국방전력발전업무훈령 제2568호, 시험평가 참고서(육군 시험평가단)

미래 무기체계 중 인공지능(AI)과 자율주행이 탑재된 무인 전투체계의 핵심분야는 감지센서와 인식 및 처리, 자율제어, 임무 및 통제, 통신 및 통신망에 대한 핵심 요구능력을 충족해야 합니다. 이러한 무인 전투체계의 핵심 요구능력에 대한 시험평가를 위해서 무인체계에 대한 시스템을 평가하고 운용의 안전성 등이 요구되고 있습니다.

또한 미래 전투수행개념에 따라 유인과 무인의 복합부대 운영개념에서 무인체계의 요구성능과 군 운용적합성 충족을 위한 검증방법(법적·윤리적 문제 포함) 및 절차에 대한 연계성이 부족하고, AI·로봇기술 등이 적용된 무기체계에 대한 시험평가 방법과 분석결과에 따라서 신뢰성에 의문을 제기할 수 있습니다.

인공지능(AI), 자율주행, 무인체계·로봇 등 첨단과학기술이 적용된 미래 첨단무기체계 운용시험평가를 수행하는데 있어 중점적으로 검토가 필요한 시험평가 항목은 **군 운용 적합성**이라고 판단 할 수 있습니다.

군 운용적합성 세부 시험평가 항목의 정의와 운용시험평가 항목의 예는 다음과 같습니다.

〈 군 운용적합성의 세부항목 및 정의 〉

구 분	정의 / 세부 운용시험평가 항목(예)
① 운용 및 조작 적합성, 안정성	체계(System)를 적절하게 사용하고 다루어 움직임과 관련된 요구 사항을 충족하는 정도 예) 운용공간의 적합성, 자동사격통제장치 조작 및 편의성, 장전기 운용 적합성, 조종수 장비운용 적합성 및 안정성 등
② 전술적 운용의 적합성	체계(System)를 전술 상황에서 적절하게 사용함에 있어 요구사항을 충족하는 정도 예) 조준(방렬), 계획사격 수행, 야전기동, 항법정확도, 사격준비, 사격임무수행, 자체방호 및 생존성, 공역통제 등
③ 기존 무기체계와의 상호운용 적합성	신규 개발되는 체계(System)과 기존에 전력화된 무기체계간 상호 간섭없이 품목, 장비, 제품 등이 존재하거나 기능을 발휘하는 정도 예) 유·무선통신, 물리적 호환성, 전기적 호환성
④ 환경 적응성	인적 및 자연환경 요인들의 효과나 작용이 체계(System)에 미치는 정도 예) 고온/저온, 강우, 침수, 습도, 결빙, 강설, 낙뢰, 진동, 화염, 진흙, 염수분무 등

* 출처 : 국방전력발전업무훈령 제2568호, 시험평가 참고서 별책부록(육군 시험평가단)
미 국방획득대학교(DAU) Defense Acquisition Guidebook(2017)

본 설문에서는 4차 산업혁명 시대에 인공지능(AI), 자율주행, 무인체계·로봇 등 첨단과학기술이 적용된 미래 첨단무기체계 **운용시험평가에 있어 개선분야를 분석하고자** 합니다.

7. 소요군에서 실시하고 있는 운용시험평가 중 미래 첨단 무기체계 연구개발을 위해 개선해야 필요하다고 생각하십니까?
가. 개선이 필요하다 나. 불필요하다 다. 기타()

7-1. “개선이 필요하다, 기타”를 선택하셨다면 그 이유는 무엇이라고 생각하십니까?

173) ‘국방전력발전업무훈령 제59조(시험평가 구분) 제2568호(‘21. 6. 30.)

174) ‘총수명주기관리업무훈령’, 제2502호(‘21. 2. 8.) 제정에 따라 종합군수지원(ILS)에서 통합체계 지원(IPS)로 변경됨

7-2. “불필요하다”를 선택하셨다면 그 이유는 무엇이라고 생각하십니까?

8. 첨단 무기체계 시험평가를 위해 현 운용시험평가 항목 중 개선이 필요한 항목은 무엇이라고 생각하십니까? (복수선택 가능)

- 가. 작전운용성능의 충족성(주요 작전운용성능, 기술적 / 부수적 성능)
- 나. 군 운용적합성(운용 및 조작 적합성, 안정성, 전술적 운용의 적합성, 기존 무기체계와의 상호운용 적합성, 환경 적응성)
- 다. 전력화지원요소의 실용성(전투발전지원요소, 통합체계지원(IPS)요소)
- 라. 기타()

8-1. 개선이 필요한 분야(기타 포함)를 선택하셨다면 그 이유는 무엇이라고 생각하십니까.



< 8번 문항 ‘가’항 선택 시 >

9. 첨단 무기체계 시험평가를 위해 현 운용시험평가 항목의 작전운용성능의 충족성 중 개선이 필요한 항목은 무엇이라고 생각하십니까? (복수선택 가능)

- 가. 주요 작전운용성능
- 나. 기술적 / 부수적 성능
- 다. 기타()

9-1. 개선이 필요한 항목(기타 포함)을 선택하셨다면 그 이유는 무엇이라고 생각하십니까.

