



# 한국의 데이터 기반 디지털 경쟁력 변화추이 통계적 분석과 데이터 강국 실현 전략 연구\*

박 성 현\*\*

## Statistical Analysis of Data-based Digital Competitiveness Trend of Korea, and Strategic Study to Achieve a Strong Data Power Country\*

Sung Hyun Park\*\*

---

### ABSTRACT

This study statistically investigated the world digital competitiveness ranking and the global AI index ranking in order to understand the strong and weak points of Korea, and to secure digital competitiveness in the digital transformation era. Man power education and data strong country are necessary conditions to secure digital and AI competitiveness. In this study, the man power education strategies are suggested, and the principles and promotional polices of data science are suggested.

Next, the basic ideas and current status of data industry are investigated, and the promotional strategies of data industry are shown. Finally, in order to be a strong data power country for Korea, seven strategies are scrutinized which are as follows. Firstly, promotion of data science

---

\* 본 논문은 2023년도 대한민국학술원 전문학술활동 연구비 지원으로 이루어짐

\*\* 대한민국학술원 자연과학부 제1분과 회원

education and cultivation of one million data·AI manpower, secondly, activation of the data basic law and achievement of 100 billion won of data industry sales, thirdly, realization of digital platform government and digital new-deal, fourthly, activation of my-data business, fifthly, enlargement of data use in medical field, sixthly, installation of data science·AI international cooperation center, and finally, construction of data·AI research fusion cluster. If a data strong power country is realized in Korea, the basic strengths for digital and AI competitiveness are prepared, and Korea can stand as a world leading country in this digital transformation era.

Key words : AI competitiveness, Data industry, Data science, Data strong country, Digital competitiveness

## 초 록

본 연구는 디지털 전환기를 맞아 디지털 경쟁력을 확보하기 위해서 한국이 가진 강점과 약점을 파악하기 위해 한국에 대한 세계 디지털 경쟁력 평가와 글로벌 인공지능(AI) 경쟁력 평가 변화 추이를 통계적으로 조사한다. 디지털과 AI 경쟁력은 충분한 인재 양성과 데이터 강국이 되는 것이 필수적인 요건이다. 본 연구에서는 인재 양성 전략을 살펴보고, 데이터 강국의 중요한 인프라인 데이터과학의 본질과 진흥 정책과 전략을 살펴본다.

다음으로 데이터 산업의 본질과 현황을 살펴보고, 데이터 산업의 육성 및 진흥 전략을 조사한다. 마지막으로 데이터 강국으로 발전하기 위한 전략으로 일곱 가지 전략을 제시한다. 첫째, 데이터 과학 교육 진흥과 데이터·AI 인력 100만 명 양성, 둘째, 데이터 기본법 활성화와 데이터 산업 매출액 100조 원 달성, 셋째, 디지털 뉴딜과 디지털 플랫폼 정부 실현, 네 번째, 마이데이터 사업 활성화, 다섯 번째, 의료부문 데이터 활용 극대화, 여섯 번째, 데이터과학·AI 국제협력센터 설치, 일곱 번째, 데이터·AI 연구 융합 클러스터 조성 등이다. 데이터 강국이 실현될 때 우리나라는 디지털과 AI 경쟁력의 기본 체력이 강화될 것이며, 디지털 전환기인 지금 세계 선도 국가로 발돋움할 수 있을 것이다.

주제어 : 디지털 경쟁력, 인공지능 경쟁력, 데이터과학, 데이터 산업, 데이터 강국

---

## 목 차

I. 서론	V. 데이터 산업 현황과 진흥 전략
II. 디지털 경쟁력 변화 추이의 통계적 분석	VI. 데이터 강국 발전 전략
III. AI 경쟁력 변화 추이 분석	참고문헌
IV. 인재 양성 전략과 데이터과학	

---

## 1. 서론

### 1. 디지털 전환 시대

지금은 4차 산업혁명 시대이며, 급변하는 디지털 전환(digital transformation) 시대이다. 이 시대의 핵심 경쟁력은 데이터 기반의 빅데이터, 인공지능(AI), 사물 인터넷, 메타버스, 자율주행자동차 등의 기술이다. 우리나라는 현재 세계 10위권의 경제국으로 많은 분야에서 선진국들과 경쟁하고 있으나, 앞으로의 국가 간의 핵심 경쟁력은 데이터 기반의 디지털 경쟁력(digital competitiveness)과 인공지능 경쟁력(AI competitiveness)이라고 말할 수 있다. 앞으로의 10년이 우리나라의 디지털과 인공지능 경쟁력 진흥에 위기와 기회가 공존하며, 위기의 원인을 밝히고 글로벌 중추 국가(global pivot state)로 도약하기 위한 준비를 하는 것은 매우 중요하다.

디지털 경쟁력을 키우기 위해서 우선 몇 년간에 걸친 우리나라의 디지털 경쟁력의 통계적 변화 추이를 자세히 살펴보고, 우리나라의 강점과 취약점이 무엇인지 조사 분석할 필요가 있다. 디지털 경쟁력 평가 자료로 가장 신뢰성이 있는 자료는 스위스 국제경영개발원(IMD)이 발표하는 세계 디지털 경쟁력 순위(IMD World Digital Competitiveness Ranking)가 있다. 이 자료 속에는 ‘빅데이터 및 분석 기술 활용’ 순위도 나와 있다. 우선적으로 이 자료를 구체적으로 분석하면서 디지털 경쟁력이 강한 나라들과 한국을 비교 연구한다.

다음으로 디지털 경쟁력에서 AI가 매우 중요하다. 우리나라의 AI 경쟁력이 어느 수준인지는 영국의 데이터 분석업체 토터스 인텔리전스(Tortoise Intelligence)

가 발표하는 글로벌 AI 지수(Global AI Index)가 있다. 본 연구에서는 이 지수의 변동 추이도 조사하고 우리나라 AI 기술의 강점과 취약점 등의 변화 추이를 주요 국가들과 비교 조사 분석할 필요가 있다. 디지털/AI 경쟁력을 진흥시키기 위해서는 데이터 강국(data power country)이 되어야 한다. 데이터 강국은 유용한 데이터의 양이 많고 질이 우수하며, 데이터 활용 능력이 뛰어난 나라이다. 데이터 강국이 되는 것이 디지털/AI 경쟁력을 제고하는 데 핵심적인 요소이다. 따라서 데이터 강국으로 가는 길이 디지털 전환 시대에 앞서가는 나라이다.

## 2. 연구의 목적

본 연구의 목적은 우리나라의 디지털 경쟁력과 AI 경쟁력의 취약점을 상세히 분석한 후에, 이를 보강하여 데이터 강국으로 발전하기 위한 전략을 연구하는 것이다. 데이터 강국 발전 전략에는 공공데이터의 활용 능력 제고, 데이터 컨트롤타워의 효율적인 작동, 데이터 산업(data industry)의 진흥, AI 산업의 진흥, 각종 규제 여건의 완화, 개인 데이터의 적극적 활용, 초·중·고등학교에서의 데이터 교육 진흥 전략, 대학에서의 데이터과학(data science) 교육을 통한 인재 양성 방안 등을 연구하고 제안한다.

한국이 데이터 강국으로 부상하게 되면, 자연스럽게 관련 학문 분야인 통계학, 데이터과학, AI학 등이 발전하게 되며, 결국 관련된 학문 분야의 학문적 소통이 증진되고, 국가 발전에 기여하게 될 것이다.

## 3. 연구의 내용 및 범위

위의 연구의 목적을 달성하기 위해 본 연구과제의 연구 내용과 범위는 다음과 같이 크게 네 가지로 나누어져 있다.

- 1) 스위스 국제경영개발원(IMD)이 발표하는 세계 디지털 경쟁력 순위에 근거하여, 지난 수년간에 걸친 한국의 디지털 경쟁력을 주요 국가들(미국, 중국, 일본, 독일 등)과 비교 분석한다. 디지털 경쟁력 평가 기준은 3대 분야, 9개 부문, 52개 세부 지표로 구성되어 있고, 3대 각 분야에는 다음과 같이 각각 3개 부문이 있다.

- 지식 (인재, 교육훈련, 과학기술)
  - 기술 (규제 여건, 자본 여건, 기술 여건)
  - 미래준비도 (신기술 적응도, 사업 능력, IT 통합)
- 이들 9개 부문과, 주요 세부 지표들에 대하여 한국의 현황이 수년간에 어떻게 변화하는지 조사 분석한다.

2) 또한 AI가 매우 중요한 기술이므로, 토머스 인텔리전스의 ‘글로벌 AI 지수’에 의한 한국의 AI 경쟁력 평가를 분석한다. 이 평가는 7개 부문(개발 능력, AI 인프라, 연구 수준, 정부 전략, 운영환경, 인재, 산업화)으로 나누어진다. 각 부문에 대하여 한국의 강점과 취약점이 무엇이며, 주요 국가와 비교 분석한다.

3) 데이터 강국으로 가는 가장 중요한 인프라는 인재 양성이다. 이를 위한 연구의 내용과 범위는 다음과 같다.

- 초·중·고등학교에서의 데이터 교육 진흥과 대학에서의 데이터과학(data science) 교육의 보편화 전략
- 데이터과학 국내외 인재 양성 현황 비교와 인재 양성 전략
- 2027년까지 데이터 기반 디지털 인재 100만 명 양성 전략 연구
- 데이터·AI 연구원 설립과 연구 클러스터 조성
- 데이터·AI 국제협력센터 설치

4) 앞에서 분석한 디지털/AI 경쟁력에서 한국이 취약한 부분을 보강하며, 앞으로 데이터 강국으로 부상하기 위한 다음과 같은 전략들을 구체적으로 연구하고 제안한다.

- 공공데이터의 활용 능력 제고, 데이터 컨트롤타워의 효율적인 작동 방안
- 데이터 기본법 활용과 디지털 플랫폼 정부 실천 방안
- 데이터 산업(data industry)의 진흥 방안 (2030년에 데이터 산업 매출액 100조 달성 목표)
- AI 산업의 진흥 방안
- 각종 규제 여건의 완화
- 기업(기관)에서의 빅데이터 도입률 제고와 클라우드 시장 국내 기업 점유

을 제고

- 개인 데이터의 적극적 활용 방안, 마이데이터 사업 활성화
- 데이터 품질관리 및 표준화 제고
- 국내 데이터 산업에 참여하는 기업의 해외 시장 진출 확대 방안

## II. 디지털 경쟁력 변화 추이의 통계적 분석

### 1. 2022년 세계 디지털 경쟁력 순위

국가 간에 빅데이터, 인공지능, 사물인터넷 기술 등의 신기술 적응도를 포함하는 디지털 경쟁력을 평가하는 기준으로 스위스 국제경영개발대학원(IMD)이 매년 발표하는 'IMD 세계 디지털 경쟁력 순위(IMD World Digital Competitiveness Ranking)'가 있다. 이 평가는 기술변화에 대한 한 국가의 적응력, 대응력, 기술 개발 능력 등을 평가하며, 3대 분야(지식, 기술, 미래 준비도), 9개 부문, 52개 세부 지표로 구성되어 있다.

2022년 9월에 발표한 IMD의 '2022년 세계 디지털 경쟁력 평가' 결과를 보면, 우리나라는 64개 평가대상 국가 중에서 8위를 차지하여 이 평가가 시작된 2015년부터 2022년 사이에 연차적으로 18위 → 17위 → 19위 → 14위 → 10위 → 8위 → 12위 → 8위로, 2017년에 잠시 19위로 밀렸으나, 곧이어 상승하다가 현재 8위를 유지하고 있다. 평가 결과는 9개 부문별로, 52개 세부 지표별로는 큰 차이가 있어 한국이 취약한 부분도 상당히 많이 있다. 3대 분야, 9개 부문별로 우리나라의 평가 순위와, 각 부문별로 특별히 유의할 특징적인 세부 지표를 적어 보면 <표 2.1>과 같다.

〈표 2.1〉 한국의 2022년 세계 디지털 경쟁력 평가 결과

(괄호 속은 우리나라의 순위)

3대 분야	9개 부문	52개 세부 지표 중 특징적인 일부 세부 지표
지식 (knowledge): 16위	인재 (talent): 33위	- 디지털 기술 능력 (46위) - 외국인 숙련직 직원 (49위) - 국제 학생의 순 유입 (38위)
	교육훈련 (training & education) : 16위	- 고등교육 성취도 (4위) - 대학 이상 전문교육 학생:교수 비율 (30위) - 여성 학위 소지자 (20위)
	과학기술 (scientific concentration) : 3위	- 1인당 R&D 총액 (2위) - 1인당 총 연구개발 인력 (3위) - 여성 연구원 (53위)
기술 (technology): 13위	규제 여건 (regulatory framework) : 23위	- 창업 (19위) - 지적 재산권 (37위) - 기술 개발 및 적용 (48위)
	자본 여건 (capital) : 15위	- IT&미디어 주식시장 자본화 (4위) - 벤처 자본 (35위)
	기술 여건 (technological framework) : 7위	- 인터넷 대역폭 속도 (12위) - 고도기술 수출 비중(%) (6위) - 인터넷 사용자 (8위)
미래 준비도 (future readiness): 2위	신기술 적응도 (adaptive attitudes) : 1위	- 전자 참여 지수 (1위) - 인터넷 소매업 매출액 (1위) - 스마트폰 보유율 (4위) - 세계화에 대한 태도 (11위)
	사업 능력 (business agility) : 2위	- 로봇 활용도 (3위) - 기업의 실패 두려움(2위) - 빅데이터 및 분석 기술 활용 (34위) - 지식 전달 (30위)
	IT 통합 (IT integration) : 14위	- 전자정부 (2위) - 사이버 보안 능력 (28위) - 소프트웨어 불법복제 (20위)

‘지식’은 새로운 기술을 이해·발견·확장할 수 있는 역량을 측정하는 분야로, 우리나라의 경제력 규모(2021년 국내총생산(GDP) 기준으로 1조 8,068만 달러로 세계 10위)보다 좀 떨어진 16위이다. 특히 ‘인재’ 부문에서 뒤처지는데, 이는 주로 디지털 기술 능력(46위)이 아직도 불충분하고, 외국인 숙련직 직원(49위)이 모자라고, 국제 학생의 순 유입(38위)이 떨어지기 때문이다. 인재 부문에서 더 적극적인 국제화가 필요하다고 하겠다. 또한 ‘교육훈련’ 부문에서 대학에서의 학생:교수 비율(30위)이 낮은 수준이고, 여성 학위 소지자(20위)가 충분하지 못하기 때

문이다. ‘과학기술’ 부문이 3위로 우수한데, 이는 R&D 총액(2위)과 1인당 총연구개발 인력(3위)이 많기 때문이다.

