훈련결과보고서

인공지능(AI)의 개발 추이와 활용 가능한 국방사업 분야에 관한 연구 (국방로봇 윤리에 대한 고찰 포함)

방 위 사 업 청
이 정 휘

<목 차>

<국외훈련 개요>

1. 훈련국 : 캐나다

2. 훈련기관명 : RIFKIND LAW

3. 훈련분야 : 직무

4. 훈련기간 : 2019. 11. 15. ~ 2020. 11. 14.

<훈련기관 현황>

1. 주소 : 5001 Yonge Street, Suite 301, Toronto, ON M2N 6P6

2. 전화번호 : +1 416-222-4597

3. 이메일 : Bradley@yigalrifkind.com

4. 훈련기간 성격 : 전문 로펌 및 인공지능 관련 컨설팅

제1장. 연구의 배경 및 목적

이미 인공지능(AI) 기술은 현재 세계 각국에서 4차 산업혁명에 있어서 핵심 중의 핵심으로 각 분야에서의 적용이 되고 있어, 마치 그 파고가 쓰나미처럼 우리에게 다가오고 있다. 특히 2006년 기나긴 인공지능 분야의 겨울을 봄으로 바꾸어 놓은 캐나다 제프리 힌튼 교수의 딥러닝 개념이 처음 소개된 이후 불과 15년이 지났지만, 인공지능 분야는 국방 분야에서 뿐만 아니라 모든 산업 분야에 있어서 필수적인 기술이 되었고, 이에 대한 연구가 더욱 절실히 필요한 때이다.

2012년 미국 네바다(Nevada)주에서 처음으로 일반도로에서 자율주행차의 합법화(2014년 운전자의 통제 가능성을 전제로 개정함)를 시작으로하여, 2020년 5월 1일 우리나라는 자율주행차법(자율주행자동차 상용화 촉진 및 지원에 관한 법률)을 시행하여 법적으로 자율주행 택시 서비스도 가능하게 되었으며, 중국은 2035년 완공을 목표로 베이징 근교에 자동차 전용 도시를 건설 중이다. 이렇듯 세계는 지금 무인자동차 뿐만 무인항공기와 드론과 같은 무인장비들이 개발되고 있고, 이의 핵심 기술은 다름이 아닌 인공지능(AI) 그 중에서도 데이터를 통해 스스로 학습 하는 머신러닝(ML) 분야이다.

해마다 혁신적인 전략 기술을 조사하고 발표하고 있는 시장조사기관 가트너의 '2020년 10대 기술 트렌드(Gartner Top Strategic Technology Trends)' [1] 를 보면 올해는 예전과는 다르게 두 가지 분류 즉 사람 중심(people-centric, 고객이나 임직원 등에 영향을 주는 기술)과 스마트 공간(smart spaces, 집, 사무실, 자동차 등에 영향을 주는 기술)로 나누어서 소개를 하고 있는 것이 특징이며, 그 기술들을 살펴보면 다음과 같다.

1) 가트너 선정 '2020년 10대 기술 트렌드 동향
www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020

- 사람 중심(people-centric) 기술

초자동화 (Hyperautomation)	**인공지능(AI)와 머신러닝(ML)을 포함한 기술**들을 이용하여 기존 자동화의 범주를 넓혀 주며 더욱 정교한 작업(발견, 재평가, 디자인 등)도 가능함. 예를들면 RPA(Robot Process Automation), iBPMS(intelligent business management software), Digital Twin[2] 등이 있음.
다중 경험 (Multiexperience)	다중경험이 기술을 이해하는 사람(technology-literate people)을 사람을 이해하는 기술(people-lilterate technology)이 대체할 것으로 전망함. 가상현실(AR), 증강현실(AR), 혼합현실(MR) 뿐만 아니라 다중채널 휴먼머신 인터페이스(HMI) 등이 있음.
민주화 (Democratization)	기술의 민주화는 사람들에게 기술이나 사업분야에 집중적인(비용이 드는) 교육 없이도 쉽게 접근할 수 있도록 하는 것이며, 4대(응용 개발, 데이터와 분석, 디자인 그리고 지식) 분야에 집중될 것이라 전망하며, 이것은 **코드와 자동 시험을 생성하는 인공지능(AI) 주도의 개발**에 의존될 것임.
휴먼 증강 (Human Augmentation)	휴먼 증강은 사람의 인지와 물리적인 경험을 강화하기 위한 기술로 사용되며, 웨어러블 장치들을 사용하여 물리적인 증강을 가질 수 있어 작업자의 생산성을 증가시킨다. 물리적인 증강의 주요 4대 카테고리는 1.센서 증강 2.바이오 기능 증강 3.뇌 증강 4.유전적 증강임.
투명성과 추적성 (Transparency and Traceability)	기술의 진화는 사용하는 데이터에 대한 신뢰의 문제로 이어지게 되며, 기관들은 데이터와 저장과 수집에 책임을 지고 있습니다. 게다가 **인공지능(AI)와 머신러닝(ML)은 갈수록 인간을 대신하여 결정을 내리는데 사용되므로, 이는 설명이 가능한 인공지능(explainable AI)와 AI 거버넌스**와 같은 방식이 필요해짐. 6대 요소는 윤리, 통합, 공개, 신뢰, 권한 및 일관성임.

- 스마트 공간(smart space) 기술

강화/자율권을 가진 에지 (The Empowered Edge)	정보 처리, 콘텐츠 수집과 전달이 해당 정보가 발생하는 장소, 소비자와 가까운 곳에서 처리되는 컴퓨팅 토폴로지(Topology)임. 즉 중앙 집중식 이 아닌 데이터가 생성되는 네트워크 엣지와 가까운 곳에서 데이터를 처리하는 방식이며, IoT과 5G 상용화에 따라 급속이 확산될 것임.

2) 디지털 트윈(Digital Twin)의 개념은 GE社에서 만들었으며, 물리적인 사물과 컴퓨터에 동일하게 표현되는 가상 모델을 만들어 모의실험을 함으로써 실제 상황의 특성에 대해서 정확히 예측을 할 수 있다. 에너지, 항공, 헬스케어, 자동차, 국방 등 여러 산업 분야에서 자산 최적화, 돌발 사고 최소화, 생산성 증가 등 설계부터 제조, 서비스에 이르는 모든 과정의 효율성을 향상시킬 수 있다.

분산 클라우드 (Distributed Cloud)	클라우드 공급자의 물리적인 데이터 센터를 벗어나 다양한 장소로 분산되는 클라우드 서비스이며, 이러한 변화는 클라우드 컴퓨팅 시대를 견인하나, 보안문제가 이슈가 되고 있음.
자율 사물 (Autonomous Things)	로봇, 자율주행 자동차, 드론 등 자율적으로 행동사물이며, 이는 **인공지능(AI)을 활용**한 이전과는 다르게 보다 자연스럽게 상호 작용하는 고급 행동을 구현하여 점점 더 공개된 공공장소에 배치될 것임.
실용적인 블록체인 (Practical Blockchain)	신뢰성 구축, 비용절약, 시간단축, 등의 장점 등으로 블록체인은 많은 기업들에 의해 도입이 되고 있으며, **AI, IoT와 통합이 시작되면** 점차적으로 “진정한 블록체인”이 산업계를 변화시킬 가능성이 있음.
AI 보안 (AI Security)	초자동화와 자율 사물과 같은 기술이 진화하면서 보안 취약성 역시 새로운 공격 대상이 되며, 보안 담당자들은 **AI 기술을 기반으로 한 시스템 보호, AI를 활용한 보안방어 기술향상, AI를 통한 공격자의 범죄예측 및 예방등에 더욱 집중하고 대비**해야함.

위에서 보는 바와 같이 향후 인공지능(AI) 기술은 10가지 핵심 기술 트렌드 중 6가지에는 직접적으로 그 외에는 간접적으로 영향을 주고 있으며, 이는 현재 뿐만 아니라 향후에도 사회 전반에 걸쳐 매우 중요한 기술로 국가 경쟁력과도 직결이 되고 있음을 알 수 있다.

또한 여러 악조건과 특수상황 등에서 적용되는 국방분야에서의 첨단화된 기술은 항상 기술을 선도해 오고 있으며, 인공지능(AI)기술 역시 이미 많은 국가에서 그 적용 및 활용에 관한 연구가 되고 있으며, 이에 대한 기술 개발 및 적용 수준은 곧바로 국가를 안전하게 뒷받침해주는 국방력으로 이어질 것이며, 또한 민간분야로 활용이 되며 국가 산업 경쟁력으로도 이어져 선순환을 하게 된다.

이미 미국의 경우 국방부의 인공지능 전략(Department of Defense Artificial Intelligence Strategy:Harnessting AI to Advance Our Security and Prosperity)의 제목에서 알 수 있듯, AI 기술에 박차를 가하면서 이는 안전 뿐만 아니라 번영으로 이끌 것으로 기대하며 국방부에서 전략을 짜고 있는 것이다.

또한 인공지능(AI) 기술은 그 자체로 어떤 좋고 나쁨이 있는 것이 아닌 우리가 어떻게 사용하는가에 그 결과가 나타나게 되어있다. 최근 해외 로봇공학자들의 카이스트와 한화시스템의 인공지능(AI) 무기 개발에 대한 '보이콧 선언'으로 제기된 국방로봇에 대한 윤리적인 문제가 대두되고 있으며, 이미 UN에서는 이를 LAWS(Lethal Autonomous Weapons Systems)로 부르고 있으며, CCW(Certain Conventional Weapons) 회의에서 각 정부의 전문가 모임(GGE, Group of Governmental Experts)에서 윤리적인 문제 등을 검토를 하고 있다.

그리고 캐나다의 경우 인공지능(AI) 암흑기에도 꾸준한 투자로 인공지능 기술에 획기적인 전환점을 만들었으며, 이는 곧 캐나다를 인공지능 분야에 있어서 세계 최고 수준으로 이끌었다. 또한 2017년 약 200명의 인공지능(AI) 연구자들이 각 중요 부처(과학부, 외교부, 국방부 등)를 포함하여 직접 캐나다 총리에게 공개 서한을 통하여 LAWS에 인공지능(AI) 기술이 악용되지 못하도록 요청하는 등 윤리적인 분야에 있어서 노력을 하고 있다.

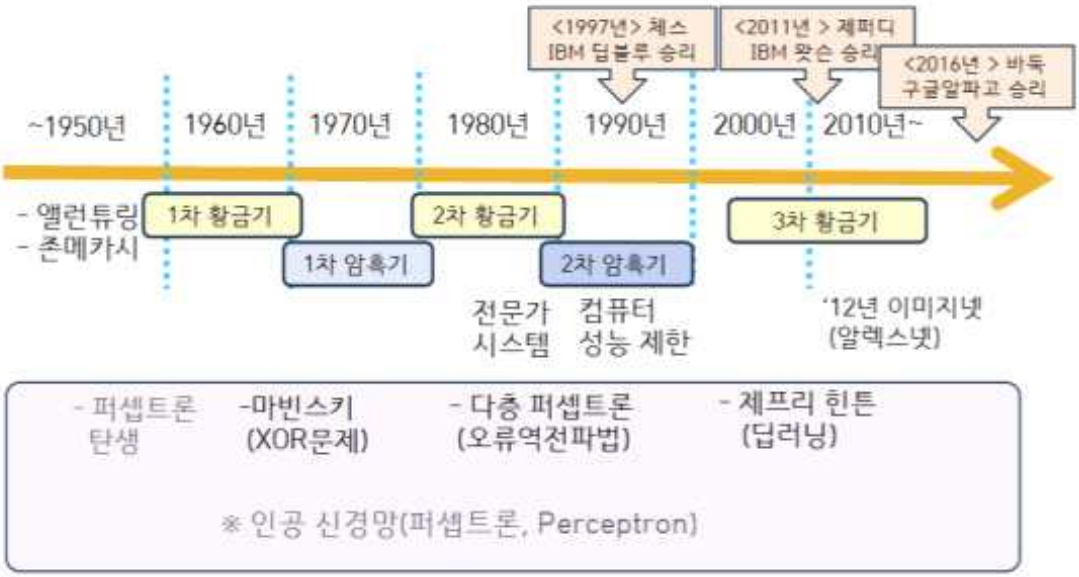
따라서 본 과제 연구를 통하여 훈련국인 캐나다와 같은 북미지역으로 연관성이 많은 미국을 중심으로 인공지능의 개발 현황 및 국방과 관련된 진행사항을 검토하여, 향후 우리나라의 인공지능 기술 국방에 적용이 될 수 있는 분야 등을 확인하며, 또한 국방 분야에 있어서 문제가 될 수 있는 윤리적인 문제 역시 같이 검토를 함으로써 튼튼한 국방과 동시에 윤리적인 부분에 있어서도 문제가 되지 않는 참고자료가 될 수 있으며, 끝으로 이런 부분들이 다시 민간 부문으로 적용되어 국가 경쟁력 육성에도 기여하고자 한다.

제2장. 인공지능 개관 및 주요 국가의 인공지능 개발 추이

1. 인공지능의 흐름

인공지능(AI)이란 개념은 오래전부터 존재하였으며, 최근의 하드웨어와 신경망 알고리즘의 발전과 더불어 미래 기술의 핵심이 되었다. 하지만 그 과정이 순탄치 않았으며, 아래와 같이 몇 번의 암흑기가 있었지만, 포기하지 않고 그 길을 개척한 선구자들에 의해 그 난관을 극복할 수 있었다.

< 인공지능(AI)의 역사 >



< 출처 : 저자 직접 작성 >

1) 1950년대

- 1950년 영국 수학자 앨런 튜링(Alan turing)은 Computing Machinery and Intelligence 논문을 통해서 학습하는 기계에 대해서 처음으로 기술함

- 1956년 다트머스대학 존 메카시는 다트머스회의에서 처음으로 AI 용어를 언급

- 1958년 코넬대 프랭크 로젠블랫의 연구에서 퍼셉트론(인공신경뉴런)이 탄생하였으며, 이는 향후 딥러닝의 기초가 된다.

2) 1960년대

- 1969년 마빈스키와 세이무어 페퍼트는 XOR문제는 선형분리가 불가함을 수학적으로 증명하는 '퍼셉트론' 이란 책을 펴냄으로써 미국방부(DARPA)는 인공지능(AI)에 연구자금을 중단하는 등 1차 암흑기를 맞는다.

3) 1980년대

- 연역적 방법의 전문가시스템(Expert System)은 전문가들의 지식을 논리적 법칙으로 변환하여 전문가 대신 답해주는 시스템으로 1980년대에 도입되었지만, 관리적 문제와 효용성의 문제 발생함

- 그리고 이시기에 지금의 딥러닝을 이루고 있는 다층 퍼셉트론 등의 혁신적인 연구들이 발표되었으나, 컴퓨터 성능과 알고리즘의 제한으로 빛을 보지 못하였다.

4) 2000년대 이후

- 2006년 캐나다 토론토대학의 제프리힌튼 교수는 두 번째 암흑기에 대한 해결책이었던 Deep belief nets에 관한 논문을 내면서 신경망 이론으로 다시 한번 부각시키며 딥러닝에 대한 부흥을 일으킴

- 2012년 국제이미지인식기술대회(ILSVRC)에서 토론토대 제프리 힌튼 교수팀은 딥러닝 기반으로 이미지 분석을 하는 '알렉스넷(Alexnet)'으로 우승을 차지[3]하였다.

3) 국제이미지인식기술대회는 1000개의 칱테고리와 100만개의 이미지로 이미지인식의 정확도를 겨루는 대회로, 2012년이전까지는 신경망 방식이 아닌 다른 기계학습 방법이 대다수 우승을 하였으나, 2012년 토론토대 제프리 힌튼 교수팀은 나선형 신경망(CNN)을 이용하여 알렉스넷(Alexnet)으로 우승을 차지하였다. 이는 약 84.7%의 정확도였으며, 2등은 74% 밖에 되지 않았다. 이후 2015년에는 MS팀이 딥러닝 방법으로 96%의 정확도로 우승을 차지하였다. 이는 인간의 오차율인 5%보다 우수한 것으로 이후 인공지능(AI) 기계학습의 주류가 딥러닝으로 변화하게 만들었다.

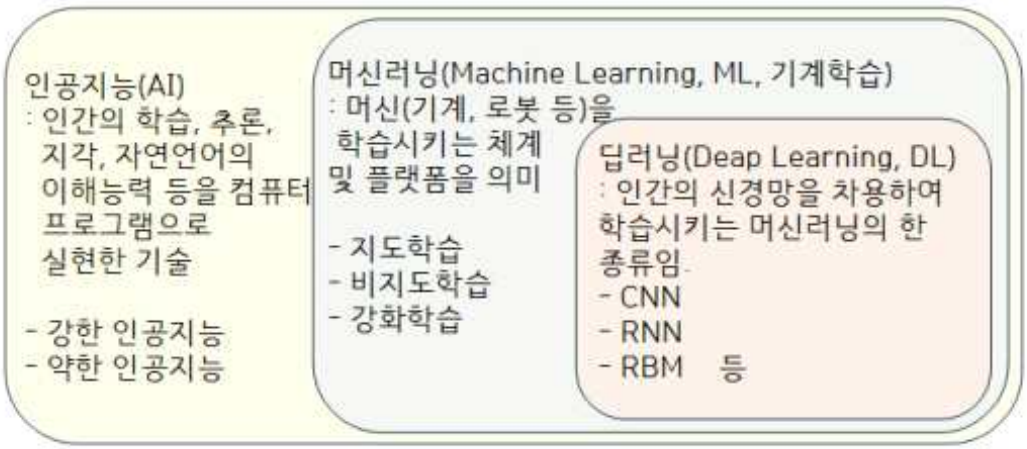
2. 인공지능의 분류

인공지능(AI)은 크게 강한 인공지능(Strong AI)와 약한 인공지능(Weak AI, narrow AI)가 있다. 강한 인공지능은 사람과 같은 지능을 가지고, 사람처럼 느끼면서 지능적으로 행동하는 기계를 말하며, 이는 앨런 튜링이 제안한 튜링테스트[4]로 판단하고 있다. 그리고 약한 인공지능은 특정 문제를 해결하는 지능적인 행동을 하는 기계를 말하며, 예를 들어 이세돌을 이긴 알파고 역시 바둑이라는 특정 분야만 해결할 수 있으므로 약한 인공지능이라 할 수 있다. 아직은 인간과 같은 강한 인공지능을 개발하는 데는 많은 어려움이 있으나, 지속적으로 사람과 같은 인공지능을 만들기 위해서 노력하고 있으며, 이런 인공지능 개발에 적용되는 방법에 대한 분류를 알아 볼 필요가 있다.

1) 인공지능 vs 머신러닝 vs 딥러닝

보통 인공지능(AI)을 이야기할 때 인공지능, 머신러닝, 강화학습과 딥러닝에 대한 정확한 의미를 알지 못할 때, 서로 혼동하며 잘못 사용하는 경우가 있으며, 일반적으로 사용되는 관계와 범주를 아래의 그림으로 정리해 보면 다음과 같다.

< 인공지능(AI)의 분류 >



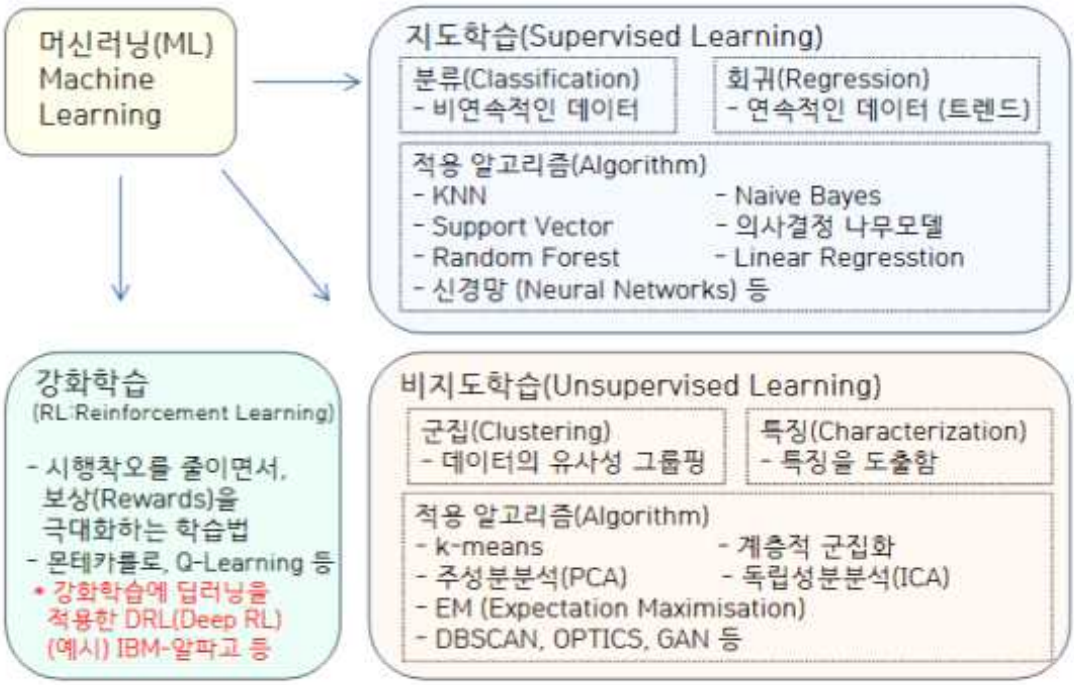
< 출처 : 저자 직접 작성 >

4) 튜링테스트(Turing Test) : 지금은 인공지능을 판별하는 기준이 된 튜링테스트는 사람을 격리된 방에 두고 상대방이 누구인지 모른 상태로(상대가 기계인지 모른 상태) 대화를 나눈 후, 상대방에 대해서 이상한 점을 발견하지 못하고 있다면 컴퓨터가 최소한 인간 정도의 지능을 가지고 있다고 판별하는 방법이다.

2) 머신러닝

머신러닝은 Learning 이라는 단어에서 알 수 있듯이 수집되어 입력된 데이터로부터 스스로 학습을 하여 성능을 향상시키는 기술이다. 이런 머신러닝은 앞의 그림에 나온 것과 같이 학습데이터의 제공 방식에 따라 일반적으로 지도학습(Supervised Learning, 정답이 있는 데이터 사용), 비지도학습(Un-supervised Learning, 정답이 없는 데이터 사용), 강화학습(Reinforcement Learning)으로 나눌 수 있다.

< 머신러닝(ML)의 분류 >



< 출처 : 저자 직접 작성 >

이러한 분류에 초창기 딥러닝은 지도학습으로 분류를 하였지만, 뒷부분에 잠시 언급이 되는 것과 같이 지금은 비지도학습과 강화학습 모두에 적용되어 머신러닝 분야의 비약적인 발전을 이루어 강한 인공지능 개발에 대한 기대를 높이고 있다. 그렇다면 이 인공신경망 기반의 딥러닝은 위의 인공지능 역사 도표에서 볼수 있듯이 오래된 개념의 기술이지만, 지금 전성기를 누리는 이유는 대략 3가지로 들 수 있다.

① Big Data - 데이터를 많이 적용할수록 높은 성능구현

② 하드웨어의 발전 - 복잡한 알고리즘의 처리속도를 가능케 함

③ 소프트웨어의 발전 - 개선된 기법과 텐서플로우 등과 같은 툴 개발

- 지도학습 (Supervised Learning)

입력된 데이터에 Labeling이 되어있고, 출력값으로 사상되는 함수(활성화함수)를 학습하여 성능을 향상시키는 방향으로 학습하는 것임.

그리고 지도학습이 다루는 문제들은 주로 분류(Classification)과 회귀(Regression)와 관련된 것들로 분류는 값의 선택 즉, 맞다, 아니다(이진 분류 문제) 혹은 개, 고양이, 토끼(다중 분류 문제) 등을 분류하는 것이며, 회귀는 어떤 데이터들의 특징(Feature)[5]를 기준으로 연속된 값(그래프)을 예측하는 문제로 주로 트렌드, 경향을 예측할 때 사용됨.

또한 다음은 지도학습에 사용되는 알고리즘의 일부이다.

알고리즘(Algorithm)	설 명
KNN (K-Nearest Neighbors)	기본 모델로 가장 근접한 값을 찾아내는 방법으로 이해하기가 용이함
Naive Bayes	분류를 하기 위한 확률적 모델로써 인과관계를 단순화 한 모델이며, 변수간의 의존관계를 고려한 것이 Bayesian Network임
SVM (Support Vector Machine)	마진(margin, 클래스 사이들의 간격)을 최대화하는 이진분류알고리즘으로, 비선형도 분류할 수 있으며, 빠른학습이 가능
의사결정 나무모델 (Decision Tree)	분류에 사용되며, 설명변수(feature)들의 관계 등으로 목표변수(label)를 분류하는 나무구조의 모델로 적용(시각화에 좋음)
Random Forest	Decision Tree의 정확도를 개선하기 위해, 여러개의 나무를 생성하여 각 나무의 결과를 종합적으로 판단하여 결론을 내리는 구조
선형 회귀 (Linear Regression)	데이터의 트렌드를 가장 잘 표현하는 1차 함수적 모델인 선을 찾는 방법
신경망 (Neural Networks)	일명 '딥러닝'으로 불리며, 인간의 뉴런(Neurun) 구조를 모방, 다층 구조의 뉴런, 활성화함수, 잡음에 견고한 구조

5) 특징(Feature) : 전통적인 머신러닝에서는 특징을 정의하는 엔지니어가 미리 정의한 후 머신러닝으로 결과를 도출하였지만, 요즘은 딥러닝 발달로 Raw data를 그대로 입력하면 DNN 기반의 알고리즘 등은 직접 그 특징을 추출하여, 분류 혹은 회귀를 수행하고 있다. 하지만 머신러닝 알고리즘에만 의존하면 자칫 잘못된 해석이 될 수 있으므로, 이 특징에 대한 처리 전,후 처리작업은 중요함

- 비지도학습 (Unsupervised Learning)

데이터의 Labeling이 되어있지 않은 데이터를 입력받아 그 데이터들의 공통적인 특성을 파악하는 것이 목적이며, 대표적으로는 군집(Clustering)과 특성(Characterization)이 있다. 정답이 없는 데이터에서 사람도 알지 못하는 본질적인 문제나 데이터에 숨어져있는 특징 등을 연구할 때 많이 활용된다. 또한 스스로 학습하는 특징에 따라 인간과 같은 수준의 인공지능(AI)을 가지기 위해서는 이 비지도학습법에 가장 적합하며, 이에 최근 연구가 많이 이루어지고 있는 상황이다.

비지도학습에는 다음과 같은 알고리즘들이 사용되고 있다.

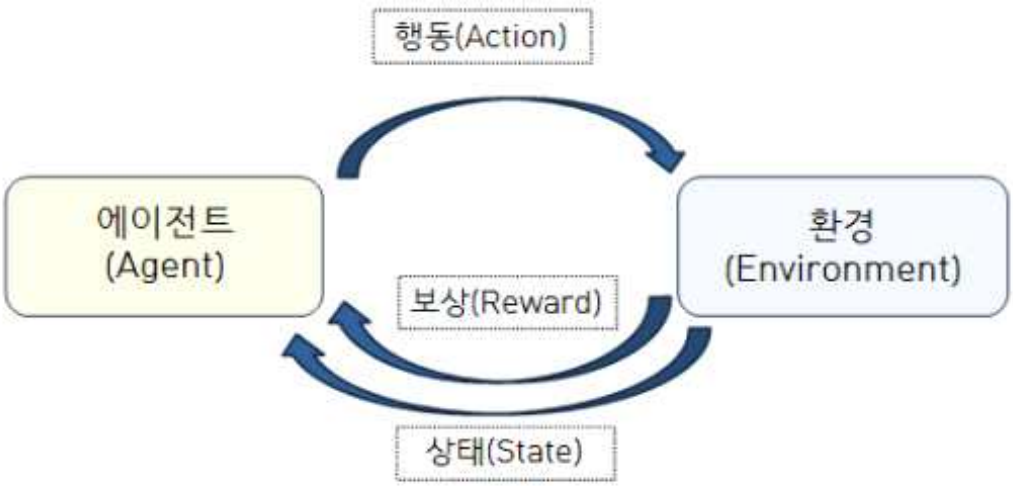
알고리즘(Algorithm)	설 명
k-means	군집별 중심에서 거리를 기반으로 그룹을 분류하여 k개의 군집으로 나누며, 특징학습에 유효함
계층적 군집화	전체 데이터를 가까운 데이터끼리 계층적으로 그룹핑하는 방법
주성분분석(PCA)	고차원의 데이터를 저차원으로 환원시키는 기법이며, 데이터마이닝에 사용
독립성분분석(ICA)	독립성분 분석 및 요인을 분석함
EM(Expectation Maximization)	통계 모델이며, 최대 가능도의 획득을 함
DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	기하학적 인덱스 구조를 사용한 비모수 데이터 알고리즘이며, 일반적으로 군집화에 사용됨
OPTICS (Ordering Points To Identify the Clustering Structure)	DBSCAN 비슷하며, 변동성 포인트 밀도를 다루는 것이 차이이며, 군집화에 사용됨
GAN(Generative Adversarial Network)[6]	임의 변수와 확률분포를 활용하여 새로운 데이터를 생성하는 기법임.

머신러닝의 기본 원리인 학습을 하기위해서는 많은 데이터가 필요로 하였지만, GAN은 스스로 유사 데이터를 생성할 수가 있어 스스로 진보된 학습을 할 수가 있다. 이는 생성자(Generator)가 새로운 이미지를 생성하면 판별자(Discriminator)가 진위 여부를 확인하는 과정이 반복하게 된다. 활용사례로는 가짜 지폐, 진짜와 같은 사람 사진 등이 있다.

6) 2014년 GAN이 처음 논문으로 나왔을 때, 딥러닝 분야의 최고 석학인 페이스북 AI 연구팀의 얀 르쿤(Yann LeCun) 교수는 GAN을 최근 10년간 머신러닝 연구 중 가장 혁신적인 아이디어라 함.

- 강화학습 (Reinforcement Learning)

입력된 데이터(행동)의 선택에 따라 보상을 줌으로써, 시행착오를 통해 보상을 극대화하는 것을 목표로 하는 학습방법으로 강화학습에 필요한 개념으로는 5가지(에이전트, 환경, 상태, 행동, 보상)가 있다.



< 출처 : 저자 직접 작성 >

강화학습이 자주 사용되는 게임을 보자면, 에이전트(Agent)가 게임 환경(Environment)에서 현재상태(State) 보다 높은 점수(Reward)를 얻어가는 행동(Action)을 학습 방법으로 일정 반복을 통해서 높은 보상을 얻는 전략이 만들어진다. 그리고 이세돌을 이긴 IBM의 알파고는 강화학습(몬테카를로)기법과 함께 효율적으로 할 수 있도록 딥러닝 기법을 결합한 DRL(Deep Reinforcement Learning)을 적용하여, 획기적으로 성능을 향상시켰다.

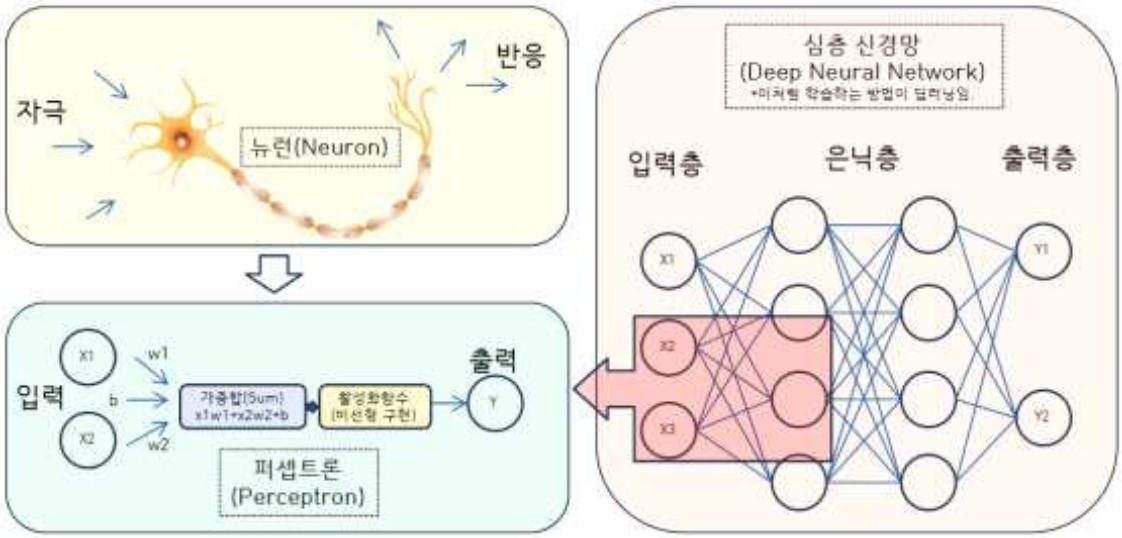
강화학습에는 다음과 같은 알고리즘들이 사용되고 있다.

알고리즘(Algorithm)	설 명
몬테카를로	난수를 사용하여 함수의 값을 확률적으로 계산함
Q-Learning	유한마르코프 결정과정(Finite Markov Decision Process) 적용
DQN (Deep Neural Network)	Q-Learning을 'Deep Neural Network'로 구성한 학습 방법
A3C (Asynchronous Advantage Actor - Critic)	여러개의 agent를 만들어 DQN의 단점인 시간의존 등이 개선됨

3) 딥러닝

딥러닝은 강화학습과 더불어 머신러닝의 르네상스를 오게 하였으며, 이는 인간의 신경망(뉴런) 구조를 그대로 컴퓨터에 적용 하여 학습능력을 극대화하는 기술이다. 다른 머신러닝 기법과는 다르게 많은 양의 데이터를 필요로 하며, 그럴 경우 더욱 좋은 결과를 얻을 수 있다. 이러한 딥러닝 기술은 컴퓨터 비전(컴퓨터를 사용하여 사람의 눈과 같은 기능을 하도록 구현하는 기술, 이미지 인식 등), 음성인식, 자연어처리, 음성 및 신호 처리 등의 분야에 적용되어 놀라운 발전을 이루고 있다.

< 뉴런, 퍼셉트론, 딥러닝 >

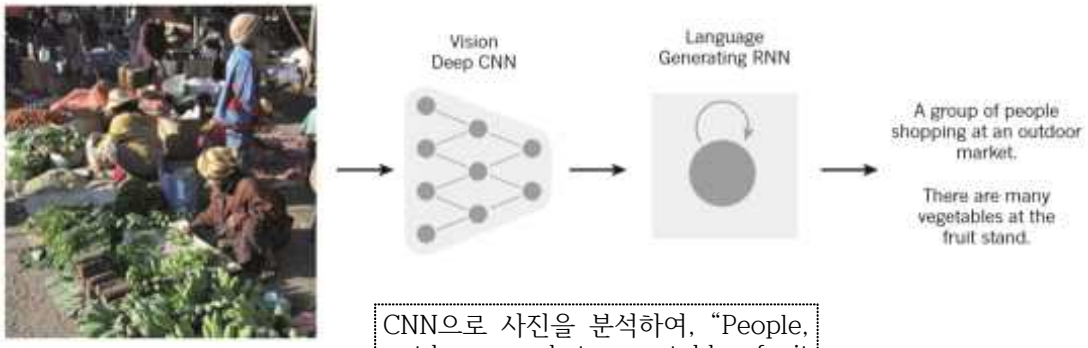


< 출처 : 저자 직접 작성 >

딥러닝에는 다음과 같은 알고리즘들이 사용되고 있다.

알고리즘 (Algorithm)	설 명	적 용
DNN (Deep Neural Network)	입력층과 출력층 사이에 다중의 은닉층으로 구성된 신경망	범용, 이미지 인식
CNN (Convolution Neural Network)	인간의 시신경구조를 모방한 vision처리 수행 모델	컴퓨터 비전
RNN (Recurrent Neural Network)	은닉층에서 출력층간 데이터의 저장 및 흐름 가능한 신경망	언어모델링, 기계번역, 이미지캡션 생성
RBM (Restricted Bolzman Machine)	Hidden node와 Visual node로 구성된 무방향 그래프	DBN의 기본단위
DBN (Deep Belief Network)	RBM이 적층되어있는 구조	MNIST(손글씨) 분류

그리고 각 알고리즘들은 단독으로만 사용되지 않고, 좋은 결과를 얻기 위하여 서로 연관되어 사용되어진다. 아래는 사진을 이미지 인식으로(CNN 사용) 주요한 특징을 추출한 후, 그 특징을 바탕으로 언어모델링으로(RNN 사용) 문장을 출력한 예를 보여준다.[7)]



CNN으로 사진을 분석하여, "People, outdoor market, vegetable, fruit stand" 와 같은 단어를 추출하여 RNN으로 전해주면 RNN에서 문장을 최종적으로 생성한다.

