

공공분야 인공지능 플랫폼 분석 연구

윤 창 희*, 김 재 일^o

A Systematic Review on AI Platform for Public Sector

Chang-Hee Yun*, Jaeil Kim^o

요 약

최근 수년간 ICBM (사물인터넷(IoT), 클라우드 컴퓨팅(Cloud Computing), 빅데이터(Big Data), 모바일(Mobile))은 높은 수준으로 발전하였으며, ICBM 위에 빠르게 발전하고 있는 인공지능은 다양한 분야에서 인간의 능력을 넘어서는 수준으로 구현되고 있다. 더불어 인공지능 기술은 엄청난 데이터를 학습하여 정보화할 수 있는 빅데이터를 토대로 눈부신 발전을 하고 있다. 본 연구는 체계적인 분석기법을 이용하여 공공분야에서 데이터 기반으로 인공지능 기술력을 제고하기 위한 인공지능 플랫폼의 현황과 특징, 공공분야 적용을 위한 기술 요소에 대한 분석 결과를 제시한다. 특히, 논문 및 기술 문서, 기업 공개 정보를 기초로 체계적 분석 연구를 수행하여, 국내외 인공지능 플랫폼의 구성 요소와 특징을 알아보고, 공공분야에서 요구하는 인공지능 플랫폼의 조건에 대해 고찰한다. 본 연구 결과는 공공분야에서의 인공지능 플랫폼 구축 및 확산 방안 수립에 도움이 될 것으로 생각되며, 또한 인공지능 분야에 선진국과의 격차 해소를 위한 정책 및 전략수립 방안에 활용할 것으로 기대한다.

Key Words : Artificial Intelligence, AI Platform, Public Sector, Public Platform, Systematic Review

ABSTRACT

In recent years, ICBM (Internet of Things (IoT), Cloud Computing, Big Data, and Mobile) has developed to a high level, and the artificial intelligence, which is rapidly developing on top of ICBM, is being implemented at a level beyond human capabilities in a variety of fields. In addition, artificial intelligence technology is making remarkable developments based on big data that can learn enormous amounts of data and convert it into meaningful information. This study presents the results of the analysis of the current status and features of AI platforms to enhance AI technology based on data in the public sector, and technical factors for application in the public sector using systematic analysis methods. In particular, the AI platform's conditions required in the public sector are examined through investigating the components and characteristics of the AI platform at domestic and abroad by conducting systematic analysis research based on thesis and technical documents and public information of the companies. The results of this study will help in establishing and spreading AI platforms in the public sector, and it is also expected to be used in policies and strategies for reducing gaps with advanced countries in the AI sector.

* 이 연구는 2020년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임('20011875')

• First Author : National Information Society Agency, Kyungpook National University Department of Information Science, yunch@nia.or.kr, 석사연구원/학생(박사과정), 정회원

^o Corresponding Author : Kyungpook National University, School of Computer Science and Engineering, jaeilkim@knu.ac.kr 조교수, 정회원

논문번호 : 202007-140-0-SE, Received June 30, 2020; Revised August 23, 2020; Accepted August 27, 2020

I. 서 론

2016년 알파고 이벤트 이후 우리나라는 인공지능 기술에 대한 관심이 급증했다. 이에 따라 다양한 민간 분야에서는 산업 고도화를 위한 인공지능 기술 개발에 집중하기 시작하였고, 공공 분야 역시 인공지능을 4차 산업혁명 시대의 핵심 기술로 판단하고, 기술력 확보를 통한 인공지능 선진국 진입을 위한 다양한 전략을 수행하고 있다. 특히, 최근 빅데이터에 기반한 인공지능 기술 개발의 토대가 되는 인공지능 플랫폼의 중요성이 강조되고 있다¹⁾. 하지만 인공지능 선진국에 비해 인공지능 및 플랫폼 구축 기술력이 부족한 상황이며, 인공지능 플랫폼 기반의 산업 기술 발전 방향 수립에 대해 여러 시행착오를 겪고 있다²⁾.

플랫폼은 공급자와 수요자의 다양한 그룹이 참여하여, 공정한 거래를 통해 서로 가치를 교환하여 새로운 가치 창출을 이끌어 내는 생태계로 정의할 수 있다³⁾. 이러한 측면에서 인공지능 플랫폼은 인공지능 기술 개발자나 공급자가 단일 방향으로 기술을 제공하는 형태가 아닌 다양한 기술 수요자가 누구나 쉽게 인공지능 솔루션을 활용해서 새로운 가치 및 시장을 창출하고, 공급자는 수요에 맞추어 기술 고도화 및 다양성을 갖추는 선순환 구조를 갖추어야 한다. 특히, 공공분야에서 수집 가능한 사회 전반의 빅데이터를 인공지능 플랫폼과 유기적으로 결합하여, 경제적, 사회적 인 효과가 공공을 넘어 민간으로 확대될 수 있는 방안 마련이 필요하다.

이러한 인공지능 플랫폼의 기대 효과는 플랫폼 기반의 사업 구조 혁신 사례를 통해서도 예상할 수 있다. 구글, 아마존, 애플, 페이스북, 우버, 에어비엔비 등 글로벌 시장을 장악하고 선도하고 있는 기업들 대부분이 기존 시장의 일방향적 비즈니스 형태를 플랫폼을 활용하여 다방향적 플랫폼 비즈니스 모델로 바꾸는 파괴적 혁신을 통해 급성장하였다. 2019년 글로벌 시가 총액 상위 10개 기업 중 6개가 마이크로소프트, 애플 등 플랫폼 기업이었으며(그림 1), 이는 기업의 미래 성장 전략이 과거의 외형적 규모 확대에서 플랫폼 기반의 사업모델로 전환되는 추세임을 나타낸다⁴⁾.

현재 인공지능 기술에 대한 민간⁵⁾, 공공⁶⁾ 분야에서의 수요를 고려하였을 때, 인공지능 플랫폼 구축은 비용 절감, 데이터 보안 및 공유, 인공지능 기술 개발 및 제품화의 편의성 등 다양한 장점을 이끌어 낼 수 있으며⁷⁾, 기술 배포 및 활용 단계에서의 공간적·시간적 제약을 축소하여, 공급자와 소비자 모두 경제적 혜택을 줄 수 있을 것이다. 특히, 공급자와 수요자가 양

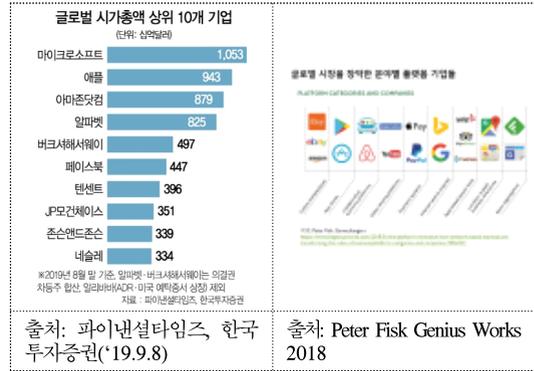


그림 1. 글로벌 시가총액 상위기업 및 글로벌 시장 분야별 플랫폼 기업
Fig. 1. Top Companies in Global Market Cap and Platform Companies by Global Market Segment

방향으로 교류하는 플랫폼 생태계를 형성하여, 공급자와 소비자 모두에게 탐색과 신뢰의 비용을 줄여주는 경제적 혜택을 줄 수 있다.