13. 설문 응답하시느라 수고 많으셨습니다. 설문 구성 가운데 수정하거나 추가 혹은 불필요한 설문 내용이 있다면 의견을 제시해 주시기 바랍니다.

소속 기관	병과	현 직책/성명 (직위/계급)	연락처 (H.P)
현재 담당하고 있거나 과거 담당했던 업무를 선택해 주세요. (복수선택 가능)			
① 국방정책	② 무기체계(전력지원체계) 소요기획	③ 사업관리	④ 시험평가
⑤ 분석평가	⑥ 연구개발주관기관(방산업체)	⑦ 기타()	
해당분야 근무경력은?			
① 1년 미만	② 1년~3년 미만	③ 3년~5년 미만	
④ 5년~10년 미만	⑤ 10년~20년 미만	⑥ 20년 이상	

#2 계층적 분석기법(AHP) 설문지

첨단 무기체계 시험평가 개선방안에 관한 연구

안녕하십니까?

저는 광운대학교 일반대학원 방위사업학과 박사과정을 이수중에 있습니다.

본 설문은 「미래 첨단 무기체계 시험평가 발전방안에 관한 연구」를 수행하기 위한 설문이며, 여러분들께서 바쁘신 시간을 할애하여 답변해주신 내용은 본 연구에 많은 도움이 될 것입니다.

본 설문을 통하여 얻은 자료는 **설문의 목적인 미래 첨단 무기체계 시험평가 개선을 위한 문제점 진단 및 개선방안을 분석하기 위한 자료**로 사용될 것이며, 수집된 자료는 하단에 명시된 범률에 의거 **연구조사 목적**으로만 사용됩니다.

그러므로 각 문항에 정오답은 없으며, 귀하가 느끼신 바를 말씀해주시면 됩니다. 귀하의 소중한 응답이 저의 연구에 많은 도움이 될 것입니다.

설문과 관련하여 **의문사항**이 있으시면 **아래 연락처로 연락**주시면 성실히 답변 드리겠습니다.

설문에 응해 주신 전문가 분들의 관심과 정성어린 답변에 감사의 말씀을 드립니다.

2022년 9월

지도교수 : 광운대학교 대학원 방위사업학과

교수 정석재

공동지도교수 : 광운대학교 대학원 방위사업학과

교수 박찬봉

연구자 : 광운대학교 대학원 방위사업학과 박사과정

김백중

1. 본 연구는 **개인정보 보호법 [시행 '20. 8. 5, 법률 제 16930]에 의거 통계조사 목적으로만 사용됩니다.**
2. 본 연구의 내용은 통계법 [시행 '20. 6. 9, 법률 제 17339] 제5장 33조(비밀의 보호)에 의거 **비밀이 보장되며 통계조사 목적외에는 사용되지 않습니다.**

I. 연구의 개요

1. 연구배경 및 필요성

- 가. 4차 산업혁명의 첨단과학기술을 적용한 무기체계 획득을 위한 시험평가 환경구축 필요
(국방혁신 4.0, 육군비전2030·2050-미래전 대비 첨단과학기술 기반의 전력 구축)
- 나. 국방부는 빠른 기술발전 속도에 부합할 수 있도록 무기체계 획득절차 개선중
- 다. 육군은 Army Tiger 기반전투체계, 드론봇 전투체계, 워리어플랫폼 등 첨단무기체계 전력화를 통한 미래전 준비 중
- 라. 기존 연구들은 선진 군(미군, 이스라엘 등)의 시험평가 방법 및 절차 등을 참고하여 발전시키는 것에 초점을 맞추고 있어 4차 산업혁명의 시대 첨단 무기체계 시험평가를 위한 시험평가 방법 및 절차 정책/제도, 조직/인력 보강 개선 필요
- 마. 국방전력발전업무 훈령에 제시된 운용시험평가의 시험평가 항목(범위)이 미래전장 환경에서 임무수행 달성 및 소요군의 운용자 요구사항을 충족하기에 부족하고, 시험평가 항목(범위)에 대한 개념과 세부 척도 등 재정립 필요