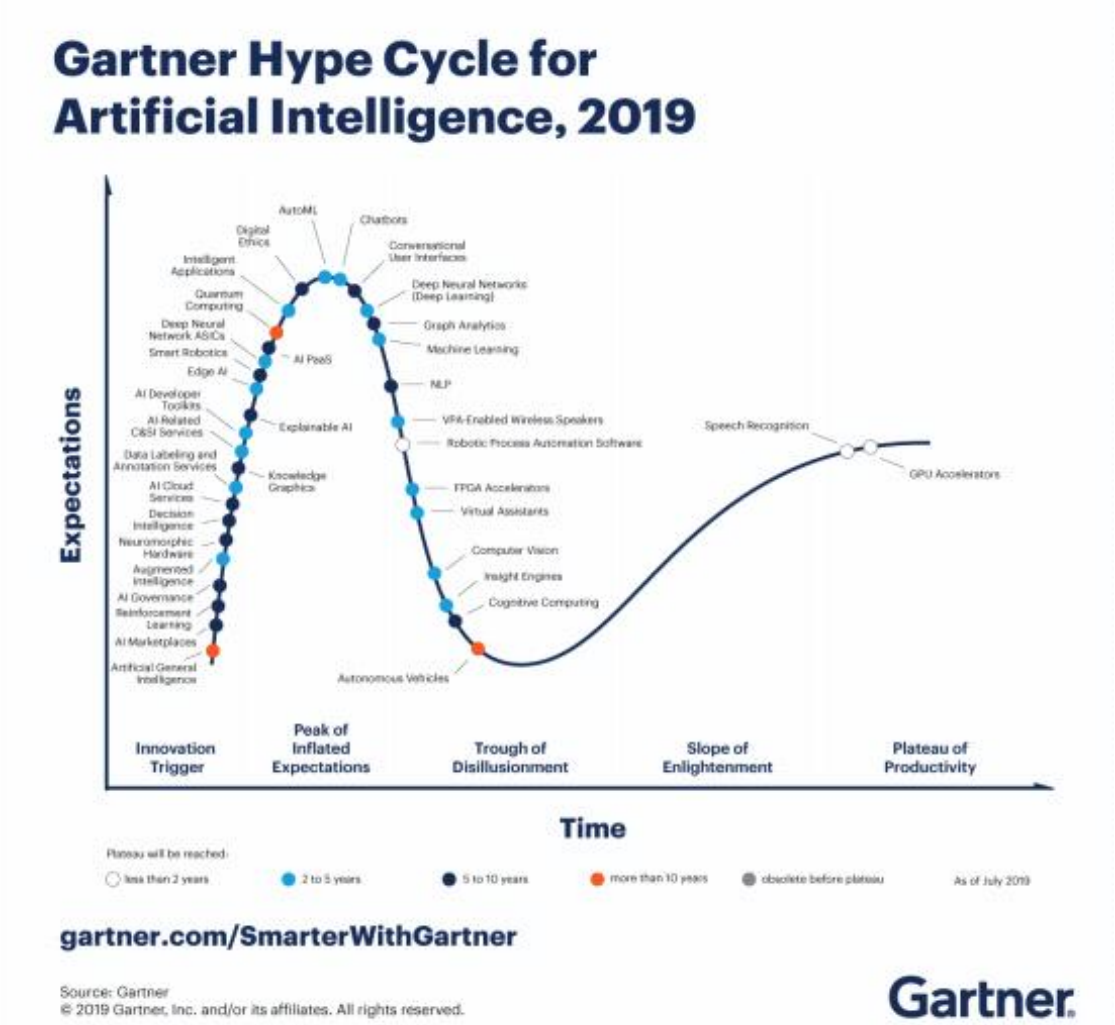
< 출처 : Deep learning, Yann LeCun 등 3인, 각주 7 참조 >

이러한 인공신경망 기반의 딥러닝 학습법은 지도학습, 비지도학습, 강화학습 등에 다른 알고리즘과 함께 두루 적용이 되고 있으며, 높은 수준의 정확도를 가지지만, 딥러닝의 근본적인 문제점 중 하나는 바로 결과값에 대한 구체적인 매커니즘을 이해하거나 과정에 대한 설명을 하기 어렵다는 것이다. (반면 예를들어 위의 지도학습 방법 중 의사결정 나무트리의 경우에는 결과값에 대한 높은 수준의 설명이 가능하다.) 그렇기에 앞으로는 이런 설명이 가능한 알고리즘에 대한 기술 개발이 필요하며, 많은 연구가 되고 있다.

7) 인공지능(AI) 분야의 최고 전문가인 Yann LeCun, Yoshua Bengio, Geoffrey Hinton의 Nature의 논문 참고, 원문보기 : https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf

3. 주요 국가의 인공지능 개발 동향

인공지능 기술에 대한 가트너社의 Hype Cycle 기술 수준을 통해서 안정기에 접어들 기간별로 주요기술을 살펴보면 다음과 같다.

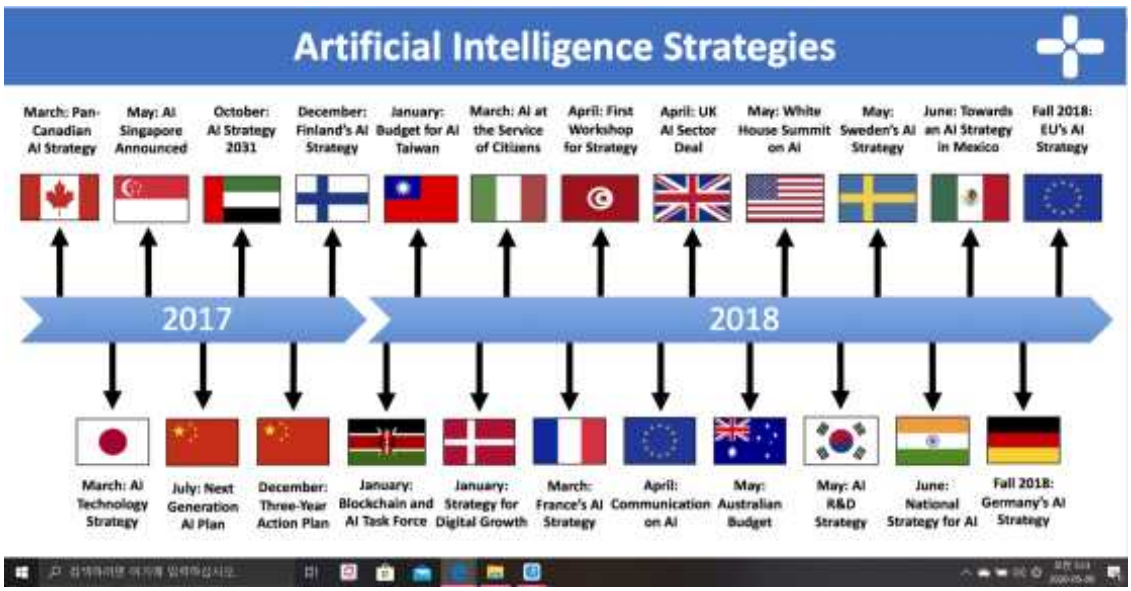


< 출처 : Gartner 홈페이지, Hype Cycle for AI, 2019>

- 2년 이내 : 음성 인식, 로봇 처리 자동화 소포트웨어 등
- 2년 ~ 5년 이내 : 컴퓨터 비전, 딥러닝, 챗봇, 머신러닝 등
- 5년 ~ 10년 이내 : 디지털 윤리, 강화학습, AI Paas[8], 의사결정 등
- 10년 이상 : 퀀텀 컴퓨팅, 자율주행차, 보편적 인공지능

8) 클라우드 서비서는 지원 자원에 따라 다음의 3가지로 1.SaaS(Software as a Service, 소포트웨어 제공), 2.IaaS(Infrastructure as a Service, 장비와 같은 인프라 제공), 3.PaaS(Platform as a Service, 플랫폼 제공) 나뉘며, AI PaaS(AI Platform as a Service)는 AI 플랫폼을 빌려주는 클라우드 서비스 형태임. 따라서 신생 스타트업도 적은 비용으로 쉽게 인공지능 분야에 연구할 수 있음.

위의 가트너 기술 수준에 따라 전세계적인 추세로 인공지능 기술은 개발되고 있을 것과 동시에, 대대수의 주요 국가들에서는 인공지능(AI) 분야에 대한 정책과 비전을 담은 발전전략을 2017~2018년에 걸쳐 아래 그림과 같이 인공지능 국가 정책을 발표하였다. 그 중 훈련국인 캐나다와 그 외 우리나라의 국방과 관련하여 중요한 역학적 관계인 미국, 중국 등의 주요 국가에 대한 인공지능 개발 정책[9]들을 살펴보고자 한다.



< 출처 : An Overview of National AI Strategies (2018) > [10]

1) 캐나다

위의 그림에서와 같이 캐나다 정부는 2017년 세계 최초로 국가차원의 인공지능(AI) 전략을 수립하였으며, 이는 국가고등연구원(CIFAR) 주도의 범캐나다 AI 전략(Pan-Canadian AI Strategy)으로써 이를 통해 인공지능(AI) 분야에 있어서 세계적으로 선도하는 국가들 중 하나로 만드는 데 큰 기여를 하였으며, 기술적인 분야 뿐만 아니라 윤리적인 측면에서도 선도하기 위해 노력을 하고 있다.

9) 인공지능(AI) 시대 주요국의 인재양성 정책 (2019, 소프트웨어정책연구소)
10) 출처 : https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd

이는 몬트리올, 토론토, 워터루 등 중요 거점 도시단위로 생태계(아래 그림 참조, 지속적인 거점 도시를 넘나드는 워크숍, 훈련 프로그램, 컨퍼런스 그리고 다른 많은 활동하고 있음)를 구축하여, 파편화된 지원책이 아닌 세계에 열려있는 국제적인 AI 연구 센터를 만드는 클러스터 육성 전략을 통한 신생 스타트업을 포함한 산·학·연 공동 협업으로 각각의 특성화된 AI 생태계를 만들어 이를 효율적으로 활용한 선순환 구조를 가질 수 있도록 한 것이 특징이다.

< 출처 : 2019 Canadian AI Ecosystem (2019, jfgangne) > [11]

- Pan-Canadian AI Strategy 주요 내용

  · 2017년 3월 국가적 차원의 AI 전략을 선언한 첫 번째 나라였던 캐나다의 Pan-Canadian AI Strategy을 수행한 연간 보고서[12]는 매년 CIFAR 홈페이지에 개재되고 있으며, 이를 통하여 그동안 진행된 내용을 알 수 있음.

<출처 : Annual Report of the CIFAR Pan-Canadian AI Stragtegy (2019, CIFAR) >

11) 출처 : https://jfgagne.ai/2019-canadian-ai-ecosystem/
12) 출처:https://www.cifar.ca/docs/default-source/ai-reports/ai_annualreport2019_web.pdf?sfvrsn=244ded44_17

· 캐나다의 전체 AI 생태계의 성장과 발전을 위하여 최상위의 연방정부 차원에서 $125M을 투자하고 있으며, 각 주(온타리오주 $80M, 알버타주 $100M, 퀘벡주 $100M)에서도 추가적으로 투자를 함으로써 국가 차원의 AI 전략을 성공적으로 추진하기 위해 노력하고 있어 전체적으로 $405M가 투자되고 있다.

· 국내 및 국외 연구 수상 실적은 Best Paper Award, IEEE Winter Conference of Applications of Computer Vision, Facebook Research Award 및 Bloomberg 50 등 저명한 실적들이 27건이며, 국가 전체적으로 AI 스타트업은 매년 27% 증가하여 650개 이상으로 이는 매년 51%가 증가하여 $418M으로 지원되고 있는 벤처 투자 자본의 영향이 크다고 할 것임

· 주요 AI 독립 연구 기관

<출처 : Annual Report of the CIFAR Pan-Canadian AI Stragtegy (2019, CIFAR) >

① AMII (Alberta Machine Intelligence Institute)

: 알버타주 에드먼튼시의 University of Alberta에 위치하고 있으며,

특히 기계학습(Machine Intelligence) 분야의 이해와 혁신을 증진시키기 위해 특화된 기관으로, 주요 연구 분야는 선진 기계학습 접근 및 적용, 게임 이론, 강화 학습, 건강 응용프로그램, 언어학습, 로봇 및 데이터 마이닝과 시각화 등이 있음.

※ AMII 내에서의 최신 발행된 주요 인공지능 관련 실적

1. 조현병에 대한 진단. (Diagnosing Schizophrenia)[13)]

- AMII 연구소는 컴퓨터정신학이라 불리는 컴퓨터공학과 정신학을 함께 연구하고 있고, 이는 머신러닝과 같은 컴퓨터 공학 도구를 사용하여 정신적인 문제를 더 잘 이해하고, 병의 진단과 증상 측정을 도와주는 도구와 방법을 데이터를 통하여 개발하고 있다.

- 이번 연구실적은 Nature지에도 실려져 있으며, 머신러닝 알고리즘을 통해서 74%의 정확도로 조현병을 예측하는데, 특히 특정 증상의 악화와 조현병 환자의 심각한 정도의 관계를 뇌의 여러 구역에 걸쳐 관찰된 활동들 사이의 관계를 기반으로 하였으며, 연구원들은 불특정한 뇌 기능적 자기 공명 영상(fMRI, Functional magnetic resonance imaging) 데이터를 분석하였다.

- 이 연구에 사용된 머신러닝 알고리즘들은 Baseline, Nearest Neighbors, Linea SVM, Logistic Regresssion, Decision Tree, Random Forest, Naive Bayes, LDA을 사용하였으며, 이 중 Logistic regression을 사용했을 때, 가장 높은 74%의 정확도를 나타내었음을 알 수 있었으며, 향후 보다 많은 데이터셋으로 확장을 하며 우울증이나 외상 후 스트레스 장애와 같은 분야도 진행예정이다.

☞ 이와 같은 기능적 부분들이 국방분야에 적용이 된다면, 군 장병에 대한 의료진단에 사용하여 보다 빠르고, 정확한 진단으로 장병에 대한 건강을 위해 기여할 것으로 보인다.

2. 전략적 행동과 비전략적 행동의 공식적 구분.[14)]

- 개요 : "전략적" 행동과 다른 형태의 의도적이지만 "비전략적" 행동을 구별하는 것은 멀티에이전트 시스템에서는 흔히 볼 수 있는데, 일반적으로 비전략적인 에이전트는 그렇지 않지만 전략적인 에이전트

는 다른 에이전트들을 만든다. 그러나 이러한 개념들 사이의 명확한 경계는 이해하기 어려운 것으로 입증되었다.

- 문제점 : 특히 이 문제를 해결하기 위한 "비전략적" 레벨 0 에이전트와 "전략적" 상위 레벨 에이전트(예: 레벨-k 및 인지 계층 모델)의 행동을 명시적으로 구분하는 행동 게임 이론에 있어서 특히 중요하다. 하지만 전체적으로 한정적 합리성을 논하는 작업은 비전략적 대리인의 합리성이 어떻게 제한되어야 하는지에 대한 명확한 지침을 제공하는 경우가 거의 없으며, 그 대신 일반적으로 특정 의사결정 규칙을 선별하여 비공식적으로 비전략적(예를 들어, 사적 정보를 진실하게 공개, 일률적으로 임의화)이라고 주장하기만 한다.

- 방안 : 이 연구에서는 비전략적 행동의 새롭고 형식적인 특성화를 제안하고 있으며, 이는 그것이 다음의 두 가지 속성을 만족시킨다는 것이다. (1) 행동 게임 이론 문헌에서 우리가 알고 있는 모든 "비전략적" 결정 규칙을 포착하는 것은 충분히 일반적이다. (2) 이 특성화에 따르는 행동은 정확한 의미에서 전략적인 행동과 구별된다.

☞ 이러한 구분에 대한 특성화는 다양한 게임 이론이 적용되는 분야에 사용되어질 수 있을 것이다.

② MILA

: 퀘벡에 있는 기관으로 University of Montreal and Mc-Gill과 파트너쉽을 맺고 있으며, 새로운 딥러닝 알고리즘과 다양한 분야에 적용함에 있어서 많은 선구자적인 업적으로 이루어서 세계적으로 알려진 기관으로, 주요 연구 분야로는 뉴런 언어 모델링, 뉴런 기계 번역, 객체 인식 등이 있음.

※ MILA 내에서의 최신 발행된 주요 인공지능 관련 실적

1. 부분적 데이터 비편견화를 위한 적대적 훈련 접근[15)]

- 최근 사회의 많은 영역에 있어서 자동의사결정 프로세스가 널리

13) Diagnosing Schizophrenia, AMII (2020년, Spencer Murray)

14) A Formal Separation Between Strategic and Nonstrategic Behavior, AMII (2020년, James Wright 등 2인)

사용되고 있으나 프로세스의 공정과 차별 가능성에 대한 심각한 윤리적 문제를 또한 야기하고 있다.

- 이러한 문제를 해결하기 위해서 GANSan(Generative Adversarial Network Sanitizer)이라는 나머지 속성과의 기존 상관 관계뿐만 아니라 속성 자체를 제거함으로써 직접적이든 간접적이든 어떤 차별에 대한 가능성을 예방하기 위한 목적의 정제(Sanitizer)를 학습하는 방안을 이 연구에서는 제시하고 있다.

- 이 GANSan은 강력한 프레임워크인 생성적 적대적 네트워크(특히 Cycle-GANs)로부터 큰 영향을 받았으며, 두 가지 다른 분포 사이에서 분포를 경험적으로 학습하거나 해석할 수 있는 유연한 방법을 제공하고 있다. 실제 데이터셋으로 실험을 통해 이 방안의 효과와 공정성과 효용성 사이의 충분한 절충점을 입증하였다.

☞ 이와 같은 윤리적인 문제들을 보완할 수 있는 알고리즘은 의사결정 과정 뿐만 아니라 많은 부분에 있어서 적용이 가능하다 할 것이다.

2. 음성 합성을 사용한 End to End 구어 이해 모델 훈련[16)]

- End to End 모델은 별도로 훈련된 음성 인식기와 자연어 이해 모듈로 구성된 표준 파이프라인을 채택하지 않고, 발음의 의미를 원시 오디에에서 직접 추출하는 SLU(Spoken Language Understanding, 구어 이해)대한 매력적인 새로운 접근방식이다.

- 하지만 End to End SLU의 단점은 모델을 훈련시키기 위해 도메인 내 음성 데이터를 녹음해야 한다는 것이다.

- 이 연구에서는 여러 인공 스피커에서 많은 인조 훈련 데이트 세트를 생성하기 위해 음성 합성을 사용하여 그 단점을 극복하는 방안에 대해서 제안하였고, 음성 합성을 사용하여 End to End SLU 모델을 양성하는 것이 가능하다는 것을 보여주었다.

☞ 이 연구는 다양한 음성 이해 모델 등에 사용이 되어 질 수 있으며, 특히 많은 데이터를 확보하지 못한 상황에서도 음성 합성(speech synthesis)을 통하여 많은 양의 훈련 데이터를 확보할 수 있는 것으로 제한된 국방 분야에도 많이 활용이 가능할 것으로 보인다.

③ Vector Institute

: 2017년에 온타리오주 토론토에 설립된 기관이며, 토론토 대학 등과 협업을 하고 있다. 딥러닝과 머신러닝에 있어서 세계 최고의 연구와 응용을 하고 있으며 교육 자금 등을 통해 세계적인 석학들을 캐나다로 이끌고 있으며, 주요 연구 분야로 Machine Learning, Deep Learning, Quantum computing 등이 있음.

※ Vector 내에서의 최신 발행된 주요 인공지능 관련 실적

1. 안전한 수술 중 의사결정 지원을 위한 설명가능한 인공지능[17)]

- 수술 중 부작용은 외과적 질병의 공통적이고 중요한 원인이며, 이러한 부작용을 줄이고 그 결과를 완하하기 위한 전력으로 전통적으로 외과적 교육, 구조화된 의사소통 및 이상 증상 관리에 초점을 맞춰왔다.

- 하지만 지금까지 수술실에서 이러한 사건들을 미리 예상할 수 있는 방법은 거의 없었다. 이러한 문제들을 해결하기 위한 한 방안으로 이 연구에서는 수술실에서의 데이터 캡처와 이러한 데이터를 처리하기 위한 설명 가능한 인공지능(XAI) 기법의 적용을 통해 수술 팀이 수술 중 사건을 예측, 이해 및 예방하는 데 도움이 될 수 있는 실시간 임상 의사결정 지원을 가능하게 할 수 있다.

☞ 의료분야 뿐만 아니라, 설명가능한 인공지능은 향후 아주 중요한 한 분야로 떠오르고 있으며, 이러한 분야에 적용하기 위한 좋은 참고 자료가 될 것으로 보인다.

2. 머신러닝을 평가하기 위한 국제 표준 방향[18)]

- 인공지능 표준화를 위한 다양한 국제적 노력이 있으며, 이러한 노력의 상당수는 사생활, 신뢰도, 안전, 공공복지와 관련된 이슈들을 포함하고 있는데, 이것은 국제적인 공감대가 형성되어 있지 않고 있다.

- 한편, 머신러닝에서 최정상의 정확도를 위한 추가는 그러한 정확도 계산의 정확성을 제한할 수 있는 경험적 방법론을 다시 임시로 적

15) Adversarial training approach for local data debiasing(2020년, Alain Tapp 등 5인)

16) Using speech synthesis to train end to end spoken language understanding models(2019년, Loren Lugosch 등)

용하는 결과를 낳았고, 결과적으로 그러한 모델의 예측 불가능한 적용 가능성을 낳았다. 따라서 이러한 객관적 정량적 성능을 신뢰하는 것이 안전을 위한 것이며 AI의 안전에 대한 기준을 빨리 정립할 hyp필요성이 있는 것이다.

- 이 연구에서는 이러한 맥락에서 두 가지 이상의 알고리즘 성능을 비교하기 위해서는 아래의 사항에 대해서 주의깊게 관리되고 보고되어야 한다고 말하고 있다.

· Implementation · Hyper-parameters
· Preprocessing · Training and testing data
· Representative data · Appropriate baselines
· Appropriate measures · Limiting information leakage
· Limiting channel effects · Implementation

☞ 앞으로 머신러닝 알고리즘 개발에 대한 안전성에 대한 기준 정립을 위한 참고자료 등으로 활용하기에 필요하며, 이러한 방법등을 통하여 머신러닝 프로그램을 정량적으로 평가하고 정확하게 제어하는 것이 중요하다 할 것이다.

· 국가적 4대 프로그램 활동들

① CIFAR AI4Good National Training Program

: 이는 차세대 인공지능(AI) 연구자들을 위한 교육 훈련 지원 프로그램으로, CIFAR와 각 파트너들은 매년 수 백명의 캐나다인과 국제 학생들을을 대상으로 인공지능 기술을 가르치고 있으며, 대상은 고등학생부터 박사 후 과정까지임

② AI4Health Task Force

: 캐나다를 건강연구(Health research)에 대한 세계 최고의 인공지능(AI) 수준에 이르게 하기 위한 것으로 CIFAR가 주도 하고 있음.

③ CIFAR AI Catalyst Grants Program (AI 장려 보조금 프로그램)

17) Explainable Artificial Intelligence for Safe Intraoperative Decision Support(2019년, Lauren Gordon 등 3인)
18) Towards international standards for evaluating machine learning (2019년, Frank Rudzicz 등)

: 혁신적이고, 고위험/큰보상인 아이디어와 프로젝트를 활성화하기 위한 여러분야의 연구 협력하는 곳을 대상으로 보조금을 지원하는 프로그램이며, 매년 $50,000를 최대 2년간 보조금을지원하고 있음

④ AICan Symposium

: 매년 열리는 인공지능 학술 토론회로 세계를 리딩하고 있는 AI 연구자들의해 머신러닝 연구 분야의 새로운 발전들을 논의함.

이러한 또한 캐나다의 인공지능에 대한 노력으로 다양한 굴지의 기업들이 연구소를 설립하고 있어 세계적인 인공지능의 기지가 되고 있으며[19], 그 대표적인 것들로는 국내의 주요 업체인 2018년 5월 삼성전자는 Vector Institute가 위치해 있는 캐나다 토론토에 'AI 센터'를 개소하였으며(2017년 9월에는 이미 몬트리올에 인공지능 연구소를 설립하여 세계적인 권위자이자, 몬트리올대학 교수인 Yoshua Benjio와 협업중임), 2018년 8월 LG전자 또한 인공지능(AI) 연구개발(R&D) 전문 연구소를 인공지능만을 연구하는 연구소를 해외에 처음 세웠으며, 구글의 자회사 딥마인드는 2017년 7월 AMMI가 있는 알버타주의 애드먼튼시에 본사인 영국이 아닌 해외에 처음으로 인공지능 연구소 '딥마인드 앨버타'를 설립하였으며, 세계에서 가장 큰 그래픽카드 생산 업체인 Envidia는 2017년 6월에 토론토에 인공지능 연구소를 열었다. 이것은 정부에서 인공지능 관련 연구를 하는 기업 및 연구소에 투자비용의 15%를 세액 공제해주며, 매해 지원하는 인공지능 연구 인력(350명) 중 절반은 국적과 관계없이 선정하는 등의 정부의 지원이 있기에 가능한 것으로 보고 있다.

· 캐나다에서 활동하고 있는 주요 인공지능 선구자

① 제프리 힌튼 (Geoffrey Hinton)

: 앞서 언급된 인공지능 역사에는 여러 번의 침체기가 있었는데,

19) 참고자료 : 소리소문없이 인공지능 기지 된 캐나다, 구글, 엔비디아, 삼성, LG 가 몰리는 이유. https://blog.naver.com/attisevery/221331157233

첫 번째 침체기 (XOR를 풀지 못하는 문제)에 대한 해결책이었던 다층 퍼셉트론과 역전파법(Back-propgation), 그리고 두 번째 침체기(Vanishing gradient)에 대한 해결책이었던 Deep belief nets에 관한 논문을 내면서 신경망 이론인 딥러닝에 대한 부흥을 일으켰으며, 2012년 국제이미지인식기술대회(ILSVRC)에서 토론토대 제프리 힌튼 교수팀은 딥러닝 기반으로 이미지 분석을 하는 '알렉스넷(Alexnet)'으로 우승을 차지하였다. 현재는 Vector Institute의 CSA(Chief Scientific Advisor)로 캐나다의 인공지능 발전에 크게 기여하고 있음

② 조슈아 벤지오 (Yoshua Bengio)

: 2018년 튜링상[20]을 받았으며, 딥러닝 분야에서 세계 3대 석학 중에 한 명으로, 현재 MILA Institute의 SD(Scientific Director)로 딥러닝분야 등에 많은 활동을 하며 인공지능발전을 이끌고 있음

③ 리차드 서튼 (Rechard S. Sutton)

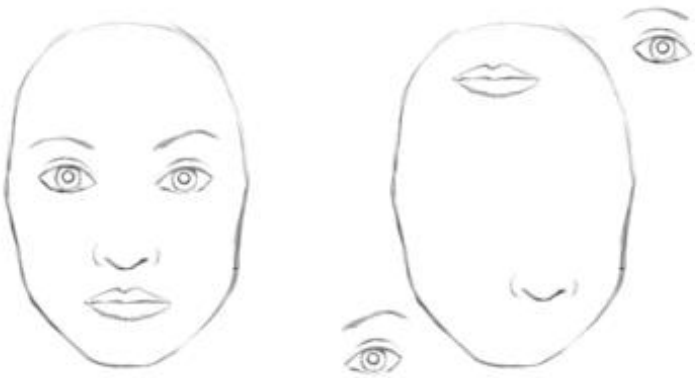
: 지금의 알파고를 만든 '강화학습'의 창시자로 알려져 있으며, 현재 AMII Institute의 CSA(Chief Scientific Advisor)로 머신러닝 특히 강화학습 분야 등의 발전에 크게 기여하고 있음

인공지능(AI) 분야에 있어서 최고 권위있는 학술단체인 인공지능진보협회(AAAI, Association for the Advancement of Artificial Intelligence)의 2020년 2월 미국에서 개최된 컨퍼런스에서 2018년 튜링상 수상자들은 특별연설[21]을 통하여 앞으로의 머신러닝의 발전방향 등에 대해서 의견을 개진 하였으며, 캐나다 제프리 힌튼 교수는 컴퓨터 인식(이미지 인식) 부분에서 가장 많이 사용되고 있는 CNN 알고리즘의 단점을 개선할 수 있으며, 인간의 시각적인 인지방법과 유사한 새로운 알고리즘인 Capsule Auto-Encoders 라는 방법을 제시하였으며, 이 방안은 장기간 제프리 힌튼 교수가 CNN 알고리즘의 문제점을 파악하고 난 뒤부터 지속적으로 연구한 방법[22]으로써, 앞으로 컴퓨터

20) 영국의 수학자이자 앨런 튜링의 업적을 기리기 위한 튜링상은 컴퓨터·과학 분야의 노벨상으로 불리는 상으로, 미국 계산기 학회(AMC)에서 뛰어난 업적을 남긴 사람에게 매년 시상하고 있다. 2018년도 튜링상 수상자로 제프리 힌튼, 조슈아 벤지오, 그리고 얀 르쿤 미국 뉴욕대 교수가 선정됐다.

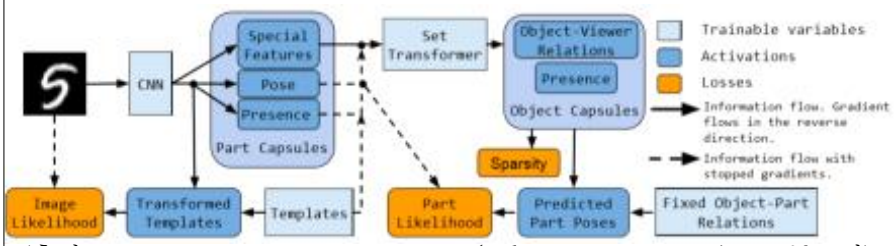
21) 출처 : AAAI 20 / AAAI 2020 Keynotes Turing Award Winners Event (2020년, Geoff Hinton, Yann Le Cunn, Yoshua Benjio)

비전에서 중요한 역할을 할 것이라 말하고 있으며, 이에 대한 내용을 정리하면 다음과 같다.

1. 기존 CNN의 특징	이미지 인식과 분류에서 탁월하며, 입력과 가까운 층에서 edge, curve 등 저수준(low level)의 특징을 학습한 뒤 점차 고수준(high level) 특징을 인식하여 이미지를 파악함. 작은 특징(Features)을 추출에 강점
2. 기존 CNN의 문제점	* 인간의 시각 인지와 여러면에서 다름. 1) 얼굴을 타원형의 얼굴과 두 개의 눈, 하나의 코, 하나의 입이 있을 때 인식할 경우, CNN은 아래의 두 그림 모두를 얼굴로 인식할 수 있다. 사람이 사물의 위치를 직관적으로 인식하는 것과는 다르다.  <출처 : 캡슐 네트워크 이해 (2017, Max Pechyonkin), 각주 참조 >[23] 2) 아래의 여러 가지 자유여신상 사진과 같은 관점 변화(viewpoint changes)에 따른 회전(Rotation)이나 크기 변화(Scailing)를 잘 대처하지 못함. ☞ 이 문제점을 해결하기 위해 2-D maps 대신 4-D나 6-D maps를 사용하지만 비용이 매우 크기 때문에 전형적으로는 한 객체에 대해서 여러 가지의 관점으로 학습을 시키고 있지만 비효율적임.

22) 처음으로 2017년 'Dynamic Routing between Capsules' NIPS-2017 논문으로 캡슐 네트워크(Capsule Network)개념을 소개하였으며, 2018년에는 'Matrix Capsules with EM Routing' ICLR-2018 논문을 내었고, 2019년에는 앞선 두 논문을 개선한(discrimninative learning/part-whole relationships → unsupervised learning/whole-part relationships) 'Stacked Capsule Autoencoders' NeurIPS2019 논문을 발표한 후 2020년에는 'Detecting and Diagnosing Adversarial Images with Class-Conditional Capsule Reconstructions' ICLR-2020 논문을 내는 등 지속적인 연구를 통해 인간과 같은 이미지 인식을 이루고자 연구하고 있다.

23) 출처 : Understanding Hinton's Capsule Networks. Part I: Intuition (2017, Max Pechyonkin)

	 〈출처 : 각주 23 참조 〉
3. 개선방안	* 적층 캡슐 오토인코더 (Stacked Capsule AutoEncoder) 1) 캡슐(Capsule)은 유사 형상이나 부분을 학습하기 위한 뉴런들의 그룹이며, 형상과 카메라 사이의 위치적인 관계와 같은 위치(pose)를 나타내는 메트릭스이며, 속도나 방향이 있는 벡터(vector)로 되었다. 2) 적층 오토인코더(Stacked AutoEncoder)는 입력과 출력 층의 차원(노드의 개수)는 동일하나, 은닉층(Hidden Layer)는 입력/출력 층보다 차원이 낮게 구성된 오토인코더를 적층한 것으로, 지도학습 없이도 입력 데이터의 표현을 효율적으로 학습하는 방법이다. 3) 이런 위치 상관관계가 포함된 캡슐이라는 새로운 개념과 내부 생성모델(Generative model)이 가능한 비지도학습법인 적층 오토인코더를 활용하여 CNN의 구조에 근본적으로 영향을 주어 문제점을 극복할 수 있으며, 사람과 유사한 이미지 인식을 할 수 있음. 아래는 Stacked Capsule Autoencoders의 논문에 나온 SCAE 구조도임.  〈출처 : Stacked Capsule Autoencoders (2019, Geoffrey E. Hinton 등 8인〉

또한 이러한 인공지능에 대한 정부의 주도적인 개발지원에 따른 기술발전을 바탕으로 현재 정부차원에서 인공지능에 의한 자동화된 의사결정시스템을 적용하고 있으며, 이는 2019년4월1일부로 Directive on Automated Decision-Making라는 훈령으로 법제화되어 있고, 정보기술관리정책(Policy on the Management of Information Technology)이 적용된 모든 기관에서 예외적인 경우를 제외하고는 사용되고 있다. 이 훈령의 목적[24]은 캐나다 법률에 부합하는 보다 효과적이고, 정확하며, 지속적인 그리고 해석가능한 결정을 내릴 수 있도록 도와주며, 캐나다인과 연방정부 기관의 위험을 줄여주기 위한 것이다. 사용 예로는 캐나다 이민국으로 인터넷으로 제출되는 방문비자에 대한 의사결정을 기존 자료등을 학습함으로써, 인공지능(AI)으로 하여금 일차적으로 판단하는데 사용하고 있다.

그리고 이 자동화된 의사결정시스템의 적용으로 인해 그 결정된 사항들은 그 결정과 관련된 사람들에게는 큰 영향을 미칠 수도 있는 것이기 때문에, 이를 적용하기 위해서는 각 기관의 차관보(Assistant Deputy Minister)가 다음의 요구조건을 만족하도록 해야한다.

- 알고리즘 영향 분석 (AIA, Algorithmic Impact Assessment)

: AI 해결방안이 윤리적이고, 인간적인 관점에서 얼마나 수용이 가능한지에 대한 평가를 할 수 있어야하며, 이의 영향 분석 수준은 Level Ⅰ(영향거의없음)에서부터 Level Ⅳ(매우 높은 영향, very high impacts)까지 총 4단계로 나타내어지고, 평가를 하게 된다. 그리고 Level Ⅰ,Ⅱ단계까지는 의사 결정과정에 사람이 개입하지 않고 결정이 주어지게되며, Level Ⅲ, Ⅳ단계는 의사 결정단계과정에서 사람이 개입하여 의사결정되며, 또한 최종의사결정은 사람이 해야만 한다.

- 투명성 (Transparency)

: 이 투명성은 인공지능(AI)에 의한 의사결정의 구체적인 근거와 그

24) 훈령(Directive on Automated Decision-Making)의 목적에 대한 영어 원문 : The objective of this Directive is to ensure that Automated Decision Systems are deployed in a manner that reduces risks to Canadians and federal institutions, and leads to more efficient, accurate, consistent, and interpretable decisions made pursuant to Canadian law.

