

인공지능과 빅데이터 기술동향

권순선 아주대학교 수학과/인공지능·데이터사이언스 대학원 교수

1. 머리말

빅데이터, 인공지능, 사물인터넷 등으로 대변되는 4차 산업혁명은 소프트웨어와 데이터 기반의 지능디지털 기술 혁명이라고 볼 수 있다. 이 혁명은 우리의 생활방식과 산업을 바꾸어 놓고 있으며, 인류 문명에 주는 충격도 엄청나다. 이러한 시대에 대비하기 위한 두 가지 핵심기술로 인공지능과 빅데이터 기술이 요구된다.

인공지능(AI, Artificial Intelligence)¹⁾ 기술은 모든 것이 연결되고 좀 더 지능적인 사회로의 진화로 전망되는 제4차 산업혁명의 주역으로, 데이터와 지식이 산업의 핵심 경쟁 원천이다. 이러한 인공지능은 인간의 인지능력, 학습능력, 추론능력, 이해능력 등과 같이 인간의 고차원적인 정보 처리 능력을 구현하기 위한 ICT(Information & Communication Technology) 기술을 말한다. 빅데이터²⁾는 기존 데이터베이스로 처리할 수 있는

역량을 넘어서는 초대용량의 정형, 비정형 데이터를 생성·수집·저장·관리·분석하여 가치를 추출하고 지능화 서비스의 기반을 지원하는 기술이다. 이러한 인공지능과 빅데이터 기술은 컴퓨터와 초연결 디바이스에서 발생하는 데이터의 폭발적 증대와 초연결 지능화를 위하여 공통으로 소요되는 대용량, 다양성, 실시간, 지능화 기능을 플랫폼으로 제공하는 SW기술을 말한다[그림 1].³⁾

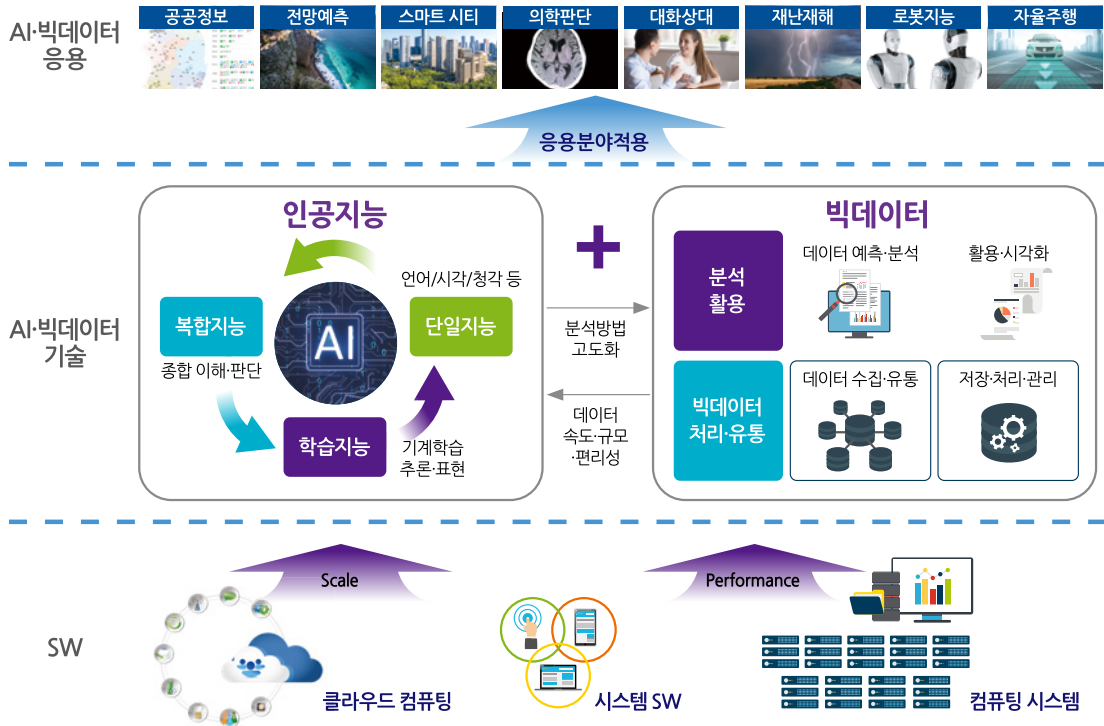
최근 5년간 인공지능은 많은 발전을 이루었으나, 현 수준에서 인공지능은 인간의 지능을 완벽하게 구현하는 데는 많은 한계를 가지고 있다. 즉, 자율적인 상황 판단과 능동적인 행동을 기반으로 하는 인간의 지능 수준과 큰 차이가 있는 것이다.