‘기술’은 디지털 혁신을 발전시킬 수 있는 역량을 나타내는 분야로, 기술의 규제 여건(23위)과 자본 여건(15위)이 안 좋은 편이다. 규제 여건 속에는 지식재산권에 관한 여건(37위), 기술 개발 및 적용에 관한 여건(48위)에서 낙후되어 있다. 그러나 고도 기술 수출 비중(6위)이 높아 기술 여건(7위)은 좋다.

‘미래 준비도’는 디지털 변화에 대한 준비 정도를 측정하는 분야로, 전자 참여 지수(1위), 인터넷 소매업 매출액(1위)과 스마트폰 보유율(4위)이 높이 평가되어 신기술 적응도(1위)가 높다. 그리고 로봇 활용도(3위)가 높고, 기업의 실패 두려움(2위)이 적어 사업 능력(2위)은 높으나, 부족한 점은 ‘빅데이터 및 분석 기술 활용(34위)’은 부족하고 지식 전달(30위) 능력이 떨어진다. 전자정부가 2위이고, 사이버 보안 능력(6위)은 좋으나 소프트웨어 불법복제(20위)가 근절되지 않아 IT 통합이 14위에 그치고 있다.

우리나라는 디지털 경쟁력을 강화하기 위해서는 결국 규제 여건, 자본 여건, 인재 부문 등에서의 경쟁력을 높이는 것이 우선적인 과제이다. 한국(8위)보다 디지털 경쟁력 높은 국가들은 1위에서 7위까지 덴마크, 미국, 스웨덴, 싱가포르, 스위스, 네덜란드, 핀란드의 순이며, 기타 주요 국가들은 영국(16위), 중국(17위), 독일(19위), 일본(29위) 등이다.

## 2. 지난 3년간 디지털 경쟁력 변화 추이

2022년에 한국의 세계 디지털 경쟁력 순위는 8위이고, 3대 분야인 지식, 기술, 미래 준비도에서는 각각 16위, 13위, 2위였다. 그리고 3대 분야에 속해 있는 9개 부문에서의 경쟁력은 <표 2.1>에서 보면 인재(33위), 교육훈련(16위), 과학기술(3위), 규제 여건(23위), 자본 여건(15위), 기술 여건(7위), 신기술 적응도(1위), 사업 능력(2위), IT 통합(14위)이었다. 그러면 지난 3년간 3대 분야, 9개 부문에서는 한국의 경쟁력은 어떤 변화를 겪고 있는가? 다음의 <표 2.2>에서 그 변화 추이를 살펴보자.

〈표 2.2〉 지난 3년간 한국의 디지털 경쟁력 변화 추이

3대 분야	순위			9개 부문	순위		
	2020년	2021년	2022년		2020년	2021년	2022년
지식	10	15	16	인재	21	26	33
				교육훈련	11	16	16
				과학기술	4	3	3
기술	12	13	13	규제 여건	26	23	23
				자본 여건	25	16	15
				기술 여건	3	7	7
미래 준비도	3	5	2	신기술 적응도	1	2	1
				사업 능력	3	5	2
				IT통합	15	16	14
종합순위	8	12	8	종합순위	8	12	8

3대 분야 중에서는 ‘미래 준비도’가 가장 좋아서 2022년에 2위로 2021년의 5위보다 더 좋아지고 있다. 고무적인 추세이다. 그러나 ‘지식’은 2022년에 16위로 그 전년보다 한 단계 하락하고 있으며, 특히 ‘인재’ 부문이 33위로 2020~2022년의 3년간에 걸쳐 점점 하락하고 있다. 인재 양성에 더 많은 노력이 필요하다고 하겠다. ‘기술’ 분야는 2021년, 2022년 모두 13위로 그렇게 나쁘지는 않지만, 기술 안에 있는 3개 부문 중 규제 여건은 23위로 나쁜 편이다. 더욱 과감한 규제 완화를 통하여 기업이 기술개발과 활용에서 장애요인이 없도록 정부의 과감한 규제 여건 완화가 필요하다.

### 3. 세부 지표가 디지털 경쟁력 순위에 주는 의미

〈표 2.1〉과 〈표 2.2〉에 있는 3대 분야, 9개 부문 안에는 총 52개의 세부 지표가 있고, 〈표 2.1〉에는 이 52개 세부 지표 중에서 중요하다고 생각되는 28개 세부 지표에 대한 순위가 나와 있다. 이 28개 세부 지표에 대하여 지난 3년간 변화를 살펴보면 〈표 2.3〉을 얻는다.

〈표 2.3〉 28개 세부 지표에 대한 한국의 순위

3대 분야	9개 부문	세부 지표	순위		
			2020년	2021년	2022년
지식	인재	디지털 기술 능력	18	33	46
		외국인 숙련직 직원	43	46	49
		국제 학생의 순유입	49	44	38
	교육훈련	고등교육 성취도	4	4	4
		대학 이상 전문교육 학생:교수 비율	33	33	30
		여성 학위 소지자	20	21	20
	과학기술	1인당 R&D 총액	2	2	2
		1인당 총 연구개발 인력	3	3	3
		여성 연구원	54	53	53
기술	규제 여건	창업	19	19	19
		지적 재산권	38	36	37
		기술 개발 및 적용	44	45	48
	자본 여건	IT&미디어 주식시장 자본화	2	2	4
		벤처 자본	41	39	35
	기술 여건	인터넷 대역폭 속도	2	12	12
		고도기술 수출비중(%)	6	7	6
		인터넷 사용자	16	7	8
	미래 준비도	신기술 적응도	전자 참여 지수	1	1
인터넷 소매업 매출액			1	2	1
스마트폰 보유율			16	16	4
세계화에 대한 태도			14	17	11
사업 능력		로봇 활용도	3	3	3
		기업의 실패 두려움	15	16	2
		빅데이터 및 분석 기술 활용	15	26	34
IT 통합		지식 전달	30	25	30
		전자정부	2	2	2
	사이버 보안 능력	21	23	28	
	소프트웨어 불법복제	20	20	20	

〈표 2.3〉을 살펴보면, ‘인재’에 속한 ‘디지털 기술 능력’이 3년간 매년 떨어지고 있는 것을 볼 수 있다. 또한 ‘외국인 숙련직 직원’이나 ‘국제 학생의 순 유입’도 3년 동안 하위권에 처져 있으며, 별로 개선의 징후가 없다. ‘인’재에 속한 세부 지표들은 좋으나 단 ‘여성 연구원’이 3년간 54위 → 53위 → 53위로 하위권

이며 개선의 징후가 안 보인다. 규제 여건에 속해 있는 ‘기술 개발 및 적용’ 지표는 3년간 44위 → 45위 → 48위로 하위권이고 좋아지지 않고 있다. 특별히 사업 능력에 속해 있는 ‘빅데이터 및 분석 기술 활용’은 3년간 15위 → 26위 → 34위로 매년 더 악화되고 있다. 빅데이터 분석과 활용을 더욱 적극적으로 장려할 필요가 있다.

### III. AI 경쟁력 변화 추이 분석

#### 1. 2021년 글로벌 AI 지수 순위

다음으로 미래 경쟁력의 핵심인 AI 경쟁력을 살펴보자. AI와 관련된 글로벌 평가 지수로는 스탠포드 AI 인덱스(미국 스탠포드 HAI(인간중심 AI) 연구소), 정부 AI 준비지수(영국 옥스퍼드 인사이트)도 있으나, 가장 권위 있는 평가 지수는 영국의 데이터 분석업체인 토터스 인텔리전스(Tortoise Intelligence)가 발표하는 ‘글로벌 AI 지수(Global AI Index)’이다. 이 지수는 2020년부터 발표가 시작되었으나 2021년도 결과부터 충실하므로 2021년부터의 결과를 살펴보기로 한다. 이 지수는 2023년에 개편되어 3년간(2021~2023)의 직접적인 비교에 어려움이 있다. 따라서 2021, 2022, 2023년 각각 조사 결과를 상세히 검토하기로 한다.

2021년도 이 글로벌 AI 지수 조사 결과에 따르면 총 7개 평가 부문에서 한국은 조사 대상 62개국 중에서 종합순위 5위로, 1~4위는 미국, 중국, 영국, 캐나다이고, 6~10위는 이스라엘, 독일, 네델란드, 싱가포르, 프랑스이다. 7개 평가부문의 한국의 순위와 내용은 다음과 같으며, 1~10위까지의 국가들의 각 부문별 순위는 <표 3.1>과 같다.

- 인재(talent) 19위: AI 인재 양성 및 현황
- AI 인프라(infrastructure) 6위: 슈퍼컴퓨팅 역량과 인터넷 기반 시설 등
- 운영환경(operating environment) 50위: 정부의 규제 수준 지표
- 연구(research) 18위: AI 특허 건수와 인용 횟수 등의 평가지표
- 개발 능력(development) 2위: AI 플랫폼과 알고리즘 등 기술 역량

- 정부 전략(government strategy) 7위: 연구 개발에 관여하고 투자하는 수준
- 상용화(commercial) 15위: AI 상용화 수준

〈표 3.1〉 2021년 '글로벌 AI 지수' 순위 (국가별, 부문별 순위)

국가	부문							종합 순위
	인재	인프라	운영환경	연구	개발능력	정부전략	상용화	
미국	1	1	35	1	1	10	1	1
중국	21	2	4	2	3	5	2	2
영국	4	19	11	3	11	14	4	3
캐나다	8	16	20	5	9	3	6	4
<b>한국</b>	<b>19</b>	<b>6</b>	<b>50</b>	<b>18</b>	<b>2</b>	<b>7</b>	<b>15</b>	<b>5</b>
이스라엘	5	28	29	7	12	51	3	6
독일	14	11	17	4	10	8	11	7
네덜란드	6	8	8	13	7	29	21	8
싱가포르	3	3	55	16	16	15	5	9
프랑스	13	12	6	11	14	6	9	10

AI 정부 전략과 관련하여 지난 문재인 정부에서는 혁신성장전략회의(2019년 1월 16일)를 통하여 '데이터·AI 경제 활성화 계획, 혁신성장 전략투자'를 발표하고, 한국판 뉴딜 국민보고대회(2020년 7월 14일)에서 '한국판 뉴딜' 정책, 그리고 뒤이어 '디지털 뉴딜 2.0'을 발표하여 데이터와 AI 활용 생태계 지원 정책을 펼쳐왔다. 또한 윤석열 정부에서도 국무회의(2022년 8월 22일)에서 관계부처 합동으로 '디지털 인재 양성 종합방안'을 발표하고, 곧이어 비상경제 민생회의(2022년 9월 28일)를 통해 '대한민국 디지털 전략'을 발표하는 등 새 정부의 미래 먹거리의 하나로 AI 육성 정책을 내 세우고 있다. 고무적인 일이며, AI 진흥이 기대된다.

총 7개 평가 부문 중 한국이 가장 높은 순위를 차지한 것은 AI 플랫폼과 알고리즘 등 기술 역량을 담은 지표인 '개발 능력' 부문으로 2위였다. 다음으로 슈퍼컴퓨팅 역량과 인터넷 기반 시설을 말하는 'AI 인프라'는 6위로 평가받았고, 국가가 연구 개발에 관여하고 투자하는 수준을 말하는 '정부 전략'에서는 7위로 평가받았다. 2016년 이후 출원된 특허 건수와 인용 횟수 등을 담은 평가지표인 '연구' 부문은 18위에 머물렀다. 정부의 데이터 관련 규제 강도와 AI에 대한 사회적 신뢰도 수준을 말하는 '운영환경'은 50위로 낮은 평가를 받았고, AI 인재 양성

및 현황을 말하는 ‘인재’ 부문에서도 19위로 비교적 나쁜 평가를 받았다. 또한 AI 산업화 수준을 말하는 ‘상용화’ 부문도 15위로 높지 않은 평가를 받았다.

종합적으로 평가할 때, 개발 능력, AI 인프라, 정부 전략은 우수한 편이나, AI 경쟁력을 한 단계 더 끌어올리기 위해서는 정부의 규제를 완화하여 운영환경을 좋게 하고, 인재 양성에 매진해야 하며, AI 상용화가 이루어지게 하고, AI 특허 건수와 인용 횟수 등을 끌어 올리는 것이 AI 경쟁력을 높이는 비결이 될 것이다.

종합순위에서는 미국과 중국이 전반적으로 주도권을 잡고 있고, 그다음으로 영국과 캐나다이다. 미국은 인재, 연구, 개발 능력, 상용화 등 4개 부문에서 1위를 차지하면서 종합순위 1위이다. 미국을 따라잡기 위한 중국의 추격도 거세다. 중국은 인프라에서 2위이고, 나머지 부문도 연구(2위), 개발 능력(3위), 정부 전략(5위), 상용화(2위) 등에서 대단한 기세이다. 만약 중국이 인재 부문(현재 21위)만 보강한다면 미국과 경쟁을 벌리게 될 것이다.

AI가 4차산업혁명 기술의 핵심인 만큼 주요국의 AI 경쟁이 한층 치열해지며, 한국도 속도를 더욱 높여야 하는 과제를 안게 되었다. 한국은 미국이나 중국에 비해 인구가 적고 시장 규모도 작고, 연구진도 불충분하다. 한국은 AI 논문 수에서 세계 9위, 논문 편당 인용수 31위, 원천기술 연구자 규모는 미국의 4%에 지나지 않는다. 그러나 IT 환경이 잘 갖춰져 있고, AI 기술이 전 산업군에 확산하는 속도가 매우 빠른 장점이 있다. 또한 반도체와 5G 통신 등 한국이 우위를 점한 분야를 중심으로 AI를 빠르게 상용화하는 전략으로 AI 경쟁력을 높여나갈 수 있는 장점을 가지고 있다. 여기에 정부의 공공데이터와 개인정보를 활용하는 데이터 산업이 성장할 수 있는 잠재력을 가지고 있고, 또한 AI 학습데이터와 AI 연구의 산출물(학습모델, 소스코드 등)에 대하여 통합, 관리, 재현할 수 있는 통합적인 체계 구축에서 순발력을 가지고 있다. 이런 인프라를 잘 활용한다면 한국의 AI 경쟁력을 앞으로 더 강화될 여지가 충분하다.

## 2. 2022년 글로벌 AI 지수 순위

2022년도에는 한국의 글로벌 AI 경쟁력은 2021년에 비하여 두 단계 하락한 7위를 기록했다. 그 자세한 내용을 요약하면 <표 3.2>와 같다.