판단과정에 대한 것으로써, 의사결정전에 통지를 해야하며, 의사결정 후에 설명을 제공하며, 소스코드(Source Code)를 보안사항이지 않으면 제공(Release)해야한다.

- 품질 보증 (Quality Assurance)

: 의도하지 않은 데이터 편향이나 다른 요소들이 불공정한 결과를 만들지 않도록 자동화된 의사결정 시스템을 적용하기 이전에 시험과 결과에 대한 모니터링을 해야 한다. 이는 데이터 품질(Data Quality), 동료 전문가 평가(Peer Review), 직원 훈련(Employee Training), 보안(Security), 법적인 문제(Legal) 등을 포함하고 있다.

2) 미국

세계 1위의 경제력을 바탕으로 세계의 인공지능 기술을 이끌고 있는 미국의 경우 2016년 오바마 정부 마지막 해에 통합된 국가 AI 전략이 아닌 3가지로 나누어진 보고서를 배포함으로써 국가 전략의 기초를 놓았다. 그 첫 번째 보고서는 10월 「Preparing for the Future of Artificial Intel」로 인공지능 적용과 관련한 공공부문, 연방정부, 규제, 연구개발, 산업계, 자동화, 윤리, 공정성 그리고 무기시스템[25] 등에 대해서 구체적인 권고사항(Recommendations) 23개를 작성하였고, 이와 같은 시기에 「The National Artificial Intelligence R&D Strategic plan」을 통해 인공지능 연구개발에 대한 공공자금 지원을 위한 전략적 계획의 윤곽을 7대 전략으로 만들었으며, 2개월 뒤 「Artificial Intelligence, Automation, and the Economy」을 통해 인공지능 발전으로 산업계의 자동화율이 증가하며 이에 대비하기 위해 작성되었으며, 3대 전략(투자, 교육, 인력배치)을 포함하고 있다.

이후 트럼프 정부에서 2018년 5월 백악관에서 100명 이상의 고위관료, 학계의 저명한 전문가, 산업계 연구수장 및 비즈니스 리더들을 초청하여

25) 무기시스템에 대한 권고사항 23(Recommendation 23) : 미 정부는 국제인도주의법을 준수하여 자율 및 반자율 무기들에 대한 독자 및 범정부적인 정책 개발을 완료해야 한다.

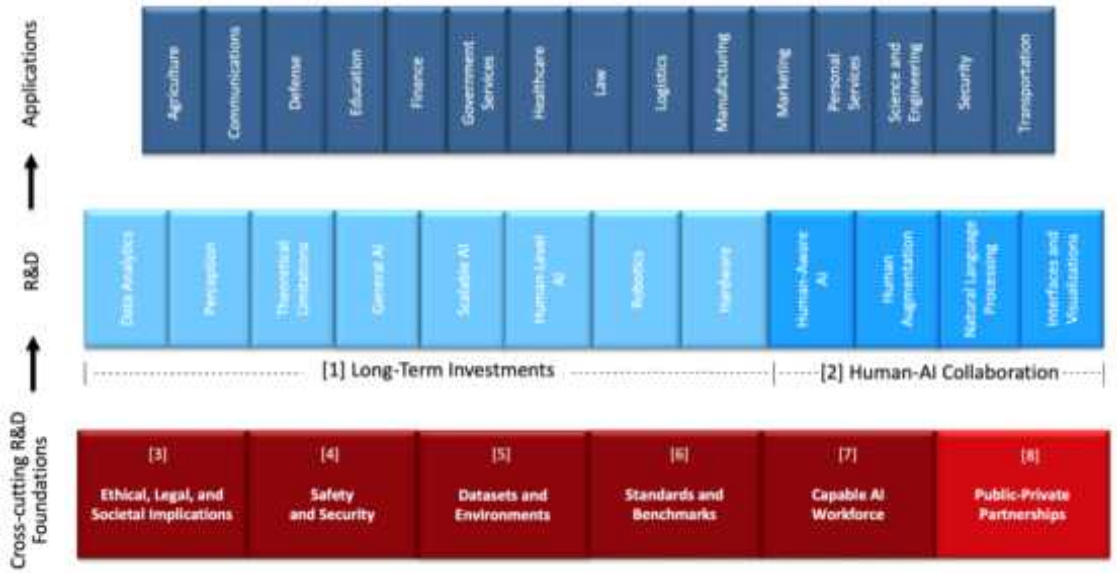
인공지능(AI) 관련 회담을 하였고, 「Summary of the 2018 White House Summit on Artificial Intelligence for American Industry」을 발표하였다. 이는 인공지능에 대한 투자, 제도적 장벽 제거 그리고 국방분야에 대한 전략적인 접근 등을 담고 있다. 그리고 2019년 2월 'Maintaining American Leadership in Aritificial Intelligence' 란 행정명령(Executive Order 13859)[26]에 서명한 후 2019년 6월 「The National Artificial Intelligence R&D Strategic plan : 2019 UPDATE」[27]로 최신화하여 재작성하였으며, 이는 기존의 7대 전략의 큰 틀은 그대로 두면서 8번째 전략(Public-Private 파트너 관계 확장)[28]을 추가하였다.

< AI R&D 8대 전략과 구조도 >

- 제1전략 : AI 연구에 장기적인 투자를 한다.
  (Make long-term investments in AI research)
- 제2전략 : 사람-AI 협업을 위한 효과적인 방법을 개발한다.
  (Develop effective methods for human-AI collaboration)
- 제3전략 : AI의 윤리적, 법적, 사회적 영향을 이해하고 고민한다..
  (Understand and address the ethical, legal, and societal implications of AI)
- 제4전략 : AI 시스템의 안전과 보안을 강화한다.
  (Ensure the safety and security of AI systems)
- 제5전략 : AI 교육과 시험관련 개방된 데이터셋과 환경을 개발한다.
  (Develop shared public datasets and environments for AI training and testing)
- 제6전략 : 기준과 벤치마킹을 통해서 AI 기술을 측정하고 평가한다.
  (Measure and evaluate AI technologies through standards and benchmarks)
- 제7전략 : 국가적 AI R&D 인력 필요를 보다 더 이해한다.
  (Better understand the national AI R&D workforce needs)
- 제8전략 : AI의 발전을 가속화하기 위해 민-관 협업을 확장한다.
  (Expand public-private partnerships to accelerate advances in AI)

---

26) Executive Order on Maintaining American Leadership in Artificial Intelligence(2019, White House)
27) THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN: 2019 UPDATE (2019)
28) 8번째 전략 민간과의 협력 파트너 관계 확장을 추가한 이유는 美연방정부의 연구개발(R&D)을 함에 있어서, 빠르게 증기하고 있는 민간부문의 인공지능 연구개발비를 고려하여, 민간 부문(Private sector)과의 협약과 산업계에 의한 인공지능(AI)의 빠른 채택에 대한 요구가 많았기 때문이다.



<출처 : The National Artificial Intelligence R&D Strategic Plan (2019 UPDATE) >

그리고 2019년 미 국방수권법에 따라 설립된 AI국가안보위원회(NSCAI)는 2019년 11월 중간 보고서(Interim report)에서 빠르게 발전하고 있는 중국의 인공지능 기술을 경계하며, 아래의 인공지능(AI)과 국가 안보 사이의 관계에 관한 7대 원칙[29]을 제시하였다.

① AI 기술에 대한 세계적 리더쉽은 국가 안보의 우선사항이다. (Global leadership in AI technology is a national security priority.)

② 국가 안보를 위해 AI 도입은 긴급한 국가적 절대사명이다. (Adopting AI for defense and security purposes is an urgent national imperative.)

③ 민관은 국민의 번영과 안보에 대한 공유된 책임을 가져야 한다. (Private sector leaders and government officials must build a shared sense of responsibility for the welfare and security of the American People.)

④ 인재 양성은 여전히 중요한 요소이다. (People are still essential.)

⑤ 기업과 사상의 자유로운 원칙은 준수되어야 한다
(The power of free inquiry must be preseved.)

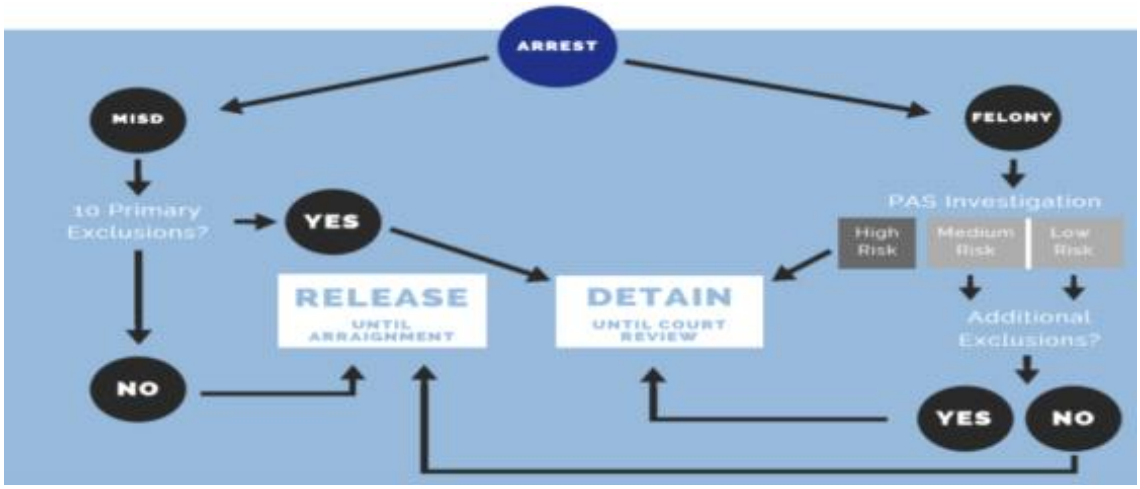
⑥ 전략적 필요와 윤리적 문제는 서로 양립될 수 있다.
(Ethics and strategic necessity are compatible with one another.)

29) 출처: Interim Report November 2019 (2019, NSC on AI)

⑦ AI의 사용은 중심에 법률과 같은 미국적 가치에 핵심을 둬야 한다. (The American way of AI must reflect American values - including having the rule of law at its core.)

인공지능(AI) 기술이 사용되어지고 있는 정부 부문 중 하나는 범죄 정의 시스템(Criminal Justice system)이며, 대부분 '사전심사 위험평가(PAS, Pretrial Assessment System)' 알고리즘으로 거의 대부분 주에서 사용이 되고 있다.[30]

이와 관련하여 캘리포니아주에서 발의된 2018년에 발의되어 2019년에 적용된 California's Senate Bill 10(SB 10)의 취지를 보면 이 법안에 의해서 돈으로 가석방되는 제도를 폐지하고 범죄 피고인에 대한 인공지능 알고리즘에 의한 위험 평가를 통해서 돈의 부족이 아닌 위험도에 따라 구속 여부를 결정하는 것이고 되어있다. 아래는 절차도이다.[31]



<출처 : Senate Bill 10(Pretrial Release and Detention, California Courts) >

범죄 정의 알고리즘(Criminal Justice Algorithms)은 범죄를 저지른 피고인이나 구속자들 미래 행동 예측을 지원하는 도구로 사용되고 있지만, 동시에 논란이 많다. 많은 위험 평가 알고리즘은 나이, 성별, 거주지, 가족 이력, 과거 범죄이력 그리고 취업상태 등의 개인 특성을 고려하고, 그 결과 동일한 범죄를 저지른 사람이라도 자신이 전혀 바꿀 수

30) 출처 : Algorithms in the Criminal Justice System : Risk Assessment Tools (2020년, epic.org)
31) 출처 : Senate Bill 10(Pretrial Release and Detention, California Courts)

없는 입력값에 근거하여 완전히 다른 사전선고(석방이나 징역)를 받을 수 있다. 또한 아래와 같이 편견(Bias)이 들어간 결과가 나올 수도 있기 때문에 2020년 5월 미국 미네소타주 미니애폴리스에서 경찰의 과잉진압으로 사망한 조지 플로이드 사건으로 인종차별에 대한 시위가 전세계적으로 번지고 있는 등 이와 관련된 투명성과 공평성 그리고 편견을 없애기 위한 방안 등이 필요하다.

* 출처-2016년 ProPublica에 의한 Florida주의 COMPAS 시스템 결과와 2년내의 재범 가능성에 대한 조사임	WHITE	AFRICAN-AMERICAN
Higher Risk로 평가되었지만, 재범을 저지르지 않음.	23.5%	44.9%
Lower Risk로 평가되었지만, 재범 저지름.	47.7%	28.0%

2016년 美 주도의 인공지능(AI) 국제 협력기구인 PAI(The Partnership on AI)는 인공지능의 윤리적인 문제 등의 아래와 같은 6대 주제를 다루기 위해 설립되었으며, 구성에는 13개 국가의 100개 이상의 파트너가 있으며, 여기에는 미국의 대표기업인 구글, 페이스북, 아마존, IBM과 캐나다의 CIFAR, Element AI 등이 참여하고 있다.

① 안전이 중요한 부문의 인공지능 (Safety-Critical AI)

- 자동결정 시스템 등에 있어서의 안전과 신뢰성 그리고 윤리적인 문제 등

② 공정하고, 투명하며 책임성있는 인공지능 (Fair, Transparent and Accountable AI)

- 공정하고, 설명가능하며 책임있는 AI 시스템

③ 인공지능, 노동 및 경제 (AI, Labor, the Economy)

④ 사람과 인공지능 사이의 협력 (Collaborations Between People and AI Systems)

⑤ 인공지능의 사회적인 영향 (Social and Societal Influences of AI)

⑥ 인공지능과 사회 선행 (AI and Social Good)

2020년 CB인사이트가 발표[32]한 Heathcare, NLP, NLG & 컴퓨터 비젼 및 Cybersecurity 등 각 분야의 대략 5000개의 스타트업 기업들을

32) 출처 : AI 100-The Artificial Intelligence Startups Redefining Industries (2020, cbinsights)

대상으로 특허 활동, 기업 관계, 투자자, 새로운 업적, 시장 가능성, 기술 참신성 등을 고려하여 인공지능 관련 100대 스타트업 기업들을 보면 미국 기업이 65개로 압도적으로 많으며, 캐나다와 영국은 8개, 중국은 6개와 일본 1개 등이 있다. (한국의 경우 없음)

<출처 : AI 100-The Artificial Intelligence Startups Redefining Industries (2020, cbinsights)>

이 중에서도 2018년, 2019년에도 선정된 Govt. & City Planning 부분에서 무인기술에 뛰어난 기업인 美Shield AI는 2015년에 네이비씰 출신의 설립자가 설립하였으며, 강력한 AI 기술이 탑재된 무인 쿼드콥트(Quadcopter, 4개의 로터를 가진 드론)로 복잡한 건물 내에서 실시간 데이터를 획득 및 제공하는 기술을 가지고 있으며, HIVEMIDD DEGE (외부의 입력없이 독립된 임무를 수행), HIVEMIND CORE(Self-directed learning), NOVA CLASS의 유-무인 협력 시스템은 특히 국방분야에서도 적용이 될 수 있는 기술들을 개발하고 있는 스타트업 기업이다.

3) 중국

중국은 앞으로 가장 핵심적인 기술로 떠오르는 인공지능(AI) 부문에 있어서 미국과 기술패권을 차지하기 위한 무한 경쟁에 돌입하고 있는 상황으로 기존의 Top-down 방식을 벗어나 민간 주도와 함께 기술발전을 꾀하고 있다. 2017년 7월 중국 국무원은 「차세대 AI 발전계획(2017.7.)[33]」을 확정하였고, 이는 4대 기본원칙, 3단계 전략목표 그리고 6대 중점임무를 통하여 인공지능에 대한 투자를 그 전략에 맞게 실행하고 있으며 그 내용은 다음과 같다.

- 4대 기본 원칙

① 기술 우선 : 퍼스트 무버 이점의 구축을 가속화 및 하이엔드 기술개발 성취

② 시스템적 청사진 : 사회주의적 시스템의 이점을 계획을 촉진시키고, 중요한 업무 수행을 위해 역량을 집중

③ 시장 우선 : 시장의 규칙을 따르면서 시장과 정부 사이의 역할을 잘 분담
* 정부의 역할 : 계획, 가이드, 정책 지원, 보안 및 보호, 시장 규제, 대외 환경 건설, 윤리적 규제 형성

④ 오픈소스 및 개방 : 오픈소스 공유 개념을 지지하며, 민·군의 혁신 자원을 공유하며, 고도의 효율적인 새로운 민·군 통합 방법을 형성하고, 국제적인 인공지능 R&D 및 관리에 적극적으로 참여함

- 3단계 전략 목표

① 1단계 (~2020년)

: 인공지능 산업의 경쟁력을 국제적으로 진입하는 첫 단계로 AI 핵심 산업 규모는 1,500억 위안(약 25조원) 이상, 연관 산업규모는 1조 위안(약 170조원) 이상으로 확대하고, AI 발전환경의 최적화, 중점 분야에서의 전면적 활용 확대, 고숙련 인재 모집, 일부 분야의 AI 윤리 규범 및 법규 구축 등을 추진한다.

33) 출처 : A Next Generation Artificial Intelligence Development Plan (2017, www.jaist.ac.jp)

② 2단계 (~2025년)

: 기본 이론 분야에 있어서 중요한 도약점을 이룩하여, 일부 기술과 응용분야는 세계 선도적인 지휘를 획득하는 단계이며, AI 핵심 산업 규모는 4,000억 위안(약 68조원)이상, 연관 산업규모는 5조 위안(약 857조원) 이상으로 확대하며, 일부 기술 및 응용 분야에서 세계적 수준에 도달하여 중국 산업발전 및 경제전환의 핵심 동력으로 활용할 계획이다.

③ 3단계 (~2030년)

: 인공지능 이론, 기술 및 응용분야 모두 세계를 리딩하는 단계에 이르고, 중국을 세계 최고의 인공지능 혁신 센터를 만들며, 지능화된 경와 지능화된 사회 응용면에 있어서 가시적인 결과를 얻는 단계이며, AI 핵심 산업 규모는 1조 위안(약 170조원) 이상, 연관 산업규모는 10조 위안(약 1700조원) 이상으로 확대하고, 뇌과학 기반 지능, 자율지능, 하이브리드 지능과 군집지능 등 주요 분야에서 국제적으로 중요한 영향력을 확보하는 것을 주요 목표로 설정하였다.

- 6대 중점임무

① 개방되고 조화로운 인공지능 과학기술 혁신시스템을 구축

② 최상위 고도의 효율적인 스마트 경제 육성

③ 안전하고 편리한 지능 사회를 구축

④ 인공지능 분야의 민·군 통합을 강화

- 공통의 중요 일반 기술에 대한 연구 개발
- 연구소, 대학, 기업 및 방산 시설의 교류 및 협력
- 지휘결심체계(command and decision-making)의 인공지능 기술 강화
- 모든 인공지능 기술이 국방혁신분야에 신속 적용을 촉진
- 민·군간의 AI 기술 표준 시스템의 구축 강화 등

⑤ 안전하고 효율적인 지능 인프라 시스템 구축

⑥ 차세대 인공지능 주요 과학기술 프로젝트에 대한 계획

아래는 주요 8대 부문에 대한 기본 이론과 구체적인 방안들에 대한 설명을 나타내고 있다.

#1. 차세대 인공지능의 기본 이론들	
· 빅데이터 지능 이론 - 불충분한 정보에서의 스마트 의사결정 시스템에 대한 기본 이론 및 체계 등	· Cross-media 센싱 및 컴퓨팅 이론 - 인간의 시각 능력을 초월하는 스마트 센싱 및 자율학습하는 사고 엔진 연구
· 하이브리드 및 향상된 인공지능 이론 - 인간과 기계의 스마트 공생 및 뇌와 기계 연결 그리고 실생활의 유·무인 협력	· 군집 지능 이론 - 군집 지능 보상 메커니즘 및 출현 매커니즘 그리고 군집 지능 학습 이론 등
· 자율 협력, 관리 및 최적 의사결정 이론 - 자율 무인 시스템과 관련한 협력 센싱과 상호작용, 지식 기반의 인간-기계-물체의 삼각 연결 및 상호성 연구 등	· 높은 수준의 머신러닝 이론 - 불확실성 속에서의 사고 및 의사결정에 관한 연구와 Small-sample learning, 딥 인텐시브 러닝, 비지도 학습 등
· 뇌 연결(inspired) 지능 컴퓨팅 이론 - 뇌 연결 센싱, 뇌 연결 학습 및 뇌 연결 회상 매커니즘 및 통제 등	· 퀀텀(Quantum) 지능 컴퓨팅 이론 - 인지 퀀텀 모델과 고유 매커니즘 등에 관한 연구

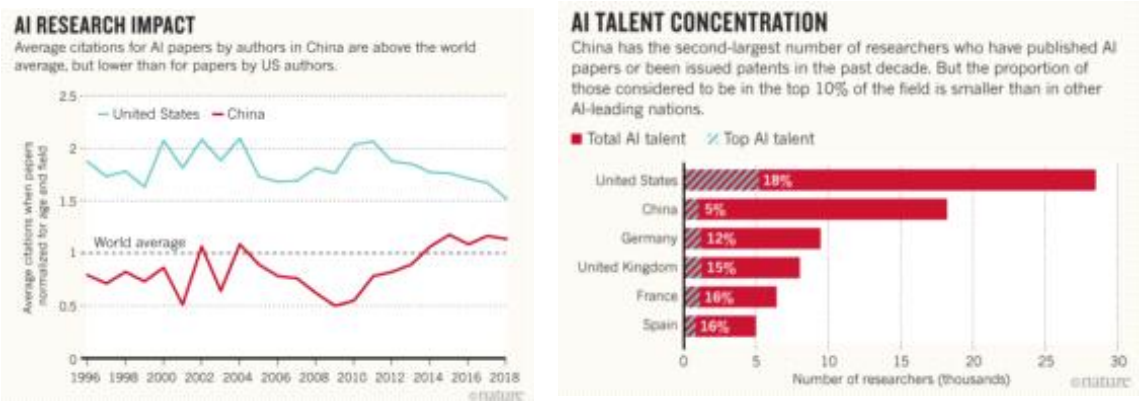
아래는 주요 8대 부문에 대한 중요 일반 기술과 구체적인 방안들에 대한 설명을 나타내고 있다.

#2. 중요 일반 기술들	
· 지식 컴퓨팅 엔진과 지식 서비스 기술 - 지식 컴퓨팅, 디지털 생성(creation)과 혁신적인 디지인 연구 등	· 크로스 미디움 분석 사고 기술 - 크로스 미디움에 관한 지식 맵 생성과 학습, 분석 사고 엔진과 검증 시스템
· 주요 군집 지능 기술 - 군집 지능의 액티브한 인지 및 발견, 이동 군집 지능의 조화된 의사결정 및 통제 기술	· 새로운 향상된 하이브리드 지능 기술 - 인간과 기계의 집합적인 유도, 온라인 지능 학습 기술 등
· 자율 무인 지능 시스템 기술 - 자동차, 선박, 수중, 우주 등의 무인 자율 인공지능 조절 및 로봇 기술 등	· 가상 현실 지능 모델링 기술 - 가상 대상의 지능적 행동을 위한 수학적 표현과 모델링 방법 및 사용자와 가상 대상의 긴밀한 상호작용 연구 등
· 지능 컴퓨팅 칩과 시스템 - 고 에너지 효율의 Neural Network 프로세스 개발 및 AI 운영 시스템 개발 등	· 자연어 처리 기술 - 기계 인지와 인간과 기계의 상호 시스템을 위한 언어 간의 의미 이해 기술 등

아래는 기본적인 인공지능 지원 플랫폼에 대한 내용이다.

#3. 기본 지원 플랫폼들
1. AI 오픈 소스를 위한 하드웨어 및 소프트웨어 인프라 및 플랫폼
2. 그룹 지능 서비스 플랫폼
3. 향상된 하이브리드 지능 지원 플랫폼
4. 자율 무인 시스템 지원 플랫폼
5. AI 기본 데이터와 보안 감지 플랫폼

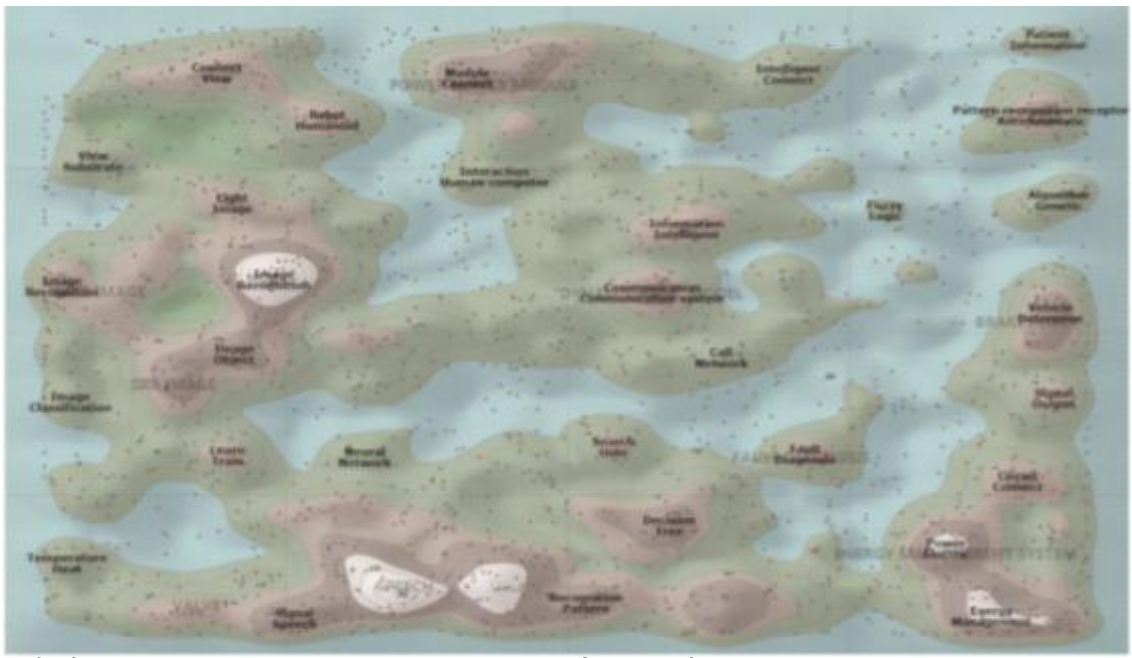
위와 같은 중국의 국가적 차원의 지원에 힘입은 결과로, 나타난 중국의 인공지능 개발 현황을 알아보면, 첫째로 중국과 미국의 AI 논문의 인용 빈도(출처-Nature[34])를 살펴보면, 최근 중국은 비록 미국에 지고 있으나 세계 평균을 넘어섰고, 또한 추세적으로는 미국은 내려오고 있는 반면 중국은 가파르게 오르고 있는 상황이다. 두번째로 인공지능 관련 논문이나 특허를 가지고 있는 전문인력수의 경우, 미국이 1위를 차지 하고 있으나, 그 뒤를 중국이 바짝 따라가고 있는 상황이라 할 수 있다. (하지만 Nature의 조사에 따르면, 아래의 그림과 같이 상위 10%에 속하는 전문인력에 대해서는 독일, 영국 등 다른 국가들에 비하여 적은 것으로 파악이 되고 있다.)



<출처 : Nature 2019년 기사, 각주 34 참조>

34) 출처 : Will China lead the world in AI by 2030? (2019, Nature)
https://www.nature.com/articles/d41586-019-02360-7

중국 정부의 막대한 투자 이외에도 이렇게 중국의 인공지능 기술이 빠르게 발전할 수 있는 사회적인 이유 중의 하나는 세계1위의 약 14억 명 이상의 인구와 이와 관련된 정보를 활용할 수 있기 때문일 것으로 판단된다. 아래 그림[35]은 Derwent Innovation을 활용한 인공지능 관련한 특허의 분야별 중점도를 나타내고 있으며, 이에 따르면, 현재 가장 중점이 되고 있는 부분은 음성 및 이미지 인식과 결정 트리와 관련된 분야이다.



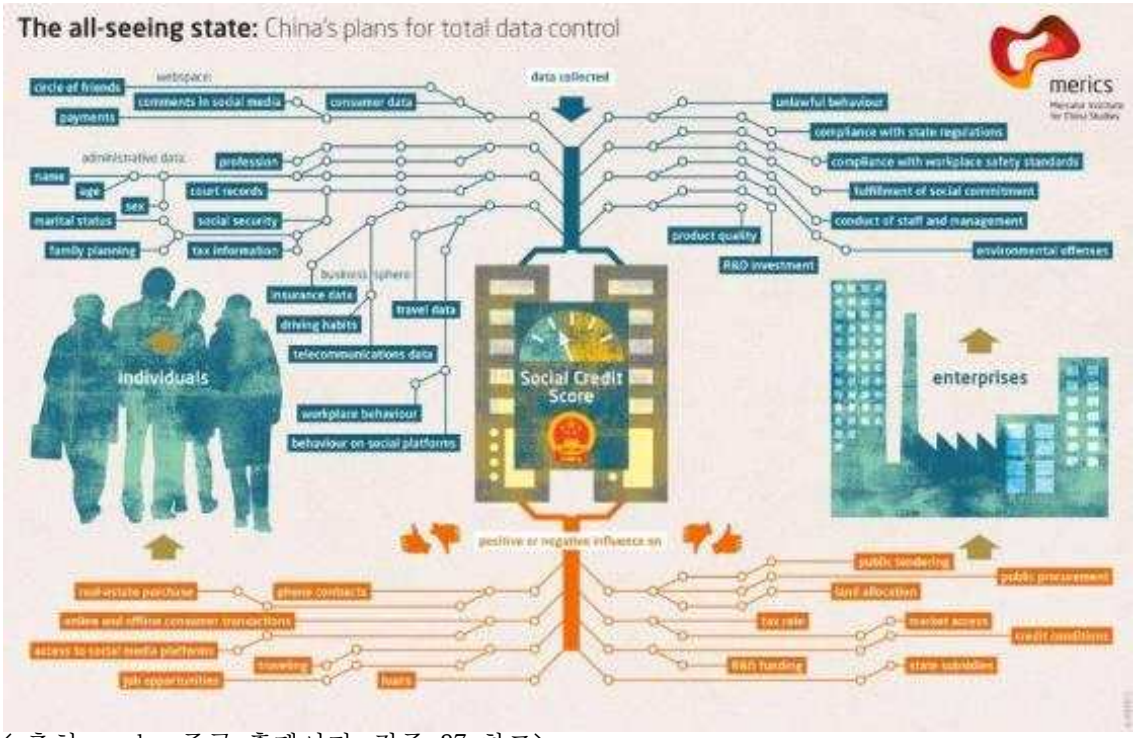
< 출처 : China AI Development Report 2018, 각주 35 참조>

이는 주로 딥러닝으로 학습이 필요한 것으로 중국의 인구는 빅데이터 및 학습을 위해 많은 데이터를 제공할 수 있는 장점으로 작용하고 있으며, 이를 활용하여 중국은 인공지능 분야의 기술 개발에 가속도를 붙이고 있다.

이러한 여건 속에서 중국은 최근 Covid-19 바이러스의 확산을 방지하는 차원에서 학교부터 관공서까지 지문인식이 아닌 안면인식을 필수적으로 시행하고 있으며, 이 안면인식시스템과 국가가 운영하는 CCTV로 저장된 모든 자료는 데이터화되어서 관리가 되고 있다. 그리고 이

35) 출처 : China AI Development Report 2018의 Figure 2-26 The ThemeScape of AI patents http://www.sppm.tsinghua.edu.cn/eWebEditor/UploadFile/China_AI_development_report_2018.pdf

전부터 계획하고 준비해온 사회신용시스템(Social Credit System) 일부 지역에 시범 적용 및 운영 후, 2020년부터 적용을 하고 있는 것으로 보인다.[36] 이는 아래 그림[37]과 같이 방대한 데이터를 정부가 직접 통제하며, 개개인의 각 활동내역에 따라서 긍정 및 부정적인 점수가 매겨지게되고, 이의 결과로 개인은 사회 신용 점수를 가지며, 국가는 이를 통하여 사회를 관리하고 있다.



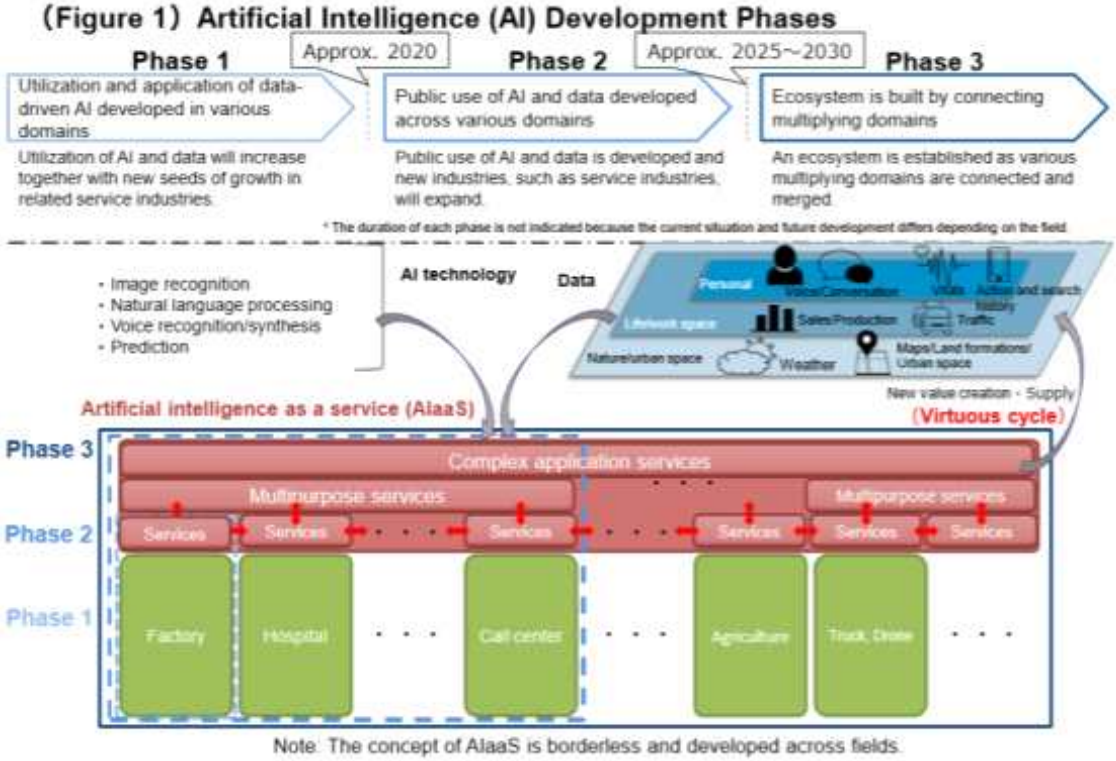
< 출처 : sohu 중국 홈페이지, 각주 37 참조>

하지만 위의 사회신용시스템은 국민의 일거수일투족을 감시하고 더 나아가서는 이를 국민에 대한 통제수단으로써 악용이 될 수 있다는 점에서 많은 비판들이 나오고 있으나, 중국 정부는 안전을 위한 조처라고 설명을 하고 있으며, 이에 대한 개인정보와 관련한 인권침해 문제 등에 대해서 대비가 필요하며, 현재 중국 상무위원회에서 개인정보보호법 및 데이터보안법 등의 제정을 검토하고 있는 것으로 파악된다.

36) 출처 : "14억 인구, CCTV 6억대로 감시" 中, 걸음걸이까지 데이터화 (20200625, 한국경제)
37) 출처 : https://www.sohu.com/a/245386306_777813

4) 일본 국가 정책

일본은 2017년 3월 인공지능기술전략회의를 통하여 인공지능의 정책과 로드맵을 발표하였다. 이는 3단계로 이루어진 발전 개념으로 다음 그림[38]과 같이 1단계-여러 분야에서 개발된 데이터 기반의 AI를 적용 및 활용, 2단계-여러 분야를 걸쳐서 개발된 AI와 데이터의 공공 사용, 3단계-다양한 분야에 연결되어 형성된 에코시스템을 적용하는 것으로 되어있다.