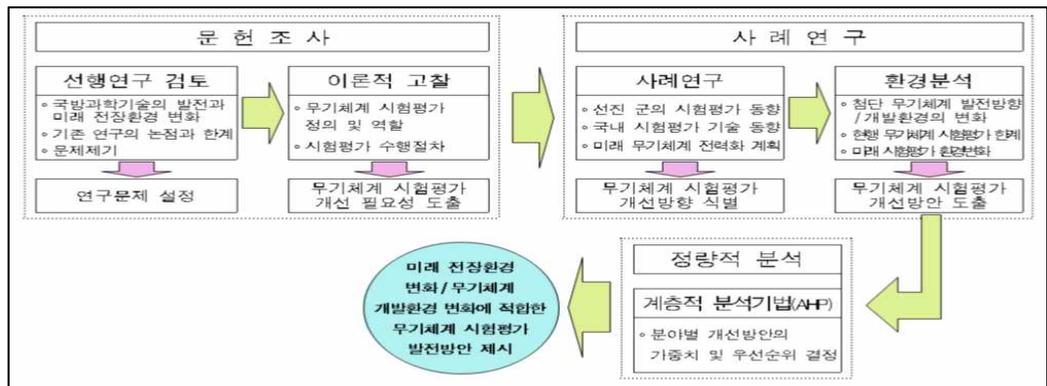
2. 연구목적

- 가. 현 무기체계 시험평가 방법 및 절차, 시험시설, 제도 등으로 Army TIGER 3대 전투체계 (기반전투체계, 드론봇 전투체계, 워리어플랫폼)를 구성하는 주요 전력의 핵심기술에 대한 적절한 시험평가가 제한됨
- 나. 4차 산업혁명 시대 첨단 무기체계 시험평가를 위해 과학적 시험평가 방법, 제도, 시설 및 인력 등 운용시험평가 발전방안 연구

3. 연구방법

- 가. 무기체계 시험평가 관련 선행연구를 통해 기존 연구들에 대한 연구문제 도출
- 나. 국내 및 선진국의 무기체계 시험평가 사례, 국방 연구개발 환경분석을 통해 첨단 무기체계 시험평가를 위한 평가지표 도출
- 다. 무기체계 연구개발 및 시험평가 전문가를 대상으로 계층구조에 대한 AHP 설문을 통해 첨단 무기체계 시험평가 개선 우선순위 도출

4. 연구모형



II. 참고자료

1. 국방과학기술의 발전과 미래전 양상

가. 4차 산업혁명의 첨단 과학기술이 적용된 무기체계 보편화

- AI, 빅데이터 등 ICT기반기술, 무인운송수단, 로봇공학, 신소재 등 신기술 결합(융복합)
- 인간의 인지, 학습, 추론 등과 같이 고차원적 정보처리할 수 있는 AI핵심기술 구현
- * AI기술의 적용 확대, 초연결 시대(다양한 계층의 체계와 연결), 무인전투체계 확대

나. 국방과학기술의 발달로 전쟁수행 패러다임 변화(미래전 양상)

- 5차원의 전장공간으로 확장(지·해·공·우주·사이버전)
- 다양한 작전공간을 넘나드는 전 영역 통합작전 수행(다영역 동시통합작전)
- 인간·드론봇·AI가 결합된 유·무인 복합전 수행
- 위협의 근원을 최단시간 내 제거하는 새로운 형태의 비대칭전(초장거리 정밀타격)
- 초연결망을 기반으로 한 초연결 전쟁(비선형·하이브리드 전쟁), 사이버전자전 일상화

*** 2030년 이후에는 4차 산업혁명 기술들이 더욱 광범위하게 실용화**

다. 국방전략기술 8대 분야 선정 및 추진

- | | |
|-----------------------|---------------------|
| ① 자율·AI 기반 감시정찰 | ⑤ 유무인 복합전투수행 |
| ② 초연결 지능형 지휘통제 | ⑥ 첨단기술 기반 개인전투체계 |
| ③ 초고속·초위력 정밀타격 | ⑦ 사이버 능동대응 및 미래형 방호 |
| ④ 미래형 추진 및 스텔스 기반 플랫폼 | ⑧ 미래형 첨단 신기술 |

라. 첨단 과학기술이 선도하는 미래 무기(12개) 예측(국방기술진흥연구소, KRIT)

- | | |
|---------------------|-----------------------|
| ① 지능형 초연결 통합 네트워크 | ⑦ AI 기반 전영역 유·무인 복합체계 |
| ② 한국형 전술작전 지능화 체계 | ⑧ 복합임무용 해양생체 모방로봇 |
| ③ 메타버스 기반 통합해상 교전체계 | ⑨ AI 기반 초연결 군집 무인기체계 |
| ④ 다중정보 기반 해양감시체계 | ⑩ 무인 스텔스 전투차량 |
| ⑤ 다계층 복합 소형위성군 | ⑪ 파워슈트 착용 슈퍼솔저 |
| ⑥ 우주자산 방어용 미사일 요격체계 | ⑫ 슈퍼 솔저용 스마트 레이저건 |

2. 現 무기체계 시험평가 사례(현실태)

가. 시험평가 제도(법령 / 훈령) 측면

- 1) 現 무기체계 획득절차는 빠른 기술 변화에 신속한 반영 제한
 - 기술 발전속도가 빠른 4차 산업혁명의 신기술 기반 첨단무기체계에 現 획득절차 적용중
 - 무기체계 소요결정~전력화까지 장시간 소요(장기신규 소요제기시 약 10~15년)
 - * K2전차 23년 소요('92 ~ '15년), TMMR 무전기 19년 소요('96 ~ '25년)
 - 2) 일부 기술변화 속도에 맞는 획득체계 정립 중 (국방부, '20년~)
 - 소요결정前 민간과 군의 요구사항을 받아 과제선정 후 단기간에 신속연구개발 및 신속구매 추진(군 시범운용 → 2년 이내 전력화)
- ☞ 기술진부화 방지 및 성능개선을 위한 진화적 국방 획득체계 활성화, 무기체계 시험평가 수행체계 개선 필요