본 논문은 체계적 논문 분석과 국내외 인공지능 플랫폼에 대한 분석을 통하여, 현재 인공지능 플랫폼의 특성과 한계점 등을 파악하고, 공공 분야에서의 인공지능 플랫폼 적용을 위한 다양한 조건과 이슈를 제안한다. 본 논문에서 2장은 인공지능 플랫폼을 인공지능 학습데이터, 인공지능 알고리즘, 컴퓨팅 인프라의 3가지 핵심 구성 요소로 정의하여 설명하고, 3장은 인공지능 플랫폼 분석 연구를 위한 논문 검색 및 자료 수집 방법과 분석 결과를 소개한다. 4장은 국내외 민간 및 공공의 인공지능 플랫폼 현황 및 분석 결과를 소개하고, 5장과 6장을 통하여 공공분야 인공지능 플랫폼 구현을 위해 고려할 주요 사항들을 소개하며 논문을 마무리한다.

II. 인공지능 플랫폼 정의

인공지능 플랫폼은 ICT 관점에서 컴퓨터 운영체제나 어떤 소프트웨어가 제공하는 공간 위에서 수요, 공급자가 상호 간에 인공지능 관련 서비스, 정보 등을 교류할 수 있는 장이라고 정의할 수 있다⁸⁾. 인공지능 플랫폼은 학습 데이터를 제공하는 역할 뿐만 아니라, 수요자가 원하는 기술과 서비스를 개발하기 위한 S/W 및 H/W 솔루션을 함께 제공한다. 수요 기술의 종류와 수준이 다양하지만, 대규모 데이터에 기반하여 의사 결정을 하는 인공지능 알고리즘 개발과 개발 환경 제공은 공통 사항이다.

플랫폼에 따라 기술 개발 시간을 단축할 수 있는

시각적 저작 도구 등을 제공하며, S/W SDK (Software Development Kit) 형태로 인공지능 기술 및 서비스 개발을 할 수 있도록 한다. 또한 이미지 인식, 자연어 처리, 음성인식 등 사전 학습된 고차원 기계학습 모듈을 제공하여, 서비스 개발에 집중할 수 있는 하이엔드 기술 솔루션을 포함하기도 한다. 다양한 소비자 수요에 대응하여, 글로벌 기업들은 다각적 인공지능 플랫폼 서비스를 운영하고 있으며, 수요자와의 상호 작용을 위한 실시간 피드백 플랫폼도 운영한다⁹⁾.

Amazon에서 운영 중인 플랫폼 AWS의 경우 H/W, S/W 솔루션으로 인공지능의 서비스를 개발하기 위해 AWS Inferentia 추론칩과 EC2 instance 등을 제공하여 데이터 기반 머신러닝 프레임워크를 제공한다 (그림2). 개발 환경으로는 세미메이커 스튜디오를 통해 데이터 전처리, 알고리즘 구축, 모델 학습 보정 등의 ML 모델개발이 가능하도록 자동화 기능을 제공한다. 또한 음성인식, 검색, 보안, 개발 등 고차원의 신규 인공지능 모듈 서비스를 제공하는 등 클라우드 기반의 AI 플랫폼을 구축하여 운영 중이다¹⁰⁾.

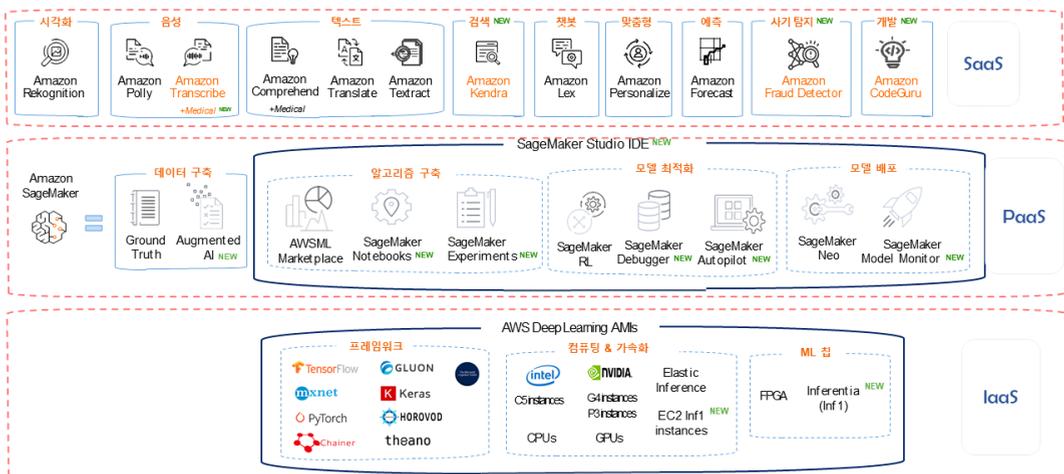
2.1 인공지능 플랫폼의 구성 요소

과거 인터넷 기반에서 2008년 스마트폰 등장 이후 개인에게 모바일 디바이스가 확산되면서, 네트워크를 통한 데이터 저장 능력의 확장에 따라 기하급수적으로 데이터가 증가하였고, 더불어 병렬처리 H/W와 클라우드 플랫폼 발전에 따른 컴퓨팅 자원의 증가, 딥러닝으로 대표되는 기계학습 분야의 혁신이 인공지능 알고리즘 기술의 보편화, 고도화의 계기가 되면서 데

이터, 알고리즘, 컴퓨팅 자원은 인공지능 플랫폼을 구성하는 기본 요소가 되었다.

(데이터) 인공지능 성능의 가장 중요한 요소는 학습 데이터이다. 학습데이터는 인공지능이 알고리즘에 따라 각각의 특징, 패턴 등을 인식할 수 있도록 전처리된 데이터로써, 공공기관이나 연구기관이 데이터베이스 형태로 보유하고 있는 공공데이터나 연구데이터로 구분할 수 있다. 인공지능의 정확성은 많은 데이터를 읽고 훈련하는 과정에서 확보되는 것이기 때문에 기계가 학습할 수 있는 형태로 가공한 인공지능 학습 데이터 확보가 필요하다.