위와 같은 한계점을 극복하기 위해 수학·과학 분야의 이론과 모델링, 특히 빅데이터에 기반으로 인공지능을 구현하려는 시도가 활발히 진행 중이다. 본고에서는 우리나라 빅데이터 기반의 인공지능 기술 발전을 도모하기 위해 국내외의 인공지능과 빅데이터 기

1) 인공지능은 사고와 지능적인 행위의 근저에 깔린 메커니즘의 이해와 그 메커니즘을 기계에 구현한 것(미국 인공지능발견협회(AAAI)의 역할 정의에서 발췌, <https://www.aaai.org/home.html>)

2) 빅데이터는 통상적으로 사용되는 데이터수집·관리·처리 SW의 수용 한계를 넘어서는 방대한 크기의 데이터로, 데이터의 양(Volume)-데이터 입출력의 속도(Velocity)-데이터 종류의 다양성(Variety)으로 정의 될 수 있는 정보자산(Doug Laney, 3D Data Management, Gartner, 2001)

3) 출처: ICT R&D 기술로드맵 2023-인공지능·빅데이터분야, 2018.9.



[그림 1] 인공지능·빅데이터 기술 분야 개념도

술 현황과 개발 동향에 대해서 살펴보고자 한다.

2. 인공지능과 빅데이터 기술 동향

2.1 인공지능에서 빅데이터의 역할 및 우리나라의 수준

4차 산업혁명의 기반기술로 불리는 인공지능 기술에 대한 관심이 급격히 높아지고 있다. 인공지능 분야의 경우 본질적으로 데이터 기반으로 기술 개발을 한다는 점에서 빅데이터 분석과 활용에 대한 수요는 이전보다 급증하고 있다. 이에 빅데이터 기술, 즉 데이터 기반 기술은 무엇이고, 현재 우리나라 수준은 어떠한지 보고자 한다.

데이터 기술(DT, Data Technology)이란 데이터의 측정, 수집, 축적 기술에서부터 시작하여, 데이터의 전송, 분석 및 해석 능력, 데이터로부터 정보와 지식을 창출하는 기술, 통계적 모형화 기술, 미래를 예측하는 기술 등을 다루는 통계 기반의 과학적 기술을 말한다. DT는 데이터의 취급, 소프트웨어의 구축, 모형화 및 미래 예측 기술을 주로 다루고 있으므로, 그 진행과 결과가 눈에 잘 띄지 않으며 간과하기 쉽다. 그러나 지식 정보화 사회에서 정보 창출을 위해서는 DT의 활용은 필수적 요소이다. 데이터 기술이란 용어는 처음으로 박성현⁴⁾에 의하여 언급되었으며, 통계학 응용으로서의 발전방

4) 박성현(2001). '데이터 기술의 경제학', 한국경제신문 다산칼럼, 2001년 12월 3일.

'지식기반 사회에서의 통계학 패러다임의 변화와 데이터 기술의 발전', 경영정보논총, 서울대학교 경영대학, 제11권, p.53-59, 2001년 12월.

5) Sung H. Park and Moon W. Suh(2008), 'Data Technology as a new discipline for broader application of statistics', Journal of Data Science, Vol.6, No.3, p.357-368, July University, New York.

향은 ‘Data Technology as a new discipline for broader application of statistics’⁶⁾와 ‘데이터 기술: 지식창조를 위한 새로운 융합과학기술’⁷⁾에서 찾아볼 수 있다. 이러한 데이터기술에 ‘빅(Big)’을 붙여 불리고 있는 빅데이터 기술은 일반적인 데이터베이스 소프트웨어가 저장·관리·분석할 수 있는 범위를 초과하는 규모의 데이터 분석 활용 기술로 인공지능 기술과 더불어 핵심 기술로 불리고 있다.