2. 현 무기체계 시험평가 사례(현실태)(계속)

나. 시험평가 구조(조직 / 평성) 측면

- 1) 합참 / 각군 시험평가 조직 및 체계는 시험(Test)에 한정적으로 운용
 - * 미국은 시험(Test)과 평가(Evaluation)를 구분하여 시행, 시험조직과 별도의 평가조직을 운용 중
 - 2) AI·로봇·자율주행·무인체계 등 첨단 무기체계 시험평가를 위한 전문인력 부족
 - 3) 시험평가 전문인력이 장기간 업무를 수행할 수 있는 인사제도 미흡
 - * 현역(중·소령) 근무기간은 1~2년 이내, 일반직 군무원은 근무기피로 임기제 군무원으로 대체 운용
- ☞ 시험평가의 신뢰성을 강화하기 위한 시험과 평가체계 분리, 평가·지원부서 편성 필요
- ☞ AI기반 유·무인 복합전투체계 등 미래 첨단 무기체계 시험평가를 위한 전문인력 양성 및 전문특기 부여 등 전문인력 풀 구축 필요

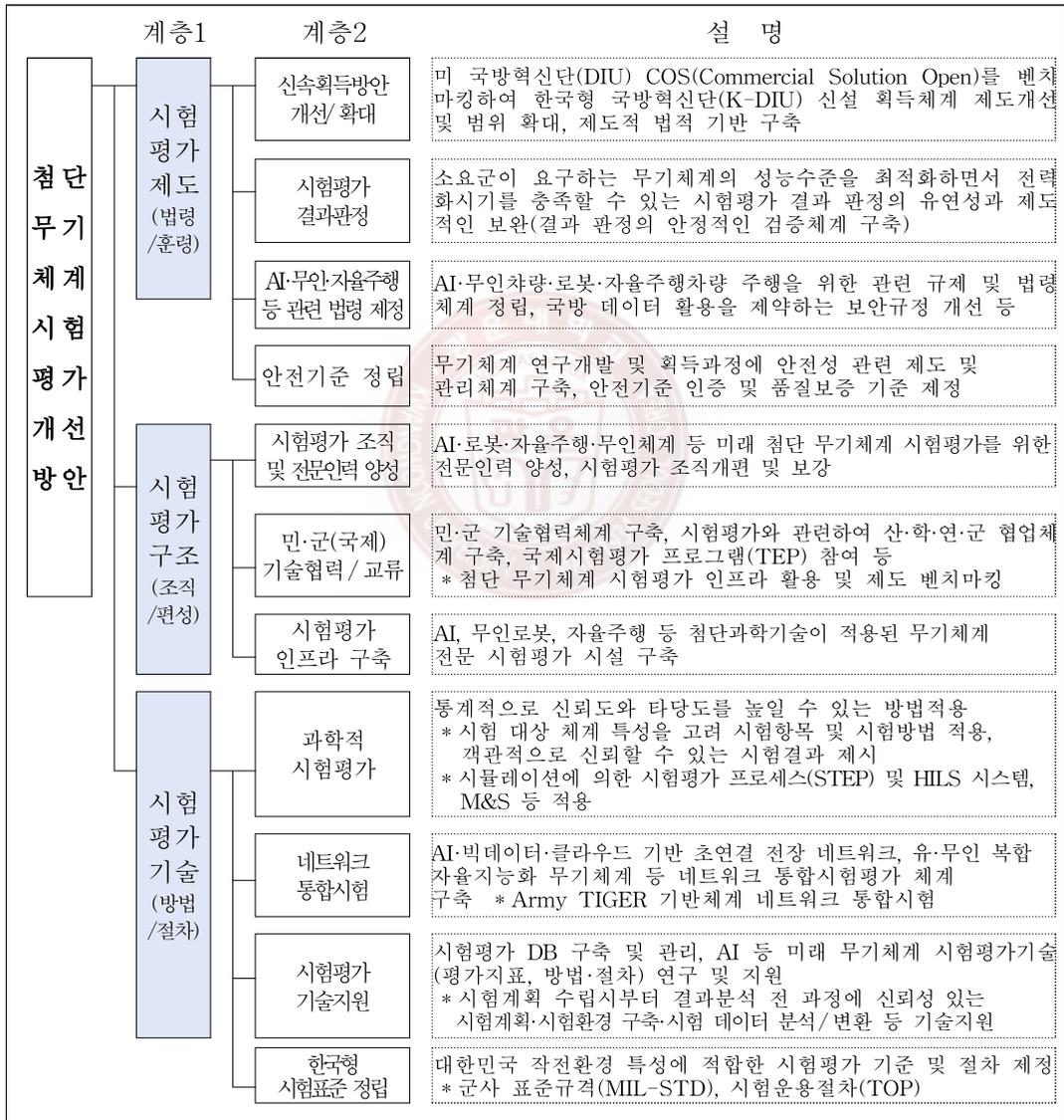
다. 시험평가 기술(방법 / 절차) 측면

- 1) 군용 장비 / 물자에 인공지능(AI) 탑재 ⇨ 인공지능(AI) 시험평가 소요 증가
 - 2) 인공지능(AI) 시험평가 방법에 대한 이견 존재
 - 인공지능(AI) 시험평가간 기존 무기체계와 동일한 시험평가 절차 적용
 - * AI에 대한 시험평가 결과의 충분성·신뢰성 측면에서 불확실성 내재
 - AI특성을 고려한 시험평가 항목에 대한 정의·방법론 등에 대한 이해 차이 발생
 - * AI에 적합한 시험평가 필요성, 시험평가 시기의 적절성, AI 시험평가 방법 등
 - AI 시험평가 결과에 대한 쟁점 제기
 - 3) M&S를 활용한 시험평가는 규정상 가능하지만 실제 시험평가간 적용은 미흡
 - 4) 現 네트워크 장비 시험평가는 시제·인원 등 제한으로 다양한 전투상황을 고려한 충분한 시험평가 제한, 시간과 비용 과다 소요
 - * '18년 네트워크 통합시험 : 기간(6주), 인원(1,100여명), 참가부대(1개 연대(+))
 - * 現 시험평가 기법으로 육군의 Army TIGER기반체계, 드론봇 전투체계는 초연결망을 기반으로 한 무기체계 네트워크 시험평가 제한
- ☞ AI기반 초연결 전투체계인 미래 지상군 Army TIGER를 추진하기 위한 네트워크 통합시험평가 체계 구축 필요
- ☞ AI체계에 대한 데이터 기반 성능평가, 시험평가 방법 및 절차 발전 필요

III. AHP 계층분석 구조

연구자는 지난 6월에 국방전력업무 전문가(45명)를 대상으로 설문을 실시한 결과, 미래 첨단무기체계 시험평가가 발전을 위해서는 시험평가가 제도(법령/훈령), 구조(조직/편성), 기술(방법/절차) 분야에 대해 아래와 같은 세부 평가항목을 도출하였습니다.