〈표 3.2〉 2022년 ‘글로벌 AI 지수’ 순위 (국가별, 부문별 순위)

국가	부문							종합 순위
	인재	인프라	운영환경	연구	개발능력	정부전략	상용화	
미국	1	4	35	1	1	17	1	1
중국	24	1	6	2	2	2	2	2
영국	3	23	24	5	11	11	4	3
캐나다	7	15	5	10	10	1	6	4
이스라엘	5	29	14	7	9	45	3	5
싱가포르	4	8	55	4	14	15	5	6
<b>한국</b>	<b>28</b>	<b>6</b>	<b>32</b>	<b>12</b>	<b>3</b>	<b>7</b>	<b>15</b>	<b>7</b>
네덜란드	6	9	10	15	8	33	18	8
독일	11	13	30	6	12	10	8	9
프랑스	9	14	7	16	15	5	10	10

총 7개 평가 부문 중 한국이 가장 높은 순위를 차지한 것은 AI 플랫폼과 알고리즘 등 기술 역량을 담은 지표인 ‘개발 능력’ 부문으로 3위였다. 이 순위도 2021년에 2위였지만 중국에 추월당했다. 다음으로 슈퍼컴퓨팅 역량과 인터넷 기반 시설을 말하는 ‘AI 인프라’는 6위로 2021년과 동일하게 평가받았고, 국가가 연구 개발에 관여하고 투자하는 수준을 말하는 ‘정부 전략’에서는 7위로 2021년과 동일하게 평가받았다. 2017년 이후 출원된 특허 건수와 인용 횟수 등을 담은 평가지표인 ‘연구’ 부문은 12위로 2021년의 18위에서 6단계 상승했다. 정부의 데이터 관련 규제 강도와 AI에 대한 사회적 신뢰도 수준을 말하는 ‘운영환경’은 32위로 2021년의 50위보다는 많이 상승했으나 아직도 낮은 평가를 받았고, AI 인재 양성 및 현황을 말하는 ‘인재’ 부문에서도 28위로 2021년의 19위에서 9단계 하락하며 나쁜 평가를 받았다. 또한 AI 산업화 수준을 말하는 ‘상용화’ 부문도 15위로 2021년과 동일하게 높지 않은 평가를 받았다.

종합적으로 평가할 때, 2021년과 유사하게 개발 능력, AI 인프라, 정부 전략은 우수한 편이나, AI 경쟁력을 한 단계 더 끌어올리기 위해서는 정부의 규제를 완화하여 운영환경을 좋게 하고, 인재 양성에 매진해야 하며, AI 상용화가 이루어지게 하고, AI 특허 건수와 인용 횟수 등을 끌어 올리는 것이 AI 경쟁력을 높이는 비결이 될 것이다. 한국보다는 AI 경쟁력에서 뒤져 있지만 호주(11위), 일본(16위), 인도(17위), 대만(24위) 등의 추격도 앞으로 주목해 보아야 할 것이다.

종합순위에서는 미국과 중국이 각각 1, 2위로 전반적으로 주도권을 잡고 있고,

그다음으로 영국과 캐나다로 2021년과 같은 순위이다. 미국은 인재, 연구, 개발 능력, 상용화 등 4개 부문에서 1위를 차지하면서 종합순위 1위이다. 미국을 따라잡기 위한 중국의 추격도 거세다. 중국은 인프라에서 1위이고, 나머지 4개 부문(연구, 개발 능력, 정부 전략, 상용화)에서 2위를 달리는 등 대단한 기세이다. 특히 중국 정부의 전폭적인 지원(정부 전략 2위)에 힘입어 최강국 미국과 함께 'AI 굴기'를 과시하고 있다. 만약 중국이 인재 부문(현재 24위)만 보강한다면 미국을 가까이 따라잡을 수도 있을 것이다.

### 3. 2023년 글로벌 AI 지수 순위

토터스 인텔리전스는 최근(2023년 6월 28일)에 제4차 '글로벌 AI 지수'를 발표했다. 이 지수에 따르면 2023년도에 한국은 62개국 중 종합 AI 경쟁력은 6위를 기록해 2022년보다 한 단계 상승했다. 2022년까지는 7개 부문만 평가하였으나, 2023년부터는 두 개의 부문인 규모(scale)과 강도(intensity)를 추가하여 평가하였다. 규모는 한 국가의 전반적인 절대 AI 능력(absolute AI capacity)을 말하고, 강도는 한 국가의 인구와 경제의 크기에 비교한 AI 능력(AI capacity relative to the size of a country's population and economy)을 말한다. 규모로 보면 큰 나라인 미국이나 중국이 유리하고, 강도로 보면 작은 나라인 싱가포르나 이스라엘 같은 나라가 좋은 평가를 받는다. 평가 기준이 변경되어 2021~2023년 평가를 나라별로 직접 비교하기는 적절치 않다.

2023년부터 실시된 9개 부문에 대한 평가에 의한 글로벌 AI 지수 순위는 <표 3.3>과 같다. 2022년 평가 순위와 비교하면 2022년에 10위 안에 들어 있던 나라 중 2023년에 10위 안에서 빠진 나라는 네델란드(11위)와 프랑스(13위)이고, 2023년에 새로이 10위 안으로 들어온 나라는 스위스(9위)와 필란드(10위)이다. 이 두 나라는 '강도'에서 높은 평가를 받아 순위가 상승한 나라이다. 미국과 중국이 2022년과 2023년에서 모두 압도적인 1위와 2위를 차지하였는데, 9개 부문 중에서 6개 부문인 인프라, 연구, 개발 능력, 상용화, 규모에서 1위와 2위를 차지했다. 앞으로 상당 기간 이 두 나라는 AI 순위에서 높은 경쟁력을 가질 것으로 평가된다.

〈표 3.3〉 2023년 ‘글로벌 AI 지수’ 순위 (국가별, 부문별 순위)

국가	부문									종합 순위
	인재	인프라	운영 환경	연구	개발 능력	정부 전략	상용화	규모	강도	
미국	1	1	28	1	1	8	1	1	5	1
중국	20	2	3	2	2	3	2	2	21	2
싱가포르	4	3	22	3	5	16	4	10	1	3
영국	5	24	40	5	8	10	5	4	10	4
캐나다	6	23	8	7	11	5	7	7	7	5
<b>한국</b>	<b>12</b>	<b>7</b>	<b>11</b>	<b>12</b>	<b>3</b>	<b>6</b>	<b>18</b>	<b>8</b>	<b>6</b>	<b>6</b>
이스라엘	7	28	23	11	7	47	3	17	2	7
독일	3	12	13	8	9	2	11	3	15	8
스위스	9	13	30	4	4	56	9	16	3	9
필란드	13	8	4	9	14	15	12	13	4	10

한국의 경우 2022년과 비교하여 순위가 상승한 부문은 인재(28위 → 12위), 운영 환경(32위 → 11위), 정부 전략(7위 → 6위)이고, 순위가 하락한 부문은 인프라(6위 → 7위), 상용화(15위 → 18위)이다. 순위가 유지된 부문은 연구(12위), 개발 능력(3위)이다. 순위가 상승한 부문들로 인하여 종합순위가 7위에서 6위로 상승하였다.

## IV. 인재 양성 전략과 데이터과학

### 1. 디지털/AI 경쟁력에 나타난 인재 양성 평가

세계 속에서 우리나라가 어느 정도 인재 양성을 위한 교육 경쟁력을 가지고 있는가를 객관적으로 평가하는 것은 쉬운 일이 아니다. 앞에서 조사한 디지털/AI 경쟁력에 나타난 교육(인재 양성) 경쟁력 평가를 우선 살펴보자.

#### 1) 디지털 경쟁력에서 보는 인재 양성과 교육훈련

〈표 2.1〉을 보면 세계 종합적인 디지털 경쟁력에서는 8위이나 디지털 인재(talent) 부문은 33위에 그치고 있고, 교육훈련 부문은 16위를 차지하고 있다. 디

디지털 인재에서 떨어지는 세부 지표들은 디지털 기술 능력(46위), 외국인 숙련직 직원(49위), 국제 학생의 순 유입(38위) 등이다. 즉, 우리나라는 아직 디지털 분야에서 일하는 인력들이 디지털 기술 능력이 떨어지고, 우수한 외국인 숙련직 직원이 부족하며, 국제 학생의 순 유입이 부족하다. 따라서 우리나라는 저출산으로 고민이 많은 나라인 만큼 이민정책을 완화하여 외국인 전문가를 숙련직 직원으로 유치하는 것이 바람직하고, 또한 외국 유학생을 많이 초빙하여 외국 학생의 순 유입을 늘리고 이들을 교육시켜 국내에서 일하게 하는 것이 바람직하다.

〈표 2.1〉에서 보면 교육훈련은 한국이 세계 10위권의 경제 대국임을 감안할 때, 16위는 좋은 편은 아니다. 주요한 세부 지표를 보면 고등교육 성취도(4위)는 좋으나, 대학 이상 전문교육 학생:교수 비율은 30위로 좋지 않고, 여성 학위 소지자도 20위로 만족스럽지 않다. 따라서 전문교육에서 교수를 더 충원해야 하고, 여성 학위 소지자도 더 늘리는 교육정책을 실시해야 할 것이다.

## 2) 글로벌 AI 경쟁력에서의 교육 경쟁력

한국은 〈표 3.1〉에서 보는 바와 같이 2022년 ‘글로벌 AI 지수’에서 62개국 중 7위로 종합적으로 좋은 편이나, 조사 7개 부문 중에서 취약한 부문은 AI 인재 양성(28위), 정부의 규제 수준을 나타내는 운영환경(32위)이다. 증기기관이 산업혁명을 폭발시켰듯이 AI 혁명은 지금 인류 문명과 산업을 통째로 뒤흔들고 있다. 지난 10여 년간 AI 기술은 비약적으로 발전해 IT, 제조 산업, 의료, 제약, 식품 등 다양한 분야로 확산되고 있다. AI 분야 인재에서 뒤지면 순식간에 선진국 대열에서 이탈할 수도 있으므로 AI 교육 경쟁력에 철저히 대비해야 한다.

다행스럽게도 〈표 3.3〉을 보면 2023년 ‘글로벌 AI 지수’에서 인재 양성이 2022년의 28위에서 2023년에는 12위로 대폭 상승하고, 운영환경도 32위에서 11위로 대폭 상승했다. 이는 매우 바람직한 현상으로 AI 인재 양성이 잘 되고, 정부의 각종 규제도 완화되는 것은 AI 경쟁력을 키우는 데 큰 도움이 된다.

## 2. 디지털 전환 시대를 대비하는 정보 교육 현황

우리나라는 초·중·고등학교 교육에 많은 문제점을 노출하고 있다. 교육부·한국교육과정평가원(2022)에 발표한 ‘2021년 국가수준 학업성위도 평가’를 보면

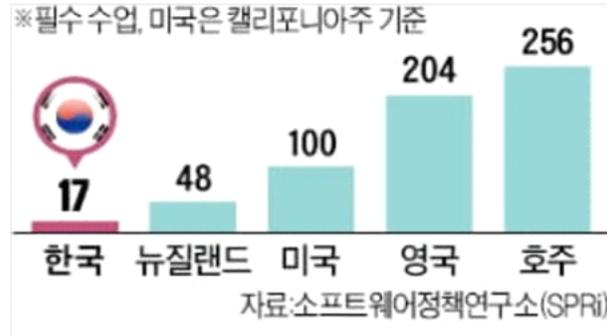
기초학력 미달자가 의외로 매우 많다. 또한 교육부·통계청(2023)이 발표한 '2022년 초·중·고등학교 사교육비 조사' 결과를 보면 사교육비가 급증하고 있다. 이런 문제점들은 빨리 그 보완책을 찾아야 할 것이다.

지금은 디지털 전환 시대이며, 데이터·AI 기반 초연결사회인 지금은 미래를 준비하기 위해서 국가적인 디지털 전략이 필요하며, 이를 뒷받침할 초·중·고등학교에서의 정보 교육은 무엇보다 중요하다. 정부는 지난 2022년 9월 28일에 광주 김대중 컨벤션센터에서 제8차 비상경제민생회의를 열고, 윤석열 대통령은 “디지털을 기반으로 다시 도약하는 대한민국을 만들겠다”며 ‘대한민국 디지털 전략’을 발표했다. 이 전략은 국가 디지털 정책을 이끌 핵심 방안을 제시한 것이다.

이 디지털 전략에서 초중고의 정보 교육과 관련된 중요한 내용은 디지털 인재를 양성하기 위해 정보·컴퓨터 교육 수업시수를 현재보다 두 배가량 늘려 초등학교 34시간(현재 17시간), 중학교 68시간 이상(현재 34시간 이상) 교육을 받을 수 있도록 한다는 것이다. 또한 디지털 6대 분야(AI, AI 반도체, 데이터, 사이버 보안, 메타버스, 개인정보 보호)에 대한 소프트웨어 중심대학과 대학원을 확대해 디지털 인재 100만 명을 양성할 계획이며, 2027년까지 데이터 시장(2021년 23조 원)을 두 배 이상인 50조 원 이상으로 확대 발전시키겠다는 것이다.

선진국을 중심으로 요즘 전 세계 학생들은 AI와 SW를 ‘열공’하고 있다고 볼 수 있다. 미국, 영국뿐 아니라 중국, 인도 등 신흥국 학생도 국어, 영어, 수학에 맞먹을 만큼 AI·SW 교육을 받는다. 디지털 전환 속도가 빨라지면서 미래 인재를 확보를 위해 각국이 앞다퉀 교육 시간을 대폭 늘렸기 때문이다. 반면 한국 초등학생의 AI·SW 교육은 한 달 1시간 남짓에 불과하고, 정보 교육 교사도 충분하지 않다. 일부 학부모와 학생들이 코딩학원을 찾지만, 그나마도 고등학교에 진학하면 대부분 그만둔다. 수학능력시험에 나오지 않기 때문이다.

이처럼 ‘디지털 선진국’으로 가는 우리나라의 AI·SW 교육 현실은 열악하다. 이러한 현실은 통계(<그림 4.1> 참조)에서 뚜렷하게 나타난다. 소프트웨어정책연구소(SPRi)에 따르면, 한국 초등학교 정보 교육 시간은 17시간에 불과하지만 미국(캘리포니아 기준)은 100시간, 영국은 204시간, 호주는 256시간에 달한다. 중학교의 경우 한국은 34시간인데, 영국(102시간), 이스라엘(110시간), 미국(135시간) 등 한국의 3배가 넘는다. 지난 2022년 8월에 교육부가 뒤늦게 2025년부터 정보 수업 시간을 초등학교와 중학교에서 현행보다 두 배로 늘리겠다고 발표했지만, 이조차 주요국에 비해 부족한 수준이다.