< 출처 : Artificial Intelligence Technology Strategy (2017), 각주 38 참조>

그 후 2019년 3월에 새로운 '인공지능 전략'을 '인공지능 사회 원칙'과 함께 발표하였다.[39] 이 발표는 일본이 지향하고 있는 Society 5.0[40]의 실현을 목표로 하고 있으며, 이를 위해 교육혁신에 방점을 둔

38) 출처 : Artificial Intelligence Technology Strategy (2017, Stragegic Council for AI Technology) https://www.nedo.go.jp/content/100865202.pdf
39) 참고문헌 : 일본의 인공지능(AI) 전략 동향 : AI 전략 2019 보고서 (2019, 소프트웨어정책연구소)
40) 스웨덴의 DX(Digital Transformation), 독일의 Industry4.0과 유사한 용어로 2015년 일본에서 소개된 개념으로 정보사회(Society 4.0) 이후의 미래사회를 말하며, 이는 인공지(AI), IoT, 로봇 등 새로운 기술들로 새로운 가치가 창출되어 다양한 사람들이 자신만의 행복을 서로 존중하면서 실현할 수 있고 또 지속가능한 인간중심의 사회를 말한다.

인재육성 및 확보 계획을 수립하였다. 특히 인공지능(AI)의 성공을 위해서 특정부처가 아닌 범부처간협업을 통하여 단순한 신기술이 아닌 장기적인 국가적 과제로 접근하고 있다. 특히 이런 역할을 통하여 국제사회에서 인공지능(AI)관련 리더십 확보하기 위해 노력하고 있으며, 인공지능(AI) 유리적인 문제도 '인공지능 사회 원칙'을 지속적으로 다자간 협의 체계를 통해 국제적 논의를 하고 있다.

그리고 '인공지능 사회 원칙'에는 아래와 같은 7대 원칙이 있다.

① 인간중심의 원칙 : 인간의 기본적 인권을 침해해서는 안된다.

② 교육,리터리시 원칙

: 양극화나 약자가 발생하지 않도록 폭넓은 교육 실시

③ 프라이버시 확보 원칙

: 개인에게 위해를 방지하기 위한 기술적 틀을 마련해야 함

④ 보안확보원칙

: 위험 평가와 위험 저감을 위한 연구개발 및 리스크 관리

⑤ 공정경쟁확보원칙 : 부와 사회적 영향력의 부당한 편중 배제

⑥ 공평성, 설명 책임 투명성 원칙

: AI 동작 결과의 적절성을 담보하는 시스템에 대한 개발 등

⑦ 혁신의 원칙 : 대학, 연구기관,기업간 협업, 연계 등

이렇게 각 분야에서 인공지능 기술을 접목하고 있으며, 특히 국방분야에서 참고가 될 수 있는 부분은 국가 안전 및 재난 방지와 관련해서 디지털 트윈을 구축하는 사업으로 보인다. 앞서 서론에서 언급한 것과 같이 디지털 트윈은 실제 상황을 시뮬레이션 상의 똑같은 모형으로 재현하여 향후 미래의 예측 및 현 문제 상황에 대한 해결책 제시 등의 다양한 시나리오는 직접 적용 혹은 실제 해결을 할 수 있는 개념으로써, 일본은 로봇 및 각종 센서 등의 신기술을 도입하여 국가의 주요 인프라 시설 등에 적용하여 점검 및 진단을 할 예정이다.

5) 영국 국가 정책

영국 정부는 2018년 4월 'AI Sector Deal' 이라는 인공지능 국가정책을 가지고 AI 분야에서 First Mover를 유지하기 위한 다방면의 노력을 하고 있으며, 이는 다음의 산업 전략 다섯 가지 분야를 강화하는 것을 기본으로 하고 있다.

- Ideas : 세계 최고의 혁신적인 경제를 위한 것으로 민간 부분의 R&D 투자를 2027년까지 2.4%까지 그리고 장기적으로 3%이상을 지출하도록 지원하는 것을 포함 하고 있다. 그리고 다른 유수의 기업들은 이미 영국에 많은 투자를 하고 있으며, 이에는 Google, Element AI, Amazon, HPE, Beyond Limits, Ironfly Technologies, Astroscale, Chrysalix 등이 있다.

- People : 좋은 직업을 창충하고 훌륭한 인재를 양성하기 위해서 저명한 글로벌 튜링 Fellowhship 프로그램을 개발하고, 기초과학 및 고급 인력양성에 많은 투자를 하고 있다.

- Infrastructure : 세계 최고의 디지털 역량을 갖추기 위해서 10억유로 이상을 투자를 할 예정이며, 이에는 5G 이동네트워크에서부터 Full-fibre broadband 등이 있다.

- Business Environment : 비즈니스를 시작하고 또한 키울 수 있는 가장 좋은 곳으로 만들기 위한 노력을 하고 있고 이는 영국비즈니스은행에서 벤처투자프로그램을 통하여 많은 기술기업들에게 투자를 하고 있다.

- Places : 영국을 통하여 주위 공동체들을 번영하도록 도와줄 수 있도록 Tech City UK와 Tech North를 확장하는 등의 역할을 하고 있다.

그 중 AI 사무소(Office for Aritificial Intelligence)는 두 중앙부처인 디지털문화미디어스포츠부(DMCS, Department for Digital, Culture, Media and Sprot)와 비즈니스에너지산업정책부(BEIS, Business, Energy and Industrial Strategy)의 연합부서로 인공지능과 데이터의 중요한 임무를 추진하고 감독하기 위해 설립이 되었다. 이 부서에서는 정부의 'AI 획득을 위한 가이드라인(Guidelines for AI Procurement)[41]' 을 제시했고, 그 내용 중

41) 참고자료 : Guidelines for AI procurement Published 8 June 2020(Gov.UK 홈페이지)

가장 중요한 10가지 고려사항들에 대해서 다음과 같이 말하고 있다.

① 정부의 AI 도입 전략내에서 각자의 조달을 포함

: 각자의 기술 및 데이터 전략들이 통합적인 AI 기술 도입에 업데이트 되어야 한다. 이는 전 정부에 아우러는 AI 도입을 지원하기 위한 전략적으로 조달을 사용하라는 것이다.

② 다양하고 종합적인 팀내에서 의사결정

: AI 기술이 접목된 상호의존적 분야를 이해하는 다양한 팀과 함께 AI 프로젝트의 개발, 평가, 전달이 더욱 효과적이며, 다음의 내용이 있다.

- 모델 개발(심층 학습 등), 데이터 윤리학, 시각화/정보 디자인 등

③ 조달 프로세스를 시작하기 전에 데이터 평가를 실시

: 데이터는 현재 인공지능(AI) 기반 솔루션 대다수의 기반이 되고 있다. 따라서 관련 데이터의 가용성은 AI 시스템의 필수 조건이기 때문에 데이터가 없다면 AI 조달에 대해 논의하는 것은 시간낭비이다. 또한 시장에 출시하기 전에 데이터 내부의 결함과 잠재적인 편견을 해결하거나 해결하기 위한 계획을 세워야 한다.

④ AI 적용에 대한 혜택과 위험 평가

: 공공의 혜택 목표를 설정하는 것은 AI 시스템을 통해 얻고자 하는 전체적인 프로젝트와 조달 프로세스를 위한 기준을 제시합니다.

⑤ 처음부터 시장에 효과적으로 참여

: 정부 지출은 공정하고 경쟁적인 시장을 만드는데 사용이 되고, 이것은 더 나은 AI 시스템으로 이어진다. AI 공급업체와의 조기 협력은 보다 나은 대응으로 이어지므로, 이는 성공적인 조달과 나은 프로젝트 완료로 이어지는 것이다.

⑥ 시장에 올바른 방향을 설정하고 구체적인 해결책보다는 도전에 초점을 맞춤

: 조달되는 AI 시스템은 해결하고자하는 도전(과제)를 해결하고 시장의 책임있고 혁신적인 대응을 촉진해야 한다. 신중하게 작성된 요구사항은 공급업체가 조달하는 부서의 필요 사항을 이해하고 최적의 해결방안을 제안하는데 도움을 줄 수 있다.

⑦ 거버넌스와 정보 보증을 위한 계획 수립

: AI 시스템의 전 수명주기에 걸쳐서 정밀 조사를 할 수 있는 적절한 감독 기구를 구축해야 한다. 그리고 AI 의사 결정의 투명성을 극대화하여 사용자에게 AI 시스템이 잘 기능한다는 확신을 주어야 한다.

⑧ 블랙박스 알고리즘과 공급업체의 잠금장치 방지

: 알고리즘의 설명 가능성과 해석 가능성을 장려하고 이를 설계 기준 중 하나로 만들어야하며, 이는 조달부서가 결과를 이해할 수 있도록 하는 방법과 기술을 사용하는 것을 의미하는 것이다. 그리고 이는 또한 공급업체에 종속될 수 있는 위험을 피하여, 다른 공급업체와 협력하여 향후 AI 시스템을 지속하거나 구축할 수 있도록 해준다.

⑨ 평가 시 AI 적용의 기술적 및 윤리적 한계를 해결해야 할 필요성에 초점을 맞춤

: 입찰 평가를 수행하기 위해 광범위한 전문 지식이 있는 다양하게 구성된 팀의 경험을 활용하여 평가 프로세스를 지원해야 하여 다음의 내용 등을 확인한다. 공급업체가 데이터 내에서 편향된 문제를 확인하고 해결하였는지? 적절한 기술 표준을 준수하였는지? 등

⑩ AI 시스템의 전 수명주기관리 고려

: 공공부문의 AI로 추진된 솔루션은 윤리적인 사용을 보장하기 위해서 구현 계획, 지속가능하고 지속적인 평가 방법, 데이터 모델에 피드백하는 매커니즘이 중요하며, 노하우 이전과 교육 역시 이루어져야 한다.

그리고 인공지능과 데이터윤리 및 혁신 센터(Centre for Data Ethics and Innovation)에서는 신뢰할 수 있는 데이터를 공유할 수 있는 정책들을 만들고 있으며, 그중 일부는 Trust matrix[42] 작성 및 배포하여 신뢰할 수 있는 데이터에 대한 기준을 삼고 있으며, 이는 5가지 영역(Value, Security, Accountability, Transparency, Control)으로 구성되어 있고, 데이터가 공유되는 것으로 예를들어 가치(Value)에 대해서는 누가 이익을 얻는지? 누가 어떤 리스크를 져야하는지?를 구체적으로 판단하기 위해서 예상되는 혜택의 설명이 명확하게 되어 있는지? 사회의 다양한 단체나 개인이 어떻게 영향을 받는지? 등을 확인하게 되어 있다.

42) 출처 : Driving forward trustworthy data sharing (2020, Centre for Data Ethics and Innovation)

6) 덴마크 국가 정책

덴마크 정부는 2018년 1월 국가적 인공지능 계획인 ‘Strategy for Denmark’s Digital Growth’를 발표한 이후 2019년 3월 다시 ‘National Strategy for Artificial Intelligence[43]’라는 인공지능에 대한 정부의 국가적 전략을 내놓았다. 이 보고서에는 두서에 4가지 목적을 명시하고 있으며, 그 중 첫 번째가 바로 『덴마크는 인공지능에 대한 공통된 윤리적이고 인간 중심적인 기준을 가져야 한다.』로 다른 어느 나라 보다 더 인공지능의 윤리적인 측면을 더욱 강조하고 있다.

윤리적인 인공지능을 위한 덴마크 정부의 추진계획
1. 인공지능을 위한 윤리적 6가지 원리 ①자결권(Self-determination) ②존엄성(Diginity) ③책임성(Responsibility) ④설명가능성(Explainability) ⑤평등과 공정(Equality and justice) ⑥개발(Development)
2. 데이터 윤리 위원회의 설립
3. 보안과 인공지능
4. 인공지능의 개발과 사용에 대한 법적인 명확성
5. 공공 분야에서의 투명한 인공지능 알고리즘 사용
6. 비즈니스 분야에서의 윤리적으로 책임있고 지속가능한 데이터의 사용
7. 인공지능 표준에 대한 덴마크 각인(imprint)

그리고 양질의 많은 데이터에 대한 필요성을 윤리성에 이어 두 번째로 중요한 추진과제로 내세우고 있으며, 이를 위해서 날씨, 기후 해양 데이터에 대한 무상 제공하고 있으며, 특히 공통된 덴마크 언어에 대한 데이터 확보를 위한 다양한 방법으로 정부에서 추진하고 있다.

< 출처 : 덴마크 인공지능 국가정책(2019), 각주 43 참조>

43) 출처 : National Strategy for Artificial Intelligence published Mar 2019 (The Danish Government)

4. 일반 민간 인공지능 플랫폼

아마존 '알렉사', 구글 '어시스턴트' 등과 같은 인공지능 플랫폼들을 각 기업들이 앞다투어 출시하고 있으며, 이는 많은 데이터를 학습하면서 다른 기업들보다 기술을 앞서기 위한 것으로 마이크로소프트의 window OS와 같이 플랫폼을 선점하는 것은 향후 국가 경쟁력에 있어 중요하다. 이와 관련하여 훈련국 캐나다에 있는 Element AI社에서 개발한 AI OS 플랫폼 등을 확인해 보고자 한다.

일반적으로 인공지능 플랫폼은 아래와 같은 기능이 필요하다.

< 플랫폼에 필요한 기능들 >

- 머신러닝 (Machine Learning)
- 자동화 (Automation) : End to End Auto ML[44)]
- 자연어 처리와 이해 (Natural Language Process and understanding)
- 클라우드 환경 (Cloud infrastructure)

캐나다의 인공지능 기술에 큰 역할을 기여하고 있으며, 죠슈아 벤지오 등이 공동창업한 Element AI社는 우수한 인력으로 최첨단 인공지능 기술을 가지고 있으며, 각 기업의 필요 및 요구와 특수한 상황도 문제없이 맞춤 설계를 하고 있고 End to End Auto ML 플랫폼을 제공하고 있는 기업이며, 인공지능 AI OS의 특징으로 아래의 8가지가 있다.

- Optical Character Recognition(OCR) : 광 문자 인식[45)]

  : 저품질, 손글씨 및 움직이는 문자도 인식을 쉽게 하도록 설계

< 출처 : Element AI 社 홈페이지, 각주 45 참조>

44) End to End Auto ML 플랫폼은 고객이 어렵고 복잡한 중간과정 없이 데이터만 준비하면 다양한 머신러닝 및 딥러닝 학습을 위한 알고리즘을 쉽게 쓸 수 있도록 만들어진 AI 전용 플랫폼임.

45) 사진 출처 : https://www.elementai.com/api/ocr

- Time Series Forecasting

: 미래 예측 알고리즘으로 유명한 M4 예측 대회[46] 우승 시스템인 우버의 ES-RNN 보다 더 우수한 예측 성능을 보인 순수 딥러닝 알고리즘인 N-Beats를 논문을 통해 소개하고 있다.

- Object Detection & Counting

: 재고 관리 등에 필수적인 시스템인 물체 식별 및 개수 계산

- Optimization

: 매장 물품, 자산 관리 및 트럭 배차 등에 최적화된 결정을 할 수 있도록 지원해 주는 시스템

- Explainability

: 설명 가능한 AI 시스템으로써 인공지능의 윤리적인 측면에서 바라볼 때 향후 인공지능 개발에 필수적으로 중요한 두가지 요소인 결과에 대한 설명가능성 및 공정성에 대한 내용으로 참고 예시 화면은 다음과 같다.

< 설명 가능성 > * 출처:ElementAI社 홈페이지 < 공정성 분석 >

- Recommender System : AI 기반의 강력한 의사결정 시스템임.

- Anomaly Detection (변칙 식별)

: 사이버보안, 위조식별, 예방점검으로써, 기록 등을 모니터링하여 비정상적인 사용자의 행동이나 사이버공격을 인지하여 막아주는 시스템임

46) M-시리즈 예측 대회는 예측 전문가인 Spyros Makridakis에 의해 견인된 대회들로 다양한 예측 방법들을 정확도를 가지고 비교 및 평가를 하는 대회이며, 총 4회(M1-1982,M2-1993,M3-2000, M4-2018~2020)가 완료되었으며, 최근 M4 대회에서는 머신러닝과 통계적 기법을 동시에 사용한 우버 테크놀로지가 우승을 차지하였다. 참고로 우버는 600개 이상의 도시에서 실시간으로 예상 고객 수와 가능한 운전자 수 등에 대한 정확한 예측 시스템 기반으로 하는 사업이며, 이를 위해서 우수한 예상 시스템을 개발하기 위해서 많은 비용을 투자를 하고 있다.

- Text Extraction & Analysis

: 방대한 양의 문서등에서 필요한 정보를 가진 텍스트를 추출해 주고 자료를 분석해주는 시스템으로써, 업무 효율이 획기적으로 높아질 뿐만 아니라, 고객들의 반응에 대한 감성분석(Sentiment analysis) 등에도 효과적으로 사용된다.

이외의 주요 플랫폼과 특징을 살펴보면 다음과 같다.

- Google의 양방향 언어모델 : Bert
- OpenAI의 강력한 언어 모델 : GPT-3(Generative Pre-Training 3) 이는 몇 개의 키워드만 입력하면 작문을 해주는 혁신적인 AI 언어생성 모델이며, 알고리즘으로 알려져 있다.
- TensorFlow : Deep flexibility, True portablility, Auto- differentiation, Connect research and production
- Rainbird : Visual User Interface, RBLang - An Intuitive Language, Controlled Learning Algorithms,
- Inforsys Nia (지식기반의 AI 플랫폼) : Infosys Information/ Automation/ Knowledge Platform
- Wipro HOLMES (로봇, 드론 등의 인지적인 과정을 자동화하는 AI 플랫폼) : Digital Virtual Agents, Predictive Systems, Cognitive Process Automation
- Dialogflow (자연어 대화에 특화되어 있음.) : Machine Learning, Integrations, Conversation Support
- Premonition (법률적인 AI 프랫폼) : Know the track record of your Attorney, Select Co-Counsel who have never lost in front of certain Judges, Analyze the Court, Judge and Opposing Counsel by their Win Rates and Results.
- MindMeld : Discover on-demand music and video, Enable

위에서 살펴본 바와 같이 비국방분야에서 많은 국가들이 인공지능 기술을 정부의 의사결정 시스템 및 국민들의 신용도에 대한 관리 등과 같은 곳에 직접적으로 사용하고 있다. 그리고 일반 기업에서 제공하고 있는 인공지능 플랫폼을 살펴보며 인공지능 알고리즘에 있어서 이미 상당한 기술이 진척이 있는 것을 알 수 있다. 이와 같은 인공지능 알고리즘은 비단 민간부문 뿐만 아니라 국방 분야에서도 활용이 되고 있으며, 다음 장에서는 훈련국인 캐나다와 주요 국가에서 국방 분야에 사용하고 있는 인공지능 기술을 확인해보고자 한다.

제3장. 훈련국 캐나다와 주요국의 인공지능 국방분야 적용 추이 및 개발 현황

4차 산업혁명시대에서의 전쟁 양상은 인공지능(AI)의 등장과 함께 이전과는 전혀 다를 것이며, 이런 패러다임 변화를 인지하고, 훈련국인 캐나다 및 주요국가에서 인공지능 기술을 국방분야에 적용하고 있는 사례 및 수준 등을 파악함으로써 우리나라에 향후 적용하거나 참고할 수 있는 부분을 확인하여 미래의 전장에 자칫 뒤처지지 않도록 해야 할 필요가 있다.

1. 캐나다

면적에서는 큰 차이가 있지만, 국토의 삼면이 바다로 둘러싸인 우리나라와 대서양과 북극해를 길게 접하고 있는 캐나다와 유사점이 많아 국방 관련 조직 및 개발 중인 4차 산업혁명과 관련한 프로젝트 등을 연구 및 참고해서 우리나라에 적용가능한 요소들을 발견할 수 있으며, 우선은 다시 생소한 캐나다의 국방 조직 및 획득 관련 업무를 살펴보면 다음과 같다.

1) 캐나다 정부의 국방정책

현 캐나다 정부는 2017년 국방정책서[47]를 발간하면서, 표어이자 정책 비전으로 아래의 세 가지를 내세웠다.

① Strong (at home) : 지속적인 국내 방어 및 신속한 자연재해 복구 등 국가 비상상황에서의 강한 캐나다 군대 육성 등

② Secure (in North America) : 북미방공사령부(NORAD)[48]를 통한 미국과의 강한 연대와 첨단 국방기술 연구 등

③ Engaged (in the world) : NATO, UN 및 동맹에 대한 견고한 헌신 등

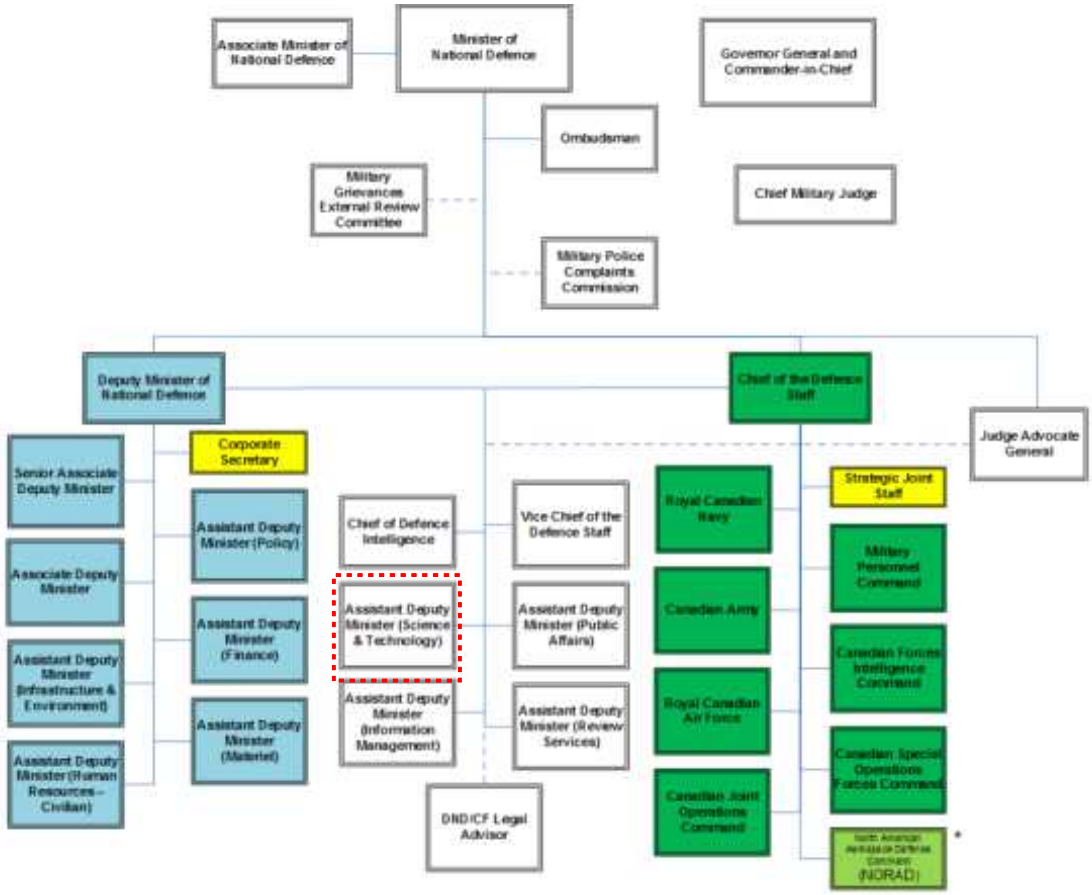
---

47) 출처 : 캐나다 국방정책서 (Strong, Secure, Engaged)
http://dgpaapp.forces.gc.ca/en/canada-defence-policy/docs/canada-defence-policy-report.pdf
2015년 연방선거를 통하여 이전 집권당인 보수당에서 현 집권당인 자유당으로 바뀌며, 정책기조도 변화함.

48) 북미 방공 사령부(NORAD, North American Air Defense Command) : 북극을 가로질러 공격하는 소련 폭격기와 미사일의 위협을 대응하기 위해 창설됨

국방정책서에는 4차산업혁명과 관련한 Deep Learning, Autonomous systems 등의 신기술에 대해서 국내와 국제적인 법률을 철저히 지키면서 새로운 신기술들을 적용한다는 내용이 있다.

2) 국방부 내 조직도[49]



< 출처 : 캐나다 국방부 홈페이지, 각주 49 참조>

위 조직도의 붉은색 점선으로 된 Assistant Deputy Minister (Science and Technology)는 국방부(DND)와 캐나다군(CAF)의 과학과 기술에 투자되는 부분을 담당하고 있으며, 그 조직의 산하에는 국방과학기술의 전문기관인 DRDC[50]를 두고 있으며, 이 곳에서 신기술 등에 대한 연구가 이루어지는 것으로 보여진다.

---

49) 출처 : 캐나다 국방부 홈페이지
https://www.canada.ca/en/department-national-defence/corporate/organizational-structure.html

50) 캐나다 국방과학기술소(DRDC, Defence Research and Development Canada) : 국방부의 산하기관(Agency)로 되어 있으며, 큰 캐나다 국토 전역에 8군데에 약 1,300여명의 직원들이 일하고 있다. 우리나라의 국방과학연구소와 같은 기능을 하는 것으로 판단되며, 현 Assitant Deputy Minister(Science & Technology)는 DRDC의 대표로도 활동하고 있다.

3) 캐나다의 국방 획득 절차

캐나다의 국방 획득 절차는 아래의 5단계가 있고, 대략 10년 소요됨.

① 1단계 : 식별 (Identification)

- 첫 번째 단계에서는 캐나다군에 필요한 소요식별 및 확정을 한다.

- 2단계로 진행되기 전에 국방전력위원회(Defence Capabilities Board)와 국방획득을 위한 독립적인 검토자문단[51](Independent Review Panel for Defence Acquisitions)의 의한 높은 수준의 검토를 거쳐야 한다.

② 2단계 : 방안 분석 (Options analysis)

- 두 번째 단계에서는 프로젝트팀이 작전요구성능 초안과 최적의 방안을 포함한 획득방안들에 대한 완성된 사업분석보고서를 준비하여 국방전력위원회와 사업관리위원회의 검토 및 승인을 받아 다음 단계로 진행하기 위한 자금을 할당받는다.

③ 3단계 : 정의 (Definition)

- 세 번째 단계에서는 필요 전력을 달성하기 위해 해야할 것들에서부터 어떻게 최적의 방안을 실행되도록 해야할지를 결정한다.

- 프로젝트가 사업관리위원회의 승인을 얻으면, 프로젝트팀은 국방부장관이나 예산위원회로 Corporate Submisstion을 준비하여 다음단계로 가기위한 지출승인을 얻는다.

④ 4단계 : 실행 (Implementation)

- 네 번째 단계에서는 예산내에서 적기에 전력을 맞추기 위해서 캐나다 공공서비스조달청(Public Services and Procurement Canada)와 협력하여 일반적으로 경쟁을 통해 납품할 업체와 계약을 하고 조달하는 것이다.

51) 제3 자의 관점에서 사업을 검토하고 있으며, 연간 총사업비 $100M (약 900억) 이상의 사업에 대해서 검토함.

- 복잡한 장비에 대해서는 일반적으로 국방부와 소요군이 시험과 검증을 하고 있다.

* 경쟁 입찰 평가[52)]

- 일반적으로 평가 점수는 아래와 같이 가격, 기술적 이익 그리고 가치제안(10% 이상) 이라는 3가지 영역으로 구성되어 있다.

(가치제안 관련 추가 설명)

- $100M 이상의 계약 금액($20M~$100M에 대해서는 적정성을 판단한 뒤 적용)에 대해서 업체의 가치제안(Value Propositon)을 의무적으로 적용하게 되며, 이는 'Industrial and Technological Benefits(IBT) policy'에 의해 규정이 되어 있으며, 크게 5가지 영역에 있어서 캐나다의 산업과 기술적 이익을 위해 계약당사자가 이행해야 할 것들에 대해서 평가를 하는 것이다. 우리나라의 절충교역과 유사한 성격으로 볼 수 있음.

ⅰ. 캐나다 방산 업계에서의 작업 수행

ⅱ. 중소업체를 포함한 캐나다 국적의 협력업체 발전

ⅲ. 캐나다 내에서의 연구개발(R&D)

ⅳ. 캐나다에서의 해외수출 실적 및 전략

ⅴ. 혁신적인 작업기술 개발과 훈련

⑤ 5단계 : 종결 (Closeout)

- 다섯 번째 단계는 대상 서비스나 장비가 충분히 성능을 구현할 때 시작하며, 이는 프로젝트팀이 국방부로 전력이 모든 조건과 제한을 만족하도록 획득되었다는(획득시 배운 교훈 포함) 공식적인 통보를 한다.

- 이 종결단계는 보통 3개월이내에 완료된다.

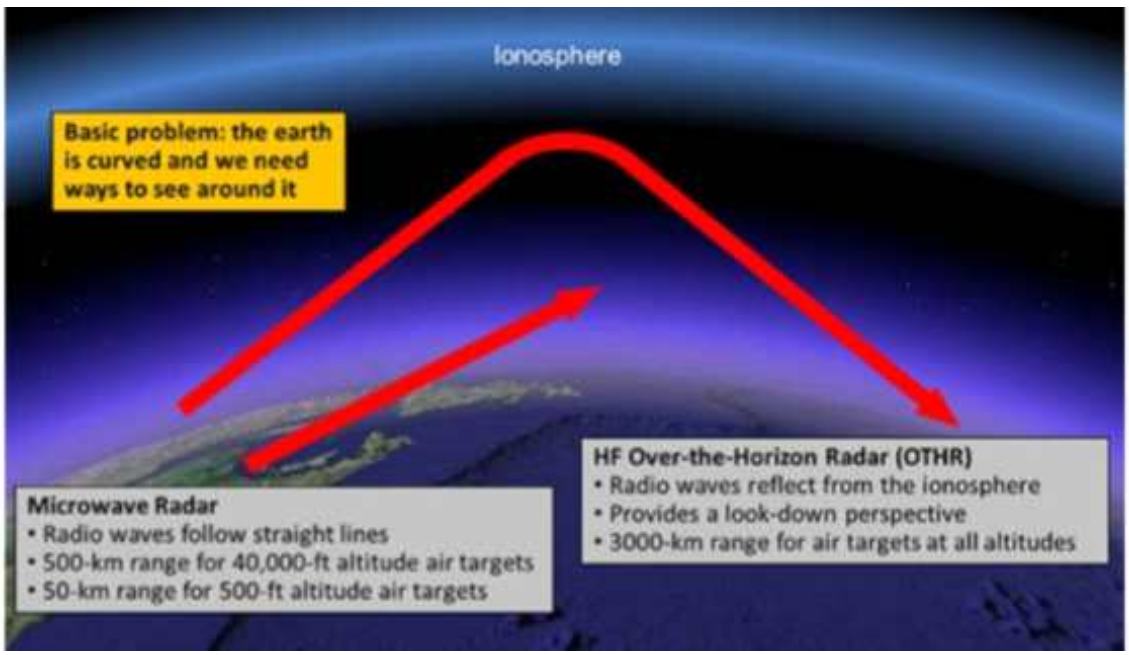
52) 출처 : Industrial and Technological Benefits Policy: Value Proposition Guide (2018, 캐나다 정부)

4) ADSA S&T[53] 사업

캐나다 국방과학기술소(DRDC)에서 진행 중인 사업 중 하나로, 아래는 ADSA S&T 사업으로 수행되고 있는 큰 4가지 카테고리이다.

① Over the Horizon Radar(OTHR) Project

지평선 너머를 탐지하기 위한 초지평선 레이더(OTHR) 연구로써, 아래의 그림(출처-각주)과 같이 Microwave Radar는 직진성이 강하여 곡선인 지구 형상으로 500km 까지만 탐지가 가능하지만 OTHR은 이온층(Ionosphere)에 반사되기 때문에 3000km까지 탐지를 할 수 있다.



< 출처 : 캐나다 국방부 홈페이지, 각주 53 참조>

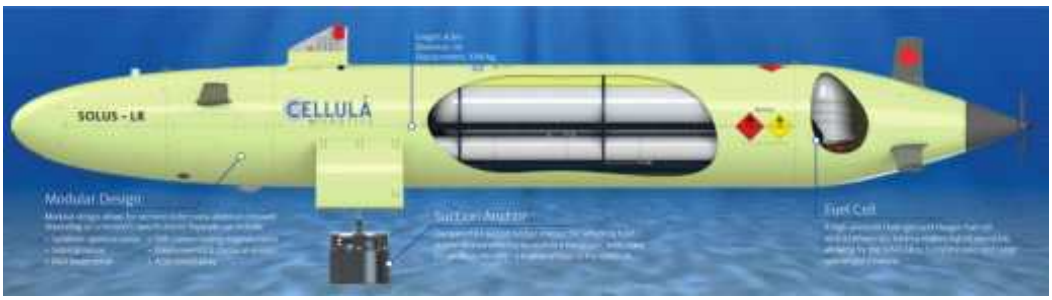
② Canadian Arctic Underwater Sentinel Experimentation (CAUSE) Project

캐나다 북극 해저 경계에 관한 연구이며, 무인잠수정(UUV) 등이 있음

53) ADSA S&T(All Domain Situational Awareness Science and Technology Program) : 캐나다 전역의 상황을 인식하는 과학 기술 프로그램으로 2015년부터 2020년까지 $133M (약 1,200억원)을 투입하여 북미지역의 경계활동을 위한 혁신적인 방안을 개발하기 위해 지원하는 사업임
https://www.canada.ca/en/defence-research-development/programs/all-domain-situational-awareness-program.html

* 무인잠수정 사업, Unmanned Underwater Vehicle (UUV)[54]

- 캐나다의 해저시스템 전문기업인 셀룰라 로보틱스社(CELLULA Robotics)에서 개발 중으로,
- 최근 1차('19.12) 및 2차(' 20.03) 시제 시험을 성공적으로 완료하여 '20년 여름에 프로젝트 완료 예정이며,
- 고압력의 수소-산소 연료전지(리튬이온베터리)를 사용한 작전반경 2,000km의 장거리 무인잠수정임
- 해저 고정(Suction Anchor) 기능을 통해 저전력 모드에서 감시 위치를 몇 주 혹은 몇 달 동안 지키며 작전 수행이 가능함



< 출처 : 셀룰라 로보틱스社 홈페이지>

③ Threat, Requirement and Gap (TRG) Analysis Project

캐나다 북쪽의 공중, 해상 및 해저에 대한 위협에 관한 연구로 아래와 같은 프로젝트들이 있다.

- 지능 감시, 로봇 및 위성시스템 전문 캐나다 기업인 MDA社와 계약한 All Domain Sensor Mix Evaluation Tool (ADSMET)
- 인공지능 알고리즘 개발 업체인 Complex Systems Inc와 계약한 Context-aware sensor selection layered architecture for Arctic surveillance

④ Compression of the Tasking, Collection, Processing, Experimentation and Dissemination (TCPED) Cycle Project

54) Unmanned Underwater Vehicle(UUV)는 Autonomous Underwater Vehicle(AUV)라고도 하며, 일반적으로 무인장비는 미지의 환경에 대한 지도를 생성을 해야하므로 SLAM(Simultaneous Localization and Mapping) 기능이 중요하며, 게임 등에는 SLAM에 딥러닝을 적용한 기술이 많이 사용되고 있음.

위성을 활용한 국토 경계 능력을 향상시키기 방안들에 관한 연구이며, 작전 선박 탐색, 식별 및 추적을 위한 멀티 위성데이터 통합 등이 있음

* 작전 선박 탐색, 식별 및 추적을 위한 멀티 위성데이터 통합 사업, Multi-satellite data integration for operational ship detection, identification and tracking[55)]

- 캐나다의 지상관측 및 원격측정 전문기업인 C-Core社에서 개발 중인 프로젝트로 선박 식별, 확인 및 추적을 위성 자료를 활용하여 머신러닝 방법으로 분석하는 연구임
- 위성 데이터는 Sentinel-2 영상으로 L-1C와 L-2A를 사용함[56)]

Table 1: Sentinel-2 product types

Name	High-level Description	Production & Distribution	Data Volume
Level-1C	Top-of-atmosphere reflectances in cartographic geometry	Systematic generation and on-line distribution	600 MB (each 100x100 km2)
Level-2A	Bottom-of-atmosphere reflectance in cartographic geometry	Systematic generation and on-line distribution and generation on user side (using Sentinel-2 Toolbox)	800 MB (each 100x100 km2)

< 출처 : Sentinel.esa.int 홈페이지 >

- 자료분석은 머신러닝 방식 중 딥러닝을 포함한 지도방식의 여러 가지 알고리즘(LDA, QDA, SVM, KNN, NNet, DT, Combined Classifiers)을 적용하였으며, 일부 선박과 빙하 사진을 분류하는 정확도는 그 중에서 QDA가 가장 높은 결과를 보임.