이번 설문은 미래 무기체계 시험평가 발전방안에 대해 계층적 분석기법에 의한 각 계층별 평가항목에 대한 가중치를 산출하고자 합니다.



IV. 상대적 중요도 평가시 유의사항

- 본 설문에는 정답이 없습니다. 각 설문 문항에 대해서 너무 오랫동안 고민하지 마시고 상대적으로 중요하다고 생각하시는 항목에 대한 본인의 생각을 바탕으로 솔직하게 응답하시면 됩니다.
- 본 설문은 기준 문항간 상대적으로 중요도를 비교하는 것으로 구성되어 있습니다. 각 기준 항목별 설문 작성은 아래 예문을 참고하여 평가해 주십시오.
예) 시험평가 제도 개선이 시험평가 기술 보다 '중요'하다고 생각하는 경우

왼쪽 기준이 중요 ←														동일	→ 오른쪽 기준이 중요																			
기준항목	절대중요				매우중요				중요				약간중요				같다	약간중요				중요				매우중요				절대중요				비교항목
시험평가 제도	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	시험평가 기술																

- 귀하께서는 각 요소 간 상대적인 중요성을 판단하여 아래와 같은 방법으로 해당 항목에 각각 체크하여 주십시오.

1순위(㉠) > 2순위(㉡) > 3순위(㉢)

㉠ 바나나, ㉡ 딸기, ㉢ 사과

왼쪽 기준이 중요 ←														동일	→ 오른쪽 기준이 중요																			
기준항목	절대중요				매우중요				중요				약간중요				같다	약간중요				중요				매우중요				절대중요				비교항목
㉠	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	㉡																
㉠	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	㉢																
㉡	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	㉢																

* 쌍대비교에서 ㉠은 ㉡보다 2배 좋으며, ㉠은 ㉢보다 6배 좋다는 의미임. 만약 ㉠이 ㉡보다 n배 좋다면 ㉡는 ㉠보다 1/n 배 좋아야 하며 ㉠이 ㉡보다 2배 좋고, ㉡가 ㉢보다 3배 좋다면 ㉠은 ㉢보다 5배 이상 좋아야 서수적 및 기수적 일관성이 유지된 평가라고 합니다.

* 예) ㉠, ㉡, ㉢의 상대적 중요도 평가의 경우 : ㉠이 ㉡보다 2만큼 중요하고, ㉡가 ㉢보다 3만큼 중요하다면? ㉠은 ㉢보다 5~9만큼 중요하다고 평가해야 합니다.

그런데 1~4에 체크를 한다면 일관성에 문제가 있어 Data 활용이 제한될 수 있습니다.

V. 상대적 중요도 평가

1. 4차 산업혁명의 시대, 미래 첨단 무기체계 시험평가 발전을 위해 어떤 요소가 상대적으로 중요하다고 생각하십니까?