우리나라는 세계적 수준의 ICT 인프라를 보유하고 있으나, 빅데이터를 활용할 수 있는 데이터

의 양이 부족하고, 기술수준도 선진국에 비해 낮은 수준으로 평가된다. 우리나라는 빅데이터 인프라인 공공데이터 개방이 아직 미흡하다. 행정안전부(2018)⁷⁾에 따르면 영국의 56.6%, 프랑스의 73.8% 수준으로 아직 양적으로 미흡하며, 인공지능데이터, 공간·위치데이터 등 4차 산업혁명 관련 데이터 개방도 선진국에 비하여 부족한 수준이다. 빅데이터의 양적 수준뿐 아니라 기술 수준도 높지 않다. 한국정보화진흥원(2017)⁸⁾에 따르면 2015년 기준으로 선진국 기술수준을 100으로 볼 때, 우리나라는 65.7로 기술수준 격차는 3.1년, 선진기술

〈표 1〉 인공지능과 빅데이터 기술 분류

중분류	소분류	세분류	요소기술
인공지능	학습지능	머신러닝	베이지안 학습, 인공신경망, 딥러닝, 강화학습, 앙상블러닝, 판단근거 설명
		추론·지식표현	추론, 지식표현 및 온톨로지, 지식처리
	단일지능	언어지능	언어분석, 의미이해, 대화 이해 및 생성, 자동 통역·번역, 질의응답(Q/A), 텍스트 요약·생성
		시각지능	영상 처리 및 패턴 인식, 객체 인식, 객체 탐지, 행동 이해, 장소·장면 이해, 비디오 분석 및 예측, 시공간 영상 이해, 비디오 요약
		청각지능	음성분석, 음성인식, 화자인식·적응, 음성합성, 오디오 색인 및 검색, 잡음처리 및 음원분리, 음향인식
	복합지능	행동·소셜지능	공간 지능, 운동 지능, 소셜 지능, 협업 지능
		상황·감정이해	감정 이해, 사용자 의도 이해, 뇌신호인지, 센서 데이터 이해, 오감인지, 다중 상황 판단
		지능형 에이전트	에이전트 플랫폼, 에이전트 기술, 게임 지능, 모방창작 지능
		범용 인공지능(AGI)	상식 학습, 범용 문제 해결, 평생 학습, 도덕·윤리·법 지능
빅데이터	빅데이터 처리·유통	빅데이터 수집·유통기술	빅데이터 수집·정제·융합·가공, 데이터 품질관리, 실시간 ETL·ELT, 데이터 생성·증강, 비식별화 및 필터링, 데이터 마켓, 데이터 레이크, 데이터 프리퍼레이션, 마스터 데이터관리
		빅데이터 저장·처리·관리 기술	실시간 스트림 처리, 데이터 처리 및 관리, 데이터 라이프사이클, 가상 데이터맵 관리, 데이터 처리 프레임워크
	빅데이터 분석·활용	빅데이터 분석·예측 기술	심층 분석, 실시간 분석, 그래프 분석, 예지분석, 예측분석, 지시적 분석, 인지 분석, 시공간복합분석, 분석 모델링, 시뮬레이션, 데이터 기계학습·딥러닝, 이상치 검출
		빅데이터 활용·시각화 기술	빅데이터 서비스, 빅데이터 응용, 사회변화 예측, 데이터 시각화, 분석 시각화, 대시보드

6) 박성현(2010). ‘데이터 기술: 지식창조를 위한 새로운 융합과학기술’, 한국품질경영학회지, Vol.38, No.3, p.294-304, 2010년 9월호.

7) 행정안전부(2018). ‘정부의 빅데이터 추진체계 및 정책방향’, ‘빅데이터 추진체계 현황 및 향후과제’ 세미나 자료, 공공데이터정책과, 2018.3.22.

8) 한국정보화진흥원(2017). ‘2017 국가정보화백서’, 2017.

도달 시간은 3.4년이 걸릴 것으로 평가되었다.

미국, 영국, 일본 등 빅데이터 관련 정책을 제대로 추진하고 있는 대부분의 국가에서는 공공데이터 개방, 빅데이터의 공공서비스 확대, 빅데이터 전문 인력 양성 등에 대한 진흥정책과 개인정보보호 등에 관련된 법제도 개선을 중심으로 빅데이터 정책이 균형있게 추진되고 있다.

국내외적으로 빅데이터 정책 방향은 공공데이터 개방을 우선적으로 확대하고, 민간 기업 등으로 하여금 공공데이터의 활용을 제고하는 방향으로 추진되고 있다. 개방된 공공데이터는 민간데이터 등과의 융·복합을 통해 새로운 비즈니스 모델 창출 등에 활용되고 있다.

2.2 인공지능과 빅데이터의 기술 분류 및 개발 동향

인공지능 기술은 지식을 확보하는 알고리즘을 연구하는 학습지능, 시각, 청각, 언어 등과 같은 한 종류의 입력을 가지고 지식을 확보하는 단일지능과 여러 형태의 입력을 통합하여 이해 판단하는 복합지능 기술을 포함한다.