또한 드론 및 무인정찰기등의 연합무인감시장비 사업인 JUSTAS(Joint Unmanned Surveillance Target Acquisition System) 프로젝트는 최초 2005년에 Option Analysis 단계에 있었지만, 요구성능 미충족 및 모든 요구성능 충족하는 업체의 제한 등의 사유로 지연이 발생하였으며, 이후 2017년 동맹국들의 체계 시스템 용어 변경에 따라 프로젝트명을

55) 출처 : Multi-Satellite Data Integration for Operational Ship Detection, Identification and Tracking, Progress Report 2(2019, DRDC)

56) Sentinel 위성은 유럽우주국(ESA, European Space Agency)에서 개발 및 담당하고 있으며, 획득된 위성영상을 무료로 배포하고 있다. 위의 표에 나와있듯이 Sentinel-2는 두가지로 제공되는데, Level-1C는 대기상태가 보정되지 않은 TOA(Top-of-atmosphere) 영상이며, Level-2A는 대기 보정이 된 BOA(Bottom of atmosphere) 자료임.

RPAS(Remotely Piloted Aircraft Systems)로 바꾸었으며 이는 국방정책서(The Strong, Secure, Engaged) 실행계획[57] 50과 91에 따라 다시 사업을 착수하게 되었고, 대략적인 사업개요[58]는 다음과 같다.

< RPAS 사업개요 >

1) 예상 총사업비 : 10억 달러 ~ 50억 달러

* 포함내용 : 프로젝트 관리비용, 인프라, 계약 및 비상예비비(Contingency)

2) 장비의 예상 기대 수명 : 약 25년

3) 현 단계 : 정의(Definition Phase)

* 현재는 캐나다의 항공기 관련 전문 기업인 L3 테크놀로지 MAS 기업과 미정부/제너럴 아토믹스 에어로티컬 시스템기업(General Atomics Aeronatical Systems, 미국 캘리포니아에 본사가 있으며 무인정찰기인 MQ-1 Predator로 알려져 있음)을 압축되었으며, 2021년에 최종 업체를 선정할 예정이다.

4) 리스크 요인 및 방안

: 프로젝트 팀에서 예상하고 있는 리스크 요인으로는 납기를 맞추기 위한 고도로 숙련된 전문프로젝트 관리인원이 불충분 하다는 것이며, 이는 계약상대의 인력(contractor resources)을 고용하여 리스크를 완화할 수 있을 것으로 판단하고 있으며,

또한 armed RPAS를 제작 할 수 있는 업체가 제한되어 있어 가격 협상에서의 어려움을 적시하고 있다.

이렇듯 캐나다는 4차 산업혁명에 발맞추어 국방과학기술소(DRDC)를 중심으로 하여 머신러닝 등과 관련된 신기술 및 선행연구를 수행하고 있으며, 국방부 내 조직인 Assitant Deputy Minister(Science & Technology)와 DRDC의 대표를 공동 역임하고 있는 것과 같이, 계획과

57) 실행계획 50 (New Initiatives 50) “Invest in medium altitude remotely piloted systems.”
- 실행계획 91 (New Initiatives 91) “Invest in a range of remotely piloted systems, including an armed aerial system capable of conductiong surveilance and precision strikes.”

58) 출처 : Remotely Piloted Aircraft System (RPAS), (2019, 캐나다 국방부 홈페이지)

실행과의 긴밀한 관계를 맺고 있으며, 이러한 관계는 현재 우리나라의 방위사업청과 국방과학연구소와의 역할 부분에 있어서도 신속한 의사결정과 실행이 될 수 있는 유기적인 관계를 통한 기술혁신과 무인잠수정 적용, 선박의 식별 및 추적 등에 머신러닝(Machine Learning)과 같은 신기술의 활용하고 있는 점과 컴퓨터 비전 등에 머신러닝 알고리즘이 사용되는 무인정찰기 등의 획득을 위해 큰 규모의 사업예산을 가지고 RPAS 사업을 추진하는 점과 이를 추진하는 과정에서 성공적인 사업관리를 위한 관련 전문가들을 고용하기 위한 계획 등은 우리나라의 국방관련 인공지능 관련 기술 개발 계획하는데 참고할 수 있을 것으로 보인다.

2. 주요 국가의 국방분야 인공지능 적용 현황

1) 미국

美국방부는 인공지능(AI)와 머신러닝 분야를 각 분야에 접목시킴에 있어서, 양대 두 축인 「National Defense Strategy」[59]와 「Summary of the 2018 Department of Defense Artificial Intelligence Strategy : Harnessing AI to Advance Our Security and Prosperity」[60]에 따라 업무를 추진하고 있으며, 국방뿐만 아니라 국가를 이롭게 하기 위한 전략으로 다음과 같은 전략적 접근을 하고 있다.

① 핵심 임무를 해결하기 위한 AI가 가능한 역량 제공
(Delivering AI-enabled capabilities that address key missions)

※ 적용 예시 : 상황 인식 및 의사 결정 개선, 장비 작업 안전 제고, 예방적 정비 및 부품공급 적용, 사업 절차 간소화 등 (새로운 방법의 작업 방법을 도입하고, 지루한 인지적 혹은 육체적 업무를 제거함으로써 인력의 역량을 증강시키는 AI 시스템 영역에 최우선으로 집중할 예정.)

② 탈중앙화된 개발과 실험을 가능케 하는 공통된 기초를 통해서 국방 전영역에 AI의 영향 증대 (Scaling AI's impact across DoD through a common foundation that enables decentralized development and experimentation)

※ 실험실이나 중앙정부 사무실이 아닌 실제 사용자들에 의해 발견된 최전선에서의 실험적인 방식으로 AI가 가능한 역량이 이끌어질 것이며, 이는 AI에 접근을 증대시키고 민주화시키는 중요 구성블록과 플랫폼을 통해서 구현되며, 동시에 디지털화와 스마트 자동화를 통한 AI 적용을 위한 준비를 할 예정임.

③ 업계 최고의 AI 인력 양성 (Cultivating a leading AI workforce)

※ MOOCs, e-Book, online 비디오와 전문가 수업 등

④ 상업, 학술 및 국제 동맹과 파트너들과의 협력 (Engaging with commercial, academic, and international allies and partners.)

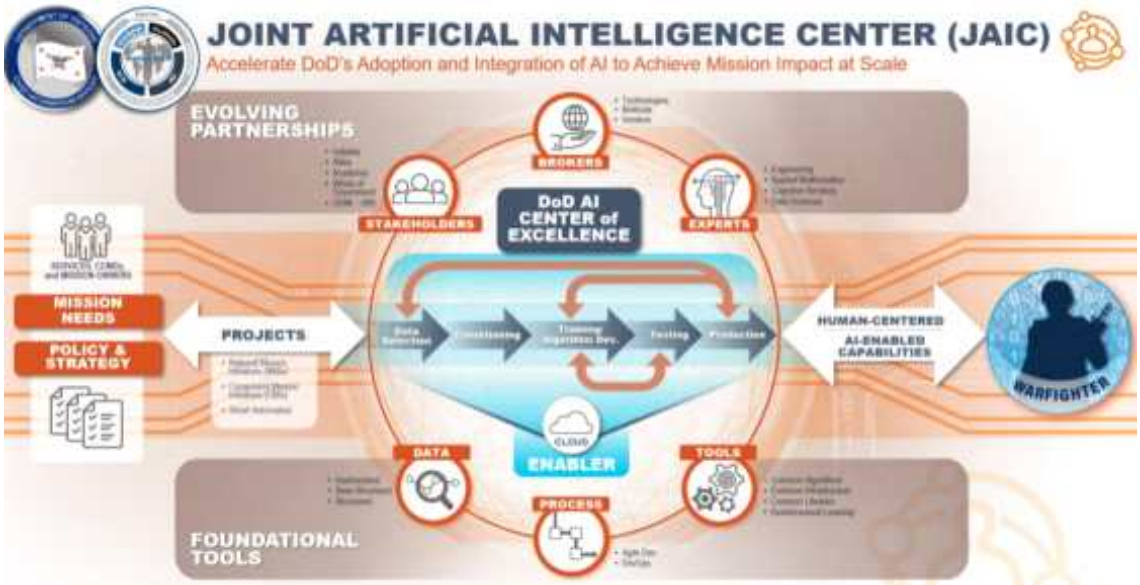
---

59) 원문 : https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf
60) 원문 : https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

⑤ 군사적 윤리와 AI 안전성에 있어서의 선도
(Leading in military ethics and AI safety)

※ 국방부는 AI의 가치를 촉진하기위해 합법적이고 윤리적인 방법으로 AI를 사용하기 위한 원리를 가이드하고 그에 대한 비전을 명확히 할 것이며, 이는 회복이가능하며(resilient), 강인하고(robust), 신뢰할만하며(reliable), 그리고 안전한(secure) AI 시스템을 연구개발하기위해 투자하고 있음. 또한 보다 더 설명이 가능한 AI(Explainable AI)를 가능케 하는 기술들에 대한 연구를 위해 자금지원을 지속적으로 할 예정임.

그리고 이 전략에는 크게 두가지 카테고리, 즉 국가적 임무과제(NMIs, National Mission Initiatives)와 구성단위 임무과제(CMIs, Component Mission Initiatives),로 나누어서 실행과제들을 관리하고 있으며, 이를 추진할 기관으로 美국방부는 연합AI센터(JAIC, Joint Artificial Intelligence Center)를 설립하였고, JAIC는 DARPA, DIU, 기타 국방부 연구소 및 다른 기관들과 함께 협력하여 국방분야의 인공지능 기술 발전을 이끌고 있으며, 이는 아래의 JAIC의 역할을 나타낸 그림과 같이 위에서 언급한 5가지 전략의 이행 및 AI 계획, 정책, 거버넌스 윤리, 사이버 보안 및 다자간(연구소와 일선 부대와의) 조정 등의 역할을 할 것으로 기대하고 있으며 주요 참고할 만한 진행 사항들은 다음과 같다.



< 출처 : JAIC AUSA 발표 자료, 주석 61 참조> [61]

61) 출처 : DoD Tech Talk :JAIC (2019), https://meetings.ausa.org/autonomy/pdf/JAIC.pdf

현재 연합AI센터는 NMIs 프로젝트의 하나로 국방혁신단(DIU, Defense Innovation Unit)[62] 및 美공군과 함께 민간에서 개발된 AI 기반의 응용 프로그램을 E-3 Sentry, F-16 Fighting Falcon, F-35 Lightning Ⅱ와 Bradley Fighting Vehicle에 적용하여 정확한 정비 소요를 통하여 운용률을 높이고, 비용을 감소시키는 예방정비(Predictive Maintenance) 프로젝트를 성공리에 진행 중이며, 2019년 DIU 보고서[63]에 따르면 지도학습 머신러닝 알고리즘을 적용하여 부품이 고장 나기전에 미리 정비기술자에게 알려주어 사전에 고장 가능성이 높은 부품을 교체함으로써 아래와 같은 결과를 얻었고, 이와 같은 결과에 힘입어 DIU는 C3.ai社와 2020년 1월 5년간 $95M의 후속계약을 체결하는 등 예방정비 프로젝트를 계속 확대 적용해 나가고 있다.

- 적용 시제품 : C3.ai社 (미국 실리콘밸리에 위치한 스타업기업)
- 임무 능력(Mission Capability) : 3~6% 개선
- 부품 대기하는 기본 수준의 항공기 그라운드 빈도 : 35%까지 감소
- 비계획 정비소요 : 40%까지 감소

또한 연합AI센터는 국가전략의 세 번째인 인력양성과 관련하여 美국방부의 리더들을 위한 “AI 기술 이해하기[64]”란 가이드자료를 작성 및 배포하여 인공지능 기술에 대한 정확하고 신속한 이해를 돕고 있다.

美국방 관련 인공지능 기술에 있어서 중추적인 역할과 혁신적인 기술들을 개발하고 있고, 60년 넘는 역사와 이미 많은 인공지능 기술 관련 프로젝트를 수행하고 있는 고등연구계획국(DARPA, Defense Advanced Research Projects Agency)에서는 2018년 9월 다년간 동안 약 $20억달러 이상을 차세대 인공지능(AI Next)라는 캠페인을 통해 인공지능 기술 발전에 투자를 진행하고 있으며, 다음의 5가지 영역에서 특별히 강력한 역량을 가질 수 있도록 노력하고 있다.

---

62) DIU(Defense Innovation Unit)는 빠르게 변화하고 있는 민간 기술을 국방력을 강화하기 위하여 빠른 속도로 미군에 적용하기 위해서 설립된 기관으로, 일반적으로 국방부의 계약 과정은 18개월 이상이 걸리나, DIU는 60~90일 정도로 단축할 수 있으며, 계약 후 통상 12~24개월의 시제품 프로젝트를 진행한 후, 성공적인 시제품의 경우에는 표준 국방부 계약 규정이나 후속 양산 계약 단계로 전환하고 있으며, 주로 5가지 분야(AI, Autonomy, Cyber, Human System, Space)에 집중하고 있다.

63) 출처 : Annual Report 2019 of Defense Innovation Unit (2019, DIU)

64) 출처 : Understanding AI Technology published April 2020 (2020, JAIC)

① 새로운 역량 (New Capabilities)

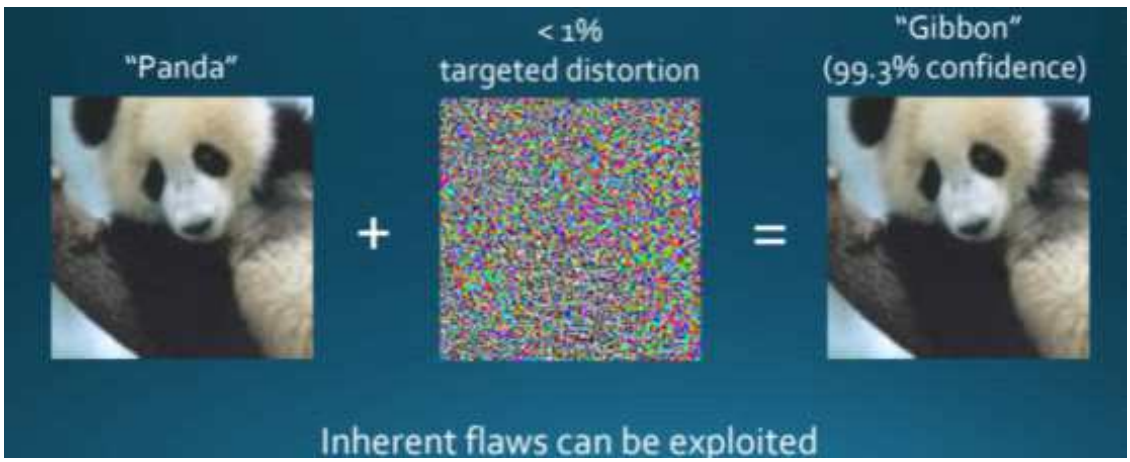
: 인공지능 기술들은 복잡한 사이버 공격의 실시간 분석, 위조된 이미지 식별, 자연어 기술들에 적용되고, 중요한 국방부 사업 진행의 자동화 등이 가능하도록 하는 하나의 새로운 역량이 되고 있음.

② 강인한 인공지능 (Robust AI)

: 인공지능 기술들은 물류공급망, 마이크로바이오 시스템 및 사이버공격 알림 등과 같은 다양한 분야에 큰 비중으로 적용되고 있지만, 동시에 실패 모드에 대해서는 이해하기가 어렵다. 따라서 DARPA는 인공지능 기술의 단점을 분석적이고도 실증적인 연구개발을 통해 개선을 할 예정이며, 이는 신뢰성있는 임무수행이 필요한 전략적 작전에 인공지능 기술을 사용하기 위한 필수적인 과제임.

③ 적대적인 인공지능 (Adversarial AI)

: 오늘날 가장 강력한 인공지능 도구는 머신러닝(ML)이나, 머신러닝은 아래 그림과 같이 사람은 결코 속지 않을 입력에도 쉽게 오인을 할 수 있으며, 훈련되는 데이터도 손상될 수 있다. 이와 같이 사이버 공격에 그 소프트웨어 자체는 영향을 받기 쉽기 때문에 이 분야에 대한 준비가 필요한 것임.



< 출처 : A DARPA Perspective on Artificial Intelligence, 주석 65 참조 > [65]

④ 높은 성능의 인공지능 (High Performance AI)

: 최근 머신러닝의 성공은 빅데이터와 소프트웨어 라이브러리와 더불어 지난 10여년 간의 컴퓨터 성능이 발달된 결과였다. 하지만 데이터에

65) 출처 : A DARPA Perspective on Artificial Intelligence (DARPA, John Launchbury)

접근하고 전략적 배치를 위해서는 저전력에서의 높은 성능이 필수적이다. DARPA는 최첨단 디지털 프로세서를 통해서 1000배 속도와 1000배 전력효율을 가진 인공지능 알고리즘의 아날로그 처리를 시현하기도 했다. 또한 지도학습 데이터에 필요한 요구조건을 상당히 줄이는 방법을 연구함으로써, 현재의 비효율적인 머신러닝을 개선하고 있음.

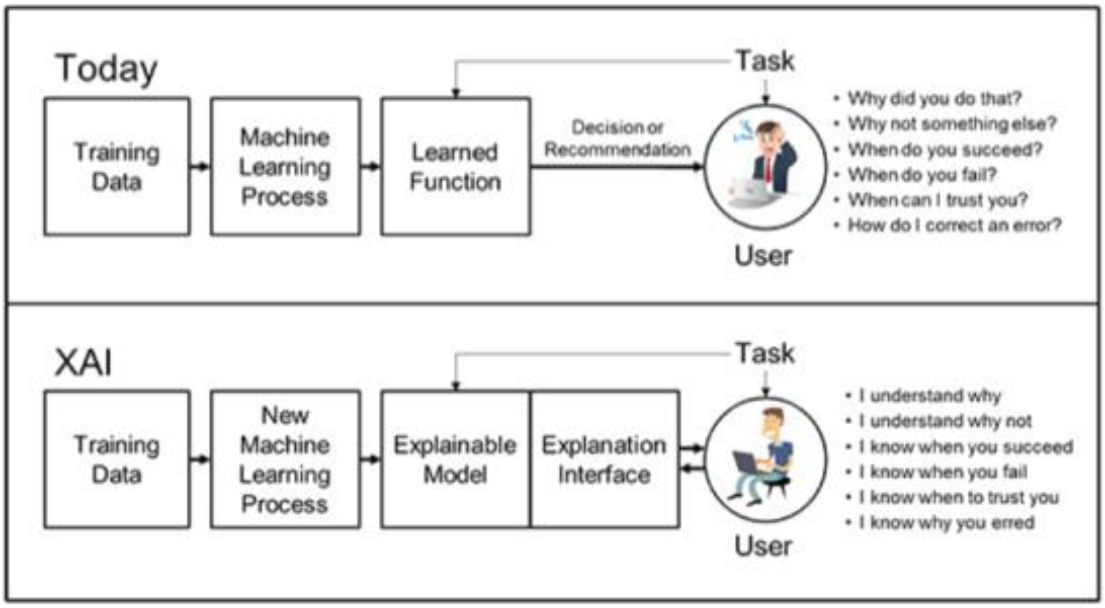
⑤ 차세대 인공지능 (Next Generation AI)

: DARPA는 인공지능 기술의 새로운 물결인 '인공지능 시스템이 도구가 아닌 문제를 해결하는 파트너'로 나아가기 위해 선구자적인 연구를 하고 있으며, 인공지능 시스템이 결과에 대한 설명과 일반 상식을 가지고 사고를 할 수 있도록 하는 것이 연구의 목표이며, 인공지능의 3가지 물결로는 다음과 같음.

DARPA에서는 인공지능의 3가지 물결을 통해 인공지능 기술의 발전을 인지하고 있으며, 1차 물결은 수작업에 의한 방식으로 좁게 정의된 문제에 대해서는 잘 작동하지만, 불확실한 영역에 대해서는 학습능력하기가 어려웠던 때이며, 2차 물결은 통계적인 학습방법으로 빅데이터를 가지고 특정 문제 영역에서 학습을 시키는 통계적 모델을 만들었고, 미세한 분류와 예측 능력은 가졌지만, 맥락적 지능과 최소한의 사고 기능은 없다. 이후의 3차 물결은 맥락적 적응(Contextual adaptation)이 핵심으로써, 이는 인간과 기계 사이의 신뢰가 기반된 협력적인 관계를 이끌 수 있으며, 인공지능의 결정 및 행동에 대한 설명 등을 포함한다. 다음은 인공지능의 3가지 물결의 도식화된 그림이다.

<그림 : DARPA AI Next Campaign > [66]

또한 DARPA에서는 국가 인공지능 전략 중 다섯 번째인 윤리적인 인공지능 기술개발의 일환이며, 차세대 인공지능 인공지능(AI Next)의 캠페인의 한 부분으로써, Explainable Aritificial Intelligentce(XAI)[67]라는 프로그램을 추진하고 있으며, 이는 갈수록 복잡해지고 깊어지고 있는 인공신경망 기반의 딥러닝 알고리즘에 따른 결과는 높은 수준의 정확도를 가질 수 있으나, 어떻게 결과를 도출했는지에 대한 정확한 과정은 알 수 없기 때문이다. 이는 인공지능(AI) 프로그램이 마치 블랙박스와 같기 때문에, 그에 대한 해석이 불가능하고, 문제가 있는지 아닌지를 확인할 길이 없다. 이에 대해서 설명가능한 인공지능(XAI)이라는 프로그램을 추진함으로써, 보다 설명이 가능한 알고리즘 모델을 만드는 것을 목표로 하고 있으며, 다음은 XAI의 개념이다.



<그림 : DARPA 홈페이지의 explainable artificial intelligence >

이외에도 인공지능 기술은 정보감시정찰(ISP, Intelligence, Surveillance and Reconnaissance), 군수(Logistics), 정보 작전(Information operations), 지휘통제(Command and control), 반자율/자율 시스템(Semiautonomous and autonomous systems) 등의 많은 분야에 적용되고 있으며, 이에 대하여 참고될 만한 시스템

66) 출처 : AI Next Campaign (DARPA 홈페이지)
67) 출처 : Explainable Artificial Intelligence (XAI) (DARPA, Dr. Matt Turek)

을 알아보면 다음과 같다.

이미 전세계적으로 자율주행차, 정찰 드론와 무인기와 같은 무인 자율분야에 인공지능 기술이 많이 적용이 되고 있고, 이는 작전에 있어서 획기적인 변화를 가져다 줄 것으로 보인다. 자율 정보감시정찰 시스템에 있어서, 인공지능 기술은 특히 A2/AD[68]와 같이 접근이 거부되는 상황에서, 목표물의 위치를 식별하고, 자동적으로 그 목표물로 가는 경로를 파악하며 그리고 목표물을 최종 확인할 수 있도록 해주고 있으며, 미공군은 저가 소모성 비행기술(LCAAT, Low-Cost Attritable Aircraft Technology) 프로젝트에 의해 Kratos社가 개발한 스텔스 무인 공격기 XQ-58A(별명 : 발키리/Valkyrie, 북유럽 신화에서 주신인 '오딘'을 섬기는 전투여신임)는 스텔스 기능 뿐만 아니라 내부와 날개에 공격무기의 장착도 가능하다. 미공군은 XQ-58A가 2023년에 전력화 될 것으로 기대하고 있으며, 1차와 2차 비행을 아래의 사진처럼 성공리에 완료하였고, 로열(충실한)윙맨 개념의 유-무인 편대 비행을 할 예정이다.

<table>
<tr><td>  </td><td>  </td></tr>
<tr><td>2019년 3월 - 1차 비행</td><td>2019년 6월 2차 비행</td></tr>
<tr><td colspan="2">* 사진 출처 - 미공군 홈페이지(www.af.mil/News)
* 길이 : 30ft(9.1m), 날개길이 : 27ft(8.2m), 속도 : 0.72Mach,
최대 비행거리 : 3,000NM(약 5,500km) (출처 : Kratos社 홈페이지)</td></tr>
</table>

이는 일반적으로 하나의 비행편대는 한대의 리더와 함께 편대를 이

68) 반접근/지역거부(A2/AD, Anti Access/Area Denial) 무기 시스템 : 상대방을 지상, 해양 및 공중 지역의 점유나 침범을 하지 못하도록하는데 사용되는 장비나 전략이며, 대표적인 예로는 대함 및 대공 미사일 등을 말한다. (참고 : Wikipedia)

루는 윙맨들로 이루어져 있으며, 이 윙맨들을 무인기로 대체하여 리더 전투기 한 대와 로열윙맨 무인기들이 함께 한 편대를 이루는 것이다. 그 모습은 아래의 그림과 같을 것으로 보이며, F-35A와 F-15EX를 무인기와 연동이 될 수 있도록 개조를 하고 있는 것으로 알려져 있다.

< 출처 : F-35 리더 한 대와 여섯 대의 로열 윙맨 드론 컨셉, 주석 69 참조 [69] >

이와함께 선상에서 미공군은 자율 무인 전투항공기 시제품을 초도전력으로 이르면 2023년 내놓을 계획이며, 이는 '스카이보그(Skyborg)' 프로그램으로도 알려져 있다. 이 프로그램에 대해 미공군연구실험실(AFRL)의 한 엔지니어는 다음과 같이 말했다.[70]

"스카이보그는 항공기를 띄워 영공에서 조종하는 다소 단순한 알고리즘부터 임무의 특정 과제나 하위 과제를 수행하기 위한 보다 복잡한 수준의 인공지능(AI)의 도입에 이르기까지 인공지능 기술의 집합체이다."

美보잉社도 호주 공군과 함께 인공지능(AI) 알고리즘이 탑재된 로열윙맨 시제기를 최근 공개했으며, 이외의 국가도 유사 개념의 무인기를 개발하고 있어, 드론 군집(swarming) 기술과 함께 공중을 차지하기 위한 기술경쟁은 더욱 치열해질 것으로 보이고, 비단 항공기 뿐만 아니

69) 출처 : 美공군 2030 비디오의 한 장면, "This Is What the US Air Force Wants You To Think Air Combat Will Look Like in 2030 (20180326, www.thedrive.com)

70) 출처 : Skyborg program seeks industry input for artificial intelligence initiative(2019, 美공군 홈페이지)

라 이미 2018년 초에 이미 잠수함 추적용 무인함정(Anti-Submarine Warfare Continuous Trail Unmanned Vessel) 시제기인 '씨헌터(Sea Hunter)'를 완료한 것과 같이 해양의 무인함정과 육지의 무인차량 등도 로열웡맨 혹은 군집활동과 같은 개념으로 개발 및 활용될 것으로 보인다. 더 나아가 미육군연구소(ARL, Army Research Laboratory)는 분산-협업 지능 시스템 및 기술(DCIST, Distributed and Collaborative Intelligent Systems and Technology) 협업 연구 연합(CRA, Collaborative Research Alliance)를 통해서 전장상황에서의 임무를 수행하기 우한 종합적인 통제 방안으로 아래의 6가지 주요 분야에 대한 기술을 개발하고 있고, 이는 서로 다른 종류(Heterogeneous) 그룹에 대한 통제에 대한 알고리즘 등으로 인공지능 기술이 핵심이다.[71]

① 군집(Swarms) ② 네트워킹(Networking) ③ 지능(Intellignece)
④ 자율(Autonomy) ⑤ 복원력(Resilience) ⑥ 협력(Collaboration)

이러한 혁신적인 전쟁양상의 변화를 근본적으로 뒷받침하는 프로그램으로 DARPA에서는 GARD(Guaranteeing AI Robustness against Deception)라는 보안에 취약할 수 있는 머신러닝 알고리즘에 대한 공격으로부터 AI를 강건하게 지킬 수 있는 방안을 연구 중이며, 아래는 그 세가지 주요 목표이다.[72]

① 방어가능한 머신러닝을 위한 이론적 기초 및 이를 기반으로 하는 새로운 방어 매커니즘의 용어의 개발
② 다양한 범위의 설정에서 방어가능한 시스템의 개발 및 시험
③ 위협적인 시나리오에 대비하여 머신러닝 방어능력을 특성화하기 위한 새로운 테스트베드를 구축

상기의 상호 의존적인 요소들을 통해, 견고성을 평가하는 엄격한 기준을 가진 기만에 강한 머신러닝 기술을 개발하고 있는 것이다.

그리고 미국은 앞에서 언급한 공격용 무인기와 드론 등을 방어하기

71) 출처 : https://www.dcist.org/ 홈페이지
72) 출처 : Defending Against Adversarial Artificial Intelligence (2019, DARPA)

위한 기술개발에 더욱 가속도를 붙일 것으로 보인다. 2020년 초 美방산업체인 레이시온社의 한 간부는 현재 무인기의 소형 쿼드콥트와 같은 드론(Group 1)이나 MQ-9 리퍼와 같은 큰 드론(Group 5)에 대한 방어는 많은 개발이 이루어져 있지만, 중간 그룹에 대해서는 아직 개발이 많이 부족한 상황이라고 말한바 있다. 참고로 美공군은 무인기를 그 무게, 운용고도 및 속도 등의 특징에 따라서 5개 그룹으로 분류를 하기 있으며(아래 그림 참조), 2019년 9월 사우디아라비아의 정유 시설을 폭파한 드론은 그 중에서 Group 3에 해당한다.[73] 이러한 안티 드론 시스템을 식별 하기 위한 방안으로 최근 인공지능 기술을 사용되고 있으며 앞으로 더욱 발전될 예정이다.

UAS Groups	Maximum Weight (lbs) (MGTOW)	Normal Operating Altitude (ft)	Speed (kts)	Representative UAS	
Group 1	0 – 20	<1200 AGL	100	Raven (RQ-11), WASP	
Group 2	21 – 55	<3500 AGL	< 250	ScanEagle	
Group 3	< 1320	< FL 180		Shadow (RQ-7B), Tier II / STUAS	
Group 4	>1320		Any Airspeed	Fire Scout (MQ-8B, RQ-8B), Predator (MQ-1A/B), Sky Warrior ERMP (MQ-1C)	
Group 5		> FL 180		Reaper (MQ-9A), Global Hawk (RQ-4), BAMS (RQ-4N)	

< 출처 : DoD의 무인기 그룹 분류 설명, 주석 74 참조 [74] >

73) 출처 : Raytheon : DOD Needs More Research on Stopping Medium-Size Drones (20200123,www.airforcemag.com)

74) 출처 : Unmanned Aircraft System Airspace Integration Plan (March 2011-Version 2.0, DoD)

2) 러시아

러시아는 미국의 XQ-58A와 유사한 무인전투항공기(UCAV, Unmanned Combat Aerial Vehicle)인 S-70(Okhotnik-B, 별명:사냥꾼-Hunter-B)라는 무인전투기의 초도비행을 지난 2019년 8월에 성공적으로 완료하였다. 그리고 2019년 9월 러시아 국방부는 5세대의 스텔스 유인전투기인 Su-57와 함께 S-70을 비행하는 영상(아래 그림 참고)을 공개하였다. 이는 앞선 미국의 로열윙맨 개념과 같으며, 정확히 서로 통신을 하면서 비행하였는데 알려져있진 않지만 Su-57과 S-70 모두 스텔스 전투기로써 무인기이며 저가인 S-70를 먼저 나아가면서 작전반경을 넓히면, 인명피해를 최소화 하면서 적진으로 침투하기 위한 최적의 상황이 되는 것이다. 현재 S-70은 개발 진행중이며, 2025년에 완료할 예정이다.

* 2019년 9월 - 합동비행 영상
* 날개길이 : 65ft (20m),
최대속도 : 1,000km/h,
최대 비행거리 : 6,000km
(출처 : Wikipedia)
* 사진 출처[75]
: 위 - Su57, 아래 - S-70

또한 러시아군은 인공지능 기능이 탑재된 무인지상전투차량인 Uran-9 (UCGV, Unmanned Combat Ground Vehicle)를 2019년 1월 정식 채택하여 실전에 배치하였고, Uran-9는 7.62mm 머신건과 대전차로켓 그리고 30mm 기관포로 무장되어 있는 러시아의 혁신적인 무인전투차량이었다. 하지만 Uran-9는 2018년에 시리아로 전투 시험을 위해 배치되었으나, 2,900미터가 아닌 300~500미터로 제한된 작전반경하였, 주기적인 단기 및 장기의 통신두절 그리고 소프트웨어와 하드웨어의 목표설정에 있어서의 불일치와 같은 심각한 문제들이 파악되었다.[76]

75) 출처 : Watch Russia's S-70 Unmanned Combat Air Vehicle Fly With An Su-57 For The First Time (20190927, www.thedrive.com)
76) 출처 : Russia's Robot Tank Sucks But Its Military Is Adopting It Anyway (20190124, taskandpurpose.com)

이와 같이 지금 각국에서 개발되고 있는 새로이 적용되는 인공지능 기능을 가진 첨단 군장비들은 많은 시험을 통해서 검증되고 문제가 없이 작동이 되는 것을 확인하여야 한다. 특히 전쟁상황과 같이 실전경험을 직접 체험하는 것이 어려운 부분에 대해서는 실전과 동일한 환경의 테스트베드를 구축하여 많은 실전환경과 유사한 모의 학습을 통한 검증이 필요한 것으로 보인다.

그리고 2020년 6월 미국의 DARPA와 같은 조직인 러시아의 로봇기술및기초개발국가센터(Russian Foundation for Advanced Research Projects' National Center for Development of Technologies and Basic Elements of Robotice)의 관계자의 말을 의하면[77] 현재 러시아는 로봇들을 사람과 같이 듣고 행동할 수 있도록 음성명령으로 통제(현재는 전자 테블릿으로 통제하고 있음)하기 위한 연구 및 실험을 Marker 로봇 플랫폼[78]을 대상으로 하고 있는 중이다. 이는 위의 Uran-9에서 나온 통신 문제점들을 개선한 방법으로 작업자로부터 상당히 먼거리에서도 임무를 수행할 수 있도록 만들기 위한 것이다.

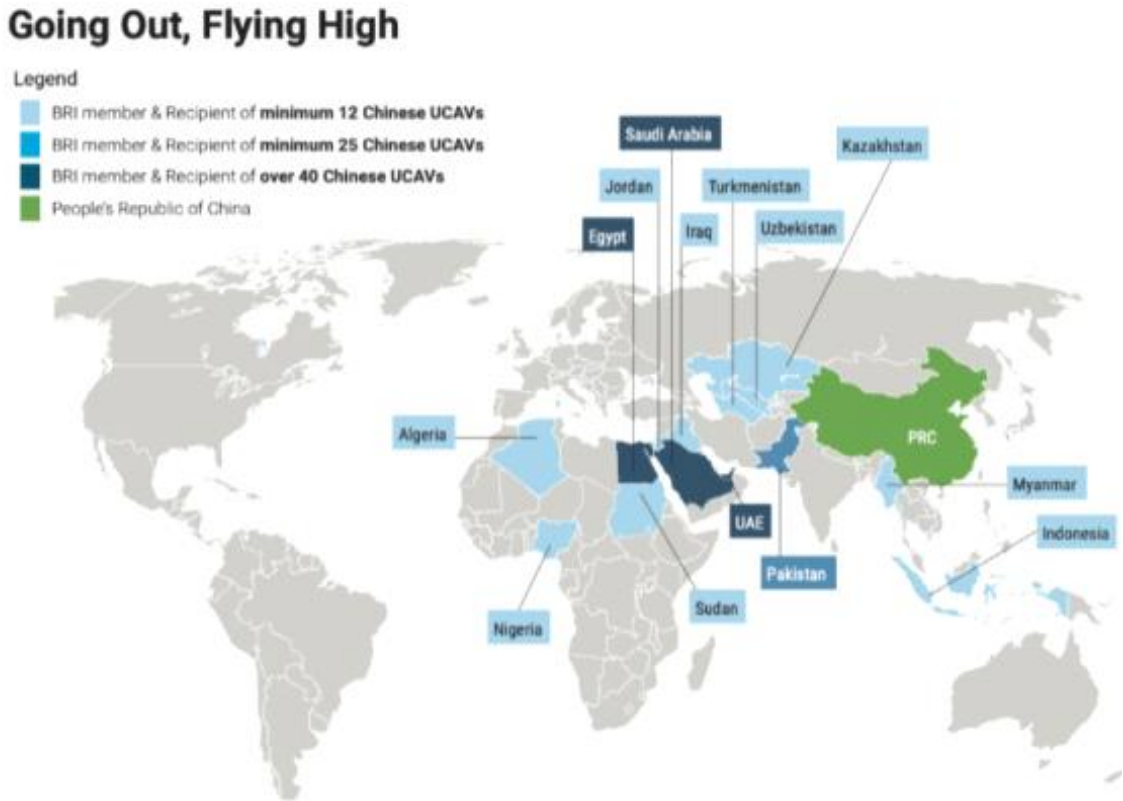
3) 중국

각종 군사력 순위 조사결과를 볼 때 일반적으로 종합적인 군사력에 있어서, 중국은 미국을 아직 추월하지 못한 것으로 평가를 받고 있다. 하지만 향후 게임체인저로써 역할을 할 인공지능 기술을 이용한 첨단 무기에 대한 투자를 많이 하고 있으며, 앞서 국가적 인공지능 전략의 중점임무에 나온 것과 같이 민간부문과 군사부문을 서로 긴밀히 협업 통하여 민간부문의 드론기업이나 안면인식과 같은 성공적인 기술이 곧바로 군사적인 용도에 활용되어질 수 가능성도 있다.