평가요소	내 용
시험평가 제도 (법령/ 훈령)	<ul style="list-style-type: none"> ◦ (사전적 의미) 제도(制度)란 관습이나 도덕, 법률 따위의 규범이나 사회구조의 체계¹⁷⁵⁾ ◦ 무기체계 시험평가 관련 법규 및 체계, 무기체계 획득정책, 대외협력, 시험평가 프로세스 및 관련 문서체계 등
시험평가 구조 (조직/ 편성)	<ul style="list-style-type: none"> ◦ (사전적 의미) 구조(構造)란 부분이나 요소가 어떤 전체를 짜 이룸, 또는 그렇게 이루어진 열개, 어떤 일을 성립시키는 것 사이의 상호 기능적 연관¹⁷⁶⁾ ◦ 무기체계 시험평가 조직 및 편성, 민·군(국제) 기술협력 및 교류, 전문인력 양성 및 시험평가 인프라 구축 등
시험평가 기술 (방법/ 절차)	<ul style="list-style-type: none"> ◦ (사전적 의미) 기술(技術)이란 과학 이론을 실제로 적용하여 사물을 인간 생활에 유용하도록 가공하는 수단, 사물을 잘 다룰 수 있는 방법이나 능력¹⁷⁷⁾ ◦ 과학적 시험평가, Army TIGER 기반체계 네트워크 통합시험, 시험평가 표준 및 절차, 시험평가 최적화를 위한 기술지원 프로그램 등

※ 순위 작성 후 순위에 맞도록 아래 설문에 응답(✓)해 주시기 바랍니다.

1순위() > 2순위() > 3순위()

① 시험평가 제도(법령/훈령), ② 시험평가 구조(조직/편성), ③ 시험평가 기술(방법/절차)

왼쪽 기준이 중요 ← → 오른쪽 기준이 중요

기준항목	절대중요		매우중요		중요		약간중요		같다	약간중요		중요		매우중요		절대중요		비교항목
① 시험평가제도	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	② 시험평가 구조
① 시험평가제도	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	③ 시험평가 기술
② 시험평가 구조	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	③ 시험평가 기술

175) 네이버 사전, 국립어학원 표준국어대사전·고려대 한국어대사전

176) 네이버 사전, 국립어학원 표준국어대사전·고려대 한국어대사전

177) 네이버 사전, 국립어학원 표준국어대사전·고려대 한국어대사전

2. 시험평가 제도(법령/훈령) 측면에서 어떤 요소가 상대적으로 중요하다고 생각하십니까?

평가요소	내용
신속획득방안 개선 / 방안	◦ 미 국방혁신단(DIU) COS(Commercial Solution Open)를 벤치마킹하여 한국형 국방혁신단(K-DIU) 신설 획득체계 제도개선 및 범위 확대, 제도적 법적 기반 구축
시험평가 결과 판정 개선	◦ 소요군이 요구하는 무기체계의 성능수준을 최적화하면서 전력화 시기를 충족할 수 있는 시험평가결과 판정의 유연성 등 제도적인 보완(결과 판정의 안정적인 검증체계 구축)
AI·무인·자율주행 등 관련법령 제정	◦ AI·무인차량·로봇·자율주행차량 주행을 위한 관련 규제 및 법령체계 정립, 국방 데이터 활용을 제약하는 보안규정 등 개선
안전기준 정립	◦ 무기체계 연구개발 및 획득과정에 안전성 관련 제도 및 관리체계 구축, 안전기준 인증 및 품질보증 기준 제정

※ 순위 작성 후 순위에 맞도록 아래 설문에 응답(✓)해 주시기 바랍니다.

1순위() > 2순위() > 3순위() > 4순위()

① 신속획득방안 개선/ 확대, ② 시험평가 결과판정, ③ AI·무인·자율주행 등 관련법령 제정, ④ 안전기준 정립

기준항목	← 왼쪽 기준이 중요									→ 오른쪽 기준이 중요									비교항목
	절대중요	매우중요	중요	약간중요	같다	약간중요	중요	매우중요	절대중요	절대중요	매우중요	중요	약간중요	같다	약간중요	중요	매우중요	절대중요	
① 신속획득방안 개선 / 확대	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	② 시험평가 결과 판정 개선	
① 신속획득방안 개선 / 확대	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	③ AI·무인·자율주행 등 관련법령 제정	
① 신속획득방안 개선 / 확대	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	④ 안전기준 정립	
② 시험평가 결과 판정 개선	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	③ AI·무인·자율주행 등 관련법령 제정	
② 시험평가 결과 판정 개선	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	④ 안전기준 정립	
③ AI·무인·자율주행 등 관련법령 제정	⑨	⑧	⑦	⑥	⑤	④	③	②	①	②	③	④	⑤	⑥	⑦	⑧	⑨	④ 안전기준 정립	

3. 시험평가 구조(조직/편성) 측면에서 어떤 요소가 상대적으로 중요하다고 생각하십니까?