빅데이터 기술은 데이터 수집·저장·처리 등에 관련된 빅데이터 플랫폼 기술과 이와 연계한 빅데이터 분석 예측 기술을 활용하여 새로운 통찰력과 비즈니스 가치를 창출하는 빅데이터 분석 활용기술을 포함한다. 인공지능과 빅데이터 기술 분류는 <표 1>⁹⁾과 같다.

세계시장의 인공지능 기술은 수많은 산업분야에 적용될 수 있어 전체적인 시장규모를 측정하기 어려우나, 대략 2017년 48억 달러에서 2023년 532억 달러 규모로 연평균 49.2%의 고도성장이 예상된다<표 2>.¹⁰⁾ 이들 중, AI SW 플랫폼 시장을 2017년 23억 달러 규모에서 2021년 84억 달러 규모로 연평균 약 40%의 고속 성장이 예상된다.¹¹⁾ 또한, AI 개인비서(IVA, Intelligent Virtual Assistant) 시장은 AI 분야 중 급속히 확산되고 있는 분야로서, 2017년 25억 달러 규모에서 2023년 252억 달러 규모로 연평균 약 47%의 고속 성장이 예상된다.¹²⁾

세계 빅데이터 시장은 연평균 10.6%로 성장하고 있으며, 2017년 534억 달러에서 2021년 814억 달

<표 2> 인공지능과 빅데이터 기술 개발 시장규모

[단위: 세계시장 백만달러, 국내시장 십억 원]

구분		2017년	2018년	2019년	2020년	2021년	2022년	2023년	CAGR
인공지능	세계 ¹³⁾	4,819	7,345	11,283	17,267	25,995	37,987	53,231	49.2%
	국내 ¹⁴⁾	6,400	7,500	9,100	11,100	13,320	15,984	19,180	20.1%
빅데이터	세계 ¹⁵⁾	53,440	59,040	65,450	72,890	81,380	90,006	99,547	10.6%
	국내 ¹⁶⁾	244.2	298.7	358.3	446.8	557.2	694.8	866.4	24.7%

9), 10) ICT R&D 기술로드맵 2023-인공지능-빅데이터분야-, 2018.9.

11) IDC(International Data Corporation), 2017.

12) Research and Market, 2018.

13) [세계시장] Statista's AI market revenue worldwide 2015-2025, 2018.

14) [국내시장] 2017년 국가정보화에 관한 연차보고서 2017(미래창조과학부 자료 인용 2021~2023 CAGR 적용)

15) [세계시장] IDS's Seminannual Dig Data and Analytics Software Tracker, May 017 92022~2023 CAGR 유추)

16) [국내시장] IDC Korea Big Data Technology and Services 2015-2019 Forecast, Feb 2016. (2020~2023 CAGR 적용)

리로 성장할 전망이다<표 2>. 빅데이터 시장은 빅데이터 인프라 시장, 빅데이터 소프트웨어 시장, 빅데이터 서비스 시장으로 나누어 집계된다. 빅데이터 인프라 시장이 가장 큰 비중을 점할 것으로 예상되며, 2016년 136억 달러에서 2020년 277억 달러로 연평균 20.3%의 성장을 보일 전망이다. 빅데이터 소프트웨어 시장은 2020년까지 159억 달러로 성장하여 서비스 시장과 유사할 전망이다, 빅데이터 서비스 시장도 24%(15~20)의 성장세를 보인다. 소프트웨어와 유사한 규모인 152억 달러의 시장을 형성할 것으로 보인다. 빅데이터 소프트웨어에서 주목할 점은 탐색 및 분석 SW가 연평균 27.2%로 빠른 성장 전망을 나타내고 있다는 것이다.¹⁷⁾

국내 인공지능 시장은 2017년 6.4조 원에서 2023년 19.2조 원 규모로, 세계시장 연평균 성장률과는 차이가 있지만 그럼에도 연평균 20.1%의 고성장이 예측된다.¹⁸⁾ 2020년 기준 음성인식 및 통번역 시장 4.2조 원, 영상처리 및 영상인식 시장 3.5조 원, 기타 SW 및 알고리즘 관련 시장 3.4조 원으로 예상된다. 빅데이터 시장은 연평균 24.7%의 높은 성장률로 확대되고 있으며, 2017년 2,442억 원에서 2020년 4,468억 원으로 성장할 전망¹⁹⁾이고, 국내 빅데이터 인프라 시장은 2017년 1,222억 원에서 2019년 1,788억 원으로 연평균 24.9%로 성장할 전망이다. 빅데이터 소프트웨어 시장은 2017년 539억 원에서 2019년 795억 원으로 연평균 23.4%로 성장할 전망이고, 빅데이터 서비스 시장은 2017년 681억 원에서 2019년 1,000억 원으로 연평균 25.4%로 가장 높은 비율로 성장할 전망이다.