인공지능 사회가 도래함에 따라 기존의 구조적이고 정형화된 데이터뿐만 아니라 비정형적인 방대한 민간 정보나 소셜 데이터가 급증하고 있어 먼저 데이터를 표준화된 체계로 저장 및 분석할 수 있는 빅데이터 기반 환경이 필수적으로 요구된다. 이러한 데이터 환경 변화에 맞춘 인공지능 서비스를 제공하기 위한 데이터의 유형, API 제공방식 등을 고려한 표준아키텍처가 뒷받침하고¹¹⁾, 인공지능 데이터 관점에서 다양한 입력 데이터를 분석·예측할 수 있는 인공지능 모델을 구축할 수 있도록 인공지능 플랫폼을 구성해야 한다. 더불어 데이터가 가지고 있는 속성값 자체가 인공지능 학습에 중요한 자원이 될 수 있기 때문에, 풍부한 데이터 확보 및 메타데이터를 기반으로 인공지능의 데이터 확장성을 지원할 수 있는 체계가 필요하다. 데이터를 통합하기 위해 RDF의 유연성과 메타데이터 기능을 SQL 데이터베이스의 일반적인 구조와 쿼리 메커니즘과 결합하는 방식도 대안으로 고려할 수 있



출처: NIA Special Report, AWS Re:Invent 2019 인공지능 플랫폼의 현재와 미래, 2019

그림 2. AWS 클라우드 기반 AI 및 ML 서비스 모델 플랫폼
Fig. 2. AWS AI and ML Service Model Platform based Cloud

다¹²⁾.

많은 학습 데이터가 대부분 외부의 스토리지(Luster, NAS, HDFS, etc.) 디바이스에 존재하는데 이와 같은 외부의 많은 데이터를 활용하기 위해서는 점차 대용량 학습 데이터를 사용하여 딥러닝 모델을 개발하고 플랫폼 기능을 고도화할 것이다. Amazon의 AWS(Amazon Web Service), MicroSoft의 Azure, Google의 GCP(Google Cloud Platform), IBM의 Bluemix Service 등 글로벌 인공지능 선도기업들은 데이터 및 컴퓨팅 클라우드에 기반한 인공지능 솔루션 서비스에 앞장서고 있다¹³⁾.

(알고리즘) 대용량 원시 데이터와 전처리를 통한 학습용 데이터를 구축한 뒤, 구축 데이터를 기반으로 각종 인공지능 기술을 개발할 수 있는 SW 솔루션의 지원 또한 인공지능 플랫폼의 중요한 요소이다. TensorFlow, Keras, PyTorch, Caffe2 등 다양한 오픈소스 딥러닝 라이브러리들의 개발이 활발히 진행되고 있으며, 이에 힘입어 과거에 비해 인공지능 기술의 개발 속도 및 비용은 크게 감소하였다¹⁴⁾. 이러한 오픈소스 딥러닝 라이브러리는 Google, Facebook, Microsoft 등 주요 IT 민간 기업과 오픈소스 개발자 커뮤니티를 중심으로 개발되고 있으며, 오픈소스 SW에 대한 투자 및 공헌은 민간 기업의 인공지능 기술의 우수성을 나타내는 대표적 척도로 활용되고 있다¹⁵⁾. Microsoft사가 최대 오픈소스 플랫폼인 GitHub를 거액에 인수한 배경도 SW 개발자 간에 자발적으로 오픈소스를 업데이트하며 폭발적으로 성장하는 인공지능 플랫폼 알고리즘의 잠재력을 높이 평가한 것이다.

공공 분야에서 인공지능 알고리즘에 대한 지원은 자국 기업에 대한 지원과 국가적 인공지능 기술 보유 측면에서 이루어 지고 있다. 현재 한국 정부는 인공지능 알고리즘 지원을 위해 언어처리, 음성지능, 시각지능, 대화처리 등을 위한 오픈 API를 제공하고 있으며¹⁶⁾, 유럽¹⁷⁾ 및 미국¹⁸⁾ 등 선진국에서도 정부 차원에서 민간 기업과 연계한 인공지능 솔루션 구축에 앞장서고 있다. 인공지능 플랫폼에서 인공지능 알고리즘은 인공지능 기술이 구현될 H/W 플랫폼에 따라 다양한 형태를 띠게 된다. 최근 지능형 임베디드 기기, 자율형 이동기기(드론, 스마트카 등), 지능형 IoT(엣지기기) 등 다양한 H/W 플랫폼에서의 인공지능 SW가 중요한 기술로써 부상함에 따라, 임베디드 소프트웨어 환경을 고려한 인공지능 알고리즘 지원체계 구축도 인공지능 플랫폼 구축 시 고려되어야 한다.

(컴퓨팅 인프라) 다양한 인공지능 기술을 원활히 개발하기 위해서는 GPU 기반의 대용량 컴퓨팅 자원

이 필요하다. 인공지능 학습 과정에서 대규모 데이터를 처리하고, 수 백만개 이상의 모델 파라미터에 대한 업데이트를 CPU로 처리하는데 한계가 있기 때문이다. 특히 다양한 요소기술이 필요한 인공지능의 특성을 고려할 때 여러 분야에서 다양한 요소 기술력을 보유하고 있으나 비용의 문제로 인공지능 개발에 제약을 받고 있는 중소·벤처 기업을 위한 고가의 GPU 컴퓨팅 자원의 공유가 중요하다. 이를 위해서는 향후 컴퓨팅 자원의 확장성을 고려하여 클라우드 기반의 인공지능 플랫폼을 구축하는 것이 효율적일 것이다. 글로벌 인공지능 선도기업들은 인공지능 기술 개발과 서비스가 손쉽게 구현할 수 있도록 클라우드 AIaaS(AI as a service) 기반의 인공지능 플랫폼 고도화에 역량을 집중하고 있는 추세이다.

아울러 인공지능 기술이 급격히 발전함에 따라 인공지능 처리기술은 점차 자율주행 자동차, 드론, 인공지능 반도체 등 실시간으로 직접 연산이 필요한 환경으로 확대되는 만큼 기존 중앙집중형 클라우드 기반에서 사용자 주변에서 데이터를 분산 처리할 수 있는 엣지 컴퓨팅으로의 확장된 컴퓨팅 인프라 구축 방향에 대한 고려도 필요하다¹⁹⁾. 이는 실시간 대응이 필요한 인공지능 분야의 핵심 기술로서 5G 엣지 기반으로 적용하면 초기 인프라 구축에 많은 비용과 시간이 소요되는 기존의 컴퓨터 버전의 머신러닝 분야에 효율성을 높일 수 있다.