평가요소	내용
시험평가 조직 및 전문인력 양성	<ul style="list-style-type: none"> ◦ 시험평가 신뢰성을 강화하기 위한 시험과 평가체계 분리, 평가지원부서 편성 ◦ AI·로봇·자율주행·무인체계 등 미래 첨단 무기체계 시험평가를 위한 전문인력 양성, 시험평가 조직개편 및 보강
민·군·국제 기술협력 / 교류	<ul style="list-style-type: none"> ◦ 민·군 기술협력체계 구축, 시험평가와 관련하여 산·학·연·군 협업체계 구축, 국제시험평가 프로그램(TEP)에 참여 *첨단 무기체계시험평가 인프라 활용 및 제도 벤치마킹
시험평가 인프라 구축	<ul style="list-style-type: none"> ◦ AI, 무인로봇, 자율주행 등 첨단과학기술이 적용된 무기체계 전문 시험평가 시설 구축

※ 순위 작성 후 순위에 맞도록 아래 설문에 응답(✓)해 주시기 바랍니다.

1순위() > 2순위() > 3순위()

① 시험평가 조직 및 전문인력 양성, ② 민·군(국제) 기술협력 및 교류, ③ 시험평가 인프라 구축

왼쪽 기준이 중요 ←

→ 오른쪽 기준이 중요

기준항목	절대중요			매우중요			중요			약간중요			중요			매우중요			절대중요			비교항목													
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	9	8	7															
① 시험평가 조직 및 전문인력 양성	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	② 민·군·국제 기술협력 / 교류
① 시험평가 조직 및 전문인력 양성	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	③ 시험평가 인프라 구축
② 민·군·국제 기술협력 / 교류	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	③ 시험평가 인프라 구축

4. 시험평가 기술(방법/절차) 측면에서 어떤 요소가 상대적으로 중요하다고 생각하십니까?

평가요소	내 용
과학적 시험평가체계 정립	<ul style="list-style-type: none"> 통계적으로 신뢰도와 타당도를 높일 수 있는 방법 등을 통해 구현 시험대상 체계 특성을 고려하여 시험항목 및 시험방법 적용, 객관적으로 신뢰할 수 있는 시험결과 제시 * 시뮬레이션에 의한 시험평가 프로세스(STEP) 및 HILS 시스템, M&S 등 적용
네트워크 통합시험체계 구축	<ul style="list-style-type: none"> AI·빅데이터·클라우드 기반 초연결 전장 네트워크, 유·무인 복합·자율지능화 무기체계 등 네트워크 통합시험평가 체계 구축 * Army TIGER 기반체계 네트워크 통합시험
시험평가 기술지원	<ul style="list-style-type: none"> 시험평가 DB 구축 및 관리, AI 등 미래 무기체계 시험평가 기술 (평가지표, 방법·절차) 연구 및 지원 * 시험계획 수립부터 결과분석 소과정에 신뢰성 있는 시험계획·시험환경 구축·시험데이터 분석/변환 등 기술지원
한국형 시험표준 정립	<ul style="list-style-type: none"> 한반도 작전환경에 적합한 시험평가 기준 및 절차 제정 * 군사 표준규격(MIL-STD), 시험운용절차(TOP)

※ 순위 작성 후 순위에 맞도록 아래 설문에 응답(✓)해 주시기 바랍니다.

1순위() > 2순위() > 3순위() > 4순위()

(A) 과학적 시험평가, (B) 네트워크 통합시험, (C) 시험평가 기술지원, (D) 한국형 시험표준 정립

왼쪽 기준이 중요 ← → 오른쪽 기준이 중요

기준항목	절대중요		매우중요		중요		약간중요		같다	약간중요		중요		매우중요		절대중요		비교항목
(A) 과학적 시험평가	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	(B) 네트워크 통합시험
(A) 과학적 시험평가	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	(C) 시험평가 기술지원
(A) 과학적 시험평가	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	(D) 한국형 시험표준 정립
(B) 네트워크 통합시험	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	(C) 시험평가 기술지원
(B) 네트워크 통합시험	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	(D) 한국형 시험표준 정립
(C) 시험평가 기술지원	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	(D) 한국형 시험표준 정립

5 첨단 무기체계 시험평가와 관련하여 추가의견이 있으시면 제시해 주시기 바랍니다.



소속 기관	병과	현 직책/성명 (직위/계급)	연락처 (H.P)
현재 담당하고 있거나 과거 담당했던 업무를 선택해 주세요. (복수선택 가능)			
① 국방정책 ② 무기체계(전력지원체계) 소요기획 ③ 사업관리 ④ 시험평가 ⑤ 분석평가 ⑥ 연구개발주관기관(방산업체) ⑦ 기타()			
해당분야 근무경력은?			
① 1년 미만		② 1년~3년 미만	
③ 3년~5년 미만		④ 5년~10년 미만	
⑤ 10년~20년 미만		⑥ 20년 이상	