〈그림 4.1〉 국가별 초등학교 정보 교육 필수 시간 비교

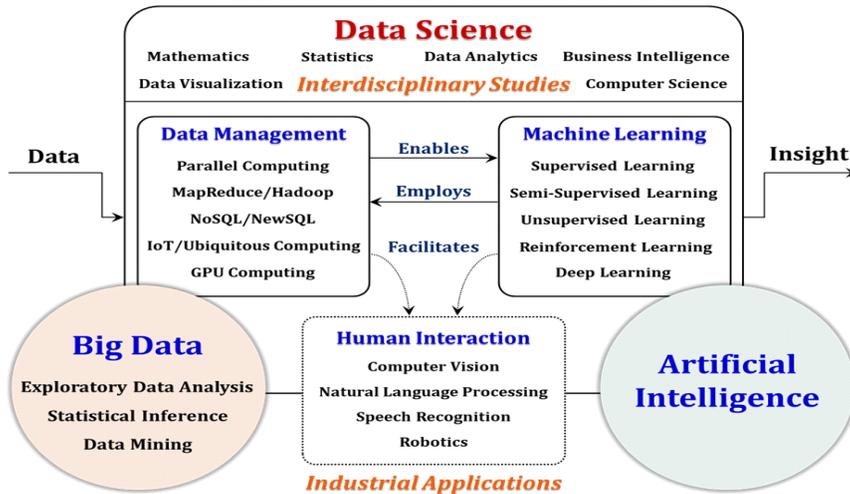
이번에 발표한 ‘대한민국 디지털 전략’으로 우리 국민의 디지털 리터러시(digital literacy)와 컴퓨팅 사고력을 중심으로 하는 교육으로 발 빠르게 전환되고, 데이터 산업, 디지털 산업, AI 산업 등이 발전하여 진정한 ‘디지털 강국’으로 한국이 부상하기를 바란다. 이 길만이 4차 산업혁명 시대인 지금 대한민국이 강국으로 발전해 갈 수 있는 길이라고 생각한다.

### 3. 데이터과학의 출현

#### 1) 데이터과학이란?

4차 산업혁명은 빅데이터, 인공지능, 사물인터넷(IoT), 메타버스 등으로 불리는 소프트웨어와 데이터 기반의 지능 디지털 기술(intelligent digital technology) 혁명을 말한다. 이 혁명은 우리의 산업과 생활방식을 급격히 바꾸어 놓고 있으며, 인류 문명에 주는 영향도 엄청나다. 4차 산업혁명의 근간에는 다량의 데이터를 소프트웨어와 연계하여 신속 정확하게 정보를 창출하고 활용하는 과학이 중요성을 띄고 있다. 이런 학문을 데이터과학(data science), 혹은 데이터 사이언스라고 부른다. 데이터과학은 데이터의 수집과 저장에 필요한 데이터 프로세싱 기술과 데이터 분석에 관한 지식(통계학, 데이터 마이닝, 기계학습 등)을 기반으로 다량의 데이터로부터 패턴을 찾아내고, 통계적 추정, 예측 모델링 등을 통하여 필요한 정보를 창출하고, 이를 실제로 활용하는 것을 연구하는 융합과학(convergence science)이다. 어쩌면 21세기 사회 발전에 가장 기여도가 큰 학문은 데이터과학이 될 것이라고 예측된다.

넓은 의미에서 데이터과학의 전체 조감도를 그려보면 <그림 4.2>와 같다. 데이터과학은 빅데이터와 인공지능(AI) 기술을 이용하여 통찰력(Insight) 있는 정보를 만들어 내기 위해 태어난 새로운 학문이다. 데이터과학의 밑바탕이 되는 기존의 학문은 수학, 통계학, 컴퓨터과학, 데이터 분석학(analytics), 데이터 시각화(visualization), 비즈니스 인텔리전스(intelligence) 등으로 명실공히 다학제 간 연구 학문이다. 데이터과학은 빅데이터를 처리·분석·저장할 수 있도록 물리적 환경을 조성해 주는 데이터관리(data management) 분야와 데이터로부터 모델을 만들어 AI로 구현할 수 있도록 해주는 기계학습(machine learning) 분야로 나누어진다. 모든 단계마다 사람이 개입하여 조절 및 통제하는 인간과의 상호작용(human interaction)이 적용되며, 이때 사용되는 기술로는 탐색적 데이터 분석, 컴퓨터 비전, 자연어 처리, 음성 인식, 로봇틱스 등이며, 궁극적으로는 다양한 산업에 적용되어 새로운 제품 및 서비스를 창출함으로써 인간의 삶을 풍요롭게 해준다.



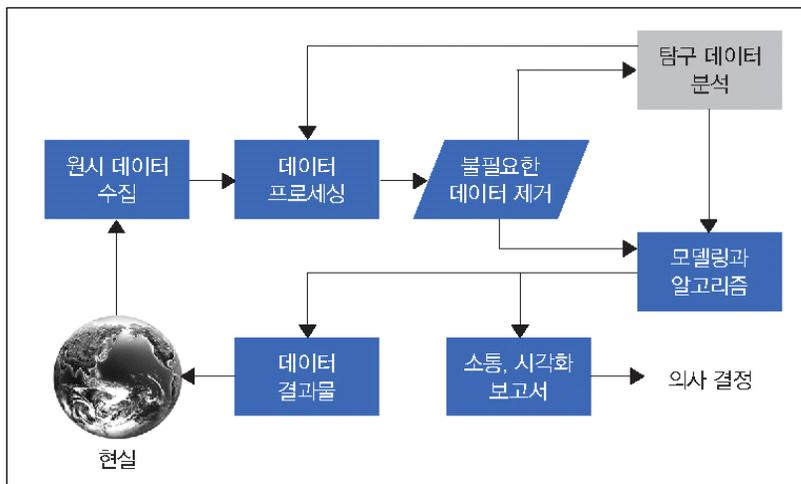
<그림 4.2> 데이터과학의 전체 조감도(자료: 한국과학기술한림원(2019))

## 2) 빅데이터란?

빅데이터(Big Data)는 데이터 형식(정형 데이터, 비정형 데이터 등)이 다양하고 생성 속도가 빨라 새로운 관리·분석 기법이 필요한 대용량의 데이터와 그 분

석 기술을 말하고, 흔히 대용량(Volume), 빠른 속도(Velocity) 및 높은 다양성(Variety)의 속성(3V)을 가진 정보 자원과 분석 기술로 정의한다. 여기에 추가하여 최근에는 데이터의 신뢰성(Veracity), 불안정성(Volatility)을 포함하여 5V의 속성으로 정의하기도 한다.

데이터과학의 핵심기술인 빅데이터는 그 분석의 흐름도로 Wikipedia 사전은 <그림 4.6>을 제시했다. 이 그림에서 보면 원시데이터를 먼저 수집(raw data collected)하고, 이를 프로세싱하여(data is processed) 정리하고 저장하고, 다음으로 잘못된 데이터를 제거하여 깨끗한 데이터(clean dataset)를 만드는 작업을 수행한다. 여기까지는 주로 전산 과학(computational science)의 연구 영역이다. 다음 단계는 탐색적 데이터 분석(exploratory data analysis)과 모델링(models & algorithms)으로 이는 통계학 및 응용수학의 연구 영역이다. 그다음 단계로는 소통하고 시각화하여 보고서(communicate, visualize and report)를 작성하고 의사결정에 중요한 정보를 제공하는 것이다. 이는 전문 학문 영역(예로, 경영학, 산업공학, 보건학 등)의 영역이라고 볼 수 있다. 데이터 결과물(data product)은 현실을 정확히 반영하여야 하며, 이는 다음 단계의 원시데이터 수집에 가이드라인 역할을 하는 것이다. 즉, 빅데이터 분석을 제대로 하려면 전산 과학, 통계학, 응용수학, 전문 학문(경영학, 산업공학, 보건학 등) 등의 협업이 필요한 경우가 대부분이다.



<그림 4.3> 빅데이터 분석의 흐름도 (출처: Wikipedia)

이제 4차 산업혁명의 정보폭발 시대를 맞아 다양한 경로로 취합된 다양한 데이터로부터 빅데이터 기술을 활용하여 숨은 정보와 새로운 지식을 발굴하고 혁신을 도모하려는 노력이 모든 분야에 확산되면서, 빅데이터와 데이터과학이 차세대 지식혁명을 이끌 주역으로 부상하고 있다. 이미 오래전 미국의 컨설팅 업체 McKinsey(2011)는 기업의 빅데이터 활용은 고객의 행동을 미리 예측하고 대처 방안을 마련해 기업 경쟁력을 강화시키고, 생산성 향상과 비즈니스 혁신을 가능하게 한다고 말했다. 이 주장은 정확히 맞아 들어가고 있으며, 지금은 기업 경영에서 빅데이터를 빼놓고 생각할 수 없는 시대가 되었다. 대표적인 빅데이터 분석 사례를 들어보자. 구글은 독감과 관련된 검색어 빈도를 분석해 독감 환자 수와 유행 지역을 예측하는 독감 동향 서비스를 개발했다. 이는 미국 질병통제본부보다 예측력이 뛰어난 것으로 밝혀졌다. 또한 구글은 수천만 권의 도서 정보와 유엔과 유럽의회, 웹 사이트의 자료를 활용해 64개 언어 간 자동번역 시스템 개발에 성공했다. 대단한 빅데이터 활용이 아닐 수 없다.

## V. 데이터 산업 현황과 진흥 전략

### 1. 데이터 산업이란?

한국데이터산업진흥원(2022)에서 발표한 「2021 데이터 산업 현황조사」에 의하면, 데이터 산업(data industry)은 '데이터의 생산, 수집, 처리, 분석, 유통, 활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업'으로 정의한다. 데이터의 가치사슬(혹은 생명주기) 상에 나타난 데이터와 관련된 제반 활동을 포함해 데이터로부터 가치가 창출되는 일련의 모든 과정, 이와 연관된 활동을 포함한다.

데이터 산업의 정의에 따라 데이터 산업 분류는 <표 5.1>과 같이 대분류는 3가지로, 데이터와 관련한 제품을 판매하거나 기술을 제공하는 데이터 처리 및 관리 솔루션 개발·공급업(빅데이터 포함), 데이터 구축 및 컨설팅 서비스업, 그리고 데이터를 판매하거나 이를 기반으로 정보제공 및 분석 서비스(클라우드 사업 포함)를 제공하는 데이터 판매 및 제공 서비스업이다.

〈표 5.1〉 데이터 산업 분류 (자료: 「2021 데이터 산업 현황조사」, 한국데이터산업진흥원)

대분류	중분류	내용
데이터 처리 및 관리 솔루션개발·공급업	11. 데이터 수집·연계 솔루션 개발·공급업	데이터 검색 솔루션, 데이터 수집 솔루션, 데이터 통합·연계 솔루션, 데이터 교환·개방 솔루션
	12. 데이터베이스 관리시스템 솔루션 개발·공급업	관계형 데이터베이스 관리 솔루션, 비관계형 데이터베이스 관리 솔루션, 인메모리 데이터베이스 관리 솔루션, 데이터베이스 어플라이언스 솔루션, 기타 데이터베이스 솔루션
	13. 데이터분석 솔루션 개발·공급업	정형·비정형 데이터분석 솔루션, 실시간 데이터분석 솔루션, 데이터 시각화분석 솔루션, 데이터처리 솔루션
	14. 데이터관리 솔루션 개발·공급업	데이터 모델관리 솔루션, 데이터 표준관리 솔루션, 데이터 품질관리 솔루션, 데이터베이스 운영·성능관리 솔루션
	15. 데이터 보안 솔루션 개발·공급업	데이터 저장소 보안 솔루션, 개인 데이터 보안 솔루션
	16. 빅데이터 통합 플랫폼 솔루션 개발·공급업	빅데이터 플랫폼 솔루션
2. 데이터 구축 및 컨설팅 서비스업	21. 데이터 구축·가공 서비스업	데이터베이스 설계·구축 서비스, 데이터 이행 서비스, 데이터 가공 서비스
	22. 데이터 관련 컨설팅 서비스업	데이터 설계 컨설팅, 데이터 품질 컨설팅, 데이터베이스 성능개선 컨설팅, 데이터 거버넌스 컨설팅, 데이터 분석·활용 컨설팅
3. 데이터 판매 및 제공 서비스업	31. 데이터 판매·중개 서비스업	데이터 판매 서비스, 데이터 중개 서비스, 분석 데이터 제공 서비스
	32. 정보제공 서비스업	포털·정보매개 서비스, 정보제공 서비스

## 2. 데이터 산업의 국내외 현황

### 1) 국내 데이터 산업의 시장 규모와 인력

한국데이터산업진흥원(2023)이 2023년 4월 18일 발표한 「2022 데이터산업 현황조사」에 따르면, 데이터 산업의 시장 규모는 〈표 5.2〉에서 지난 6년간 매년 급성장하고 있으며, 이런 추세는 계속 이어질 것으로 보인다. 특히 최근 3년간('20~22(E)년)의 연평균성장률(CAGR)은 11.9%로 매우 유망한 산업 분야이다. 2022년 기준 데이터 산업의 부문별 시장 규모(E는 추정치)를 보면 데이터 판매 및 제공 서비스업이 12조 5,235억 원(비중 50.0%), 데이터 구축 및 컨설팅 서비스업이 9조 2,570억 원(37.0%), 그리고 데이터 처리 및 관리 솔루션 개발·공급

업이 3조 2,723억 원(13.0%)을 차지하고 있다.