2.2 인공지능 플랫폼의 기대 효과

딥러닝으로 대표되는 인공지능 기술의 발전은 음성 및 영상인식, 자연어 처리, 시계열 예측 등 주요 응용 분야에서 큰 진전을 보이고 있으며, 새로운 비즈니스를 창출하고 있다. 이러한 인공지능 기술은 인공지능 플랫폼을 기반으로 미래의 진화에 핵심적인 역할을 할 것으로 기대한다. 특히, 인공지능 플랫폼은 유통업, 제조업 등 기존 시장의 전통적 환경을 혁신하는 도구로써 사용되고 있으며, 이에 힘입어 시장의 규모가 점차 커지는 현상까지 나타나고 있다. 이와 같이 4차 산업 혁명 시대에서 인공지능 플랫폼은 가장 대표적인 키워드로써 앞으로 산업과 시장을 견인할 핵심 요소로 자리매김 할 것으로 기대한다.

III. 연구방법

현재 인공지능 플랫폼의 분석 연구를 위하여, 먼저 데이터, AI, 플랫폼 분야의 학술논문 총 6,578건을 수집하였다. 논문 수집을 위한 검색 주제어들은

“Data-based AI platform” 와 공공분야 (Public Sector) 개념을 반영할 수 있도록 선정하였고, IEEE Library의 최근 5년간 Journal/Magazine으로 한정하여 조사하였다. 수집 학술 논문의 키워드별 분포 결과는 그림 3에 나타내었다. 조사결과에서 알 수 있듯이 ‘데이터 플랫폼(6,337건)’으로 연구된 논문이 ‘인공지능 플랫폼(241건)’보다 월등히 많은 것을 알 수 있다.

241건의 인공지능 플랫폼 논문 주제를 분석해 보면 대부분 HW(6건), SW기술(220건)에 대한 논문이 주를 이루고 있으며, 주제를 분석해 보면 GPU, FPGA 등 HW 기술 및 네트워크, 게임, 블록체인 등에 적용하는 SW를 각각 다루고 있다. 인공지능 플랫폼 관련 논문(15건)도 있지만 엣지, 모바일 엣지 등 클라우드 기능을 다루고 있어 공공분야에 인공지능 플랫폼을 구축하기 위해 참조할 정책적 논문은 부족하다는 것을 알 수 있다. 이는 데이터 플랫폼에 비해 인공지능 플랫폼에 대한 연구는 미미한 상태임을 의미한다. 데이터를 기반으로 한 인공지능 플랫폼에 대해 연구가 절실함에 따라 본 논문에서는 공공분야에 적용하기 위한 인공지능 플랫폼을 선별하여 분석하였다.

더불어 현재 인공지능 기술 및 민간-공공 분야 플랫폼에 대한 정보를 제안하는 정보 사이트를 이용하여, 자료 조사를 시행하였다. 표 1은 본 논문에서 활용한 플랫폼 정보 사이트의 내용을 요약한 것이다. 정보 사이트 중 PAT Research는 18개의 인공지능 플랫폼들을 구독자들의 평점에 기반하여 선정하고, 개념과 장점, 도입 시의 장점과 한계, 사용성, 제공되는 기술과 성능 등을 비교해 볼 수 있는 웹사이트를 운영하고 있고 60 Second Marketer는 157개의 인공지능 플랫폼들을 소셜미디어, 챗봇, CRM, 마케팅, SEO, 웹사이트 디자인 등 17개 분야로 분류하여 소개하고 있다. 또한 G2 Crowd는 기계학습 기술 개발 용이성, 학습

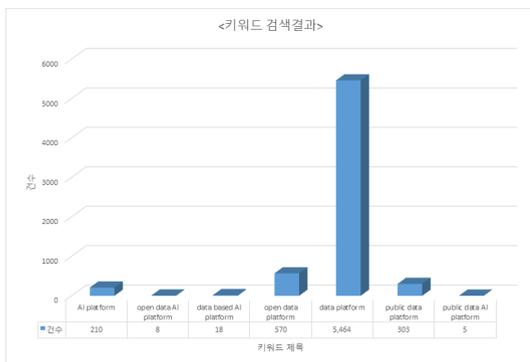


그림 3. A.A.A.L.I, 데이터, 플랫폼 키워드 검색결과
Fig. 3. AI, Data, Platform Keyword Searching Results

표 1. 인공지능 플랫폼 분석 사이트 비교
Table 1. AI platform analysis site comparison

구분	서비스 내용
PAT Research	- 18개의 인공지능 플랫폼(마이크로소프트의 Azure ML, 구글의 Cloud Prediction API와 Tensorflow, InfoSys의 HOLMES 등) - 15개 데이터분석 소프트웨어(Sisense, Periscope data 등)에 대한 구독자들의 평점으로 선정 ※ 서비스 개념과 장·단점, 제공되는 기술, 성능 등을 비교
60 Second Marketer	- 157개의 인공지능 플랫폼들을 소셜미디어, 챗봇, CRM, 마케팅, SEO, 웹사이트 디자인 등 17개 분야로 분류하여 소개 ※ Botify, Microsoft의 Lui API(이상 챗봇), Emotient(얼굴 감성 인식), IBM Watson(범용 인공지능 플랫폼으로 분류) 등을 포함
G2Crowd	- 기계학습 모델 설계의 용이성, 제공되는 기계학습 라이브러리, 학습 방법, 비전 및 자연어 처리 등 8개 분야로 26개의 인공지능 플랫폼 비교

출처: AT Research, 60 Second Marketer, G2 Crowd 각 사의 홈페이지 참조(2020. 2. 2기준)

방법, 비전 및 자연어 처리 응용 솔루션 등 8개 분야로 26개 인공지능 S/W 플랫폼을 비교하는 서비스를 제공하고 있다²⁰⁾.

IV. 인공지능 플랫폼 현황

4.1 국외 현황

글로벌 인공지능 리딩기업들은 개방형 플랫폼을 구축하여 개발자와 사용자가 손쉽게 활용하고 원하는 바를 얻을 수 있는 공간이 되도록 노력을 아끼지 않고 있다. 플랫폼에서 관리의 편의성, 사용의 용이성, 비즈니스 용이성, 지원품질 등에 만족도가 높은 플랫폼일수록 플랫폼의 사용자가 확대되면서 더욱 경쟁력 있는 인공지능 플랫폼으로 성장하기 때문에, 최근에는 인공지능 플랫폼의 만족도 및 설치·지원 항목으로 나누는 사용자 조사를 통해 플랫폼을 비교한 결과를 공개함으로써 사용자가 최상의 제품을 파악할 수 있도록 운영하고 있다. 이러한 환경은 플랫폼 간에 경쟁을 유도하여 인공지능 선도기업들의 서비스가 더욱 강화되도록 할 것이다. 이는 인공지능 기술이 기업의 개발자가 단방향으로 공급하는 형태로는 인공지능을 선도할 수 없다는 특징이 반영된 변화라고 보여진다.