77) 출처 : Russia Begins Trials of New Generation Marker Unmanned Ground System (20200630, internationalinsider.ogr)

78) 2021년에 최종완료될 예정인 지상로봇으로써, 지상로봇 연합작전, 무인 항공기 그리고 특수작전부대의 훈련을 위한 필수적인 요소들 중의 하나로 러시아 고등연구소와 안드로이드테크닉 Scientific Produciton Association과 합작하여 개발중임.

우선 최근 중국은 무인전투항공기(UCAVs, Unmanned Combat Aerial Vehicles)를 가장 많이 수출하는 국가 중의 하나로 급성장하고 있다. 중국과는 다르게 미국은 기업에서 해외로 중요 제품을 수출하기 위해서는 미국무부의 수출승인을 받아야 하고, 민감한 기술이 포함되어 있는 경우에는 수출승인이 되지 않거나, 그 기술을 제외하고 수출을 하고 있기 때문에 미국 기업이 해외에 무인전투항공기를 수출하는데는 많은 제약이 있다. 참고로 중국이 주로 무인전투항공기를 수출하고 있는 나라와 대수는 다음 그림을 통하여 알 수 있다. 이와 같이 많은 수출을 통하여 실전배치된 장비들을 직접 시험 및 개선을 하며 더욱 품질이 높은 제품을 만들어 가고 있다.



< 출처 : 각주 79 참조 [79] >

한 예로써 Ziyan UAV란 중국 기업에서 생산 및 수출하고 있는 'Blowfish A2'라는 전투헬리콥터드론(길이-1.87m, 높이-0.62m)은 시간

79) Wings Along the BRI : Exporting Chinese UCAVs and Security?
https://medium.com/@lseideas/wings-along-the-bri-exporting-chinese-ucavs-and-security-a4bf7a3324df

당 약 130km를 이동할 수 있으며, 라디오 재밍 장비와 총 혹은 폭탄을 장착할 수 있는 것이 특징이며,[80] 이러한 기능을 가진 전투드론들은 레이더로 식별이 잘 안되어, 추적 및 격추가 어려운 것이 특징이다. 또한 군집 기술와 함께 사용이 된다면 보다 더 가공할만한 파괴력을 지닐 수도 있을 것이다.

그리고 이미 중국의 일부 도시들은 안면인식 기술이 적용된 인공지능 기반의 감시를 통하여 불법 통행, 실종된 사람 찾기 및 테러리스트 색출 등을 하고 있다. 이러한 방대한 실제 사람에 적용한 안면인식 빅데이터를 통하여 인공지능 회사인 SenseTime이나 Megvii와 같은 기업들은 아주 빠른 속도로 성장하고 있으며, 여기서 개발된 안면인식 기술들은 국방분야의 드론의 이미지처리와 각종 관련 분야에 활용이 가능하다.

4) 영국

영국 국방부는 2016년 12월 범정부 기관인 방위안보촉진청(DASA, The Defence and Security Accelerator)을 발족하였으며, 대략 50여 명의 다양한 경력을 가진 직원이 근무하고 있으며, 11개의 지역 혁신 파트너들과 협력하고 있고, 국방과 안보 분야에 있어서 빠르고 효과적으로 지원하기 위한 개발 가능한 혁신적인 방안을 찾고 또한 자금 지원을 하기 위하여 설립되었다. 그리고 2020년 1월 DASA는 선박을 차세대 지능형 선박(Intelligent Ship)으로 변경하기 위해 400만 유로를 투자한다는 방안을 발표하였으며, 이는 인공지능 기술을 사용하여 수많은 정보와 데이터들 속에서 의사결정 및 처리를 효율적으로 할 수 있는 진보적인 방안을 찾고 적용한다는 개념으로, 군함, 항공기 및 지상차량과 같은 Human to AI와 AI to AI가 팀을 이루는 국방 플랫폼을 위한 혁신적인 방안에 초점이 맞추어져 있다.[81]

---

80) 출처 : Oddly shaped Chinese combat-ready helicopter drone popular in international market (2019, GlobalTimes.cn)

81) 출처 : Revolutionary Artificial Intelligence warship contrasts announced (2020, MoD Gov.uk)

방산기업 분야에서는 세계적으로도 안티드론과 관련하여 잘 알려져 있는 Operational Solutions이라는 기업은 앞선 전투 드론과 같은 위협에 대비한 장비를 생산하고 있으며, 이는 드론 탐지, 추적 및 의사결정과 같은 기능에 인공지능 기술을 활용하고 있다.

상기에서 살펴본 바와 같이 훈련국인 캐나다를 포함한 주요 국가에서 적용중인 국방분야의 인공지능 관련 기술에 대해서 살펴보았다. 이는 해양, 인공위성 자료, 무인 드론 및 전투기 등과 앞으로 이러한 인공지능을 활용하기 위한 근본적인 방안인 보안 강화 및 설명가능한 인공지능 프로그램 등이다. 이러한 방향성을 참고하여 향후 우리나라의 국방분야에 어떻게 적용하고 더 나아가 새로운 부분을 고려하는 방안을 다음 장에서 살펴보고자 한다.

제4장. 국내 국방분야 적용 현황 및 활용 가능 분야 검토

1. 국내 인공지능(AI) 기술 적용 추진 중인 현황

우리나라는 아래의 그림과 같이 국가차원의 인공지능 전략에 따라 기초학문과 인공지능 SW/HW 기술을 인재와 기반을 포함하여 개발하고 있으며, 이를 AI R&D 전략범위인 의료, 안전, 제조 그리고 국방 분야에 활용하여 향후 창업과 해외 진출을 그리고 공정 생태계를 만들기 위해 노력을 하고 있다.

< 출처 : I-Korea 4.0 실현을 위한 인공지능(AI) R&D 전략 (2018.5)>

그리고 국방기술품질원에서 3년 단위로 조사하여 발표하고 있는 국가별 국방과학기술 수준조사에서 우리나라는 2018년 최고선진국(미국) 기술수준인 100점 대비 80점으로 이스라엘에 이어 아래 그림과 같이 9위로 되어있다.[82)]

< 출처 : 2018년 주요 16개 국가 국방과학기술 수준(국방기술품질원 발표 >

82) 출처 : 2019년 국방기술품질원 통계연감, [그림 2-10] 2018년 주요 16개 국가 국방과학기술 수준

이러한 국가적인 전략과 국방 기술수준을 바탕으로하여 방위사업청은 무기체계 개발에 필요한 핵심기술에 대한 개발 방향을 제시하는 문서인 ’20~’34 핵심기술기획서를 2020년 4월 확정하면서, 「’19 ~ ’33 국방과학기술진흥정책서」에서 제시한 국방전략기술 8대 분야의 140개 세부기술영역에 대한 개발 계획을 포함하였고, 이는 다음에서 보는 것과 같이 4차 산업혁명과 연계된 미래 신기술에 대한 것으로 앞서 살펴본 국외 사례를 많이 포함한 것으로써 주로 인공지능 기술과 연관이 많은 분야라고 할 수 있다.

- 국방전략기술 8대 분야

① 자율・인공지능 기반 감시정찰, ② 초연결 지능형 지휘통제,

③ 초고속・고위력 정밀타격, ④ 미래형 추진 및 스텔스 기반 플랫폼,

⑤ 유・무인 복합 전투수행, ⑥ 첨단기술 기반 개인전투체계,

⑦ 사이버 능동대응 및 미래형 방호, ⑧ 미래형 첨단 신기술

또한 2019년 12월 방위사업청은 인공지능 분야의 핵심적인 기초 및 원천 기술을 확보를 하기 위해서 한국과학기술원(KAIST) 내에 국방과학연구소의 지원와 함께 ‘미래 국방 인공지능 특화연구센터’를 개소하였으며, 이는 향후 2025년까지 121억원의 예산을 투입할 예정으로, 주관기관인 한국과학기술원과 서울대, 포항공대, 연세대 등 9개 대학과 3개 정부출연 연구기관에서 260여 명의 우수 연구진이 함께 참여하고, 주요 연구 분야로는 아래의 4개 전문 연구실을 구성하여 총 17개 과제를 수행할 예정이다.

① 군사적 설명이 가능한 인공지능 이론 연구

: 앞서 DARPA의 연구 방향과도 유사한 것으로써, 인공지능의 결과를 이해할 수 있게 하며, 또한 딥러닝 알고리즘의 신뢰성을 향상시키고, 적대적 공격에 강인한 딥러닝 프로그램을 개발하기 위함.

② 다종 국방 데이터의 융합 학습 및 탐지 연구

: 군집(SWARM) 및 유-무인 복합체계 등의 시스템에 필수적인 다종 센서 정보를 활용한 연관 및 추적 등의 딥러닝 연구 등임.

③ 열악한 환경의 국소량 국방 데이터 기반 학습 연구

: 국방 분야의 제한된 데이터를 극복하기 위한 방안으로써, GAN과 같은 비지도학습 방법에 기반한 방안 등을 연구

④ 탐지 군사적 설명 연동 방책 추천 연구

: 다른 나라의 Decision Making 알고리즘과 유사한 의사결정 지원을 위한 인공지능 알고리즘에 대한 연구로, 상황을 정확히 파악하고 최적의 설명이 가능한 방책 추천을 강화학습 방법 등으로 연구

또한 2020년 3월 '국방과학기술혁신촉진법' 제정 및 국방개혁 2.0 등 국방 연구개발의 환경변화를 반영하고 효율적인 핵심기술 개발 수행을 위하여 '핵심기술 연구개발 업무처리지침'을 개정하였다. 이는 인공지능과 같은 핵심기술 과제에 대하여 기존 최소 3년이 소요되는 절차는 1년으로 단축하는 등 급변하는 기술개발 환경에 신속히 대응하기 위함이었으며, 이의 일환으로 첫 미래도전국방기술 사업[83]으로써 해군의 '스마트 네이비(Smart Navy)' 미래 전략에 따라 『군집 무인수상정 운용기술』을 개발하는 190억 규모의 사업이 국방과학연구소를 주관으로 한화시스템(주), KAIST, 선박해양플랜트연구소(KRISO), 동국대 등과 함께 2024년까지 추진되며, 이는 군집 통신 네트워크 및 AI 알고리즘 시연 플랫폼 구축 및 AI 강화학습 기반으로 개발된 실시간 상황 인지와 인간지능과 유사한 교전임무 수행 등이 가능하게 할 예정이다.

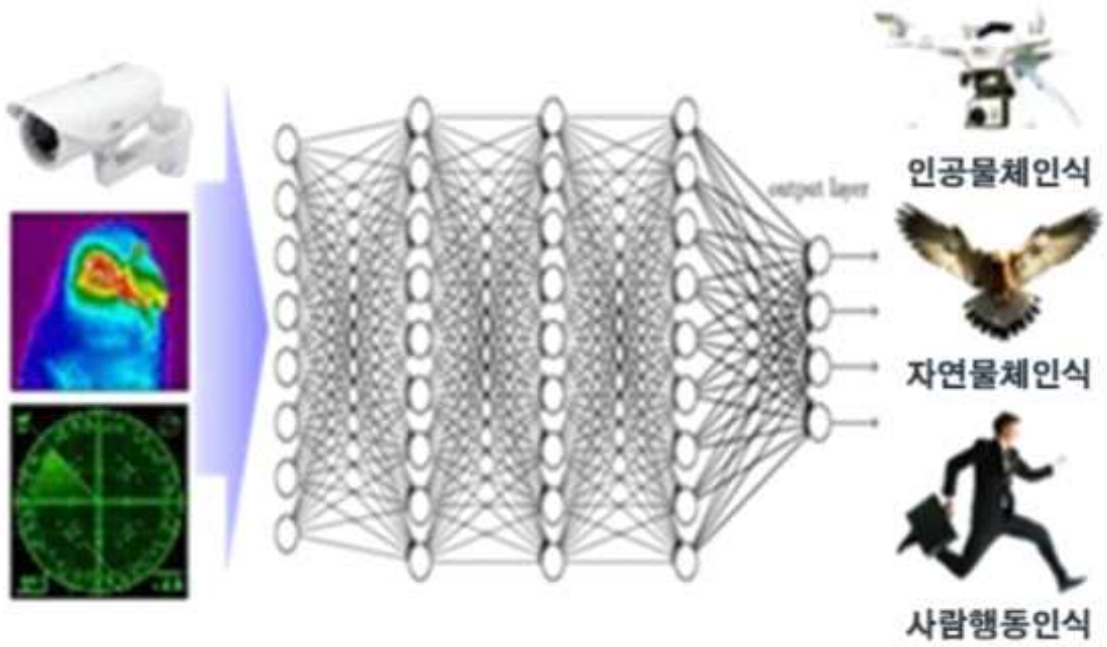
그리고 육군은 4차산업혁명 시대를 맞이하여 선제적으로 2018년 기존 '육군 지상전연구소'를 개편하여 '육군 미래혁신 연구센터'로 발족하였으며, 이후 드론봇 전투단을 창설하면서 급변하는 전략환경에 발맞추어 효율적인 미래전 수행을 위한 드론봇 전투체계[84]를 발전시키고 있으며, 또한 2019년 12월에는 한국과학기술원에 '미래육군과학기술연구소'를 개소를 하였으며, 인공지능을 포함한 10대 차세대 게임

83) 미래도전국방기술 사업 : 무기체계 소요가 결정되지 않거나 소요가 예정되지 않은 무기체계에 대한 적용을 목적으로 하는 혁신적이고 도전적인 기술개발 사업

84) 드론봇 전투체계 : 정찰드론-첩보수집 및 화력 유도, 공격 드론-목표 타격, 수송 드론-공중 재보급, 경계용드론-후방의 주요시설 보호, 공중중계드론-원활한 통신여건 보장 등의 각 임무를 수행하는 드론체계

체인저[85]에 대한 비전을 제시하며 Follower가 아닌 퍼스트무버가 되기 위해서 많은 노력을 기울이고 있다.

우리나라 국방과학기술의 핵심기관인 국방과학연구소는 각군의 소요 및 선행연구를 통하여 미래기술의 하나로써 인공지능에 대한 많은 연구를 하고 있으며, 그 중 하나로 다음 그림과 같이 각종 센서 및 장치로부터 입력된 데이터를 딥러닝 기술 등을 활용하여 물체, 환경 및 행동으로 정확 식별하는 알고리즘이 있으며, 이는 향후 유-무인 체계의 핵심이 되는 기술이다.



< 출처 : 국방과학연구소 홈페이지 >

그리고 방산업체의 경우에도 이와 관련한 많은 연구를 하고 있으며, 몇 가지 예로써, K2 전차를 개발하고 생산하고 있는 현대로템은 ADEX2019에 전시한 K2 전차를 이을 차세대 전차인 K3의 모형을 전시하였으며, 개념으로 무기와 같은 하드웨어적인 업그레이드는 물론이고, 향후 유-무인 복합운용 개념이 적용되는 네트워크 중심의 미래전장에 부합하는 인공지능 기반의 차량운용체계를 적용할 예정으로 알려져 있

85) 육군의 10대 게임체인저 : 레이저, 초장사정 무기, 유-무인 복합전투체계, 지상무기의 스텔스화, 고기동화, 양자기술, 생체모방 로봇, 사이버 및 전자전, 인공지능(AI), 차세대 워리어플랫폼

다. 또한 LIG넥스원은 해양분야에서는 복함임무 무인수상정(USV)과 무인 잠수정을 통하여 서해북방한계선(NLL) 등의 해역의 감시정찰과 수중탐색을 자율적인 운행을 통하여 장기간 할 수 있으며, 이는 우리 군의 인명손실 예방을 통한 작전능력 향상으로 이어질 수 있으며, 지상분야에서는 휴대용 감시정찰로봇(PUGV) 등을 통하여 사람의 접근이 어려운 특수환경에서도 감시정찰 및 위험물 제거/조작(EOD) 등을 개발하고 있으며 이는 인공지능 기술을 통하여 구현될 것으로 판단이 되고 있다.[86]

86) 참고자료 : LIG텍스원 홈페이지

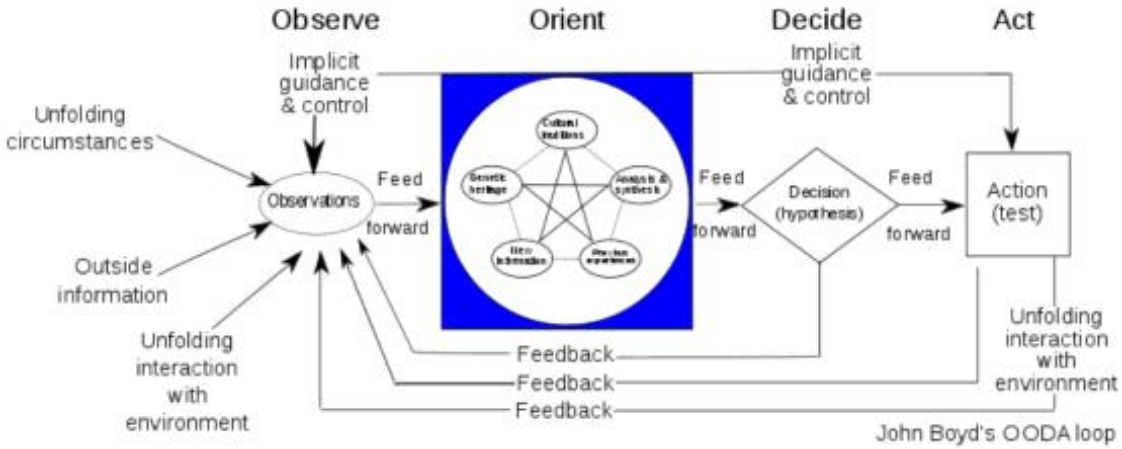
2. 향후 활용 가능 분야

앞서 훈련국인 캐나다 및 주요 국가에서 진행 중인 인공지능 기술과 관련하여 최신 개발되고 있는 신규 논문 및 기술개발 동향과 또한 국방분야에서 적용되고 있는 사례들을 살펴보았고, 우리나라에서도 퍼스트 무버가 되기 위한 많은 방안들과 노력들이 진행되고 있음을 알 수 있었다. 이를 바탕으로 우리나라 국방에 인공지능 기술을 적용 및 활용 가능한 분야를 살펴보았다.

1) OODA Loop

한국전쟁에도 참전한 미공군 대령이었던 John Richard Boyd에 의해서 고안된 OODA(Observation, Orientation, Decision, Action) Loop는 오늘날까지도 국방 뿐만아니라 경영과 같은 비즈니스 등의 많은 분야에서 이미 의사결정 및 운용 도구 등으로 널리 활용이 되고 있다. 이는 크게 4가지의 기능이 계속적으로 순환하며 작동하며, 그 주요 기능과 도식화 자료는 다음과 같다.

- 관찰 (Observation) : 정찰, 센서 등에의한 자료를 수집하는 단계
- 방향설정 (Orientation) : 다양한 기인을 바탕으로하여 수집된 자료의 분석과 종합
- 의사결정 (Decision) : 현재의 상황을 바탕으로한 최적의 실행과정 결정
- 실행 (Action) : 의사결정에 대한 실제적인 실행 단계



< 출처 : Wikipedia >

이와 같은 군사전략에 맞추어 인공지능 기술을 OODA LOOP의 각 단계에 대략적으로 어떤 기술들이 적용이 되는지에 대한 개관적인 도식을 살펴보면 다음과 같다.[87]

< 출처 : 각주 87의 내용을 참고하여 저자 직접 작성 >

위와같은 OODA Loop는 단독으로 사용되는 것이 아닌, 복합적으로 서로 연관을 가지면서 적용이 되고 있으며, 감시·정찰과 지휘통제 그리고 물류 등의 각 분야별로 세부적인 사항들이 또한 적용이 될 수 있다. 또한 OODA Loop의 인공지능 기술을 활용하여 순환 주기를 빠르게 할 수 있게 한다면 선제적인 대응 및 방어를 할 수 있다. 그리고 이러한 인공지능 기반의 OODA Loop내에서 사람이 최종 결정을 내리고 실행에 옮기기 위해서 가장 중요한 부분이 바로 과정을 설명할 수 있는 인공지능이 되어야 할 것이다. 근거를 정확히 이해하지 못하고 단순히 AI의 결과만을 믿는다면, 잘못된 결정을 할 수도 있기 때문이다. 앞서 美 DARPA의 Explainable AI Project 등을 참고하여, 우리나라의 KAIST에서 진행중인 설명가능한 인공지능 연구에 대한 개발이 더욱이나 중요하게 생각되어지는 부분이 되는 것이다. 또한 오작동 및 테러와 같은 치명적인 문제를 예방할 수 있는 Robust 및 Adversarial AI가 중요한 부분이 될 것이다.

87) 출처 : How Artificial Intelligence is Closing the Loop with Better Predictions (20180726, medium.com)
https://medium.com/hackernoon/how-artificial-intelligence-is-closing-the-loop-with-better-predictions-1e8b50df3655

그리고 의사결정(Decision) 부분에 대해서는 소위 ‘킬러로봇’으로 불리는 인명 살상용 자동로봇(LAWS, Lethal Autonomous Weapons)이 바로 사람의 개입이 전혀없는 AI 알고리즘에 의한 의사결정을 하는 것이 문제이기 때문에, 다음 장에서 다룰 로봇윤리와 연계하여 향후 무기 개발에 관심을 가져야 할 것이다.

이와 관련하여 미국의 자율 무기에 대한 규정인 DoD Directive 3000.09[88] 에서는 그 정책으로써, 아래와 같이 자율 및 반자율 시스템은 무기를 사용함에 있어서 지휘관이나 작업자들에게 적절한 사람의 판단을 수행할 수 있도록 허용하도록 설계되어야 한다.

“Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgement over the use of force.”

(참고로 긴급한 군사적 작전이 필요한 경우에는 이에 대한 면제가 가능하다는 조항도 같이 포함이 되어있다.)

그리고 시스템은 엄격한 하드웨어 및 소프트웨어 확인 및 검정(V&V, Verification and Validation)와 현실적인 환경에서 시험 평가(T&E, Test and Evaluation)을 거쳐야 한다. 또한 작업자들이 충분한 정보를 가지고 목표물에 대한 적절한 의사결정을 내리기 위해서 사람과 자율 및 반자율 무기시스템 사이의 인터페이스는 아래의 3가지 요건이 충족이 되어야 한다.

① 훈련된 작업자가 쉽게 이해하도록 되어야 한다.
(Be readily understandable to trained operators.)

② 시스템 상태에 대한 추적가능한 피드백을 제공해야한다.
(Provide traceable feedback on system status.)

③ 훈련된 작업자가 시스템 기능을 활성화/비활성하기 위한 분명한 절차를 제공해야 한다.
(Provide clear procedures for trained operators to activate and deactivate system functions.)

88) 출처 원문 : https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf

2) 국방 분야별 AI 적용 가능 방안

이 분야에서는 앞서 살펴본 훈련국인 캐나다와 주요 나라 그리고 최신 인공지능 알고리즘 경향을 바탕으로 하여 우리나라 국방 분야에 인공지능 기술을 적용할 수 있는 분야에 대해서 살펴보았다. 그리고 많은 유사성이 있어 이미 추진 중인 사례도 중복이 되어 있으며, 보안 등의 이유로 구체적인 내용 보다는 개관적인 적용 내용을 다루었다.

분야	적용
감시정찰	· 이미지 분석등에 의한 컴퓨터 비전으로, 무인비행체로 촬영한 영상 분석, CCTV 영상 및 이미지 분석을 통한 경계강화. · 철책 경계에 인공지능 기술을 효율적으로 적용 시, 지금과 같은 인구 및 복무 기간 감소에 따라 줄어든 병사 수에 필수적인 사항으로 많은 경계 인력을 줄일 수 있으나, 2020.7월 월북한 사례를 반면교사 삼아 CCTV와 음성(자연어 분석 등) 및 TOD의 복합 알로리즘 적용하고 이상징후 감시시 바로 경보를 알리는 시스템을 구축할 필요가 있으며, 무엇보다 이를 분석 및 적절히 사용할 수 있는 장병 양성이 필요함. · 미디어 포렌식 · 빅데이터 분석 · 지뢰 분석 및 제거 로봇 : 위험한 작업에 직접 사람을 투입하지 않고, 지뢰를 제거할 수 있음. · 인공위성 자료에 대하여 인공지능 이미지 분석으로 이상 동향 선제 대응
획득 및 군수	· AI를 활용한 신규 획득 사업 및 발주 시스템 · 자동 예방정비 시스템
지휘통제	· 의사결정 지원시스템 (AI-based decision making support) : 수집된 데이터를 바탕으로 한 머신러닝 알고리즘으로 최적의 의사결정을 도출하여, 신속하고 정확한 OODA 순환 수행. · 비지도학습을 통하여 사람은 파악하기 힘든 데이터들이 가진 특성 및 경향을 분석하여 앞으로의 작전 및 지휘 방안 등으로 활용

	· Mosaic Warfare · Explainable AI
무인기	· 로열윙맨 무인기 · 무인 자율 드론 군집(Swarm) 비행 · 자율 주행차량 및 선박 · 유-무인 복합 작전 임무 수행 · 강화학습 기반의 무인기 재밍공격 방어
훈련	· 모의 전장 환경에서의 AI Simulation 및 Training · 인공지능 상담사 및 교관
병영환경	· 병영생활기록에 딥러닝 기술을 적용하여 종합적으로 분석(언어 인식 및 의미 분석)하여 인공지능 관리체계 구축하여 맞춤형 병사 관리. · 병영환경을 개선하기 위한 비지도학습법의 군집 및 특징을 활용하여 식단 변화 등의 개선점을 찾아 적용하고, Health care에도 동일하게 적용하여 건강한 병영환경 구현.
보안강화	· GARD · Robust AI · Adversarial AI

3) 기타

우리나라의 앞선 ICT 기술을 바탕으로 하여 인공지능 기술의 한 분야에 국한된 것이 아닌 사물인터넷(IoT) 및 빅데이터와과 결합된 Digital Twin을 국방 시설 혹은 방산 분야의 군수품 개발 및 제조에도 등에 적용함으로써 다양한 실험과 예측을 할 수 있으며, 많은 부분을 효율적으로 처리할 수 있다. 이는 가상 물리시스템(CPS, Cyber Physical System)과 유사한 개념으로 각 전장상황을 CPS로 시뮬레이션 하여 인공지능 기반의 의사결정 시스템으로도 적용할 수도 있을 것이다.

그리고 앞으로 위에 언급된 분야를 포함하여 많은 분야에 인공지능 기술이 적용이 되어오고 있으며, 되어질 것이므로, 개발한 제품이 제대

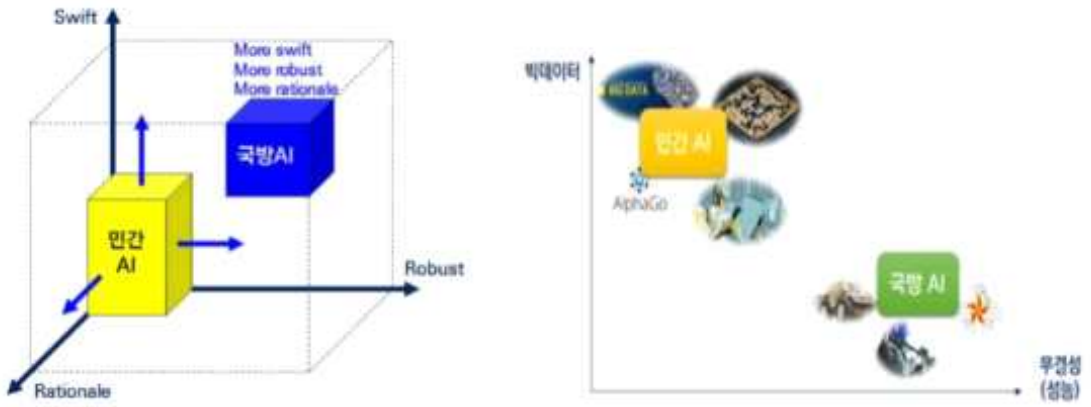
로 적용이 되는지에 대한 확인 및 검증 그리고 시험평가와 모니터링 등을 수행하기 위한 전문가 양성이 더욱 절실히 요구된다고 할 것이다. 인공지능 분야에 대한 교육 훈련 뿐만아니라 이에 대한 검증 등의 절차에 대한 마련이 중요한 사안이 되었다.

이렇듯 국방분야에 인공지능 기술은 이미 많은 분야에서 연구 개발이 되고 있고, 앞으로도 새롭게 개척해야 할 분야가 수없이 많다. 앞에서 살펴본 것들을 바탕으로 빠른 속도로 발전하고 있는 인공지능 기술을 우리나라 국방 분야에 적용하기 위해 능동적으로 늘 대처해야 하며, 이는 국방과학연구소 및 관련 기관에서 핵심 기술을 개발하는 것과 동시에 신속시범획득 사업과 같은 방안 등으로 민간의 기술을 국방분야로 빠르게 적용해야 한다(Spin on). 그리고 국방분야에서 먼저 개발된 기술은 민간으로 이전하는(Spin off)를 하여 국가 경제성장과 나아가서는 국제적인 발전에 기여하는 방법으로 나가야 할 것이다.

3. 인공지능(AI) 적용 시 고려 사항

위의 인공지능(AI) 기술을 접목할 필요성이 있는 해당 국방분야에 개발적용할 때, 인공지능의 특성과 국방 분야만의 특수성 및 폐쇄성 등을 이 있으므로, 향후 국방에 적합한 인공지능 시스템 개발하기 위해서는, 다음과 같은 고려해야 할 사항들을 알아보는 것이 필요하다.

첫째, 민간 분야에서 적용되는 인공지능과 국방 분야에서 인공지능 사이에는 그 필요성의 최저 기준 과점에서 바라볼 때, 아래의 그림과 같이 속도와 강건성, 빅데이터, 무결성 그리고 설명가능성 등에 있어서 그 차이가 존재한다. 이는 예를 들어 국방의 경우 해킹 관련 문제가 발생 시, 적국으로 기밀 누출 등의 민간분야보다도 국가 전체에 심각한 안보 문제가 될 수도 있어 보안에 특히 신경을 써야 하며, AI 기반의 지휘결정 시스템을 의지하여 결정을 내릴 경우, 향후 그러한 결정을 할 경우, 그 결정의 파급력을 고려하여 그에 대한 구체적인 근거가 필요하기 때문이다. 더욱이 국방 분야는 보안과 기밀이라는 정보의 제한이 있으므로, 머신러닝 시스템이 학습을 하기 위한 기초 데이터가 민간 분야에 비해서는 많이 부족한 실정이기에 적은 데이터로도 무결성의 시스템을 구축해야 하는 어려운 현실임이 분명하다. 따라서 항상 이 점을 염두에 두고 향후 국방 분야에 인공지능 기술을 접목할 때 주의를 기울여야 함을 알 수 있다.



< 출처 : 국방 인공지능(AI) 활용 실증기획 연구 (2018년, 국방부) >

둘째, 마이크로칩의 밀도가 24개월마다 2배로 늘어난다는 무어의 법칙(Moore’s Law)과 같이 인공지능 기술과 적용 알고리즘들은 아주 빠른 속도로 변화하고 있다. 따라서 정상적인 국방 무기 획득절차에 따르면 소요제기부터 최종 무기획득까지 많은 기간이 소요되나, 개발 중이거나, 양산 후 얼마되지 않아 새로운 고성능의 기능이 개발이 된다면 많은 문제가 발생하게 된다. 이러한 경우를 개발 시에 염두에 두어야 하며, 이럴경우를 대비하여 하드웨어적인 부분은 그대로 두고서 소프트웨어 업그레이드를 필요할 때 할 수 있도록 모듈 단위의 체계적인 소프트웨어 설계를 하는 체계가 필요하다.

셋째, 지금의 인공지능(AI) 분야의 기술 개발 양상은 이전 국방분야의 앞선 기술이 민간으로 이양이 되던 시절이 아닌 오히려 민간의 막대한 투자자금을 바탕으로 한 빠른 속도로 개발된 신기술이 거꾸로 국방분야로 적용되는 경우가 많다. 따라서 민간 부문과의 긴밀한 협력은 무엇보다도 필수적이다. 예를 들어 전자통신연구원(ETRI)에서는 AI 드론이 수행하는 임무 등의 고성능 계산 능력이 필요한 경우를 대비해서, 비행제어 SW와 AI 임무제어 SW과의 가상화기술을 이용하여 기체의 경량화와 전력소모를 줄였으며 최종 美 FAA(미 연방 항공청)의 안정성 시험 과정을 거쳐, ‘DO-178C Level A’의 인증을 국내 기관 중 최초로 받았다.[89] 이와 같이 민(산・학・연)・관의 긴밀하고도 적극적인 교류와 협력이 더욱 중요한 시대가 온 것이다.

끝으로 최근 방위사업청은 민간의 인공지능 기술과 같은 신기술을 적용한 제품을 구매하여 시범운용을 통해 군사적 활용성을 확인하고 신속히 전력화를 하기 위한 방안으로 신속획득사업팀을 신설하고, ‘신속시범획득 사업 업무관리 지침’도 새로이 제정하여 적극적으로 신기술 관련 사업을 발굴 및 군에 적용하고 있다. 이는 2015년 설립된 미국의 DIU와 같은 성격으로 DIU 본사는 기술전문 기업과 스타트업 기업들이 모여 있는 캘리포니아주의 실리콘밸리 근처에 있으며, 이사(Director)급 임원진만 13명이 있으며, 인공지능 부문의 現담당 이사인

89) 출처 : ETRI AI드론 SW, 美 항공청 최고 안전등급 획득 (2020331, 전자통신연구원)

Jeff Klugman의 경우 30년 이상의 실리콘 밸리에서의 경험과 관련 학과를 전공하였다. 이와같이 DIU는 민간의 빠르게 변화고 전문적인 기술을 이해하는 전문가를 영업하여, 인공지능 기술이 필요한 적제적소에 적용하는 프로젝트를 수행하고 있다. 이를 참고하여 우리나라도 관련 분야의 전문가 영입 등의 방안이 필요할 것으로 보인다.

제5장. 인공지능 기술에 대한 윤리적 고찰 검토

우리나라 국방 분야에 있어서 인공지능 기술에 대한 윤리적인 측면을 본격적으로 인식하기 시작한 시점은 2018년 4월 세계 30여개국의 인공지능 전문가 50여 명이 앞으로 우리나라 카이스트와는 어떠한 연구도 함께 하지 않겠다는 보이콧을 선언했을 때였던 것으로 보인다. 이러한 배경에는 카이스트가 방산 업체인 한화시스템과 '국방 인공지능 융합연구센터'를 2018년 2월 개소한 뒤 인공지능 기술이 적용된 AI 기반 지능형 항공기 훈련 시스템 및 무인 잠수형 복합항법 알고리즘 등을 연구/개발 하고 있었으며, 이를 기사화한 해외언론에서 인공지능 무기라는 단어를 사용한 것이 발단이 된 것이다. 이에 이 문제를 해결하기 위해서, 카이스트는 보이콧을 선언한 인공지능 전문가 50여 명에게 "킬러로봇 개발 의사가 전혀 없으며, 인권과 윤리를 최고의 가치로 여긴다"라는 내용의 해명서를 보내기도 하였다. 이와 같이 '킬러로봇'과 같은 국방분야의 인공지능 기술적용에 대한 윤리적인 문제는 인공지능 기술의 개발과 동시에 병렬적으로 검토해 나가야하는 중요한 사한이 되었다.