〈표 5.2〉 데이터 산업 시장 규모 (2016~2021(E)) 요약 (단위: 억 원)

구 분	2016년	2017년	2018년	2019년	2020년	2021년	2022년 (E)	증감률 '20~'21	CAGR '20~'22(E)
데이터 처리 및 관리 솔루션 개발·공급업	15,720	16,457	18,617	20,805	25,133	29,843	32,723	18.7%	14.1%
데이터 구축 및 컨설팅 서비스업	55,850	58,894	61,290	65,412	76,999	85,274	92,570	10.7%	9.6%
데이터 판매 및 제공 서비스업	65,977	68,179	75,778	82,364	97,891	113,869	125,235	16.3%	13.1%
<b>전체</b>	<b>137,547</b>	<b>143,530</b>	<b>155,684</b>	<b>168,582</b>	<b>200,024</b>	<b>228,986</b>	<b>250,527</b>	<b>14.5%</b>	<b>11.9%</b>

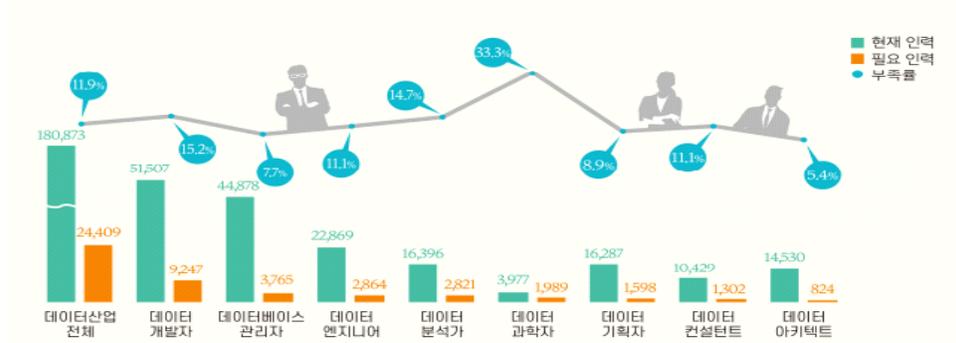
데이터 산업에 종사하는 인력은 〈표 5.3〉에 표시된 바와 같이, 2022년 기준 406,457명으로, 이 중 데이터 직무인력이 134,310명이고, 데이터 직무 외 인력이 272,147명이다. 여기서 데이터 직무인력이란 데이터 산업에 종사하는 데이터 개발자, 데이터베이스 관리자, 데이터 엔지니어, 데이터 분석가, 데이터 과학자, 데이터 기획자, 데이터 컨설턴트, 데이터 아키텍트 등 데이터 관련 업무를 수행하는 전문 인력을 말한다. 데이터 직무 외 인력이란 데이터 산업에 종사하는 인력 중에서 데이터 직무인력을 제외한 인력을 말한다.

데이터 산업 전체 인력이 2016년에는 288,621명이었으나 2022년에 406,457명으로 40.8% 증가하였다. 데이터 직무 인력만 보면 2016년에 73,256명에서 2022년에 134,310명으로 83.3%라는 경이적인 증가를 보이고 있다. 일자리 창출에서도 데이터 산업이 우리나라의 유망산업이라는 증거를 데이터로도 볼 수 있다.

〈표 5.3〉 2015~2021년 데이터 산업 인력 현황(자료: 2022 데이터 산업 현황조사)

구 분	2016년	2017년	2018년	2019년	2020년	2021년	2022년	증감률 '21~'22	CAGR '20~'22
데이터 직무인력	73,256	77,105	82,623	89,058	101,967	122,431	134,310	9.7%	14.8%
데이터 직무인력 외	215,365	217,648	235,439	255,614	264,054	272,278	272,147	0.0%	1.5%
<b>전체</b>	<b>288,621</b>	<b>294,753</b>	<b>318,062</b>	<b>344,672</b>	<b>366,021</b>	<b>394,709</b>	<b>406,457</b>	<b>3.0%</b>	<b>5.4%</b>

데이터 산업의 데이터 직무별 인력 현황과 수요를 보면 향후 5년(2022~2026년) 내 필요 인력은 24,409명으로 조사되었으며, 부족률은 11.9%에 달한다. <그림 5.1>에서 보는 바와 같이 부족률이 가장 높은 직무는 데이터 과학자로 부족률이 33.3%에 달한다. 여기서 부족률이란 [필요 인력]/[현재 인력 + 필요 인력] x 100으로 계산한 것이다. 데이터 과학자의 경우 [1,989]/[3,977 + 1,989] x 100 = 33.3%가 얻어지는 것이다. 인력이 크게 부족한 직무는 데이터 과학자 다음으로는 데이터 개발자(15.2%), 데이터 분석가(14.7%) 등의 순이다.



<그림 5.1> 데이터 직무 인력 현황 및 수요(자료: 2021 데이터산업 현황조사)

「2021 데이터 산업 현황조사」에 따르면 2020년도에 데이터 산업은 아니나 '데이터 산업 관련 업종'의 시장 규모는 106조 4,337억 원으로, 2020년 데이터 산업 시장 규모(20조 24억 원)보다 약 5.3배 정도 크며, 이들은 데이터 산업의 진흥에 큰 역할을 담당하고 있다. 데이터 산업 관련 업종으로 다음의 3가지가 있으며, 이들의 2020년도 시장 규모도 매우 크다.

- SW 개발 및 공급업 : 시장 규모 59조 6,891억 원
- 컴퓨터 프로그래밍, 시스템 통합 및 관리업 : 시장 규모 25조 9,277억 원
- 정보서비스업 : 시장 규모 20조 8,169억 원

-----  
 데이터 산업 관련 업종 : 시장 규모 합계 106조 4,337억 원

2020년도에 데이터 산업의 시장 규모(20조 24억 원)와 데이터 산업 관련 업종의 시장 규모(106조 4,337억 원)가 전체 우리나라 산업의 시장 규모(4,116조

9,953억 원)에 비하여 아직은 그 비중이 각각 0.49%, 2.59%로 작으나, 앞으로 그 비중이 더욱 커질 것이며, 이들 산업은 4차 산업혁명 기술들의 핵심적인 기초 기술에 해당하므로, 데이터산업과 그 관련 산업의 발전은 매우 중요하다.

## 2) 데이터 산업의 국제 현황

다음으로 데이터 산업의 국제 현황을 살펴보자. <표 5.4>의 주요국(미국, 유럽 연합, 일본, 브라질, 중국)의 데이터 산업 시장 규모를 보여주고 있다. 이 자료는 「2021 데이터 산업 현황조사」에서 발췌한 것이다. 원자료는 EU집행위원회(2022)가 발간한 「유럽 데이터 시장 연구(European Data Market Study)」에서 얻은 것이다.

<표 5.4> 2016~2021년 주요국 데이터 산업 시장 규모(Billion Euro)

구분	2016	2017	2018	2019	2020	2021 (E)	5년 평균 성장률(%)
미국	129.2	147.0	164.0	184.9	213.5	240.0	13.2
EU	61.8	68.8	77.3	83.5	91.3	94.4	8.8
일본	25.5	26.7	29.8	32.9	36.6	40.0	9.4
브라질	6.0	7.0	7.4	7.9	8.4	8.9	8.1
중국	-	-	-	24.2	27.5	31.7	-
합계	222.5	249.5	278.5	309.2	377.3	414.8	13.3

(자료 : EU집행위원회에서 발간한 The European Data Market Study 2021~2023  
First Report on Facts and Figures (IDC & The Lisbon Council, 2022.2.))

EU집행위원회가 발간한 이 자료에 따르면, 미국의 데이터 산업 시장 규모는 2021년 2,400억(추정치(E)) 유로로 2020년 대비 12.4% 증가했으며, 5년(2017년~2021년) 평균 성장률이 13.2%의 고속 성장이었다. EU는 2021년 기준 미국의 약 39.3% 수준의 시장 규모를 유지하고 있으며, 5년 평균 성장률이 8.8% 정도의 성장률을 유지하고 있다. 2021년 한국의 데이터 산업 규모(236조 972억 원 = 약 1,760억 유로)는 미국의 데이터 산업 시장 규모 대비 약 7.3% 수준이고, 유럽의 데이터 산업 시장 규모 대비 약 18.6%인 것으로 산출된다. 우리나라의 데이터 산업이 급성장하고 있으므로 조만간 미국에 비해 10%의 벽을 넘어설 것으로 예상된다.

다른 산업에 비하여 우리나라의 데이터 산업은 해외 진출이 적으며 주로 국내용이다. 우리나라의 데이터 산업 중에서 해외 시장에 진출한 경험이 있는 기업은 6.9%에 불과하다. 데이터 산업 부문별로는 데이터 처리 및 관리 솔루션 개발·공급업이 가장 높은 7.4%이고, 데이터 구축 및 컨설팅 서비스업이 9.2%이며, 데이터 판매 및 제공 서비스업이 가장 작은 4.3%이다. 해외시장 진출 경로는 직접 진출 유통(58.0%) 방식이 가장 높게 나타났고, 현지 법인과의 협력(25.0%), 온라인 판매(12.0%) 순으로 나타났다.

우리나라 데이터 산업의 해외 국가별 매출 발생 비중은 일본(36.6%), 베트남(32.4%)이 가장 크고, 다음으로 중국(28.2%), 미국(21.1%) 등의 순으로 나타나, 아시아 국가들의 비중이 높다. 해외 시장 진출 시 애로사항으로는 ‘해외 거래처 및 바이어 발굴 역량 부족’이 47.9%로 가장 높고, 다음으로 ‘현지 제휴 합작투자 파트너 발굴 곤란(32.4%)’, ‘해외시장 진출을 위한 기술경쟁력 확보 어려움(31.0%)’ 등의 순으로 나타났다.

### 3. 데이터 산업의 육성 및 진흥 전략

앞에서 설명된 데이터 산업의 국내외 현황 등을 감안하여 데이터 산업의 진흥 전략으로 다음의 다섯 가지를 제안한다.

#### 1) ‘데이터 컨트롤 타워’ 구축

우리나라는 공공데이터 작성에서 분산형 통계 시스템을 가지고 있어서 2023년 6월 24일 기준으로 435개 통계작성기관에서 1,306종의 승인 통계(조사통계 575종, 보고통계 481종, 가공통계 250종)를 생산하고 있다. 통계청이 통계법에 따라 승인 통계를 관리하고 있으나 효율적인 통계 관리는 역부족인 상태이며, 모든 공식 통계를 소위 ‘데이터 댐’에 모으고 사용하기 편리하도록 운영하기 위해서는 ‘데이터 컨트롤 타워’가 필요하다. 데이터 컨트롤 타워의 구축은 데이터 산업에 필요한 데이터의 보급에 중요한 인프라 역할을 할 것이다. 통계청이 현재는 기획재정부 산하의 외청이나 이를 국무총리실이나 대통령 직속 기관으로 옮겨 권한을 강화하여 실질적인 데이터 컨트롤 타워가 되도록 하는 것이 좋은 방안이 될 수 있다.

## 2) 개인정보의 활용과 빅데이터 도입률 제고

데이터 중 큰 부분인 개인정보는 그동안 데이터 3법(개인정보보호법, 정보통신망법, 신용정보법)에 막혀 개인정보 활용 사업에 발전이 없었다. 그러나 다행스런 계도 2020년 1월 9일에 데이터 3법 개정안이 국회를 통과하였고, 그 시행령이 제정되었다. 이 개정안에서는 가명 정보(특정인을 식별할 수 있는 정보의 상당 부분을 가린 정보)를 통계 작성, 공익적 기록 보존, 과학적 연구 등에 정보 소유자의 사전 동의 없이 사용할 수 있도록 했다. 또한 개정안에서는 익명 정보(가명 정보보다 개인을 조금이라도 식별할 수 있는 정보는 모두 삭제한 정보)는 개인정보 보호 대상에서 완전히 제외했다. 그러나 아직도 개정된 데이터 3법이 제대로 시행되고 있지 않으며, 개인정보 활용에 어려움이 해소되고 있지 못하다. 데이터 산업의 진흥을 위해 개인정보 활용도를 제고할 필요가 있다.

한국데이터산업진흥원(2021)에서 실시한 「빅데이터 시장 현황 조사」에 의하면 2020년 국내 기업의 빅데이터 도입률은 13.4%에 불과하며, 매출 1,000억 원 이상 기업에서도 그 도입률은 35.0% 수준이다. 우리 기업들에서 빅데이터 도입률을 올리고, 그 활용률을 제고하는 것은 데이터 산업의 진흥에 크게 기여할 것이다.

## 3) 데이터 산업 인력 정예화

앞의 <그림 5.3>에 있는 데이터 직무별 인력 현황 및 수요를 보면, 데이터 산업에 종사하는 인력은 2021년 기준 180,873명이고 향후 5년 내 필요 인력은 24,409명으로 부족률이 11.9%에 달했다. 이 중에서 특히 부족한 인력은 데이터 과학자(31.4% 부족), 데이터 개발자(14.5% 부족), 데이터 분석가(14.5% 부족) 등이다. 이들 부족 인력은 데이터 산업의 인력 정예화에 필수적인 인력들이다. 데이터 과학자 등을 충분히 확보하여 인력 정예화를 시키고 생산성을 제고하는 것은 데이터 산업 진흥의 첩경이다.

## 4) 플랫폼 간 표준화의 필요성과 데이터 품질관리

플랫폼 간 메타데이터(meta data, 데이터를 효율적으로 이용하기 위해 구조화한 데이터 정보)의 공유·활용이 가능하도록 표준화 및 연계 규격이 필요하다. 각

분야별로 많은 종류의 빅데이터가 작성되어 있으나 이들 간에 데이터 표준화가 미비하여 통합 작업이 어렵고 정보 교환에 애로가 많다. 누구나 쉽게 플랫폼들에 축적된 데이터를 찾고 상호 비교 활용 가능하도록 통합 데이터 지도에 반영하는 방식은 데이터 산업 진흥에 필수적인 요소이다.

데이터의 중요한 품질은 정확성, 시의성, 비교성, 접근성, 편의성 등이며, 이런 품질의 가치를 유지하기 위해서는 데이터에 대한 지속적인 품질관리(quality control)는 필수적이다. 오래전부터 통계에서는 ‘쓰레기 데이터가 들어가면 쓰레기 정보만이 나온다(garbage data in, garbage information out)’란 말이 있다. 즉 품질관리가 안 된 데이터가 들어가면 좋은 정보 창출이 안 된다는 뜻이다. 따라서 데이터를 생산하는 조직은 데이터의 전 생명주기에서 데이터의 품질관리에 대한 책임을 수행해야 한다. 통계청은 오래전 정책과제(박성현(2003))을 통하여 통계품질관리 표준매뉴얼을 만들었다. 통계 데이터의 중요한 품질은 확보하기 위하여 통계 작성 과정에서 어떻게 품질관리를 해야 할지를 제시한 것이다.