표 2에서 정리한 바와 같이 세계 주요 ICT 기업들은 이미 인공지능 플랫폼을 기반으로 생태계를 만들어 가기 시작했다. 구글은 빅데이터와 알고리즘 역량에 기반해서 다양한 분야에 활용 가능한 범용 인공지능 플랫폼을 지향하고 있고 페이스북은 개별 사용자의 성향을 분석하여 맞춤형 서비스를 제공을 추구하

표 2. 국외 인공지능 플랫폼의 비교
Table 2. Comparison of overseas artificial intelligence platforms

구분	아마존	구글	MS	IBM
데이터 구축	외부 연계를 통한 수집데이터 분석기능(IoT 분석 솔루션)	데이터셋 서치 - 약 600백만개 표를 인덱스	외부 연계를 통한 수집데이터 분석 지원(Azure의 데이터팩토리) 가시화 인터페이스 제공	외부정보와 연계를 통한 수집데이터 분석기능 - IA (Information Architecture)
알고리즘 (라이브러리)	Regression, Binary, Multiclass classification 등	이미지 인식, 비디오/텍스트 분석, 번역, Tensorflow 등	Decision tree/ Bayesian recommendation/ DNN/regression clustering	IBM's data Science Experience Workbench (다양한 API 제공)
기계학습 알고리즘 자동추천	지원	지원	지원 안함	지원 안함
컴퓨팅 파워	GPU	TPU 개발(Tensorflow 통해 딥러닝 특화)	FPGA Brainwave(딥러닝 가속 플랫폼)	GPU
데이터 저장	AWS	구글 클라우드	Azure	IBM Bluemix
대화형 스피커(CAP)	Alexa	Assistant	Cortana	Watson
집중분야	빅데이터 연계, 고객의 다양한 상품구매 연계	대부분의 분야에 활용 가능한 AI 서비스 개발(데이터+AI 기반)	모바일 퍼스트, 클라우드 퍼스트	Watson으로 빅데이터+AI 사업 강화
주요 서비스	세이지 메이커 (ML개발 지원)	텐서플로우 엔터프라이즈 (ML개발도구 제공, AI확장 지원)	애저 머신러닝(ML개발, 오픈소스 프레임워크 지원)	왓슨스튜디오(ML개발 지원 API지원)

고 있다. 아마존은 실생활에 인공지능 플랫폼을 적용하여 소비자의 일상을 혁신, IBM은 의료분야에서 전문의 역할을 지원, GE는 항공, 에너지, 제조 등의 영역에서 각각 인공지능 플랫폼을 활용하여 4차 산업혁명을 주도하고자 한다²¹⁾.

인공지능 플랫폼은 기본적으로 시각지능, 언어지능 서비스 및 분석툴을 제공하며 다양한 가시화 기술, 사람 친화적인 인터페이스를 제공하는데 인공지능 개발자는 이러한 기능을 활용하여 데이터분석 기술과 도메인 지식 없이 다양한 분석 및 제어 서비스 개발이 가능하다는 아마존, 구글, MS, IBM 등에서 제공하고 있다. 또한 클라우드 기반의 서비스가 많고 물리적인 자원관리 또한 클라우드 상에서 구현되고 있다는 것을 보여주고 있다.

최근 주목을 받는 기술분야로서 Augmented Analytics는 데이터 준비, 분석, 해석의 전반에 걸친 분석 라이프사이클에 대해서 기계학습 및 자연어처리 기능을 활용하여 사용자에게 대한 조력 기능을 제공하고 있다. 또한 새로운 패러다임인 대화형 플랫폼 (CAP, Conversational AI Platform)으로서 딥러닝을 기반으로 한 언어지능 기술의 성숙 및 저변화로 많은 IT 서비스들의 인터페이스가 자연어로 대체될 것이다.

4.2 국내 현황

국내 인공지능 플랫폼 구축은 아직 초기단계이고

민간, 공공분야에서 각각 다양한 노력을 하고 있는 실정이다. 최근 4차 산업혁명에 대한 정부 지원사업 및 조직들이 설립되고 국가 차원의 로드맵이 수립되면서 민간분야에서는 스타트업 중심으로 산업 특화형 플랫폼들을 구축하는 사례들이 지속적으로 증가하고 있는 추세이다.

솔트룩스는 인공지능 플랫폼인 아담(ADAMS)은 자연어 이해, 음성인식, 기계번역 등 10개의 엔진을 제공하는 AI Suite을 기반으로 인공지능 기술을 융합하여 지능형 소프트웨어 솔루션과 서비스를 통합하는 인공지능 플랫폼 서비스를 제공하며²²⁾, 아크릴은 2018년에 인공지능 플랫폼 조나단(Jonathan)을 발표 이후, 현재 고성능 딥러닝 알고리즘 라이브러리를 제공하며 텍스트의 언어, 억양, 문맥 이해, 발화패턴 인식·답변 등 대화지능 및 표정, 언어, 음성의 감성을 이해하는 시청각 지능을 제공한다²³⁾.

한편, 특정 산업분야에 특화된 인공지능 플랫폼들은 다음과 같이 스타트업 중심으로 추진되고 있다고 볼 수 있다. 루닛은 의료영상물 통한 진단과 치료를 돕는 인공지능 기업으로서 CB Insights에서 세계 최초로 발표한 100대 AI 기업 랭킹인 ‘The AI 100’에 국내 기업으로 유일하게 선정된 바 있고, 2019년에는 세계 디지털 헬스 기업 목록인 ‘Digital Health 150’에도 포함되는 등 기술력을 입증 받았다²⁴⁾. 뷰노는 딥러

닝 기술 기반의 국내 최초의 AI 기반 골연령 진단 소프트웨어 개발 및 퇴행성 뇌질환 지원 등 다양한 의료 분야에 인공지능 솔루션을 제공한다²⁵⁾.

또한 국내 대기업들 역시 자사의 비즈니스 분야에서의 강점을 살려 시너지를 극대화할 수 있는 서비스 개발을 끊임없이 추진하고 있다. 삼성SDS의 데이터를 수집, 처리하여 클라우드 인공지능 기반으로 데이터를 빠르게 분석하고, 이해하기 쉽게 시각화해 주는 통합 AI 플랫폼인 ‘Bright AI’²⁶⁾, LG CNS의 인공지능 빅데이터 플랫폼인 ‘DAP’는 빅데이터 분석기능에 중점을 둔 기존 버전(DAP 1.0)에 머신러닝·딥러닝 기반 AI 분석 기능을 확대시켜 데이터 수집, 분석, 시각화가 즉시 가능하며 클라우드 기반의 인공지능 신규 서비스 개발 환경을 제공하는 플랫폼 DAP 2.0을 제공한다²⁷⁾.