1. 로봇 및 인공지능 윤리에 대한 이론 및 추이

로봇 윤리 원칙에 대한 고전으로 널리 알려져 있는 1942년 아이작 아시모프(Isaac Asimov)가 단편 'Runaround'에서 제안된 로봇 윤리 3원칙은 다음과 같으며, 0원칙은 1985년 단편 'Robots and Empire'에서 추가된 것으로 1원칙과 비슷하나, 단순히 한 개개의 인간에서 인류 전체로 그 의미를 확장한 것이다.

[아시모프 로봇윤리 원칙]

① 0원칙 : 로봇은 인류에게 해를 가하거나, 혹은 해야 할 행동을 하지 않음으로써 인간에게 해를 끼쳐서는 안 된다.

(A robot may not injure a human being or, by inaction, allow a human being to come to harm.)

② 1원칙 : 로봇은 인간에 해를 끼치거나, 혹은 해야 할 행동을 하지 않음으로써 인간에게 해를 끼쳐서는 안 된다.
(A robot may not injure a human being or, through inaction, allow a human being to come to harm.)

③ 2원칙: 로봇은 인간의 명령에 복종해야 한다. 단 이러한 명령들이 첫 번째 법칙에 위배될 때에는 예외로 한다.
(A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.)

④ 3원칙 : 로봇은 자신의 존재를 보호해야 하다. 단 그러한 보호가 첫 번째와 두 번째 법칙에 위배될 때에는 예외로 한다.
(A robot must protect its own existence as long as such protection does not conflict with the First or Second LAWS.)

이는 2000년 이전에 로봇이 많지도 않고, 사람과 함께 생활하던 시절이 아닌 시기로, 로봇대해서 일방적인 임무를 많이 강조하던 원칙으로, 로봇이 해야하는 일들(인간에 대한 복종 및 자신을 보호)에 대해 기술해 놓았다. 이후 2004년 일본의 후쿠오카에서 개최된 International Robot Fair에서는 아래와 같이 세계로봇선언(World Robot Declaration)을 선포하였으며, 로봇과 인간이 서로 함께 공존한다는 것을 보다 강조한 것이 이전과는 다른 것이 특징이다.

[후쿠오카 세계로봇선언]

① 차세대 로봇은 인간과 함께 공존하는 파트너가 될 것이다.

② 차세대 로봇은 인간을 육체적, 정신적으로 보조할 것이다.

③ 차세대 로봇은 안전하고 평화로운 사회구현에 기여할 것이다.

그 후 우리나라에서도 2007년에 과학자, 의사, 심리학자, 변호사 공무원 등 각계인사 12명을 중심으로 로봇윤리 협의체를 구성하고, 2008년에 로봇윤리 헌장 초안을 발표하였으나, 로봇의 지위 결정 문제가 해결되지 않아 아직 로봇윤리 헌장이 완성되지 않고 있으며, 초안을 살펴보면 다음과 같다.[90]

[우리나라 로봇윤리 헌장 - 초안]

① 1장(목표) : 로봇윤리헌장의 목표는 인간과 로봇의 공존공영을 위해 인간중심의 윤리규범을 확인하는 데 있다.

② 2장(인간, 로봇의 공동원칙) : 인간과 로봇은 상호간 생명의 존엄성과 정보, 공학적 윤리를 지켜야 한다.

③ 3장(인간 윤리) : 인간은 로봇을 제조하고 사용할 때 항상 선한 방법으로 판단하고 결정해야 한다.

④ 4장(로봇 윤리) : 로봇은 인간의 명령에 순종하는 친구 · 도우미 · 동반자로서 인간을 다치게 해서는 안 된다.

⑤ 5장(제조자 윤리) : 로봇 제조자는 인간의 존엄성을 지키는 로봇을 제조하고 로봇 재활용, 정보보호 의무를 진다.

⑥ 6장(사용자 윤리) : 로봇 사용자는 로봇을 인간의 친구로 존중해야 하며 불법개조나 로봇남용을 금한다.

⑦ 7장(실행의 약속) : 정부와 지자체는 헌장의 정신을 구현하기 위해 유효한 조치를 시행해야 한다.

그리고 2010년 영국의 공학과물리과학연구위원회(EPSRC, Engineering and Physical Sciences Research Council)에서 로봇 윤리에 관한 5원칙을 발표하였고, 앞선 윤리원칙들과 다른 점은 책임 주체와 개인의 사생활 보호와 관련된 사항은 포함한 것이며, 이는 인공지능 학습을 위해서는 빅데이터와 같은 많은 데이터가 필요하기 때문에 우려했던 것으로 보인다.

90) 출처 : 로봇윤리 이론, http://robofriend.kr/skin_mw3/sub_page.php?page_idx=89

[EPSRC의 로봇 윤리][91)]

① 로봇은 국가안보 사유 외에는 무기 용도로 설계되어서는 안된다.
(Robots should not be designed as weapons, except for national security reasons.)

② 로봇은 사생활을 포함하여 現법체계를 준수하도록 설계되고 운용되어야 한다.
(Robots should be designed and operated to comply with existing law, including privacy.)

③ 로봇은 안전과 보안 보장에 적합하도록 설계되어야 한다.
(Robots are products: as with other products, they should be designed to be safe and secure.)

④ 로봇은 사람에게 착각이나 환상을 불러일으키도록 사용되어서는 안된다.
(Robots are manufactured artefacts: the illusion of emotions and intent should not be used to exploit vulnerable users.)

⑤ 모든 로봇에 대해서 (법적)책임있는 사람이 명시되어야 한다.
(It should be possible to find out who is responsible for any robot.)

2017년 스카이프 창업자인 얀 탈린 등 5명이 설립하였으며, 저명한 인공지능 학자인 스튜어트 러셀과 테슬라 CEO인 일런 머스크 등이 자문 역할을 하고 있는 민간 단체인 삶의미래연구소(FLI, Future of Life Institute)에서 크게 연구 이슈, 윤리 및 가치 그리고 장기 이슈의 3가지 범주에서 23개의 '아실로마 AI 원칙(Asilomar AI Principles)[92)]'을 아래의 표와 같이 발표했다. FLI는 인공지능 기술이 인간을 번성하게도 하지만 킬러로봇과 같이 잘못 사용될 경우 위험한 결과를 초래할 수 있음을 강조하며 나은 미래를 만들기 위한 노력을 해야함을 강조하고 있다.

분야	적용
연구 이슈 (Research Issues)	**1) 연구목표** : 인공지능 연구의 목표는 방향성이 없는 지능을 개발하는 것이 아니라 인간에게 유용하고 이로운 혜택을 주는 지능을 개발하는 것이다.

91) 출처 : https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/
92) 출처 : 아실로마 AI 원칙(한국어) https://futureoflife.org/ai-principles-korean/

연구 이슈 (Research Issues)	**2) 연구비 지원** : 인공지능에 대한 투자에는 컴퓨터 과학, 경제, 법, 윤리 및 사회 연구 등의 어려운 질문을 포함해 유익한 이용을 보장하기 위한 연구비 지원이 수반되어야 한다: ・어떻게 미래의 인공지능 시스템을 강력하게 만들어 오작동이나 해킹 피해 없이 우리가 원하는 대로 작업을 수행하도록 할 수 있나? ・사람들의 자원과 목적을 유지하면서 자동화를 통해 우리 번영을 어떻게 성장시킬 수 있나? ・인공지능과 보조를 맞추고 인공지능과 관련된 위험을 통제하기 위해, 보다 공정하고 효율적으로 법률 시스템을 개선할 수 있는 방법은 무엇인가? ・인공지능은 어떤 가치를 갖추어야 하며, 어떤 법적 또는 윤리적인 자세를 가져야 하는가? **3) 과학정책 연계** : 인공지능 연구자와 정책 입안자 간에 건설적이고 건전한 교류가 있어야 한다. **4) 연구 문화** : 인공지능 연구자와 개발자 간에 협력, 신뢰, 투명성의 문화가 조성되어야 한다. **5) 경쟁 회피** : 인공지능 시스템 개발팀들은 안전기준에 대비해 부실한 개발을 피하고자 적극적으로 협력해야 한다.
윤리 및 가치 (Ethics and Values)	**6) 안전** : 인공지능 시스템은 작동 수명 전반에 걸쳐 안전하고 또 안전해야 하며, 적용 가능하고 실현 가능할 경우 그 안전을 검증할 수 있어야 한다. **7) 실패의 투명성** : 인공지능 시스템이 손상을 일으킬 경우 그 이유를 확인할 수 있어야 한다. **8) 사법적 투명성** : 사법제도 결정에 있어 자율시스템이 사용된다면, 권위 있는 인권기구가 감사 할 경우 만족스러운 설명을 제공할 수 있어야 한다. **9) 책임성** : 고급 인공지능 시스템의 디자이너와 설계자는 인공지능의 사용, 오용 및 행동의 도덕적 영향에 관한 이해관계자이며, 이에 따라 그 영향을 형성하는 책임과 기회를 가진다. **10) 가치 일치** : 고도로 자율적인 인공지능 시스템은 작동하는 동안 그의 목표와 행동이 인간의 가치와 일치하도록 설계되어야 한다. **11) 인간의 가치** : 인공지능 시스템은 인간의 존엄성, 권리, 자유 및 문화적 다양성의 이상에 적합하도록 설계되어 운용되어야 한다. **12) 개인정보 보호** : 인공지능 시스템의 데이터를 분석 및 활용능력의 전제하에, 사람들은 그 자신들이 생산한 데이터를 액세스, 관리 및 통제할 수 있는 권리를 가져야 한다. **13) 자유와 개인정보** : 개인정보에 관한 인공지능의 쓰임이 사람들의 실제 또는 인지된 자유를 부당하게 축소해서는 안된다. **14) 공동이익** : 인공지능 기술은 최대한 많은 사람에게 혜택을 주고 힘을 실어주어야 한다. **15) 공동번영** : AI에 의해 이루어진 경제적 번영은 인류의 모든 혜택을 위해 널리 공유되어야 한다. **16) 인간의 통제력** : 인간이 선택한 목표를 달성하기 위해 인간은 의

윤리 및 가치 (Ethics and Values)	사결정을 인공지능 시스템에 위임하는 방법 및 여부를 선택해야 한다. **17) 비파괴** : 고도화된 인공지능 시스템의 통제로 주어진 능력은 건강한 사회가 지향하는 사회적 및 시정 과정을 뒤엎는 것이 아니라 그 과정을 존중하고 개선해야 한다. **18) 인공지능 무기 경쟁** : 치명적인 인공지능 무기의 군비 경쟁은 피해야 한다.
장기 이슈 (Longer-term Issues)	**19) 인공지능 능력에 관한 주의** : 합의가 없으므로 향후 인공지능 능력의 상한치에 관한 굳은 전제는 피해야 한다. **20) 중요성** : 고급 AI는 지구 생명의 역사에 심각한 변화를 가져올 수 있으므로, 그에 상응 한 관심과 자원을 계획하고 관리해야 한다. **21) 위험성** : 인공지능 시스템이 초래하는 위험, 특히 치명적인 또는 실존적 위험에는, 예상된 영향에 맞는 계획 및 완화 노력이 뒷받침되어야 한다. **22) 재귀적 자기 개선** : 인공지능 시스템이 재귀적 자기 복제나 자기 개선을 통하여 빠른 수적 또는 품질 증가를 초래한다면, 설계된 시스템은 엄격한 안전 및 통제 조치를 받아야 한다. **23) 공동의 선** : 초지능은 널리 공유되는 윤리적 이상을 위해, 그리고 몇몇 국가나 조직이 아닌 모든 인류의 이익을 위해 개발되어야 한다.

또한 로봇윤리 외에 인공지능 기술에 대해서 국제표준기구인 ISO/IEC의 JTC1/SC42(Working group3)에서는 2020년 5월 '인공지능의 신뢰성 개관(Overview of trustworthiness)' 라는 규격(ISO/IEC TR 24028:2020)을 제정하였으며, 이는 인공지능의 신뢰성을 위한 방안으로 투명성, 설명가능성, 제어가능성, 개인보호, 기능적 안전 및 시험과 평가 등의 내용을 다루고 있다. 그리고 인공지능 윤리(AI Ethics)에 대해서도 사안의 중요성을 인식하고 있으며, 이와 관련한 표준을 제정하기 위하여 연구가 활발히 진행 중이다.

그리고 유네스코(UNESCO, 국제연합 교육과학문화기구) 역시 인공지능 윤리에 대하여 이에 맞는 적절한 전문가 그룹(AHEG, Ad Hoc Expert Group)을 구성하여, 2020년 5월 AI윤리 권고안에 대한 5가지의 주요목적(Action Goal)과 11개의 정책과제(Policy action)를 담은 초안[93]을 작성 및 공개하였으며, 의견 등을 검토하여 2021년 11월 41차 컨퍼런스에서 최종 채택 및 배포할 계획을 가지고 있다.

93) 출처 : Outcome document:first draft of the Recommendation on the Ethics of Artificial Intelligence (2020, UNESCO), https://unesdoc.unesco.org/ark:/48223/pf0000373434

이러한 인공지능 윤리에 관하여 많은 국가들이 참여하여 공동된 인식과 윤리 체계를 정립하고자 최초 훈련국인 캐나다와 프랑스를 중심으로 추진되어 2020년 6월 15일 우리나라를 비롯한 14개 국가[94]는 인간 중심의 인공지능, 책임성 있는 인공지능을 위한 국제협의체인 '인공지능에 대한 글로벌 파트너십(GPAI, Global Partnership on Artificial Intelligence)'를[95] 공식적으로 창립하였으며, '책임성 있는 인공지능을 위한 공동선언문'을 발표했다. 조직으로는 GPAI 회원국 대표(장・차관급)로 구성된 Council의 연 1회 회의와 100~150명의 각 기관 전문가로 이루어진 다중이해관계자 전문가그룹(Multistakeholder Experts Group) 그리고 캐나다 몬트리올과 프랑스 파리에 전문센터(Centre of Expertise)로 이루어져 있다. GPAI는 업계・시민사회・정부기관・학계 등 다방면의 전문가가 참여하여 1)책임성있는 인공지능, 2)데이터 거버넌스, 3)미래의 일자리, 4)혁신과 상업화 주제의 전문가 그룹을 운영할 계획할 예정이며, 책임성있는 인공지능과 관련하여 인공지능 윤리 등과 관련한 논의가 많은 진전이 있을 것으로 생각된다.
다음은 GPAI 조직도이며, 4개의 워킹그룹으로 구성되어 있다.

94) GPAI(Global Partnership on AI) 참여국 : 한국, 프랑스, 캐나다, 호주, 미국, EU, 독일, 이탈리아, 일본, 뉴질랜드, 싱가폴, 슬로베니아, 영국, 멕시코 14개국
95) 출처 : 세계최초 인공지능 협의체 GPAI 공식 창립 (20200615, 과학기술정보통신부 보도자료)

그리고 아래는 공동선언문(한글) 전문으로 제 1차 총회는 2020년 12월에 훈련국인 캐나다에서 주최할 예정으로 되어 있다.

글로벌 인공지능 파트너십(GPAI) 창립회원국 공동선언문

우리, 호주, 캐나다, 프랑스, 독일, 이탈리아, 일본, 멕시코, 뉴질랜드, 대한민국, 싱가포르, 슬로베니아, 영국, 미국, 유럽연합(EU)은 함께 모여 '글로벌 인공지능 파트너십(Global Partnership on Artificial Intelligence, GPAI)'을 마련하였다. 창립회원국으로서, 우리는 OECD 인공지능(AI) 권고안에 기술된 바와 같이 인권과 기본적 자유 및 우리가 공유하는 민주적 가치에 부합하는 방식으로 인공지능의 책임 있고 인간 중심적인 발전과 사용을 지원할 것이다. 이를 위해, 우리는 다른 관심 있는 국가들 및 파트너들과 협업할 것을 기대한다.

GPAI는 인권, 포용, 다양성, 혁신, 경제성장에 기반을 둔 인공지능의 책임 있는 개발과 사용을 인도하기 위해 다자이해관계자가 참여하는 국제적 구상(initiative)이다. 이 목표를 달성하기 위해, GPAI는 인공지능 관련 우선순위에 대한 첨단 연구와 응용 활동을 지원하여 인공지능에 대한 이론과 실천 간 격차를 좁히고자 할 것이다.

GPAI 는 파트너 및 국제기구들과 협력하여, 산업계·시민사회·정부·학계 주요 전문가들을 한데 모아 1)책임 있는 인공지능, 2) 데이터 거버넌스, 3) 일의 미래, 4) 혁신과 상업화의 네 가지 작업반 주제에 걸쳐 협업할 것이다. 특히 단기적으로는 GPAI 전문가들은 어떻게 인공지능을 활용하여 코로나19에 보다 잘 대응하고 회복할 수 있는지에 대해서도 연구할 예정이다.

GPAI 는 파리에 위치한 OECD 사무국과 몬트리올과 파리에 각 각 자리할 두 개의 전문지식 센터(Centre of Expertise)에 의해 지원될 것이다. OECD와 GPAI의 관계는 GPAI의 과학기술 업무와 OECD가 제공하는 국제 인공지능 정책 리더십의 강력한 시너지를 탄생시킬 것이며, 이로 인해 책임 있는 AI를 지향하는 정책을 위한 증거 기반이 강화될 것이다. 전문지식 센터는 다양한 부문과 학문분야의 작업반 전문가가 수행하거나 평가한 실제 프로젝트에 대해 행정 및 연구 지원을 제공할 것이다. 또한 전문지식 센터는 다자이해관계자 전문가그룹 연차총회도 계획할 예정이며 제 1차 총회는 2020년 12월 캐나다가 처음 주최할 예정이다.

2. 킬러로봇에 대한 국제적인 여론

앞서 언급한 우리나라 카이스트와 한화의 국방분야의 인공지능 기술 개발과 관련한 50여명의 인공지능학자들의 보이콧 선언과 같이 이미 각국의 인공지능 학자를 포함하여 국제 인권단체들과 같은 많은 기관 및 단체 그리고 개인들은 킬러로봇 개발 금지를 위하여 목소리를 내고 있다. 이들은 공통적으로 이미 용납될 수 있는 수준을 넘어서고 있고, 많은 법적, 윤리적인 문제들을 일이키는 치명적 자율무기 시스템(LAWS)나 킬러로봇으로 불리는 완전자율무기들(Fully autonomous weapons)을 새로운 국제적인 법률을 통해 예방하거나 금지해야 한다고 주장하고 있다.

1) 킬러로봇 금지 캠페인 (Campaign to stop Killer Robots)

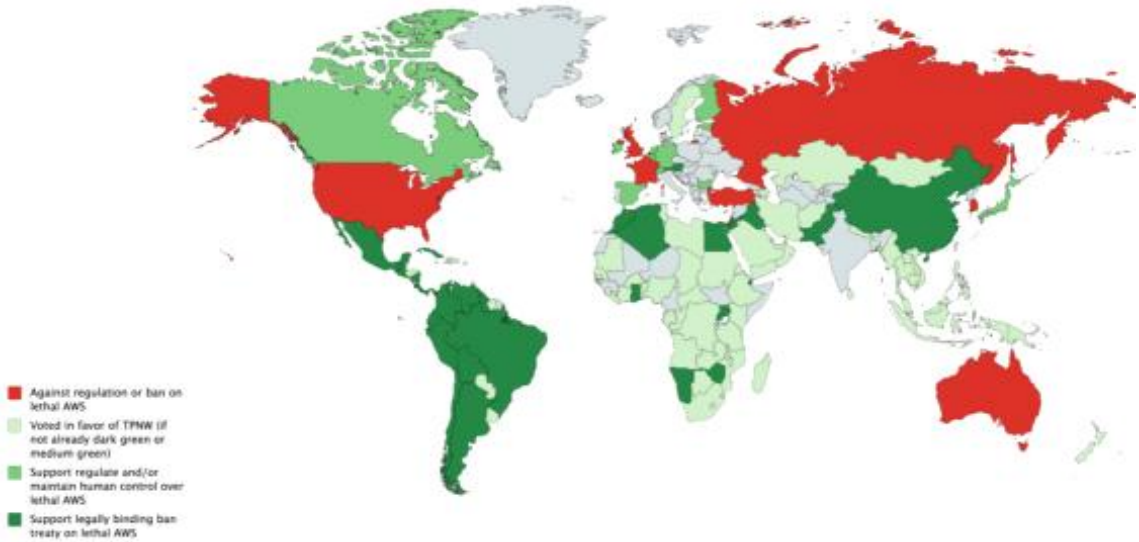
킬러로봇의 금지를 주장하며, 이를 저지하기 위해 적극적으로 활동하고 있는 이 캠페인을 추진하고 있는 단체는 NGOs이 연합으로써, 2012년 10월 설립이 되었고, 홈페이지에는 킬러로봇에 반대하는 국가와 단체 등에 대해서 다음과 같이 나와 있다.[96]

- 국가 : 브라질, 이집트 등 30개국 (각주 96 참조)
- 단체 : Human Rights Watch 등 140개 이상의 NGO 단체
- 전문가 : 4,500여명의 인공지능(AI) 전문가
- UN 사무총장 (United Nations Secretary-General) : Antonio Guterres
- The European Parliament : EU 의회
- Human Rights Council rapporteures : UN 인원 이사회 조사위원들
- 26명의 노벨 평화상 수상자들
- 일반 대중 : 약 61%의 일반인들

96) 출처 : https://www.stopkillerrobots.org/wp-content/uploads/2020/03/KRC_CountryViews_11Mar2020.pdf 이 문서에는 30개국에 대해서 명시(1. Algeria 2. Argentina 3. Austria 4. Bolivia 5. Brazil 6. Chile 7. China* 8. Colombia 9. Costa Rica 10. Cuba 11. Djibouti 12. Ecuador 13. Egypt 14. El Salvador 15. Ghana 16. Guatemala 17. Holy See 18. Iraq 19. Jordan 20. Mexico 21. Morocco 22. Namibia 23. Nicaragua 24. Pakistan 25. Panama 26. Peru 27. State of Palestine 28. Uganda 29. Venezuela 30. Zimbabwe)하고 있다.

그리고 홈페이지의 동영상에는 자율무기 개발을 점점 더 가속화하는 국가들로 미국, 중국, 이스라엘, 한국, 러시아, 그리고 영국을 직접적으로 명시를 하고 있다. 이는 일반 대중들이 볼 때 한국에 대한 부정적인 시각을 가질 수도 있는 부분으로 보여진다.

※ 이는 앞에서 언급된 삶의미래연구소(FLI, Future of Life Institute)의 홈페이지에도 킬러로봇에 대한 국가별 입장에 대해서 아래와 같이 도식화 하고 있고, 우리나라는 LAWS 금지를 반대하는 국가를 의미하는 붉은색으로 표시되어 있다.



< 출처 : Future of Life institute 홈페이지 >

※ 또한 Human Rights Watch 단체의 홈페이지에는 각국의 입장을 정리[97]하고 있으며, 우리나라는 아래와 같이 되어있다. 따라서 상기의 정확한 정보가 아닌 추측이나 오해 등은 해결해야 할 문제로 보이며, 그로써 잘못된 불신을 해소하는 것이 필요할 것이다.

"South Korea is researching, developing, and investing in military applications of artificial intelligence and weapons systems with autonomy in their functions, but says it does not possess lethal autonomous weapons systems and does not intend to develop or acquire them.(한국은 LAWS를 소유하고 있지 않으며 개발하거나 획득할 의도가 없다.) South Korea participated in every CCW meeting on killer robots in 2014-2019."

97) 출처: https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and#_ftn236

이 캠페인에서는 UN에서 LAWS와 관련하여 진행되고 있는 희의에 매번 참석하고 있으며, 2020년 6월에는 UN으로 앞으로 진행되어야할 방안에 대해서 의견서(commentary)[98]를 제출하였고, 크게 사람의 통제를 유지하는 것과 규범적인 틀이 필요하다는 두 가지로 나뉘고 있으며, 그 내용은 다음과 같다.

- 사람의 통제 유지(Retaining human control)

* 사람의 역할이 상대적 덜할 것 같은 판단(judgement)과 개입(intervention)과 같은 단어 보다는 **통제(control)라는** 단어를 더욱 선호함.

① 의사결정에서의 법적인 규직과 윤리적인 원리를 준하면서 취해질 수 있는 의미있는 사람의 통제가 필요하며, 작업자는 운용환경과 시스템의 작동 원리를 이해해야하고, 목표물을 확인하고 숙고할 수 있는 시간이 있어야 한다.

② 무기시스템의 기술적 특징으로 예상이 가능하고 신뢰성을 말하며, 이는 작업자에게 관련 정보의 제공하는 것과 시스템의 활성화 후 사람이 개입이 가능한 것을 말한다.

③ 운용상의 요소로써 사람의 통제가 더욱 의미있도록 하기위해서 무기시스템이 작동되는 시간과 장소 그리고 목표로하는 대상을 제한하는 것이다.

- 규범적인 체계의 필요(Normative framework needed)

① 행위(무력의 사용, use of force)에 대한 통제를 유지하기 위한 일반적인 의무(General obligation)가 필요하며, 이는 군사충돌과 법 집행의 두 상황 모두에 적용.

② 법적인 금지수단으로 의미있는 사람의 통제 없이 목표물을 식별 및 교전하는 무기시스템의 개발, 생산 및 사용을 금지하는 것임.

③ 목표물을 식별 및 교전하는 모든 다른 시스템을 사용함에 있어서 의미있는 사람의 통제를 가능하도록 하는 구체적인 긍정적인 의무

98) 출처 : Commentary for the Convention on Conventional Weapons Group of Governmental Experts on lethal autonomous weapons systems (20200605, Campaign to stop killer robots)

(Positive obligations)가 있어야 하며, 이는 완전 자율적 무기로 운용되는 시스템의 사용을 불법으로 만드는 요건을 확립하는데 도움이 될 것임.

2) 캐나다

앞서 언급한 캐나다를 이끌어가고 있으며, 세계적으로도 인공지능(AI) 분야에서 제프리 힌튼, 죠수아 벤지오 그리고 리치 서튼 등의 저명한 인공지능학자들과 함께 캐나다 연구 위원회(Canada Research Chair in Ethics, Law and Technology at the University of Ottawa)를 맡고있는 Ian Kerr는 직접 캐나다 총리에게 2017년 11월 2일 '인공지능(AI) 기술이 악용되지 않도록 캐나다의 역할을 다 해달라' 는 내용의 서한을 보냈으며, 이러한 과정을 통하여 캐나다는 앞선 나온 GPAI의 설립에 주도적 역할을 한 것으로 보인다. 아래는 서한의 주요 내용이다.

- 『Re:An International Ban on the Weaponization of AI』[99]

진화하고 있는 인공지능 기술 특히 머신러닝 기술들이 많은 다양한 곳에서 진전을 이루고 있으며, 더욱이 도덕적 관심이 필요하다고 하며, 무엇보다도 의미있는 사람의 통제없는 치명적인 자율무기 시스템(Lethal autonomous weapons systems that remove meaningful human control)에 대해서 사용되지 않게 하기 위하여, 또한 UN에서의 CCW(Certain Conventional Weapons) 회의에서의 자율무기 시스템에 대한 GGE(Group of Governmental Experts)를 설립한 것을 환영하며, 캐내다가 강력한 리더쉽으로 국제적인 법적, 윤리적, 사회적인 것으로 고려하여 UN에서의 활동에 있어 그 선두에 서서 그 가치를 지켜달라는 내용이다. 특히 해결되지 않을 경우의 '절망적 결과로 사람이 아닌 기계가 누구를 죽이고 살려야 할지를 결정하게 될 것(The deadly

99) 출처 : Open Letter to the Prime Minister of Canada (20171102, Ian Kerr 등) https://techlaw.uottawa.ca/bankillerai - 이 편지를 뒷받침하는 각 전문가들의 인용문구를 살펴보면 다음과 같다. "Leading in AI also means acting responsibly about it."(AI를 리딩하는 것은 그것에 대해 책임감 있는 행동하는 것이다.), "AI 기술은 선천적으로 좋거나 나쁜 것이 아니다. 그것은 지혜롭게 사용하도록 하는 우리에게 달려있다."

cnsequence of this is that machines-not people-will determine who lives and dies.)' 이라고 경고하고 있다. 또한, 이 문서는 각 과학부장관, 국방부장관, 외교부 등 주요 관련 부처에도 참고 수신인으로 지정되어 전달이 되었다.

Killer Robot에 대한 사용 금지를 요구하며, 인도주의적인 역할을 하고 있는 캐나다의 민간 단체 Mines Action Canada (MAC)는 최근 2019년 UN CCW-GGE 회의에도 참석하여 실천이 없는 회의만 할 것이 아니라 각 국가레벨의 금지 선언을 하기를 원하고 있으며, 나아가 의무적인 CCW의 LAWS에 대한 사용금지를 법적으로한 도구를 논의하고, 무기 사용에 있어 의미있는 사람의 통제를 확실히 해야한다고 공개발언[100] 하였다.

이렇듯 킬러로봇에 대한 우려는 이미 세계 곳곳에서 표명하고 있고, 이를 사전에 규제하기 위한 움직임이 나타나고 있다. 다음에서는 이를 대응하고 국제적인 규범을 제정하고 위한 노력 등에 대해서 알아보았다.

100) Missing the Forest for the Tresss at CCW (20191113, https://www.minesactioncanada.org)

3. 국방로봇에 대한 윤리적인 검토 사항

인공지능과 같은 신기술이 국방분야의 치명적자율무기시스템(LAWS)와 같은 분야에 사용이 되는 것에 대해서 '킬러로봇'과 같은 용어로 불러지며 많은 우려를 낳고 있다. 이에 대해서 UN에서 국제적인 조약과 같은 규범을 제정하기 위해 진행되고 있는 사항과 미국과 같은 주요국의 사항에 대해서 알아보았다.

1) UN의 치명적 자율무기에 대한 CCW-GGE[101] 회의

안토니우 구테흐스 UN사무총장은 2020년 1월 22일 신년 서두 연설로 2020년의 UN 사무총장의 4대 우선과제 (4대 위협)[102]를 언급하고 있으며, 이는 21세기 진전을 위태롭게 하고, 21세기 가능성을 위험에 빠뜨리게 할 떠오로는 위협이라고 하였다. 그 중 4번째는 인공지능의 발전에 의한 LAWS를 주로 의미하는 것으로 아래의 내용이 포함되어있다.

① "Lethal autonomous weapons — machines with the power to kill on their own, without human judgement and accountability — are bringing us into unacceptable moral and political territory." (인간의 판단이나 책임없이 스스로 죽일 수 있는 힘을 가진 치명적 자율 무기는 우리에게 받아들이기 힘든 도덕적이고 정치적인 영역을 안겨줄 것이다.)

---

101) CCW(Convention on Certain Conventional Weapons) : '특정재래식무기금지협약'의 정식명칭은 '과도한 상해나 무차별한 영향을 초래하는 특정재래식무기의 사용 금지 또는 제한에 과한 협약'이며, 비인도적 재래식무기협약이라고도 하며, 현재는 당사국(High Contracting Parties)이 총 125개국으로 되어 있고, 아래의 의정서 중에 2개 이상 가입하면 당사국이 되도록 규정되어 있음.

* CCW는 본문 11조 및 5개 부속의정서로 구성되어 있고, 우리나라는 제3,4의정서는 미가입함.
  - 제1의정서 : X-Ray로 탐지 불가능한 파편무기 사용 금지(1983.12월, 118개국)
  - 개정 제2의정서 : 지뢰 및 부비트랩 사용금지 또는 제한(1998.12월 106개국)
  - 제3의정서 : 소이성 무기의 사용금지 또는 제한(1983.12월, 115개국)
  - 제4의정서 : 실명 레이저무기의 사용 금지(1998.7월, 1109개국)
  - 제5의정서 : 전쟁잔류폭발물(ERW)의 제거 및 협력(2006.11월, 96개국)

* GGE(Group of Governmental Experts) : LAWS에서 새로운 기술과 관련한 각 정부의 전문가 그룹

102) 출처 : Remarks to the General Assembly on the Secretary-General's prioritiese for 2020 (20200122, UN홈페이지)
2020년의 4대 우선과제는 ① 높은 세계적 지정학적인 긴장상태 (the highest global geostrategic tensions) ② 인간의 존재와 관련된 환경적 위기 (an existential climate crisis) ③ 깊고 커져가는 세계적 불신 (deep and growing global mistrust) ④ 디지털 세계의 어두운 면 (the dark side of the digital world)

② “I have a simple and direct plea to all Member States: Ban lethal autonomous weapons now.” (나는 모든 회원국들에게 간단하고 직접적인 요청을 하고 있습니다. 지금 치명적인 자율 무기들을 금지하십시오.)

이렇듯 UN은 인공지능 기술 등과 같은 최신 기술이 활용된 LAWS에 대한 국제적인 규범 등의 제정하기 위해 노력하고 있으며, 이와 관련하여 2014년부터 2016년까지는 비공식 전문가 회의를 진행했으며, 2017년부터는 매년 LAWS에 대한 각 정부의 전문가그룹 회의를 통하여 LAWS에 대한 금지나 제한에 대한 국제적인 회의를 주최하고 있으며, 이를 통해 각국 및 NGO 단체들은 의견을 개진하고 있다. 2019년에는 3월과 8월에 제네바에서 회의가 열렸으며, 회의 내용에 대한 보고서[103]를 작성 및 공개하고 있으며, 회의에서 다루어진 안건은 아래의 5가지였다.

① An exploration of the potential challenges posed by emerging technologies in the area of Lethal Autonomous Weapons Systems to International Humanitarian Law; (국제인도법에 대한 치명적 자율 무기 시스템 분야에서 새로운 기술들이 제기할 수 있는 잠재적 도전에 대한 탐구)

② Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention; (협약의 목적과 목적과 관련된 개념과 특성에 대한 공통의 이해를 촉진하기 위해 고려 중인 시스템의 특성화)

③ Further consideration of the human element in the use of lethal force; aspects of human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems; (치사력 사용에 대한 인적 요소의 추가 고려; 치명적 자율 무기 시스템 분야에서 새로운 기술의 개발, 배치 및 사용에 있어 인간과 기계의 상호작용의 측면)

④ Review of potential military applications of related technologies in the context of the Group’s work; (그룹의 업무 맥락에서 관련 기술의 잠재적인 군사 적용에 대한 검토)

⑤ Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the Convention without prejudging policy outcomes and taking into account past, present and future proposals. (정책 결과를 선입견하지 않고 또한 과거, 현재 및 미래 제안을 고려하지 않고 협약의 목적과 목표의 맥락에서 치명적인 자율 무기 시스템 분야에서 새로운 기술들이 야기하는 인도주의적 및 국제적 보안 문제를 해결하기 위한 가능한 옵션.)

---

103) 출처 : Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems(20190925, UN홈페이지), 이 보고서에는 참가한 국가와 단체 등도 명시되어 있다.

그리고 토론 후 당사국에 의해 채택된 가이드 원리는 다음과 같다.