## 5) 클라우드 시장 국내 기업 점유율 제고

지금은 데이터 홍수 시대로 개별 회사가 관리하고 싶은 막대한 양의 빅데이터를 모두 자체적으로 개별 전산실 서버(중앙 컴퓨터)에 보관하고 이용하는 것은 한계에 도달했다. 클라우드 컴퓨팅(cloud computing) (간단히 클라우드라 칭함)은 자체 전산실 서버를 이용하는 대신 타 회사의 대용량 데이터 센터 저장 공간을 빌려 데이터를 저장하고 관리하는 서비스로, 필요한 만큼만 빌려 쓸 수 있고, 인터넷에 접속하면 어디서나 업무가 가능하다. 클라우드는 대용량의 가상 저장 공간인 구름(cloud) 속에 데이터를 저장해 두고 필요할 때만 꺼내서 이용한다는 뜻에서 시작된 비즈니스 용어이다. 4차 산업혁명 시대에 접어들면서 인공지능(AI), 비즈니스 데이터 플랫폼, 자율주행차, 스마트시티 등으로 기업이 다뤄야 할 데이터의 양이 폭증하는 점도 클라우드 시장을 키우는 원인이 되고 있다.

과학기술정보통신부 산하에 있는 한국클라우드산업협회(KACI)가 2023년 3월에 발간한 ‘2022 국내 클라우드 산업 실태조사 결과 보고서’에 따르면, 국내 클라우드 시장 규모는 2019년에 3조 3,714억 원이었으나, 2020년에 4조 원을 돌파하고, 2021년에는 약 5조 원에 육박하는 4조 9,250억 원으로 급신장하고 있다고 한다. 클라우드 사업은 매우 중요하고 유망한 사업임에 틀림없다.

문제는 4차 산업혁명의 중요 기술들인 빅데이터, AI, 자율주행차 등의 발전에 우리가 만든 데이터가 필요하며, 이들 데이터가 대부분 클라우드에 보관되고, 국내에서 활동하는 클라우드 기업이 거의 외국 기업이라는 것이다. 우리가 만든 데이터에 대한 '데이터 주권'을 상실할 수 있고 잘못하면 '데이터 식민지'가 될 우려도 있다. 글로벌 클라우드 시장을 점유하고 있는 기업들은 빅3 클라우드 업체(미국의 아마존웹서비스(AWS), 마이크로소프트(MS)의 Azure, 구글의 GCP)와 신흥 도전자로 불리는 알리바바(중국), IBM(미국), 텐센트(중국), 바이트댄스(중국) 등이다. 물론 한 기업이 외국 클라우드 기업에 맡긴 빅데이터를 그 기업이 자유롭게 사용할 수는 있으나, 외국 클라우드 기업들은 자국 정부에서 특별한 이유를 들어 요구할 시 자기가 관리하는 빅데이터를 자국 정부에게 보여줄 의무도 가지고 있다. 따라서 한 나라의 빅데이터를 다른 나라 정부에서 자세히 들여다보는 것이 가능하다.

국내 클라우드 시장 점유율은 빅3 외국 클라우드 업체가 국내 시장의 상당 부분을 잠식하고 있는 것으로 알려져 있다. 그러나 국내 클라우드 기업들도 조금씩 점유율을 높여가고 있다. 국내 클라우드 업체로는 KT가 앞서가고 있으며, 그 뒤에 네이버, 카카오, NHN, 삼성 SDS, LG유플러스, 가비아 등의 국내 IT 업체들이 클라우드 서비스를 지속적으로 키워나가고 있다.

클라우드 시장 선점을 위해 중요한 것 중 하나는 데이터를 저장하고 관리하는 물리적 장소인 데이터 센터이다. 아마존의 AWS는 가장 일찍이 2016년 1월에 서울 가산동에 데이터 센터를 지어 시장을 선점하고 있고, MS와 구글도 2016년에 국내 클라우드 시장에 진출했다. 구글은 2016년에 LG 유플러스와 데이터 센터 임대 계약을 맺어 클라우드 서비스를 시작했다. 국내 클라우드 시장을 장악하고 있는 외국 기업에 맞서려면 국내 업체의 데이터 센터 구축 확대가 필수적이다. 국내 클라우드 기업들의 데이터 센터 설치와 좀 더 적극적인 클라우드 사업을 확대해 나가기를 바란다.

## VI. 데이터 강국 발전 전략

데이터 강국으로 발돋움하기 위해서는 앞 장에서 논의한 데이터 산업의 육성 및 진흥 전략이 중요하다. 그러나 데이터 산업을 육성하면서 큰 틀에서 데이터 강국으로 발전하기 위해서는 다음과 같이 일곱 가지 전략이 필요하다. 이들에 대해서 하나씩 살펴보자.

### 1. 데이터과학 교육 진흥과 데이터·AI 인력 100만 명 양성

#### 1) 데이터과학 교육의 진흥

데이터과학(data science)은 <그림 4.2>에서 설명한 바와 같이 데이터의 수집과 저장에 필요한 데이터 프로세싱 기술과 데이터 분석에 관한 지식을 기반으로 다량의 데이터로부터 패턴을 찾아내고 필요한 정보를 창출하며 이를 인공지능(AI) 등 실제 데이터 기반 기술에 활용을 연구하는 융합과학이다. 따라서 데이터과학의 핵심기술은 빅데이터와 AI이고, 데이터과학은 데이터·AI 전문 인력을 양성하는 학문이라고 간단히 말할 수 있다.

디지털 전환 시대의 디지털 인재는 디지털 신기술(빅데이터, AI, 클라우드, 일반 소프트웨어(SW), 메타버스, 사물인터넷, 사이버 보안, 5G/6G 등)을 다룰 수 있는 인재를 말한다. 데이터과학 인재는 주로 빅데이터, AI, 클라우드 중심으로 데이터 기반 기술 인재이므로, 디지털 인재는 데이터과학 인재를 포함하는 개념이다. 즉, 데이터과학 인재는 디지털 인재 중에서 핵심적인 데이터 기반 기술을 다루는 인재를 말한다.

지난 2022년 8월 22일에 윤석열 정부에서 관계부처합동(2022)으로 발표한 '디지털 인재 양성 종합방안'은 매우 시의적절하고 중요한 디지털 인재 양성 전략으로, 2026년까지 100만 디지털 인재 양성 방안을 제시하고 있다. 디지털 인재 중 핵심 인력을 데이터과학 인재라고 볼 때 2026년에는 데이터과학 인재가 최소 60만 명 정도가 필요하고, 더 나아가 2030년에는 100만 명 정도가 필요하게 된다.

데이터과학 교육은 대학에서 인문·사회과학 분야를 포함한 모든 신입생에게 가르치는 주요 교양과목이 되어야 한다. '디지털 인재 양성 종합방안'에서 밝힌 바와 같이 데이터과학 관련 신규 학과도 많이 증설되어야 하고, 데이터과학 관련

대학원도 더 많이 증설되어야 한다. 그러나 대학 안에서의 인재 교육만으로는 데이터과학 인재의 수요를 채울 수 없다. 따라서 대기업의 연구소에서도 자체 교육 프로그램을 가동할 필요가 있으며, 기업 내에서 직원들을 대상으로 하는 데이터과학 교육이 광범위하게 이루어져야 한다.

또한 중요한 것은 일반인들을 상대로 한 ‘데이터 사회교육 프로그램’이 작동되어야 한다. 일반인이 취득할 수 있는 가칭 ‘데이터과학 기사(혹은 기술사)’ 자격증 제도를 신설하고, 일반인이 적절한 교육을 이수 후에 국가 공인 시험을 보아 합격하면 자격증을 주는 제도를 말한다. 이런 제도를 운영하면 사회교육 프로그램이 성공할 수 있으며, 데이터과학 인재를 양성하는 데 큰 역할을 할 것이다.

디지털 인재 양성 추진이나 데이터과학 인재 양성 추진에서 ‘디지털 전환’에 관한 마인드를 갖도록 인재를 양성하는 것이 중요하며, 이를 위해서는 ‘디지털 리터러시(digital literacy)’를 길러야 한다. 디지털 리터러시를 디지털 문해력이라고도 부르며, 이는 글을 읽고 쓸 줄 아는 것처럼, 디지털 정보 체계와 디지털 기술을 활용할 줄 아는 역량을 의미한다. 따라서 디지털 문해력이 부족하면 ‘디지털 문맹’이라고 말할 수 있다. 과거의 인재는 ‘아날로그(analog; 연속) 논리에 기반한 시스템의 이해와 활용만으로 가능했지만, 지금의 디지털 전환 시대의 인재는 이산(離散, discrete) 논리(디지털 컴퓨터가 이해하는 코딩 가능한 논리)에 다른 시스템이나 플랫폼 사용에 익숙한 인재를 양성하여야 한다. 이산 논리의 기본적인 이론을 다루는 분야는 이산수학(discrete mathematics)으로, 고등학교 수학이나 통계학 과목에서 이산수학의 개념을 강조하여 교육할 필요가 있다. 따라서 문과(인문, 사회 등), 이과(과학, 공학 등), 예체능의 교과 내용이 초등교육부터 중등학교까지 이산 논리가 바탕을 이루는 교과로 내용이 구성되는 것이 바람직하다.

## 2) 데이터·AI 전문 인력 100만 명 양성

전(全)산업에서의 데이터 직무 인력수를 살펴보면 <그림 6.1>에서 보는 바와 같이 2016년에 102,375명이었던 것이 2022년에는 197,802명으로 6년간 93.2% 증가하였다. 데이터 직무인력은 데이터 아키텍트, 데이터 개발자, 데이터 엔지니어, 데이터 분석가, 데이터베이스 관리자, 데이터 과학자, 데이터 컨설턴트, 데이터 기획자 등을 말한다. 2021년에 필요한 데이터 직무인력을 다 채우지 못한 기업이 많으므로(<그림 5.1> 참조), 추가로 필요한 데이터 직무인력을 최소 2만 명

정도로 본다면, 2022년도의 필요한 데이터 직무인력은 22만 명 정도이다. 이는 2016년과 비교하여 약 120%의 인력 증가가 필요한 것이다. 데이터 직무인력은 기하급수적으로 증가하는 추세이므로, 이런 추세로 나가면 2026년에는 60만 명 정도가 필요하고, 2028년에는 80만 명 정도가 필요하다. 따라서 지금부터 7년 후인 2030년에는 100만 명 정도의 데이터 직무인력이 필요하게 된다.

〈표 6.1〉 2016~2022 전(全)산업 데이터 직무인력 현황 (단위: 명)

(자료: 2022 데이터 산업 현황조사)

구 분	2016년	2017년	2018년	2019년	2020년	2021년	2022년	증감률 '21~'22	CAGR '20~'22(E)
일반산업	29,119	32,215	35,104	41,775	48,210	58,442	63,492	8.6%	14.8%
데이터산업	73,256	77,105	82,623	89,058	101,967	122,431	134,310	9.7%	14.8%
전산업	102,375	109,320	117,727	130,833	150,177	180,873	197,802	9.4%	14.8%

넓은 의미의 데이터과학은 AI를 포함하고 있고, 빅데이터와 AI가 연계된 기술로 발전하고 있으므로, 데이터 직무인력은 빅데이터, AI 등을 포함하여 데이터와 관련된 모든 지식을 갖추어야 할 것이다. 따라서 데이터 직무인력은 데이터에 기반한 AI 기술도 다룰 수 있어야 한다. 즉, 데이터 직무인력은 통계, 컴퓨터, 빅데이터, AI 등 넓은 의미의 데이터과학에 관한 지식을 가진 인력을 말한다. 이런 인력을 데이터·AI 전문 인력이라고 부를 수 있고, 이런 인력을 2030년(8년 후)까지 100만 명을 양성해야 할 것이다.

데이터·AI 전문 인력은 데이터·AI 연구전문가와 데이터·AI 활용전문가로 나눌 수 있다. 전자는 대학이나 연구소 등에서 데이터·AI 이론연구나 핵심 기술을 연구하는 인력이다. 그리고 후자는 기업(산업 현장)이나 사회에서 데이터·AI 기술을 활용하는 전문가이다. 전자의 양성은 시간이 소요되나, 후자의 양성은 비교적 단기간에 양성 가능하다. 단기간 양성은 기업에서의 인력 재교육 프로그램, 사회에서는 전문 학원이 일반인을 모집하여 훈련 프로그램을 운영하면 가능하다. 기업에서는 직접 자체 인력을 단기간에 훈련시키는 각종의 프로그램을 운영할 수도 있고, 대학에 위탁하여 교육시킬 수도 있다. 이미 삼성그룹이나 LG 그룹은 이러한 자체 양성 프로그램을 운영하고 있다. 이런 양성 과정에서 교수를 확보하기 어려우면 MOOC를 이용한 강의 프로그램을 운영하는 것도 좋을 것이다. 교육

방식(대학/연구소 교육, 사회교육, 기업 내 교육)으로 나누어 2023년부터 2030년까지 대략적인 인력 양성 로드맵을 제시해 보면 <표 6.2>와 같다.

<표 6.2> 데이터·AI 전문 인력 100만 명 양성 로드맵

(7년 후인 2030년까지, 단위: 명)

교육 방식/연도	2023년	2024년	2026년	2028년	2030년
대학/연구소 교육	7만	12만	16만	20만	25만
사회교육 프로그램	8만	18만	30만	44만	55만
기업 내 프로그램	5만	10만	14만	16만	20만
합계	20만	40만	60만	80만	100만

대학/연구소에서 데이터·AI 전문가 인력 양성은 데이터사이언스대학원, AI 대학원, 컴퓨터공학과, 통계학과 등에서 학사 및 석사, 박사 등의 인력을 양성하는 것을 말한다. 사회교육 프로그램을 가동시키기 위해서는 정부 기관에서 가칭 ‘데이터·AI 전문기사’ 자격증 제도를 만드는 것이 빠른 방법이다. 희망하는 인력이 민간 학원에서 이와 관련된 교육을 이수하고 자격증 시험을 보아 합격하면 전문기사 자격증을 주는 제도를 말한다. 기업이 이들 전문기사를 채용하면 민간인들에게 인기가 있을 것이며, 인재 양성의 지름길이 될 것이다. 기업 내 프로그램은 대기업들이 자체 수요를 충족시키기 위하여 실시하는 사내 교육을 말한다.

전산업에 종사하는 데이터 직무인력 수는 2022년 기준 197,802명이고, 데이터 산업에서의 직무인력 수는 134,310명이다. 그러나 전산업이나 데이터 산업에서의 우수 직무인력은 많이 부족한 상황이며, <그림 5.1>에서 보는 바와 같이 데이터 과학자의 부족률이 33.3%로 가장 높고, 다음으로 데이터 개발자(15.2%), 데이터 분석가(14.7%) 등의 순으로 나타났다. 따라서 데이터 전문 인력의 양성이 시급하다고 하겠다.