SK 텔레콤은 Android OS, iOS, Linux OS의 다양한 운영체제를 지원하며, 음성 인식, 음성 합성, 자연어 이해 등을 기반으로 대화 기반 인공지능 서비스를 개발할 수 있는 ‘NUGU’ 인공지능 플랫폼을 서비스하고 있다. KT는 기가지니 인사인드는 음성인식, 음성합성, 자연어 처리 등 인공지능 기반기술을 제공해 파트너사가 KT의 축적된 인공지능 기술과 서비스를 바로 이용할 수 있도록 했다. 클라우드 인공지능 플랫폼을 파트너사에게 공개하고 안드로이드, 리눅스, 윈도우 등 다양한 단말 운영체제에 탑재 가능한 S/W SDK를 다양하게 제공한다²⁸⁾.

이와 같이 국내의 인공지능 플랫폼은 클라우드 기반의 플랫폼 Bright AI, DAP, NUGU, Giga Genie, 산업특화 플랫폼인 루닛, 뷰노 및 공통 서비스 제공 플랫폼인 ADAMs, 조나단 등 다양한 형태로 개발이 진행되고 있어 서비스 도입을 검토중인 기업들은 자사의 비즈니스 기회에 맞는 플랫폼 선택을 위한 신중한 검토가 필요하다.

4.3 공공분야 인공지능 플랫폼 현황

(해외 정부 인공지능 플랫폼 운영현황) 해외 정부에서는 데이터 수집 및 분석능력의 중요성을 인식하여 공공데이터 활용을 위한 포털운영 등의 노력을 기울이고 있으나 인공지능 기술개발을 위해 학습데이터, 알고리즘 및 컴퓨팅 자원을 모두 지원하는 인공지능 중심의 공공 플랫폼 모델은 흔하지 않다. 다만 데이터를 기반으로 클라우드 컴퓨팅 환경을 구축하여 개방형 플랫폼의 강점을 활용하는 추세이다. 아래 해외 스마트도시 공공 플랫폼에서 알 수 있듯이 인공지능 적용으로 확대 가능하도록 프라이빗 또는 퍼블릭 클라

표 3. 해외 데이터 기반 스마트시티 공공플랫폼 현황
Table 3. Status of overseas data-based smart city public platforms

구분	플랫폼 방식	적용 서비스	데이터 유형	데이터 개방유무
중국	On-premise (알리바바)	교통문제, 모바일 결제 등	차량경로, 센서, 카메라, 지도, 통신사	시티브레인 플랫폼 개방
코펜 하겐	MS Azure 플랫폼	유동인구 분석, 쓰레기 처리, LED조명, 스마트 교통문제 등	교통, 주차, 센서, 금융거래, 전력데이터, 통신, 소셜미디어 등	시민 개방
영국	13개 기관 컨소시엄 데이터 허브	에너지, 수자원, 교통문제 해결	도시데이터, SNS데이터, 클라우드 소싱	시민 개방(데이터 소유자 보상)
스페인	유럽 9개국, 호주 19개 기관 협업 플랫폼	교통, 강우, 인구이동, 주차장, 장소 AR 등	도심지역에 센서를 통한 시민데이터 소싱	시민 참여형 개방
EU	다양한 도메인 통합 데이터 허브	교통, 통신, 에너지, 주차, 기후 등	도시 방문객, 시민 활동 데이터, 도메인별 서버 수집	개발자 개방
네덜란드	민간-공공 9개기관 협력 플랫폼	식물관리, 태양광, 에너지, 교통, 기후, 배송 등	암스테르담 도시관리 데이터 수집	시민 개방
미국	AWS 일부 활용한 통합데이터 플랫폼	신호, 키오스크, 주차, RFID, 커넥티드 교통네트워크	통신, 모바일, 교통 등 데이터 소싱	민간, 공공 개방

우드를 기반으로 관계기관 간의 통합 플랫폼을 구축하면서 여러 경로를 통해 수집된 데이터를 일반 시민에 개방하는 과정을 통해 서비스를 고도화하고 있다는 공통점을 찾을 수 있다.

인공지능을 구현할 수 있는 다양한 알고리즘까지 탑재한 인공지능 플랫폼은 공공플랫폼이 뒷받침되면 보다 수월하게 인공지능을 진화시킬 수 있는 기반이 되기 때문에 인공지능 플랫폼의 핵심요소(데이터, 알고리즘, 컴퓨팅)와의 연계를 고려해서 공공플랫폼을 설계하면 보다 효율적으로 플랫폼을 구축할 수 있다.

현재는 대부분의 국가에서는 공공 데이터 기반의 공공플랫폼을 구축하는 경우가 많으나 독일, 중국은 인공지능이 미래사회의 핵심기술이 될 것으로 판단하여 정부주도 인공지능 플랫폼을 기반으로 연관산업의 동반성장을 추진하는 전략을 추진하고 있는 한편²⁹⁾, 클라우드 플랫폼 구축에 있어서는 독일, 프랑스가 선도적으로 Gaia-X 프로젝트 기획하여 유럽의 디지털 주권, 데이터 이용, 수집, 공유 생태계 형성을 추진하는 등 인공지능 플랫폼으로의 확산을 염두한 국가 인공지능 플랫폼 전략을 추진하고 있다³⁰⁾.

(국내 정부 인공지능 플랫폼 운영현황) 정부에서는 R&D 과제를 중심으로 운영중인데 과학기술정보통신부의 엑소브레인(Exobrain), 딥뷰(Deep view), 산업통상자원부의 산업특화 개방형 AI 클라우드 개발 등을 시작으로 현재는 지능정보산업인프라조성 사업의 일환으로 인공지능 학습데이터를 구축 및 관련 알고리즘, 컴퓨팅 자원을 제공하는 인공지능 플랫폼을 구축하여 운영 중이다.

현재 과학기술정보통신부가 AI 오픈 이노베이션 허브(www.aihub.or.kr)를 구축하여 CCTV영상, 복합영상 등 인공지능 학습용 데이터셋을 국내 인공지능 관련 중소·벤처기업의 경쟁력을 높이기 위하여 제공하고 있다. 현재 이상행동 CCTV영상, 한국인 안면이미지, 법률, 특허 등 AI 학습용 데이터를 21종 4,650만 건을 제공('19년 기준) 중이며, 2022년까지 한국어 음성, 글자체 이미지 등 범용데이터를 비롯하여 엑스레이 영상, 자율주행자동차영상 등 특정 분야 데이터 등 75종, 2억 건을 구축·개방 예정이다^{31,32)}. 이는 한국 정부에서 인공지능 고도화를 위한 3대 인프라로서 ① 데이터(AI 학습용 데이터), ② 알고리즘(Open API(14종) 지원), ③ 컴퓨팅 자원(중소벤처 대상 고성능 GPU 컴퓨팅 자원)을 서비스 중이며 지속적으로 구축범위를 확대하고 있다.