< 가이드 원리 >	< Guiding Principles >
국제법,관련 윤리적 측면 뿐만 아니라 특히 유엔헌장과 국제인도법(IHL),은 이 이 그룹의 지속적인 업무를 가이드해야 한다는 것이 확인되었다. 국제인도법에 대한 치명적인 자율무기시스템 분야에서 새로운 기술에 의한 잠재적 도전에 주목하면서, 향후 논의 결과에 대한 편견 없이 다음과 같이 확인되었다.	It was affirmed that international law, in particular the United Nations Charter and International Humanitarian Law (IHL) as well as relevant ethical perspectives, should guide the continued work of the Group. Noting the potential challenges posed by emerging technologies in the area of lethal autonomous weapons systems to IHL,1 the following were affirmed, without prejudice to the result of future discussions:
(a) 국제인도법은 치명적인 자율 무기 시스템의 개발 및 사용을 포함한 모든 무기 시스템에 지속적으로 적용된다.	(a) International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems;
(b) 무기체계 사용에 대한 의사결정에 대한 인간의 책임은 기계에 이전될 수 없으므로 유지되어야 한다. 이는 무기 시스템의 전 수명 주기에 걸쳐 고려되어야 한다.	(b) Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapons system;
(c) 다양한 형태를 취하며 무기의 라이프 사이클의 다양한 단계에서 구현될 수 있는 인간-기계 상호작용은 치명적인 자율 무기 시스템 영역에서 새로운 기술에 기초한 무기 시스템의 잠재적 사용이 적용 가능한 국제법, 특히 국제인도법을 준수하는지 확인해야 한다. 인간과 기계의 상호작용의 품질과 정도를 결정할 때 운용 상황, 그리고 무기체계 전체의 특성과 능력을 포함한 다양한 요인을 고려해야 한다.	(c) Human-machine interaction, which may take various forms and be implemented at various stages of the life cycle of a weapon, should ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with applicable international law, in particular IHL. In determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context, and the characteristics and capabilities of the weapons system as a whole
(d) CCW 프레임워크의 신흥 무기 시스템의 개발, 배치 및 사용에 대한 책임은 인간 지휘 및 통제라는 책임 있는 체인 내에서 그러한 시스템의 운영을 포함하여 적용 가능한 국제법에 따라 보장되어야 한다.	(d) Accountability for developing, deploying and using any emerging weapons system in the framework of the CCW must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control;

(e) 국제법에 따른 국가의 의무에 따라, 전쟁의 수단 또는 방법이 되는 새로운 무기의 연구, 개발, 획득 또는 채택에 있어, 그 고용이 일부 또는 모든 상황에서 국제법에 의해 금지되는지 여부를 결정해야 한다.	(e) In accordance with States' obligations under international law, in the study, development, acquisition, or adoption of a new weapon, means or method of warfare, determination must be made whether its employment would, in some or all circumstances, be prohibited by international law;
(f) 치명적인 자율 무기 시스템, 물리적 보안, 적절한 비물리적 안전장치(해킹 또는 데이터 스푸핑에 대한 사이버 보안 포함) 분야에서 새로운 기술에 기반한 신무기 시스템을 개발 또는 획득할 때 테러집단에 의한 획득 위험과 확산 위험을 고려해야 한다.	(f) When developing or acquiring new weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, physical security, appropriate non-physical safeguards (including cyber-security against hacking or data spoofing), the risk of acquisition by terrorist groups and the risk of proliferation should be considered;
(g) 위험 평가 및 완화 조치는 모든 무기 시스템에서 새로운 기술의 설계, 개발, 시험 및 배치 순환의 일부가 되어야 한다.	(g) Risk assessments and mitigation measures should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems;
(h) 국제인도법 및 기타 적용 가능한 국제법적 의무 준수를 유지하기 위해 치명적인 자율 무기 시스템 분야에서 새로운 기술의 사용은 고려되어야 한다.	(h) Consideration should be given to the use of emerging technologies in the area of lethal autonomous weapons systems in upholding compliance with IHL and other applicable international legal obligations;
(i) 잠재적 정책 조치를 마련함에 있어 치명적인 자율 무기 시스템 분야의 새로운 기술을 의인화해서는 안 된다.	(i) In crafting potential policy measures, emerging technologies in the area of lethal autonomous weapons systems should not be anthropomorphized;
(j) CCW의 맥락 내에서 이루어지는 논의와 잠재적인 정책 조치는 지능형 자율 기술의 평화적 사용에 대한 진보나 접근을 방해해서는 안 된다.	(j) Discussions and any potential policy measures taken within the context of the CCW should not hamper progress in or access to peaceful uses of intelligent autonomous technologies;
(k) CCW는 군사상의 필요성과 인도주의적 고려사항 사이의 균형을 맞추기 위해 노력하는 협약의 목표과 목적의 맥락 내에서 치명적인 자율 무기 시스템 분야에서 새로운 기술 문제를 다루기 위한 적절한 프레임워크를 제공한다.	(k) The CCW offers an appropriate framework for dealing with the issue of emerging technologies in the area of lethal autonomous weapons systems within the context of the objectives and purposes of the Convention, which seeks to strike a balance between military necessity and humanitarian considerations.

2) 미국

美국방부는 2018년에 인공지능 기술을 이용해 이미지 인식 기술을 향상해 무인 항공기 타격률을 높이는 메이븐 프로젝트(Project Maven)을 구글과 계약하여 진행하던 중, 군사적인 목적으로 인공지능이 약용될 수 있다는 우려를 표명한 Google 직원들의 청원으로 인해서 Google은 메이븐 프로젝트를 철회하였다. 이러한 일련의 과정을 통하여 2018년 7월 국방부는 국방혁신위원회(DIB, Defense Innovation Board)[104]'s로 산업계, 학계 및 비영리 단체들과의 협력을 통하여 국방분야의 인공지능 윤리원리를 정립하기 위한 대화와 일련의 조언을 촉진하기 위한 노력을 요청하였으며, DIB에서는 국방분야의 인공지능에 대한 윤리적 측면에 대해 자문 역할을 진행하였고, 그 추진 배경에 대한 일부 내용[105]은 다음과 같다. 이를 통해서 美국방부가 인공지능의 윤리적인 측면을 중요하게 고려하고 있다는 것을 알 수 있다.

> *We have witnessed how deeply committed the women and men who work in the Department are to ethics: avoiding civilian casualties, adhering to international humanitarian law, and collaborating with allies in international fora to advance international law and norms. Additionally, the Department extensively tests all its systems, especially weapons systems, and systems employing AI will likely be subjected to more scrutiny than ever before.*
> *(우리는 국방부에서 일하는 직원들이 얼마나 깊이 윤리에 대해 헌신하고 있는지 알고 있으며, 이는 민간인 사상자를 피하고, 국제인도법을 준수하며, 국제법과 규범을 발전시키기 위한 국제적 포럼속에서 동맹국들과 협력하는 것이다. 또한 국방부는 모든 시스템을 광범위하게 시험하고 있으며, 특히 무기 시스템과 인공지능이 적용된 시스템에 대해서는 그 어느 때보다도 더 많은 정밀 조사를 하고 있다.*

> *The DIB noted that -- as with all new technologies -- rigorous work is needed to ensure new tools are used responsibly and ethically. The stakes are high in fields such as medicine or banking, but nowhere are they higher than in national security.*
> *(DIB는 모든 새로운 기술과 같이 새로운 도구가 책임감 있고 윤리적으로 사용되도록 하기 위해서는 엄격한 작업이 필요다고 지적했다. 의료나 금융과 같은 분야에서의 위험도 높지만, 국방 안보 분야만큼 높은 곳은 없다.)*

---

104) 국방혁신위원회(DIB)는 국방부의 국방장관, 국방차관 및 고위관료에게 3가지 중점 분야(사람과 문화, 기술과 역량, 훈련과 작전)의 관점에서 미래의 문제를 해결하기 위한 혁신적인 방법을 독립적인 조언과 추천을 하기 위한 목적으로 구성된 위원회이며, 인공지능 윤리적인 문제도 다루고 있다.

105) 출처 : Defense Innovation Board's AI Principles Project (https://innovation.defense.gov/ai/)

그리고 2020년 2월 24일 美국방부는 약 15개월간 산업계, 학계 등의 많은 인공지능 전문가들과의 협력을 통하여 도출된 인공지능에 대해 다음과 같은 윤리적 원칙 5가지[106]를 분명히 설정 및 배포하였으며, 이 원리들을 AI 분야에서의 미군이 법적, 윤리적 그리고 정책적 약속을 지키기 위해 전투 및 비전투 기능 모두에 적용할 예정이다.

1. **책임감이 있다(Responsible).** DoD personnel will exercise appropriate levels of judgment and care, while remaining responsible for the development, deployment, and use of AI capabilities. (DoD 담당자는 AI 역량 개발, 배치 및 사용에 대한 책임을 유지하면서 적절한 수준의 판단과 관리를 수행할 것이다.)

2. **공평하다(Equitable).** The Department will take deliberate steps to minimize unintended bias in AI capabilities. (국방부는 AI 역량의 의도하지 않은 편견을 최소화하기 위해 신중한 조치를 취할 것이다)

3. **추적가능하다(Traceable).** The Department's AI capabilities will be developed and deployed such that relevant personnel possess an appropriate understanding of the technology, development processes, and operational methods applicable to AI capabilities, including with transparent and auditable methodologies, data sources, and design procedure and documentation.(국방부의 AI 역량은 관련 직원들이 투명하고 감사 가능한 방법론, 데이터 출처, 설계 절차 및 문서화를 포함하여 AI 역량에 적용할 수 있는 기술, 개발 프로세스 및 운영 방법에 대한 적절한 이해를 가질 수 있도록 개발되고 배치된다는 것이다.)

4. **믿을 수 있다(Reliable).** The Department's AI capabilities will have explicit, well-defined uses, and the safety, security, and effectiveness of such capabilities will be subject to testing and assurance within those defined uses across their entire life-cycles. (국방부의 AI 역량은 명확하고 잘 정립된 용도를 가질 것이며, 그러한 역량의 안전, 보안 및 효과성은 전 수명주기에 걸쳐서 그러한 정의된 용도 내에서 시험과 보증이 될 것이다.)

5. **통치할 수 있다(Governable).** The Department will design and engineer AI capabilities to fulfill their intended functions while possessing the ability to detect and avoid unintended consequences, and the ability to disengage or deactivate deployed systems that demonstrate unintended behavior. (국방부는 의도하지 않은 결과를 감지하고 피할 수 있는 능력과 의도하지 않은 행동을 보이는 배치된 시스템을 해제하거나 비활성화할 수 있는 능력을 보유하면서 의도된 기능을 수행하도록 AI 기능을 설계 및 엔지니어링할 것이다.)

106) 출처 : DOD Adopts Ethical Principles for Artificial Intelligence (20200224, US DoD)

4. 향후 방안 및 제언

완전 자율 무기에 대한 윤리적인 문제가 바로 로봇윤리 이자 인공지능 윤리라고 할 수 있다. 이는 로봇 자체는 이전부터 만들어져 오고 있었으나, 새로운 기술 특히 인공지능 기술이 로봇에 적용이 되면서부터 로봇이 스스로 목표물에 대한 참, 거짓 값을 분석 및 분류 즉 판단을 하면서부터 사람의 개입이 전혀없는 소위 킬러로봇으로 불려지는 완전 자율 무기가 만들어지게 되기 때문이다. 따라서 킬러로봇에 대한 반대 여론이 높은 현상황속에서도 국방 분야에서 인공지능 기술을 적용하기 위해서는 무엇보다도 개발 시에 윤리적인 측면을 같이 고려해서 진행이 되어야 한다.

2020년 7월 미국의 AI국가안보위원회(NSCAI)에서는 두 번째 분기 보고서인 'Second Quarter Recommendations[107]' 을 배포하면서 인공지능 윤리를 실천하기 위한 아래의 도표를 첨부하였고, 이는 주요 5가지 분야(핵심 가치, 엔지니어링, 시스템성능, 인간과 AI 상호 작용, 책임과 관리)에 대하여 앞서 살펴본 美국방부(DoD)의 5가지 인공지능 윤리 원칙이 필요한 사항을 정리한 체크리스트와 같은 것이라 할 수 있다.

NSCAI Recommended Practices:		DOD PRINCIPLES OF AI ETHICS				
		Responsible	Equitable	Traceable	Reliable	Governable
Core Values	A1 - Employ technologies and operational policies for privacy, fairness, inclusion, human rights, and LOAC	x	x			x
	B1 - Consider and document value considerations based on how tradeoffs with accuracy are handled	x	x	x	x	
	B2 - Consider and document value considerations in systems that rely on representations of objective or utility functions	x	x	x	x	
	B3 - Conduct documentation, reviews, and set limits on disallowed outcomes	x	x	x	x	x
Engineering	1 - Concept of operations development, and design and requirements definition and analysis	x	x	x	x	x
	2 - Documentation of the AI lifecycle			x		
	3 - Infrastructure to support traceability, including auditability and forensics		x	x		
	4 - Security and robustness: addressing intentional and unintentional failures				x	x
	5 - Conduct red teaming				x	
System Performance	A1 - Standards for metrics & reporting	x	x	x	x	
	A2 - Representativeness of data and model for the specific context at hand	x	x	x	x	
	A3 - Evaluating an AI system's performance relative to current benchmarks			x		x
	A4 - Evaluating aggregate performance of human-machine teams	x				
	A5 - Reliability and robustness	x		x	x	
	A6 - For systems of systems, testing machine-machine/multi-agent interaction	x			x	
	B1 - Specifying maintenance requirements	x	x	x	x	
	B2 - Continuously monitoring and evaluating AI system performance	x	x	x	x	x
	B3 - Iterative and automated testing and validation	x			x	x
	B4 - Monitoring and mitigating emergent behavior	x			x	x
Human-AI Interaction	A1 - Define functions and responsibilities of human operators and assign them to specific individuals	x		x		
	A2 - Policies should define the tasks of humans across the AI lifecycle	x				
	A3 - Enable feedback and oversight to ensure that systems operate as they should	x			x	
	B1 - Human-AI design guidelines	x	x	x		x
	B2 - Algorithms and functions in support of interpretability and explanation	x		x		x
	B3 - Designs that provide cues to human operator(s) about the confidence a system has in its results or behaviors	x		x		x
	B4 - Policies for machine-human handoff	x		x		x
	B5 - Leveraging traceability to assist with system development and understanding	x		x	x	x
	B6 - Training	x	x	x	x	x
Accountability/ Governance	1 - Identify responsible actors	x		x		x
	2 - Adopt technology to strengthen accountability processes and goals	x		x		x
	3 - Adopt policies to strengthen accountability	x		x		
	4 - External oversight support			x		

107) 출처 : Second Quarter Recommendations Quarterly Series, No. 2 (NSC on AI)
원문 : dnjhttps://drive.google.com/file/d/1hgiA38FcyFcVQOJhsycz0Ami4Q6VLVEU/view?usp=sharing

위의 표를 한글로 다시 정리하면 다음과 같다.

< NSCAI 권고 실행과제 >

분야	내용	적용 윤리				
		R1	E	T	R2	G
핵심 가치 (Core Values)	기술과 작전적 정책을 사생활,공정,포용,인권법 및 국제인도법을 위해 적용한다.	x	x			x
	어떻게 정확히 절충할 것인가에 근거하여 가치 중요사항을 고려하고 문서화 한다.	x	x	x	x	
	목적이나 효용함수의 표현에 의존하는 시스템의 가치 중요사항을 고려하고 문서화 한다.	x	x	x	x	
	허용하지 않는 결과에 대한 문서화, 검토 및 한계를 정한다.	x	x	x	x	x
엔지니어링 (Engineering)	운영 개발 및 설계, 요구사항 정의 및 분석의 개념	x	x	x	x	x
	AI 수명주기에 대한 문서화			x		
	감사가능성과 포렌식을 포함한 추적성을 지원하기 위한 기반시설		x	x		
	보안과 강건성 : 의도적 및 비의도적 고장에 대한 해결				x	x
	레드티밍(red-teaming) 수행 * 레드티밍:화이트해커와 같이 약점을 미리 파악하는 것.				x	
시스템 성능 (System Performance)	지표와 보고를 위한 기준들	x	x	x	x	
	당면한 특정 상황에 대한 데이터와 모델의 대표성	x	x	x	x	
	현재 기준과 비교한 AI 시스템의 성능 평가			x		x
	유-무인 복합 팀의 종합 성능 평가	x				
	신뢰성과 강건성	x		x	x	
	시스템을 위한 기계-기계/멀티 에이전트 상호 작용 시험	x			x	
	유지 및 관리 요구사항 구체화	x	x	x	x	
	지속적인 AI 시스템 성능 모니터링과 평가	x	x	x	x	x
	반복적이고 지속적인 시험과 검증	x			x	x
	비상 상황에 대한 모니터링과 완화	x			x	x
인간-AI 상호 작용 (Human-AI Interaction)	사람 작업자의 역할과 책임을 정의하고 그들을 구체적인 개개인들에게 할당하기	x		x		
	AI 수명주기에 걸쳐서 사람의 임무를 정의하는 정책을 가져야 한다.	x				
	시스템이 해야하는대로 일을 수행하는 지를 확인하기 위한 피드백과 감시를 가능하게 해야한다.	x			x	
	인간-AI 설계 가이드라인	x	x	x		x
	영상해석능력과 설명을 지원하는 알고리즘과 기능	x		x		x
	사람 작업자가 시스템이 한 결과나 행동에 대해 확신을 가지기 위한 근거를 제공하는 설계	x		x		x
	기계-인간 자리교체에 대한 정책	x		x		x
	추적 기능을 활용하여 시스템 개발 및 이해 지원	x		x	x	x
	교육	x	x	x	x	x
책임/관리 (Accountability /Governance)	책임있는 행위자를 가려내기	x		x		x
	책임 프로세스와 목적을 강화하기 위한 기술을 채택하기	x		x		x
	책임을 강화하기 위한 정책을 채택하기	x		x		
	외부 감독 지원			x		

※ 적용 윤리 : R1=Responsible, E=Equitable, T=Traceable, R2=Reliable, G=Governable

위와 같이 우리나라 국방부 내의 인공지능을 적용하기 위한 분야에서는 인공지능과 로봇 윤리에 대한 원리를 세우고, 이를 실행하기 위한 체계적인 실행과제를 작성 및 적용할 필요가 있다. 무엇보다도 국제인도법의 중요 개념인, 무력충돌의 영향력을 최소화하며, 무력충돌 상황에서도 보장되어야 하는 전쟁희생자의 기본적인 권리, 즉 유보되거나 제한할 수 없는 인권에 대한 것으로써, 인간의 생명과 존엄성을 보호 및 보장하는 것을 무기 시스템에 적용함으로써 가능할 것이다.

그리고 인공지능 기반의 무기체계 사용 시, 그 의사결정에 대한 책임을 잘 정립하는 것이 필요하며, UN 가이드 원리에 따라서 기본적으로 기계에 책임을 전가할 수 없는 것을 인식하는 것이 필요하다 할 것이다.

앞서 나온 내용들을 바탕으로 인공지능 및 로봇 윤리에 중요한 사항을 정리하면 다음과 같다.

① 강건성

: 무엇보다도 신뢰할 수 있고, 믿을 수 있는 시스템이 되어야 한다. 이를 위해서는 어떤 사이버 공격이 GPS 교란에도 문제가 없는 시스템을 구축하는 것이 필요하다.

② 투명성

: 이 부분은 다른 말로 추적이 가능하고, 설명이 가능한 인공지능 시스템을 말한다. 이는 앞서 살펴본 Explainable 및 Traceable AI 인공지능 시스템이라 할 수 있으며, 향후 인공지능 참모와 같은 전장상황에서의 중요 결정을 내리기 위한 인공지능에 의한 판단 등에 대해서 그렇게 판단한 사유 등을 사용자가 직접 이해나 설명이 가능한 것을 의미한다. 예를 들어 최근 이세돌과 인공지능(NHN 한돌)의 바둑대결에서 인공지능의 실수로 이세돌이 이긴 경우에서 볼 수 있듯이, 인공지능이 왜 실수로 그러한 결정을 하였는지에 대해서는 알 수 없으며, 단지 추측을 할 뿐이다. 국방 분야, 의료 부분 및 재판과 같은 사람의 생명과

직결된 경우 더욱 중요해지는 부분이다. 따라서 실제 전장 상황에서는 인공지능이 도출해낸 결과를 합리적으로 분석할 수 있는 플랫폼을 만드는 것은 향후 아주 중요한 부분이 될 것이다.

③ 평등성

: 사람의 통제가 없는 킬러로봇 등에 의해 발생하는 윤리적 문제가 유일한 윤리적인 이슈가 아니며, 인공지능(AI)는 기계학습 알고리즘에 사용되는 데이터를 기반으로 학습이 되는 것이며, 이는 그 데이터를 제공하는 엔지니어에 의해 결정되는 것이다. 따라서 엔지니어의 성향이 그대로 반영되게 되어있다. 즉, 성차별주의, 인종차별주의 등의 기타 여러 형태의 차별이 기계학습에 녹아들어갈 수 있으므로, 이에 대한 인식과 편견이 없는 학습이 필요하다.

④ 통제성

: 소위 킬러로봇은 사람의 통제가 전혀없이 대상 목적물을 식별하고, 교전까지 할 수 있는 시스템으로써, 많은 안전상의 문제가 내재되어 있다. 예를 들어 사람이 직접 사람의 개입이 있을 때에도 많은 에러가 발생하고 있는 현실속에서, 사람의 개입이나 판단이 전혀없다면 정말 예상치 못하는 많은 문제들이 발생하게 된다.

⑤ 책임성

: 사람의 통제 속에 있을 때, 혹여나 이러한 상황속에서 문제 발생에 대한 책임과 각종 결정정 판단의 근거로 인공지능 알고리즘에 의한 자료를 근거할 때 그 판단에 대한 책임을 잘 정립하는 것이 필요하며, 이는 앞서 나온 투명성 부분이 같이 기술발전하면서 더욱 보완 발전할 것으로 보인다.

UN의 CCW-GGE 회의 결과의 가이드 원칙 (k)에 나온 것과 같이 CCW는 군사상의 필요성과 인도주의적 고려사항 사이의 균형을 맞추기 위해서 노력하고 있으며, 이를 위한 적절한 프레임워크를 제공하려 하고 있으므로, 관련 회의를 주의 깊게 검토하여 참석하여야 하며, 회의 진행 결과에 대해서도 주요 동맹국과 함께 각국의 대응 양상을 잘 살펴야 할 것이다

제6장. 종합 및 결론

2017년 3월 국가적 차원의 AI 전략을 선언한 첫 번째 나라이자 훈련 국인 캐나다는 CIFAR라는 연구전문기관을 통해서 체계적으로 국가 단위의 인공지능(AI)을 개발하고 있으며, 눈앞의 단기성과에 대한 압력 없이 장기적인 관점에서 인공지능 분야를 지원하고 있고, 적극적으로 해외의 많은 인재들을 영입을 할 수 있도록 개방적인 이민정책을 추진함으로써 인공지능 분야에서 강국이 되고 있다. 이러한 캐나다의 사례는 우리나라의 인공지능 기술을 위한 정책적으로 조치를 취할 수 있는 방안에 대해 시사하는 바가 크다고 할 수 있다.

이미 많은 자율주행차량이 실생활에 사용되고 있지만, 가끔씩 자율주행기능을 사용하여 사고를 나는 경우가 있다. 이와 같이 잘못된 이미지 인식으로 Stop 경고문을 46mph로 잘못 해석하여 가속을 하거나, 알고리즘을 교란시키는 노이즈나 신호를 주입하여 갑자기 기능이 된다면, 머신러닝은 마치 양날의 검과 같이 심각한 결과를 초래할 수 있다. 따라서 국방 분야는 이런 취약점이 있을 경우 그 영향이 막대하므로, 인공지능 기술이 적용되는 국방분야의 장비에 대한 안전성에 대한 검증을 사전에 철저히 하는 것이 무엇보다 중요하다. 이를 위해 특히 美 국방부의 GARD(Guaranteeing AI Robustness against Deception)와 같은 프로그램을 적극 활성화하여 신뢰성 및 보안 구축에 만반의 준비를 갖추어 할 것으로 보인다. 그리고 앞서 언급된 많은 사례 등을 참고하여 앞으로 우리나라 국방 분야에 어떻게 적용할지를 고민하여야 할 것으로 보이며, 몇 가지 추가적으로 언급하고자하는 사항을 적어보았다.

첫째, 최근 '미래 국방 인공지능 특화연구센터'를 카이스트내에 설립과 같이 국방분야의 인공지능 기술개발에 많은 노력을 하고 있지만, 이런 이론적인 연구 결과를 실제 사례로 접목하고 또한 美국방부의 연합AI센터(JAIC)나 영국의 방위안보촉진청(DASA)와 같은 국방부내에 전군을 아우르며 인공지능 기술 적용이 필요한 곳을 종합적으로 식별 및

적용 할 수 있는 기관 혹은 부서의 설립이 필요할 것으로 보인다. 또한 美국방부의 『Defense Innovation Board’s AI principals project』의 역할과 같이 인공지능(AI) 분야를 국방 각 부문에 적용함에 있어서 국가의 안전을 위할 뿐만 아니라 윤리적인 부분까지도 자문을 할 수 있는 국방부의 위원회가 필요할 것으로 보인다. 또한 미국의 경우와 같이 국방인력의 인공지능에 대한 이해를 돕기위한 강의 및 자료 등을 다양하게 개발하고 배포하며, 특히 과·팀장 이상에게 빠르고 정확하게 전반적인 AI 기술을 이해하도록 관련 자료 및 교육을 추진하여 향후 업무에 있어서 적절하고 올바른 의사결정을 하도록 도와야 한다.

둘째, 딥러닝 주요 특성상 높은 완성도를 위해서는 데이터의 양이 중요하지만, 국방 분야의 보안 문제 등으로 인해서 민간부문 보다는 데이터 확보가 어려운 것이 현실이다. 그리고 美 DARPA에서 추진하고 있는 ‘차세대 인공지능(AI Next)’ 사업의 일환인 인공지능 체계에 많은 데이터가 필요하지 않는 제3의 물결(third wave)를 우리나라에 적용하도록 노력해야 하며, 이는 GAN (Generative Adversarial Newtork)나 Zero-shot Learning와 같은 비지도학습 및 준지도학습 방법의 알고리즘 등을 잘 활용한다면 보안 등의 이유로 많은 데이터를 취합하기 어려운 국방분야의 현실적인 상황속에서도 이를 극복하고 선도적인 인공지능 기술을 개발할 수 있는 도약이 될 것이다.

셋째, 또한 인공지능 기술과 관련된 프로젝트의 경우, 대부분 소프트웨어 개발에 대한 소요가 많으나, 납품물이 소프트웨어인 계약에 대한 적정한 원가 산정에 대한 계약상대자의 애로 등의 문제가 발생하지 않도록, 이에 대한 획득절차를 정립하고, 비용산정 등과 관련된 문제들을 사전에 파악하여 적정한 가치를 책정하여 계약 이행에 문제가 없도록 해야 할 것이다. 그리고 2020년 CBinsights에서 발표한 100대 인공지능 스타트업 기업에 한국 기업이 없는 것과 관련하여, 우리나라의 산업에 있어서 신생 스타트업 기업들에 대한 지원이 필요한 것으로 보인다. 이는 인공지능(AI)를 성공적으로 구현하기 위해서는 중요한 자산인 데이터, 알고리즘 그리고 컴퓨팅 인프라 등의 다양한 분야의 전문 업체

들이 필수적이므로, 이에 대해서 맞춤형으로 방산기업을 육성하고 지원하는 정책을 추진할 필요가 있으며, 또한 지속적이고, 전문성이 누적이될 수 있도록 단순한 일회성 이벤트와 포상을 하는 경진대회의 성격이 아닌, 긴 안목을 가지고 매년 점진적인 향상을 꾀할 수 있도록 장기적인 계획하에 기술적 고도화를 이룰 수 있도록 미국의 방위고등연구계획국에서 주최하는 DARPA Challenge와 유사한 사업이 그 예가 될 것이다.

끝으로 우리나라는 세계 6위의 군사력과 앞선 IT 기술을 바탕으로 하여 국방 분야에서도 이미 세계적인 추세의 설명 가능한 인공지능 기술 및 적대적 기만에 강인한 프로그램 등의 기술 발전을 따라가기 위해 노력하고 있다. 이러한 분야에 있어서 선도적인 인공지능 플랫폼 등이 개발되면 향후 국가의 군사력 뿐만 아니라 경제력에도 큰 기여를 할 수 있을 것으로 보인다. 더불어 인공지능 플랫폼이 개발될 때, 윤리적인 관점에서 볼 때에도 문제없이 개발이 되도록 정부차원의 점검 방안도 필요할 것으로 보인다. 이를 위해 앞선 5장에서 살펴본 국방 분야의 윤리적인 부분에 있어서도 더욱 관심을 기울이며, GPAI와 UN CCW-GGE 등의 국제 협의체를 통하여 인공지능 기술을 윤리적으로 국방분야에 활용할 수 있는 기준과 합의를 마련하는데 많은 노력이 필요할 것으로 보인다.

<참 고 문 헌>

- 가트너 선정 ‘2020년 10대 기술 트렌드 동향
  원문:www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020
- Deep Learning, Nature 지 (2015년, Yann LeCun, Yoshua Benjio & Geoffrey Hinton)
- Top Trends on the Gartner Hype Cycle for Atrificial Intelligence, 2019
- 인공지능(AI) 시대 주요국의 인재양성 정책 (2019, 소프트웨어정책연구소)

· An Overview of National AI Strategies (2018, Politics+AI, Tim Dutton)
· 2019 Canadian AI Ecosystem (2019, jfgangne)
· Annual Report of the CIFAR Pan-Canadian AI Stragtegy (2019, CIFAR)
· Diagnosing Schizophrenia, AMII (2020년, Spencer Murray)
· A Formal Separation Between Strategic and Nonstrategic Behavior, AMII (2020년, James Wright 등 2인)
· Adversarial training approach for local data debiasing(2020년, Alain Tapp 등 5인)
· Using speech synthesis to train end to end spoken language understanding models(2019년, Loren Lugosch 등)
· Explainable Artificial Intelligence for Safe Intraoperative Decision Support(2019년, Lauren Gordon 등 3인)
· Towards international standards for evaluating machine learning (2019년, Frank Rudzicz 등)
· 소리소문없이 인공지능 기지 된 캐나다, 구글, 엔비디아, 삼성, LG 가 몰리는 이유. 원문 : https://blog.naver.com/attisevery/221331157233
· AAAI 2020 Keynotes Turing Award Winners Event(2020년, Geoff Hinton, Yann Le Cunn, Yoshua Benjio)
· Understanding Hinton's Capsule Networks. Part I: Intuition (2017, Max Pechyonkin)
· Stacked Capsule Autoencoders (2019, Geoffrey E. Hinton 등 8인)
· Directive on Automated Decision-Making (2019, Canada Government)
· Executive Order on Maintaining American Leadership in Artificial Intelligence(2019, White House)
· THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN: 2019 UPDATE (2019)
· Interim Report November 2019 (2019, NSC on AI)
· Algorithms in the Criminal Justice System : Risk Assessment Tools (2020년, epic.org)
· Senate Bill 10(Pretrial Release and Detention, California Courts)
· AI 100-The Artificial Intelligence Startups Redefining Industries

(2020, cbinsights)
- A Next Generation Artificial Intelligence Development Plan (2017, www.jaist.ac.jp)
- Will China lead the world in AI by 2030? (2019, Nature)
  원문 : https://www.nature.com/articles/d41586-019-02360-7
- China AI Development Report (2018)
- “14억 인구, CCTV 6억대로 감시“ 中, 걸음걸이까지 데이터화 (20200625, 한국경제)
- https://www.sohu.com/a/245386306_777813
- 출처 : Artificial Intelligence Technology Strategy (20170331, Stragegic Council for AI Technology) 원문 : https://www.nedo.go.jp/content/100865202.pdf
- 일본의 인공지능(AI) 전략 동향 : AI 전략 2019 보고서 (2019, 소프트웨어정책연구소)
- Guidelines for AI procurement Published 8 June 2020(Gov.UK 홈페이지)
- Driving forward trustworthy data sharing (2020, Centre for Data Ethics and Innovation)
- National Strategy for Artificial Intelligence published Mar 2019 (The Danish Government)
- Element AI 社 홈페이지 (www.elementai.com)
- 캐나다 국방정책서 (Strong, Secure, Engaged)
- Organizational structure of the Department of National Defence and the Canadian Armed Forces
- Industrial and Technological Benefits Policy: Value Proposition Guide (2018, Canada)
- 셀룰라 로보틱스社 홈페이지 (www.cellula.com)
- Multi-Satellite Data Integration for Operational Ship Detection, Identification and Tracking, Progress Report 2(2019, DRDC)
- Remotely Piloted Aircraft System (RPAS), (2019, 캐나다 국방부 홈페이지)
- Summary of the 2018 National Defense Strategy of The United States of America
- Summary of the 2018 Department of Defense Artificial Intelligence

Strategy : Harnessing AI to Advance Our Security and Prosperity
- DoD Tech Talk : JAIC (2019, AUSA)
- Annual Report 2019 of Defense Innovation Unit (2019, DIU)
- Understanding AI Technology published April 2020 (2020, JAIC)
- A DARPA Perspective on Artificial Intelligence (DARPA, John Launchbury)
- AI Next Campaign (DARPA 홈페이지)
- Explainable Artificial Intelligence (XAI) (DARPA, Dr. Matt Turek)
- This Is What the US Air Force Wants You To Think Air Combat Will Look Like in 2030 (20180326, www.thedrive.com)
- Skyborg program seeks industry input for artificial intelligence initiative(2019, 美공군 홈페이지)
- DCIST 홈페이지 (https://www.dcist.org)
- Defending Against Adversarial Artificial Intelligence (2019, DARPA)
- Raytheon : DOD Needs More Research on Stopping Medium-Size Drones (20200123,www.airforcemag.com)
- Unmanned Aircraft System Airspace Integration Plan (March 2011-Version 2.0, DoD)
- Watch Russia's S-70 Unmanned Combat Air Vehicle Fly With An Su-57 For The First Time (20190927, www.thedrive.com)
- Russia's Robot Tank Sucks But Its Military Is Adopting It Anyway (20190124, taskandpurpose.com)
- Russia Begins Trials of New Generation Marker Unmanned Ground System (20200630, internationalinsider.ogr)
- Wings Along the BRI : Exporting Chinese UCAVs and Security?
원문 : /medium.com/@seideas/wings-along-the-bri-exporting-chinese-ucavs-and-security-a4bf7a324df
- Oddly shaped Chinese combat-ready helicopter drone popular in international market (2019, GlobalTimes.cn)
- Revolutionary Artificial Intelligence warship contrasts announced (2020, MoD Gov.uk)
- I-Korea 4.0 실현을 위한 인공지능(AI) R&D 전략 (2018.5)
- 2019년 국방기술품질원 통계연감, 2018년 주요 16개 국가 국방과학기술 수준

- 국방과학연구소 홈페이지 (www.add.re.kr)
- LIG넥스원 홈페이지 (www.lignex1.com)
- How Artificial Intelligence is Closing the Loop with Better Predictions (20180726, medium.com)
  원문: https://medium.com/hackernoon/how-artificial-intelligence-is-closing-the-loop-with-better-predictions-1e8550df3655
- Department of Defense Directive Number 3000.09 (Nov 21, 2012)
- 국방 인공지능(AI) 활용 실증기획 연구 (2018년, 국방부)
- ETRI AI드론 SW, 美 항공청 최고 안전등급 획득 (2020331, 전자통신연구원)
- 로봇윤리 이론, (http://robofriend.kr/skin_mw3/sub_page.php?page_idx=89)
- Principles of robotics (EPSRC)
  원문: https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/
- 아실로마 AI 원칙(한국어) https://futureoflife.org/ai-principles-korean/
- Outcome document:first draft of the Recommendation on the Ethics of Artificial Intelligence (2020, UNESCO)
  원문 : https://unesdoc.unesco.org/ark:/48223/pf0000373434
- 세계최초 인공지능 협의체 GPAI 공식 창립 (20200615, 과학기술정보통신부 보도자료)
- https://www.stopkillerrobots.org/wp-content/uploads/2020/03/KRC_CountryViews_11Mar2020.pdf
- https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and#_ftn236
- Commentary for the Convention on Conventional Weapons Group of Governmental Experts on lethal autonomous weapons systems (20200605, Campaign to stop killer robots)
- Open Letter to the Prime Minister of Canada (20171102, Ian Kerr 등)
  원문 : https://techlaw.uottawa.ca/bankillerai
- Missing the Forest for the Tresss at CCW (20191113, https://www.minesactioncanada.org)
- Remarks to the General Assembly on the Secretary-General's prioritiese for 2020 (20200122, UN홈페이지)
- Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous

Weapons Systems(20190925, UN홈페이지)

- Defense Innovation Board's AI Principles Project (https://innovation.defense.gov/ai/)
- DOD Adopts Ethical Principles for Artificial Intelligence (20200224, US DoD)
- Second Quarter Recommendations Quarterly Series, No. 2 (NSC on AI)

원문 : dnjhttps://drive.google.com/file/d/1hgiA38FcyFcVQOJhsycz0Ami4Q6VLVEU/view?usp=sharing