## 2. 데이터 기본법의 활성화와 데이터 산업 매출액 100조 원 달성

### 1) 주요 정책과 데이터 기본법 제정

지금은 4차 산업혁명과 디지털 대전환의 핵심이자 원유라고 불리는 데이터의 중요성이 부각되는 시대이며, 경제·사회·과학기술 등 전반에서 창출되는 데이

터가 수집·가공·생산·활용되어 혁신적인 산업과 서비스가 창출되는 데이터 경제 시대이다. 이에 따라 세계 각국은 데이터 경제 시대의 주도권을 확보하기 위하여 데이터 산업 육성에 총력을 기울이고 있다. 우리나라도 데이터 경제로의 전환에 빠르게 대응하기 위해 디지털 뉴딜(한국판 뉴딜 정책의 핵심)을 범국가적 프로젝트로 문재인 정부 시절부터 추진하기 시작했고, 윤석열 정부에서도 이를 계승하고 있다. 디지털 뉴딜 정책의 대표 과제는 ‘데이터 댐’을 중심으로 데이터를 생산·수집·가공하고 5세대 이동통신(5G) 및 인공지능(AI)과 융합·활용하는 다양한 정책을 추진하고 있으나, 이를 체계적으로 추진하기 위한 법적 근거가 부족한 상황이었다. 데이터 산업 진흥과 관련하여 중요한 법 제정으로 데이터 기본법 제정(2021년 9월 28일 국회 통과, 2022년 4월 20일부터 시행)이다.

데이터 기본법(데이터 산업 진흥 및 이용 촉진에 관한 기본법)에 대하여 그 의미를 다시 검토하고, 데이터 산업 진흥을 위한 구체적인 산업화 촉진 방안을 제시하고자 한다. 데이터 기본법은 데이터 산업을 진흥시키고 데이터 경제를 활성화하기 위한 목적으로 제정된 법으로, 데이터 산업을 2030년까지 주력 산업화하는데 크게 기여할 것으로 판단된다. 이 법에는 중앙정부와 지자체는 데이터 생산과 거래, 활용 촉진 기반 조성을 위해 노력해야 하며, 정부는 데이터 산업 진흥을 위해 3년마다 기본계획을 수립해야 하는 것을 비롯하여 다음과 같은 주요 내용을 담고 있다.

### (1) 데이터 정책의 컨트롤 타워, ‘국가데이터정책위원회’ 설립

공공부문의 데이터를 규율하는 법으로는 ‘통계법’, ‘공공데이터의 제공 및 이용 활성화에 관한 법률’, ‘데이터 기반 행정 활성화에 관한 법률’ 등 공공데이터와 관련된 법적 근거는 있으나, 민간 데이터의 경제·사회적 생산, 거래 및 활용 등을 위한 기본법제는 없는 상태였으므로, 이번 데이터 기본법이 그 역할을 하게 되는 것이다. 이 법은 데이터 거래 분석사업자에 대한 신고제 도입과 지원체계를 마련하고 있으며, 데이터를 안전하게 분석할 수 있는 데이터 안심 구역의 지정, 데이터 자산의 보호 등 데이터 시장 활성화를 위한 대책을 규정하고 있다.

데이터 안심 구역 내부에서는 별도의 서버에 가명 처리 초기 단계의 미개방 데이터를 저장해 두고, 이용자가 분석할 수 있게 허락하고 있다. 대신 분석 결과물을 외부로 반출할 때에는 다시 개인에 대한 식별 가능한 정보가 포함돼 있지 않는지 확인함으로써, 개인정보를 보호하는 동시에 공공데이터의 활용도를 높일 수

있게 했다.

민간과 공공을 아우르는 범정부 데이터 정책 컨트롤 타워로 국무총리 소속 ‘국가데이터정책위원회’가 데이터 기본법에 의해 출범했다. 이 위원회는 ‘전문위원회’와 ‘사무국’으로 구성되면, 데이터 산업 진흥 정책 전반을 총괄·조정하는 역할을 수행하며, 다양한 분야 전문가가 정책 제안 및 수립에 주도적으로 참여하고, 제시된 혁신적인 아이디어들이 실질적인 정책 집행으로 이뤄질 수 있도록 지원하는 역할을 한다. 이 위원회는 기본법에 명시된 ‘데이터 산업 진흥 기본계획’을 심의하는 역할을 한다. 이 위원회가 각계(데이터 산업계, 데이터 과학자, 데이터 행정가 등)의 유능한 인재들로 구성되어 신뢰성 있는 데이터 정책을 펴나가기를 기대한다.

### (2) 데이터 유통·활용을 촉진하는 ‘데이터 거래사’와 ‘데이터 가치평가 기관’

데이터 기본법에 의해 데이터 거래사의 등록과 데이터 가치평가가 실시된다. 데이터 거래사는 개인정보보호·저작권 등과 관련된 법적 지식과 데이터 사업화 등과 관련된 전문 지식을 바탕으로 안전한 데이터 거래를 지원한다. 데이터 가치평가 기관은 데이터 거래에 있어 시장 구성원들이 참고할 수 있는 데이터에 대한 가치 기준을 제공함으로써, 신뢰할 수 있는 데이터 유통·활용 여건을 조성하는 역할을 한다. 데이터 거래사가 새로이 신설되므로, 효율적인 운영을 위해서는 데이터 거래사의 충분한 양성 기준, 교육, 등록 등을 정해야 할 것이다. 또한 데이터 가치평가 기관의 신설에 있어서도 신뢰성 있는 기관 설립이 중요하므로, 데이터 산업계의 의견을 충분히 반영하여 신중하게 설립되어야 할 것이다.

### (3) 전문 인력 양성

데이터 산업 진흥의 핵심은 전문 인력 양성에 달려 있다. 데이터 기본법에 의해 데이터 전문 인력 양성을 위한 시책 마련 및 양성기관 지정·지원 등이 데이터 산업의 주력 산업화에 큰 역할을 담당할 것이다. 데이터 기본법에 명시된 ‘데이터 산업 진흥 기본계획’ 수립 시에는 반드시 데이터 전문 인력 양성 방안이 명시되어야 할 것이다.

### (4) 데이터 산업 사업자의 구심점인 ‘데이터 산업 협회’ 설립과 전문기관 지정

데이터 기본법은 50인 이상 데이터사업자를 발기인으로 과기정통부 장관 인가

를 받아 협회를 설립할 수 있고, 설립된 협회는 데이터 거래사 교육 등의 업무를 위탁받아 수행할 수 있도록 규정했다. 설립된 협회는 데이터 사업자들의 권익을 대변할 뿐만 아니라 각종 정책 수립에 있어 민관협력을 지원하는 등 민간 중심 데이터 산업 성장에 기여할 전망이다. 또한 이 협회는 데이터 거래사 교육 등의 업무를 위탁받아 수행할 수 있도록 했다.

지난 2022년 4월 27일 기존의 사단법인 한국데이터산업협회(2019년 1월 설립)는 자발적으로 해산하고, 데이터 관련 기업 대표 57명이 참석하여 '뉴(New) 한국데이터산업협회'를 발족했다. 이 협회는 과기정통부 승인이 나면 민간인들로 구성된 1호 인가 데이터 산업 협회가 된다. 새로이 발족한 한국데이터산업협회는 한국무역협회와 빅데이터 활성화를 위한 MOU를 체결(2022년 6월)하고, 2022년 12월에는 '2022년 제13회 대한민국 데이터대상'을 시상하는 등 활발하게 활동하기 시작하고 있다.

## 2) 데이터 산업 매출액 2030년 100조 달성

2022년 데이터 산업의 총매출액은 25조 527억 원(〈표 5.2〉 참조)이었다. 이는 2016년의 총매출액 13조 7,547억 원에 비해 6년간 82.1% 증가한 수치이다. 그리고 이 증가 속도는 더 빠를 것이다. 2022년에 비해 4년 후인 2026년에는 총매출액이 100% 증가하여 50조 원에 육박할 것이고, 그 후 4년 후에도 100% 증가를 삼는다면 2030년에는 100조 원에 육박할 것으로 수치상으로 계산된다. 따라서 2030년에 데이터 산업의 매출 목표를 100조 원으로 잡는 것도 무리는 아니다.

우리나라의 데이터 산업이 빠르게 성장하고 있으며, 이런 외적 성장과 더불어 데이터의 디지털 전환(digital transformation, DX)을 통한 새로운 가치를 개인과 기업이 체감할 수 있게 해야 한다. 이렇게 될 때 데이터 산업은 더욱 급속히 성장할 것이다. 디지털 전환이란 기업의 업무환경과 고객행동, 고객 경험의 과정이 디지털로 변환하는 것을 말한다. 데이터 기본법이 디지털 전환에 기여하도록 필요시에는 계속해서 수정 보완해 나가야 할 것이다. 우리 사회가 디지털 전환에 성공하려면 데이터 개방, 규제 완화, 플랫폼 및 인적 역량 측면에서 지속적인 투자와 노력이 필요하다. 디지털 전환에 성공하면 2030년 데이터 산업 매출액 100조 원 달성도 어려운 목표가 아닐 것이며, 데이터 산업은 우리나라 산업에서 주력 산업이 될 것이다.

### 3. 디지털 뉴딜과 디지털 플랫폼 정부 실현

#### 1) '디지털 뉴딜' 정책 추진

문재인 정부는 2020년 7월 14일에 '한국판 뉴딜'을 발표하면서, 특히 '디지털 뉴딜'을 강조했다. 이를 '디지털 뉴딜 1.0'이라고 부른다. 이 디지털 뉴딜에는 모든 데이터 수요자가 원하는 데이터를 꺼내 사용할 수 있도록 하는 '데이터 댐'을 건설하고 여기에 공공데이터 14.2만 개를 개방하여 댐에 넣어 놓겠다고 했다. 또한 AI의 저변확대를 위해 AI 학습데이터를 2025년까지 1,300종을 구축하겠다고 했다. 이러한 '데이터 댐'의 건설은 오랫동안 데이터 산업에서 희망해 온 것이다.

2021년 7월에 문재인 정부의 관련 부처는 데이터 산업 육성 정책을 포함한 '디지털 뉴딜'의 성과와 주요 정책 방향을 '디지털 뉴딜 2.0' 정책으로 발표한 바 있다. 이 정책에 의하면 2022년 말까지 데이터 댐 구축 사업에 디지털 뉴딜 예산의 51.7%인 1조 4,642억 원을 사용하고, 전체적인 '디지털 뉴딜 2.0' 사업에 2025년까지 49조 원을 투자하겠다는 것이었다.

#### 2) '디지털 플랫폼 정부' 실현

현재 전 세계적으로 디지털 전환에 대한 관심이 높다. 윤석열 정부의 핵심 공약 중 하나는 '디지털 플랫폼 정부'로서, 방대한 데이터를 디지털화하여 효율적으로 국가 운영에 사용하면서 국가 경쟁력을 높이겠다는 것이다. 문재인 정부의 '디지털 뉴딜'과 맥을 같이 하고 있는 것이다. 윤석열 정부의 110대 국정과제 가운데 디지털 및 과학기술 직간접적으로 관련된 공약만도 46개에 달할 정도로, 현 정부는 디지털화에 깊은 관심을 가지고 있다. 옳은 방향이라고 생각한다.

'디지털 플랫폼 정부'는 모든 데이터가 디지털로 연결되는 디지털 전환(DX)이 이루어지고, 디지털 전환 위에 작성되는 디지털 플랫폼으로 국민, 기업, 정부가 함께 사회문제를 해결하고, 최우선적으로 국민이 편하게 살아갈 수 있도록 새로운 가치를 창출하는 정부를 말한다. 이는 기존의 전산화된 정부를 뛰어넘어 진정으로 디지털화된 정부, 칸막이를 없앤 정부, 그리고 디지털 시대에 걸맞는 업무 방식을 갖춘 정부를 말한다. 이런 정부가 되기 위해서는 국민과 정부가 모두 "DX의 가장 중요한 driver는 데이터이다"라는 개방된 생각을 가지고, 정부 정책

입안이나 실천에서 데이터 활용이 기초 사고가 되어야 한다. 또한 데이터를 중심으로 한 기업이나 비즈니스 모델에 보다 개방적이고 우호적이어야 한다.

결국 이러한 ‘디지털 플랫폼 정부’의 실천을 위해서는 데이터 산업이 활성화되어야 하고, 다양한 소스(source)의 데이터들을 역동적으로 연계시킬 수 있는 플랫폼들을 구축해야 하고, 이를 위해 필요한 기술 역량을 확보해야 한다. 이와 함께 데이터 보안 문제에 대한 신뢰성 제고 문제도 같이 고려되어야 한다. 결론적으로, ‘디지털 플랫폼 정부’의 성공적인 실천은 데이터 활용의 가치를 극대화시킬 수 있는 국가적 데이터 생태계 활성화 여부에 달려 있다. 데이터 산업이 그 한몫을 견실하게 담당해야 한다.

#### 4. 마이데이터 사업 활성화

##### 1) 개인정보 활용

데이터 중 큰 부분인 개인정보는 그동안 데이터 3법(개인정보보호법, 정보통신망법, 신용정보법)에 막혀 개인정보 활용 사업에 발전이 없었다. 그러나 다행스럽게도 2020년 1월 9일에 데이터 3법 개정안이 국회를 통과하였고, 그 시행령이 2020년 8월 5일에 제정되어 시행되었다. 이 개정안에서는 가명정보(특정인을 식별할 수 있는 정보의 상당 부분을 가린 정보)를 통계 작성, 공익적 기록 보존, 과학적 연구 등에 정보 소유자의 사전 동의 없이 사용할 수 있도록 했다. 또한 개정안에서는 익명정보(가명정보보다 개인을 조금이라도 식별할 수 있는 정보는 모두 삭제한 정보)는 개인 정보 보호 대상에서 완전히 제외했다. 그러나 아직도 개정된 데이터 3법이 제대로 시행되고 있지 않으며, 개인정보 활용에 어려움이 해소되고 있지 못하다. 개정된 개인정보보호법에 대해 정보 주체의 권리 강화 사항 도입 미비, 온라인과 오프라인 이중 규제가 해소되지 않는 등의 미비점을 개선하기 위해 2차 개정안이 2021년 1월 개인정보보호위원회에 의하여 입법 예고되었다. 이 개정안에 의해 ‘마이데이터’ 사업 등 개인정보 활용도를 제고되기를 바란다.