그러나 인공지능을 구현할 수 있는 온전한 플랫폼이 되기 위해서는 개선해야 할 부분도 있다. 정부 주도로 인공지능 신기술이 개발되기 보다는 그간 국내·외에서 개발된 오픈소스 기반의 인공지능 기술 및 데이터의 교류가 활성화되도록 인센티브 체계가 필요해 보이며 향후 민·관·학 등에서 상호 연계를 강화하기 위해 데이터와 인공지능 기술을 연계하는 표준화 모듈, 기술 평가모듈, DL/ML 라이브러리 연계 모듈 등 인공지능 기술 검증 도구도 제공이 되어야 한다.

V. 공공 인공지능 플랫폼 구현을 위한 제언

인공지능 플랫폼은 파급력에 있어 엄청난 잠재력을 가지고 있으나 현재 민간·공공 분야를 막론하고 각기 보유하고 있는 데이터의 형태, 서비스 방향 등에 따라 다양한 형태로 플랫폼을 구축하고 있는 실정이다. 이는 고도의 기술이 집약되어 인간의 지능 수준으로 서비스를 구현하려는 인공지능 기술 개발에는 한계로 작용하게 될 것이다. 왜냐하면 4차 산업혁명 시대의 중심이 되는 인공지능 기술은 어느 특정 기술을 고도화하여 구현되는 것이 아니고 다양한 요소기술이 결합될 때 진정한 인공지능 기능이 작동될 수 있기 때문

이다³³⁾.

인공지능은 특정한 범위의 자동화를 목적으로 개발하는 것을 넘어서, 사람의 지능과 흡사한 인지와 판단을 할 수 있도록 다양한 데이터와 기술을 종합하여 고도화하는 방향으로 발전하고 있으며, 이에 따라 기술 분야 간의 결합 뿐만 아니라 민간, 공공분야에서의 협력을 기반으로 한 플랫폼 구축이 이를 가속화할 것으로 기대한다. 이를 위해 다양하게 구축되고 있는 플랫폼의 핵심적인 구성 요소를 도출하여 플랫폼을 구축하는 시작 단계에서부터 향후 플랫폼 간의 연계를 고려하여 플랫폼을 구축할 수 있는 거시적 관점에서의 플랫폼 구축전략이 공유되고 설계되어야 한다. 이를 위해 민·관·학 주도로 플랫폼 구축을 전담하는 협력체계를 만들어야 한다.

VI. 결 론

인공지능 플랫폼은 컴퓨팅 리소스를 충분히 보유한 H/W 플랫폼을 기반으로 수요자와 공급자 간에 인공지능 관련 서비스, 정보 등을 교류할 수 있는 장으로 정의할 수 있으며, 본 논문을 통해 다양한 분야의 데이터와 기술이 응집되고 지속적으로 개발되고 있는 다양한 인공지능 플랫폼과 비교하며 체계적 데이터 수집·관리, 기술 개발, 컴퓨팅 인프라 수준 등의 방향을 살펴보았다. 특히, 우리나라의 인공지능 기술력이나 자원의 활용 가능성 등을 고려할 때 역량을 집중시킬 수 있는 분야에 선택과 집중을 통해 단계별로 인공지능 경쟁력을 높이는 전략이 주요할 것으로 보인다.

먼저, 인공지능 기술에 필수적인 데이터 측면에서 살펴보면 한국은 정부 주도로 공공 데이터를 구축하고 개방하는 체계는 세계적인 수준인 만큼 데이터 인프라의 경쟁력은 갖추 수 있을 것이다. 정부에서 국내 AI스타트업, AI전문기관을 대상으로 실시한 설문('19.5)에 따르면 인공지능 기술격차 해소를 위해 필수적으로 요구되는 AI 데이터 확보는 많은 시간, 비용 등이 요구됨에 따라 기업별로 비즈니스 목적에 따라 개별 구축하는 환경에서는 데이터 확보에 한계가 있다. 따라서 기후, 환경, 교통, 복지 등 공공성이 높고 사회적 수요는 높으나 법제도의 제약으로 민간에서 추진하기 어려운 분야는 정부가 주도하여 인공지능을 위한 공공데이터를 구축하고 공개하는 것이 효율적이다.³⁴⁾ 이를 통해 공공데이터를 활용해 사회문제 해결 및 민간기업의 활용도를 높일 수 있는 영역을 중심으로 경제적 효과를 극대화할 수 있도록 역량도 집중할 수 있다.

그러나 공공 데이터만으로 인공지능 기술 개발에 활용할 수 있는 데이터 기반이 완성될 수는 없다. 공공 데이터를 기계가 학습할 수 있는 형태로 전처리하고 민간 기업에서 구축하고 있는 인공지능 데이터와의 연계가 필요하다. 인공지능을 개발하기 위해 민간 분야에서도 유사한 인공지능 데이터를 구축하기 위해 막대한 예산과 연구개발 인력을 투입하고 있는 상황을 고려할 때, 대승적 차원에서 민간합동 추진체계를 마련하는 것이 시급하다. 과기정통부와 행안부에서 각각 구축하고 있는 공공영역의 AI 학습 데이터와 공공 데이터를 기계학습이 가능한 수준으로 데이터를 구축할 수 있도록 추진체계 일원화가 필요하다. 또한 민간 기업과 학계에서도 각각 비즈니스와 연구목적으로 구축하는 AI 학습 데이터와 연구 데이터의 구축목록, 범위, 우선순위 등을 상호 공유하여 중복구축을 줄일 수 있는 민간합동 컨트롤 타워를 마련한다면 민간영역의 다양한 도메인 전문가를 공공의 인공지능 학습데이터 구축과 연계할 수 있는 협력체계도 가능할 것이다.

두 번째는 알고리즘이다. 인공지능 기술은 기계학습, 딥러닝 기술의 특징과 같이 개방성에 있다. 구글이 알고리즘을 구현할 수 있는 텐서플로우를 공개하는 이유는 보다 많은 사용자를 확보할수록 세계시장과 기반기술을 확보할 수 있기 때문이다. 이것이 앞서 강조한 플랫폼의 폭발적인 잠재력을 참여자 간에 상호작용을 통한 '네트워크 효과'를 이끌어 내는 초석이 된다.