3. 맺음말

최근 인공지능 기술의 발전 속도가 점차 가속화되면서, 향후 10여 년간의 변화는 인공지능 개념이 등장한 1950년대 중반 이후부터 현재까지 약 60년간의 변화를 압도할 것으로 전망하고 있다. 또한, 인공지능의 정확도가 높아지면서 산업 전반에 인공지능 기술 개발이 활발히 이루어지고 있다. 인공지능은 4차 산업혁명시대에 들어오면서 알파고 마스터(AlphaGo Master)와 같이 혁신적으로 발전한 알고리즘, 빅데이터, 클라우드, 소프트웨어 등이 서로 융·복합되면서 실제 구현을 통해 산업 전반에 적용되어 다양한 현실 세계의 문제를 해결하고 있다. 이러한 인공지능 기술은 머신러닝·딥러닝과 같은 학습을 통해 인간과 유사하게 행동하는 컴퓨터 프로그램을 구현할 수 있고, 무엇보다 이것을 학습할 수 있는 방대한 양의 빅데이터와 컴퓨팅 파워가 요구된다.

빅데이터는 2010년대 이후 인공지능 기술 개발을 촉진시키면서, 그 중요성이 확대되었고, 이러한 빅데이터는 지금까지 양적인 측면이 강조되어 플랫폼과 분석 기술이 발전되었으나, 데이터 개방·유통·공유할 수 있는 플랫폼의 발전과 인공지능 분석 방법론을 적용한 데이터 분석 지능의 다양화로의 발전이 필요하다. 즉, 빅데이터를 수집하는 IoT뿐만 아니라 이를 분석하기 위한 빅데이터 및 컴퓨터 용량을 제공하는 클라우드 기술, 자연어 처리기술과 인식기술 등 다양한 기술의 개발이 더불어 요구된다.


구글은 ‘분산형’ Federated Learning 기술을 도입하여, ‘AI로부터 배우는 AI’기술의 시작

17) ETRI, 빅데이터 플랫폼 현황 및 이슈 분석, 2017.

18) '2017년 국가정보화에 관한 연차보고서'의 과기정보통신부 자료 인용

19) IDC(International Data Corporation), 2016.

으로 AutoML 연구를 진행 중이며, IBM의 Big Data & Analytics Hub에서는 자사의 IBM Bluemix platform 상에 IBM Graph 서비스를 출시하고, NEC 유럽연구소는 스마트시티 최적화에 활용을 최근에 발표하고 있다. 또한, Decision Optimization이라는 수학적 모델기반 분석 솔루션을 개발하여 상업, 제조, 금융, 의료 등의 분야에 적용하고 있다. 이러한 글로벌 기업들은 연구 성과물들을 자사 제품 개발과 서비스에 응용하여 상용화하고 있으며, 세계 각국들도 다양한 인공지능 정책을 수립하여 기술 주도권을 잡기 위해 많은 노력을 하고 있다.

향후 우리나라는 무엇보다 글로벌 기업의 플랫폼과 상용서비스를 활용한 방대한 학습데이터 확보 전략에 대응하기 위해서는 주어진 학습데이터로 성능향상이 가능한 인공지능 원천기술 확보와 설명이 가능하면서 인간 수준의 인식 정확도를 갖는 인공지능 기술 개발이 필요하다. 이러한 빅데이터 기반 인공지능 기술 개발에서는 데이터 활용 니즈가 커지고 있음에 따라, 데이터 자체를 새로운 비즈니스 재화로 볼 수 있다. 타기관의 데이터를 원격으로 수집·등록·유통·거래할 수 있는 메타데이터 체계가 구축되고 데이터 품질을 강화할 수 있는 기술이 확보되어, 우리나라만의 인공지능과 빅데이터 기술이 개발되길 기대해 본다. 

참고문헌

- [1] ICT R&D 기술로드맵 2023 -인공지능·빅데이터 분야 - 2018.9, IITP
- [2] 박성현, 오진호, 권순선 (2016). 통계로 풀어가는 빅데이터, 한국표준협회미디어
- [3] 정보통신기획평가원 2018, ICT 신기술 - 인공지능 기술 및 산업 분야별 적용 사례