##### 2) 마이데이터(MyData) 사업의 활성화

마이데이터 사업 본 허가를 받은 45개 기업은 2021년 12월 1일부터 ‘오픈 베타 서비스(대고객 시범 서비스)’를 시행할 수 있고, 2022년 1월부터 본격적으로

마이데이터 사업을 시작할 수 있었다. 그러나 금융권에 따르면(2021년 11월 11일자 한국경제신문 보도) 이 베타 서비스에 참여하는 기업은 국민·신한·우리·농협은행과 키움증권,뱅크셀러드, 핀크 정도라고 한다. 네이버, 카카오, 토스 등 빅테크(대형 IT) 기업의 초기 참여가 사실상 어려울 것으로 보도했다. 금융소비자보호법 이슈가 불거지면서 서비스 범위를 확정하지 못하고 있기 때문이다. 금융소비자보호법에는 빅테크 플랫폼의 ‘금융 상품 비교·판매 행위’가 금융소비자보호법 위반에 해당한다는 금융당국의 해석이 나온 뒤 금융사들이 마이데이터 사업 구상을 재검토하지 않을 수 없기 때문이다. 마이데이터에서는 개인의 금융 상품에 대한 정보를 제공하면서 더 좋은 금융 상품을 추천해 주는 것이 알짜 정보인데, 금융 상품 비교가 금지되면 마이데이터 정보의 가치가 떨어지기 때문이다.

빅테크 기업만이 아니라 소형 핀테크의 마이데이터 서비스 사업 참여에도 빨간 불이 켜졌다. 마이데이터 관련 전산망을 구축하지 못하면 서비스가 전면 시행되는 2022년 1월부터는 스크래핑(금융 정보 자동 수집) 방식을 사용할 수 없다. 이 경우 기존에 운영하던 자산관리 서비스조차 문을 닫아야 하는 어려움에 봉착하게 된다. 소형 핀테크 회사들은 비상이 걸린 셈이다. 마이데이터 사업의 활성화는 데이터 산업의 진흥에 핵심적인 역할을 감당할 것이다. 따라서 이에 걸림돌이 되는 금융소비자보호법을 개정하여 ‘금융 상품 비교·판매 행위’도 허용하는 것이 바람직하다. 일단 마이데이터 사업을 활성화시켜 그 발전하는 상황을 면밀히 검토한 후에 특별히 문제가 발생하면 금융소비자보호법을 재개정하는 방향이 옳다고 판단된다.

## 5. 의료부문 데이터 활용 극대화

### 1) 의료 빅데이터 플랫폼 구축을 통한 의료 혁신 추진

데이터가 집중적으로 사용될 수 있는 산업 분야로는 의료, 교육, 제조생산, 교통, 미래차 등 다양하나, 의료 분야는 특히 데이터가 많이 생산되는 분야이며, 의료 빅데이터 플랫폼을 구축하면 이 플랫폼은 의료산업 혁신에 큰 역할을 할 수 있고, 이를 통해 대다수 국민에게 큰 혜택을 줄 수 있다. 문재인 정부에서는 소위 ‘Big3’라고 지칭하는 시스템 반도체, 바이오헬스, 미래차에 대한 집중적인 혁신을 강조하였다.

디지털 헬스케어(digital healthcare)는 정보통신기술(ICT)과 의료기술을 접목한 개인 맞춤 시스템으로, 질환, 질병을 예방하고 스스로 건강을 관리할 수 있도록 돕는 건강 관리 시스템이다. 즉, 디지털 헬스케어는 AI, 빅데이터, 클라우드 컴퓨팅 등 다양한 기술이 의료 산업과 융합하여 시간과 장소에 제약 없이 원격 진단 및 지속적인 건강 관리를 가능하게 한다. 2022년 2월에 문재인 정부에서 발표한 ‘혁신성장 BIG3 추진회의(2022)’의 내용(〈표 6.3〉 참조)을 보면 디지털 헬스케어 서비스 산업을 육성하기 위해서는 우선적으로 디지털 헬스케어 서비스 산업 생태계 조성이 필요하며, 이를 달성하기 위하여 10대 중점 추진 과제를 들었다. 이들은 대부분 데이터 기반 추진 과제이며, 특히 보건의료 데이터 접근성 제고, 모빌리티 기반 원격 헬스케어 서비스 등이 나열되어 있다. 의료 부분의 데이터 산업의 발전은 디지털 헬스케어 서비스 산업의 진흥을 가져오는 중요한 인프라가 될 것이다.

〈표 6.3〉 디지털 헬스케어 서비스 산업 비전 및 10대 중점 추진 과제

<b>비전</b>	<b>디지털헬스케어 서비스 산업 생태계 조성</b>	
<b>목표</b>	◇ 비용효과성 측면의 디지털헬스케어 서비스 유효성 입증 ◇ 디지털헬스케어 서비스 확산을 위한 기반 조성	
<b>10대 중점 추진과제</b>	<b>1. 시장 창출 지원 강화</b>	① 다양한 혁신서비스 개발 지원 ② 대규모 실증 지원을 통한 유효성·상업성 검증 ③ 수요기반 시장 확보
	<b>2. 데이터 기반 융복합 헬스케어 기기 개발</b>	④ 디지털치료기기 개발 촉진 ⑤ 인공지능 기반 진단 보조기기 ⑥ 모빌리티 기반 원격 헬스케어 서비스 ⑦ 신체·정서적 보조 헬스케어 기기 개발
	<b>3. 활성화 기반조성</b>	⑧ 제도적 기반 마련 ⑨ 보건의료데이터 접근성 제고 ⑩ 융복합 인력양성 확대

## 2) 윤석열 정부 육성 대상 12대 국가전략기술 중 하나; 첨단 바이오산업

과학기술정보통신부는 2022년 10월 28일 용산 대통령실에서 윤석열 대통령 주재로 개최된 국가과학기술자문회의 전원회의에서 ‘12대 국가전략기술 육성 방안’을 발표하면서, 12대 국가전략기술 (〈그림 6.1〉 참조)과 50대 세부 중점기술을

제시했다. 12대 기술은 혁신선도 기술 4개, 미래도전 기술 4개, 필수 기반 기술 4개로 나누어지며, 그 상세 설명은 <그림 6.1> 속에서 찾아 볼 수 있다.



<그림 6.1> 12대 국가전략기술 (자료: 국가과학기술자문회의)

첨단 바이오 기술은 우주항공·해양 기술, 수소 기술, 사이버보안 기술과 함께 미래 도전 기술에 속한다. 디지털 헬스케어 기술을 포함한 첨단 바이오 기술을 육성하겠다는 윤석열 정부의 의지를 읽을 수 있다.

### 3) 디지털 헬스케어 산업의 급속한 신장

디지털 헬스케어 산업은 디지털 정보기술을 활용하여 전통적인 의료·헬스케어 분야를 디지털화하는 산업으로, 네 가지 분야(모바일 헬스, 디지털 보건의료 시스템, 보건의료 분석학, 원격의료)로 나누어진다. 시장조사기관인 GIA(2020, Global Industry Analysts)에서 발간한 「Digital Health: Global Market Trajectory & Analytics」에 의하면, <그림 6.2>에서 보여주고 있는 것과 같이, 글로벌 디지털 헬스케어 산업은 2020년 1,525억 달러 규모에서 2027년 5,088

억 달러 규모로 연평균 18.8%의 성장률을 보일 것으로 전망한다. 2020년에 가장 큰 분야는 모바일 헬스 산업으로, 864억 달러(57%)로 가장 큰 규모를 차지했고, 연평균성장률(CAGR) 16.6%를 기록하여 2027년에 2,531억 달러 시장을 차지할 것으로 예측했다. 그다음은 디지털 보건의료 시스템 분야, 보건의료 분석학 분야, 그리고 원격의료 분야이다.

2020년에 국가별 시장 규모는 미국이 626억 달러로 집계돼 세계 시장에서 41.0% 이상을 차지하고, 다음은 유럽(27.3%), 중국(8.3%), 일본(5.2%), 캐나다(2.5%)의 순이다. 한국은 시장 규모가 미미하다.



출처: GIA(2020); 한국보건산업진흥원(2020) 자료 재구성

〈그림 6.2〉 글로벌 디지털 헬스케어 분야별 산업 규모 및 전망(단위: 십억 달러, %)

의료 패러다임이 치료 중심에서 사전 진단, 예방, 맞춤형 의료로 변화하고 있으며, 그 중심에 AI가 기능을 발휘하고 있다. 해외 주요국에서는 AI를 활용한 개인 맞춤형 의료 서비스 제공을 위해 다양한 정책적 지원을 강화하고 있다. 이를 위해서 방대한 의료 빅데이터(유전 정보, 진료 정보, 생활 습관 정보 등)의 출현과 이를 분석하고 활용하는 데 있어 AI 기술이 핵심 기술로 자리 잡는 AI 헬스케어 부분이 활발히 연구되고 발전되고 있다.

#### 4) 기업 주도 의료 데이터 활용

정부의 적극적인 의료데이터 정책 추진에 힘입어 기업 또한 의료데이터를 활용한 신약 개발, 질병 진단 및 치료 플랫폼 구축 등 다양한 의료데이터 사업 투자 및 개발을 진행 중이다. 예를 들면, 이노엔(inno.N)에서는 ‘암 빅데이터 플랫폼’을 구축하여 신약 개발을 추진하고 있고, NAVER에서는 의료데이터를 활용하여 연구 촉진을 위한 ‘데이터 박스’ 서비스를 제공하고 있다. 또한 뉴로펫(Neurophet)에서는 데이터를 활용한 뇌손상 치료 인공지능 솔루션을 개발하여 사용하고 있고, 서울아산병원에서는 국제 데이터 표준 기반 ‘임상시험 영상 관리 시스템’을 개발하여 영상 관리에 앞서가고 있다. 바람직한 현상이다.

우리나라는 의료 강국이라고 볼 수 있다. 건강보험제도가 잘 정착되어 있고, 병원에 가기가 수월하고 병원비도 저렴한 편이다. 여기에 의료데이터 시스템이 정착되면 한국은 더 편리한 의료 강국이 될 뿐만 아니라 의학, 보건학, 약학 등의 연구에서도 세계 속의 데이터 강국으로 거듭날 수 있을 것이다.

#### 6. 데이터과학·AI 국제협력센터 설치

데이터과학·AI 분야의 국제적인 저명한 전문가들이 한국에 단기간이라도 체류하면서 국내의 전문가들과 교류가 이루어지려면 국가에서 파격적으로 지원하는 가칭 ‘데이터과학·AI 국제협력센터(Data Science·AI International Cooperation Center)’를 만들 필요가 있다. 이 센터에서 전문가들 간에 국제교류가 이루어지도록 하고, 저명 외국인 전문가들이 국내 대학, 연구소, 기업들과 협력 교류를 할 수 있으면 국내 데이터과학·AI의 발전을 물론 데이터 산업 육성에도 촉매제 역할을 할 것이다.

초기에 ‘데이터과학·AI 국제협력센터’를 설치할 때 건물 신축, 인력 보충 등에 어려움이 있다면, 비교적 여건이 잘 갖추어진 대학을 선정하여 이 센터를 두고, 단계적으로 발전시켜 나가는 방안도 생각할 수 있다. 현시점에서 우리나라는 데이터과학과 AI 분야의 국제협력센터가 필요하며, 정부에서 이를 적극적으로 지원해 주어야 할 것이다.

## 7. 데이터·AI 연구 융합 클러스터 조성

기업, 대학, 연구소, 공공기관 등이 한 곳에 모여 데이터·AI 관련 연구 개발, 인력 양성, 창업 등을 지원할 수 있는 가치 ‘데이터·AI 연구 융합 클러스터’를 조성하여 데이터의 수집, 연구, 활용 등이 활발히 일어나도록 하고, 이를 AI 연구에 접목하는 방안 연구를 할 필요가 있다. 미국의 경우에는 국립과학재단(NSF)의 지원으로 2017년부터 4개 빅데이터 지역혁신허브(Big Data Hub)를 만들어 산업계, 학계, 지방정부 등이 적극적인 협업으로 빅데이터 허브를 만들어 운영하고 있다. 이를 벤치마킹하여 우리나라도 대략 4개권 (수도·강원권, 충청권, 경상권, 전라·제주권)으로 나누어 빅데이터·AI 지역혁신허브를 구축하는 것이 바람직하다.

2019년에 정부에서 발표한 「데이터·AI 경제 활성화 계획」을 보면, 2023년까지 전국에 빅데이터 센터 100개 구축을 목표로 하고 있다. 전국에 4개 ‘빅데이터·AI 연구 융합 클러스터’를 두고, 이들 클러스터가 지역별로 다수의 빅데이터 센터를 관리하게 하는 것이 좋은 방안이 될 것이다. 또한 4개 빅데이터 지역혁신허브는 각각 중점 관리 빅데이터 분야를 지정하여 운영하는 것이 중복을 피하는 의미에서 효율적이다. 예를 들면, 수도·강원권은 바이오 관련 빅데이터, 충청권은 연구개발 관련 빅데이터, 경상권은 제조산업 관련 빅데이터, 전라·제주권은 농업, 관광 관련 빅데이터 등으로 나누는 것이다.

## 참 고 문 헌

- 관계부처합동 (2022); “디지털 인재양성 종합방안”, 관계부처(교육부, 기재부, 과학기술정보통신부, 산업부, 고용부, 중기부) 합동 정책, 교육부 주관 발표, 2022년 8월 22일.
- 교육부·한국교육과정평가원(2022); ‘2021 국가수준 학업성취도 평가 결과 및 대응 전략’ 발표, 2022년 6월 14일.
- 교육부·통계청 (2023); ‘2022년 초중고 사교육비조사 결과’, 2023년 3월 7일 발표. 상세 내용 통계청 국가통계포털(<http://kosis.kr>)에서 볼 수 있음.
- 박성현(2003); 통계적 품질관리 표준 매뉴얼 작성, 통계청 학술연구 보고서, 2003년 11월.
- 한국과학기술한림원 (2019); “빅데이터·인공지능 산업 진흥을 위한 데이터과학의 발전 전략 연구”, 한림연구보고서 130, 이영조 외 6인 공저, 2019년 12월 발행.
- 한국데이터산업진흥원 (2022); ‘2021 데이터산업 현황조사 결과보고서’, 2022년 3월.
- 혁신성장 BIG3 추진회의 (2022); “디지털 헬스케어 서비스산업 육성 전략”, 제19차 혁신성장 BIG3 추진회의, 2022년 2월 24일.
- EU집행위원회(2022); “The European Data Market Study 2021~2023”, First Report on Facts and Figures, IDC & The Lisbon Council, 2022.2.
- McKinsey (2011); “Big data: The next frontier for innovation, competition and productivity”, McKinsey Report.