세 번째는 플랫폼 공간이 활성화되면서 인공지능 기술을 제약 없이 구현하고 교류할 수 있도록 GPU 기반의 대용량 컴퓨팅 인프라를 제공해야 한다. 한국 정부에서 중소벤처기업을 대상으로 클라우드 기반의 GPU 컴퓨팅 환경을 제공하고 있는데, 다양한 인공지능 요소기술 개발을 위해 이러한 서비스를 점차 확대할 필요가 있다.

인공지능 플랫폼 구성의 핵심인 데이터, 알고리즘, 컴퓨팅 파워를 확보하고 발전시킬 수 있는 인프라를 구축하고 민·관이 협력하여 인공지능 플랫폼 구축에 대한 합의와 과감한 투자가 필요한 때이다.



그림 4. 인공지능 플랫폼 구축을 위한 정부 및 민간, 학계 협력 체계

Fig. 4. Government, Private, and Academic Cooperation system for Building AI Platform

References

- [1] S.-H. Lee, *International Comparison of Artificial Intelligence Research Capabilities and Implications*, SPRI Insight Report, vol. 2018, no. 3, pp. 2-14, Nov. 2018.
- [2] J. Park, Ministry of Science and ICT, *AI R&D Strategy for Realization of I-Korea 4.0*, pp. 9-17, Hanyang University, May 2018.
- [3] K.-S. Noh, *What is a Platform*, Communication Books, 2014.
- [4] *UK Financial Times(FT) analysis of the top companies in global market capitalization*, <https://www.sedaily.com/NewsView/1VO600UGZK>
- [5] H.-T. Yang, B. Choi, J. Lee, and H. Jang, "AI technology prospect and innovation policy direction," *STEPI*, pp. 53-110, Dec. 2018.
- [6] J. Lee and S. Choi, "Public service artificial intelligence ML application and public value," *Korea Univ. Government Stud. Res.*, vol. 24, no. 1, pp. 4-12, 2018.
- [7] M. Davies, M. Nowotka, G. Papadatos, and F. Atkinson, "A virtual platform for distributing cheminformatics tools and open data," *Open Access Challenge*, pp. 334-335, 2014.
- [8] S. Choi, "The great dictionary of industrial safety," *Gold*, May 2004. (Oejin Park "The concept and introduction strategy of artificial intelligence platform," 4p, Re-quoted)
- [9] O. Park, "The concept and introduction strategy of artificial intelligence platform," *IITP ITFIND*, pp. 7-8, Oct. 2018.
- [10] NIA SPECIAL REPORT, *AWS Re:Invent 2019 AI platform present and future*, pp. 2-4, 2019.
- [11] S. Bang, H. Ha, and C. Kim, "A study on big data-based software architecture design for utilization of public open data," *J. Korea Inf. Technol. Soc.*, pp. 99-107, 2015.
- [12] E. K. Nelson, et al., "An open source platform for scientific data integration, analysis and collaboration," *BMC Bioinformatics*, pp. 5-6, 2011.
- [13] A. Kim and J. Lee, "Data analysis platform using open source based deep learning

engine,” *Smart Content Res. Section Electron. and Telecommun. Res. Inst.*, pp. 3-4, 2018.

[14] J.-H. Lee, *The Importance of Open Sources and Their Beginnings*, NIPA Issue Report, vol. 209-20, pp. 9-12, Jul. 2019.

[15] S. Choi, *The Hidden Driving Force of the 4th Industrial Revolution, Current Status and Implications of Open Source*, KDB Research Monthly Report, no. 770, pp. 49-56, Jan. 2020.

[16] *AI Hub Portal AI software*, http://www.aihub.or.kr/ai_software

[17] The Federal Government, AI Made in Germany, *Artificial Intelligence Strategy*, pp. 6-9, Nov. 2018.

[18] J. Alstott and H. Kautz, Executive Office of the President of the United States, *The National Artificial Intelligence Research & Development Strategy Plan: 2019 UPDATE*, pp. 40-42, Jun. 2019.

[19] S.-G. Lee, “The current status of domestic and overseas AI utilization and public application,” *IITP ICT Spot Issue*, p. 18, Dec. 2018.

[20] PAT Research, 60 Second Marketer, G2 Crowd’s website (Feb. 2, 2020) from <https://www.predictiveanalyticstoday.com/>, <https://60secondmarketer.com/>, <https://www.g2.com/>

[21] S.-H. Lee, “The AI platform competition is beginning,” *LGERI*, pp. 8-14, 2016.

[22] Saltlux homepage, <http://www.saltlux.com/index.do>

[23] Acrylic homepage, <https://www.iacryl.com/acryl/sub/jonathan.php>

[24] Lunit homepage, <https://lunit.io/ko/>

[25] Vuno homepage, <https://www.vuno.co/>

[26] Samsung SDS homepage, <https://www.samsungsds.com/global/ko/solutions/off/brig/brighticsAI.html>

[27] LG CNS DAP homepage, <https://dap.lgcns.com/index2.html?dateRangetYn=N>

[28] Maeil Business Newspaper, KT Leading AI Platform with ‘Gi Genie Inside,’ Nov. 11, 2019.

[29] S. Choi, “The status and development strategy of the artificial intelligence industrial ecosystem,” *IITP*, pp. 3-6, 2019.

[30] Federal Ministry for Economic Affairs and Energy(BMWi), *A federated data infrastructure as the cradle of a vibrant european ecosystem*, pp. 11-15, 2019.

[31] Ministry of Science and ICT press release, *Expanding the supply of AI learning data, which is an accelerator of artificial intelligence industry growth*, Mar. 19, 2020.

[32] Joint Ministry, Artificial Intelligence National Strategy, *National Strategy for Artificial Intelligence*, p. 13, Dec. 2019.

[33] H. S. Jeong, “A study on the application of cloud computing in the 4th industrial revolution,” *J. KICS*, vol. 44, no. 6, pp. 1216-1217, Jun. 2019.

[34] National Information Society Agency, AI Insight Report, “*Aim to build mid- to long-term AI data heard in the field of artificial intelligence* (in-depth interview with AI experts),” pp. 6-11, Sep. 2019.

윤 창 희 (Chan-hee Yun)



1999년 2월 : 단국대학교 정보관리학과 졸업
 2017년 7월 : University for Peace, Sustainable Peace in the Contemporary World 석사
 2019년 3월~현재 : 경북대학교 정보과학과 박사과정
 <관심분야> 인공지능, 기계학습, 컴퓨터공학

김 재 일 (Jaecil Kim)



2007년 2월 : 아주대학교 미디어학부 졸업
 2015년 2월 : 한국과학기술원 전산학과 박사
 2018년 3월~현재 : 경북대학교 컴퓨터학부 조교수

<관심분야> 기계학습, 영상분석, 인공지능
 [ORCID:0000-0002-9799-1773]