#3 약어표

1. 영문 약어

AI : Artificial Intelligence, 인공지능

ADD : Agency for Defense Development, 국방과학연구소

AEC : Army Evaluation Center, 미 육군평가센터

AHP : Analytic Hierarchy Process, 계층적 분석기법

AR : Augmented Reality, 증강현실

ATEC : Army Test and Evaluation Command, 미 육군시험평가사령부

ATL : Artificial Intelligence Technology Level, 기술수준

CEP : Circular Error Probability, 정확도

CDR : Critical Design Review, 상세설계검토

CDS-MT : Combat Development Support-Management Team,
전투발전지원요소 실무조정회의

CR : Consistency Ratio, 일관성 비율

CSDP : Common Security and Defence Policy, 공동안보방위정책

CSR : Case Study Research, 사례연구

CT : Combined Test, 통합시험

CTT : Combined Test Team, 통합시험팀

DAS : Defense Acquisition System, 국방획득체계

DIB : Defence Innovation Board, 미 국방혁신위원회

DIU : Defence Innovation Unit, 미 국방혁신단

DT&E : Developmental Test and Evaluation, 개발시험평가

DARPA : Defense Advanced Research Projects Agency, 미 고등방위연구계획국

ETRI : Electronics and Telecommunications Research Institute, 한국전자통신연구원

FT : Field Test, 야전운용시험

FSD : Full Self Driving, 완전자율주행

ICT : Integrated Concept Team, 통합개념팀

ICT : Information and Communication Technology, 정보통신기술

IIPT : Institute for Information & communication Technology Planning & evaluation, 정보통신기획평가원

IOC : Initial Operational Capability, 전력화평가

IoT : Internet of Things, 사물인터넷

IPS : Integrated Product Support, 통합체계지원

IPT : Integrated Project Team, 통합사업관리팀

ITEA : International Test & Evaluation Association, 미국 국제 시험평가 협회

ITEP : International Test & Evaluation Program, 국제 시험평가 프로그램

JADC2 : Joint All-Domain Command and Control, 합동 전영역 지휘통제

JAIC : Joint Artificial Intelligence Center, 미 합동 인공지능센터

JCIDS : Joint Capabilities Integration and Development System, 소요관리체계

JSOP : Joint Strategic Objective Plan, 합동군사전략목표기획서

KIDA : Korea Institute for Defense Analysis, 한국국방연구원

KISA : Korea Internet & Security Agency, 한국인터넷진흥원

KRIT : Korea Research Institute for defense Technology planning and advancement, 국방기술진흥연구소

LCC : Life Cycle Cost, 수명주기 비용

MECM : Mutually Exclusive·Collectively Exhaustive, 상호배제, 전체포괄

M&S : Modeling and Simulation, 모형화와 모의

MR : Mixed Reality, 혼합현실

MRA : Manufacturing Readiness Assessment, 제조성숙도

OMS/MP : Operational Mode Summary/Mission Profile, 운용형태요약/임무유형
ORD : Operational Requirements Document, 운용요구서
OTC : Operation Test Command, 미 운용시험사령부
OT&E : Operational Test & Evaluation, 운용시험평가
PDR : Preliminary Design Review, 기본설계검토
PEGASUS : Pioneer & Explore the Ground And Space for United Space
Operations, 육군 우주력 발전계획으로 미래 육군을 상징
PPBEES : Planning·Programing·Budgeting·Execution·Evaluation, System
기획·계획·예산편성·집행·분석평가
P-TEMP : Pre Test & Evaluation Mater Plan, 예비시험평가기본계획서
R&D : Research and Development, 연구개발
RFP : Request For Proposal, 제안요청서
ROC : Required Operational Capability, 작전운용성능
SCT : Standard Conformance Test, 표준적합성시험
SE : Systems Engineering, 체계공학
SFR : System Functional Review, 체계기능검토
SI : System Integration, 체계통합
SIPRI : Stockholm International Peace Reserch Institute, 스톡홀름 국제평화연구소
SRD : System Requirement Document, 체계요구서
SRR : System Requirements Review, 체계요구조건검토
SSRS : System Subsystem Requirements Specification, 체계요구사항명세서
SSS : System Subsystem Specification, 체계규격서
STAT : Scientific Test and Analysis Technique, 과학적 시험 및 분석기법
T&E : Test and Evaluation, 시험평가
TEMP : Test & Evaluation Mater Plan, 시험평가기본계획서

TIGER : Transformative Innovation of Ground forces Enhanced by the 4th industrial Revolution technology, 4차 산업혁명 기술로 강화된 지상군으로 첨단 과학기술군의 결정체를 상징

TRA : Technology Readiness Assessment, 기술성숙도 평가

TRL : Technology Readiness Level, 기술성숙도 수준

TRR : Test Readiness Review, 시험준비 검토

URD : User Requirement Document, 사용자요구서

VR : Virtual Reality, 가상현실

2. 국문 약어

국과연 : 국방과학연구소

국기연 : 국방기술진흥연구소

방사청 : 방위사업청

방산기술센터 : 방위산업기술지원센터

소요제기기관 : 무기체계 소요를 제기하는 육·해·공군을 의미함

시평단 : 육군 시험평가단

시평부 : 합참 시험평가부

체계공학 : 시스템엔지니어링

합참 : 합동참모본부

