

데이터 경제 진전에 따른 산업별 파급효과 분석과 정책적 활용방안 연구

A Study on the Economic Impact of
Data Economy and its Policy Implications

2018. 12

연구기관 : 정보통신정책연구원



방송통신정책연구 2018-0-01686

데이터 경제 진전에 따른 산업별 파급효과 분석과 정책적 활용방안 연구

(A Study on the Economic Impact of
Data Economy and its Policy Implications)

민대홍/이학기/오정숙

2018. 12

연구기관 : 정보통신정책연구원



과학기술정보통신부
Ministry of Science and ICT

IITP 정보통신기술진흥센터
Institute for Information & Communications Technology Promotion

이 보고서는 2018년도 과학기술정보통신부 방송통신발전기금 방송통신 정책연구사업의 연구결과로서 보고서 내용은 연구자의 견해이며, 과학기술정보통신부의 공식입장과 다를 수 있습니다.

제 출 문

과학기술정보통신부 장관 귀하

본 보고서를 『데이터 경제 진전에 따른 산업별 파급효과 분석과 정책적 활용방안 연구』의 연구결과보고서로 제출합니다.

2018년 12월

연구기관: 정보통신정책연구원

총괄책임자: 민대홍 부연구위원

참여연구원: 이학기 부연구위원

오정숙 부연구위원

목 차

요약문	ix
제 1 장 서 론	1
제 1 절 연구의 배경	1
제 2 절 연구의 목표 및 주요내용	4
제 2 장 국내외 데이터 경제 현황	6
제 1 절 데이터 경제의 정의	6
제 2 절 국내외 데이터 경제 현황	11
1. 해외 주요국의 데이터 경제 현황	11
2. 우리나라의 데이터산업 현황	25
제 3 장 데이터산업의 경제적 파급효과	35
제 1 절 데이터산업의 정의	35
1. 국내외 문헌에서 데이터산업의 정의	35
2. 데이터산업의 정의	37
제 2 절 연구방법론-산업연관 분석	41
제 3 절 데이터산업의 산업별 파급효과	54
제 4 장 데이터 경제 활성화 정책 방안	81
제 1 절 해외 주요국의 데이터 경제 활성화 정책	81
1. 유럽	81
2. 미국	84
3. 일본	86
4. 중국	87

제2절 국내 데이터 경제 활성화의 정책 현황 및 정책 수요	90
1. 우리나라의 데이터 경제 활성화 정책	90
2. 데이터 경제 활성화 정책수요	93
제3절 데이터 경제 활성화를 위한 정책 제언	98
제5장 결론	100
참고문헌	103
[부록 1] 외생화 이전 수요측면 산업연관 분석 결과	105
[부록 2] 품목별·부문별 공급액표 상에서 데이터산업의 부분별 재분류 결과	108
[부록 3] 수요측면 산업연관 분석 결과를 활용한 데이터 경제 규모의 국가 경제 대비 비중	113
[부록 4] 정책수요 설문조사 분석	119

표 목 차

〈표 1-1〉 산업별 데이터 활용 주요사례	2
〈표 2-1〉 유럽 데이터산업 규모 추이	12
〈표 2-2〉 유럽 데이터 경제 규모 추이 및 전망	13
〈표 2-3〉 유럽 데이터산업 고용 인원 추이	13
〈표 2-4〉 EU 회원국 별 데이터 기업 수 및 시장비중 현황	15
〈표 2-5〉 미국 데이터산업 및 데이터 경제 규모(2016~2017년)	16
〈표 2-6〉 미국 데이터 브로커 업체 현황	16
〈표 2-7〉 데이터 브로커의 분야별 제공 데이터 및 업체	17
〈표 2-8〉 미국 데이터 브로커의 분야별 수집 데이터	19
〈표 2-9〉 일본 데이터산업 및 데이터 경제 규모(2016~2017년)	20
〈표 2-10〉 일본과 미국의 데이터산업 및 데이터 경제 규모 비교(2017년)	20
〈표 2-11〉 중국 빅데이터 산업 분야별 주요 기업	23
〈표 2-12〉 중국 주요 기업의 빅데이터 활용 사례	25
〈표 2-13〉 우리나라 데이터산업 중분류 및 정의	26
〈표 2-14〉 우리나라 데이터산업 규모 추이	27
〈표 2-15〉 국내 데이터산업 대분류별 매출액 대비 투자액 비중 현황	30
〈표 2-16〉 국내 데이터산업 대분류 별 투자분야 비중 현황	31
〈표 2-17〉 국내 데이터산업 대분류 별 주요 사업 대상 현황	31
〈표 2-18〉 국내 공공데이터 개방 목록 개수 추이	33
〈표 2-19〉 국내 공공데이터 개방 다운로드 신청 건수 추이	34
〈표 2-20〉 국내 공공데이터 활용 웹/앱 서비스 개발 사례	34
〈표 3-1〉 데이터산업 생태계 분류 또는 정의	36
〈표 3-2〉 ICT 통합분류체계에서 데이터산업 분류	39
〈표 3-3〉 산업연관표 내 30대 산업	44

〈표 3-4〉 산업연관표 내의 내생부문 예시	45
〈표 3-5〉 산업연관표 내의 내생부문의 투입계수표로의 전환	46
〈표 3-6〉 품목별·부문별 공급액표	55
〈표 3-7〉 품목별·부문별 공급액표에서 데이터산업의 비중	58
〈표 3-8〉 데이터산업 포함 이전의 산업연관표 예시	59
〈표 3-9〉 데이터산업을 추가한 산업연관표 예시	60
〈표 3-10〉 31개 산업 전·후방연쇄효과	62
〈표 3-11〉 데이터산업 대분류별 전·후방연쇄효과 비교	65
〈표 3-12〉 데이터산업의 외생화 예시	67
〈표 3-13〉 데이터산업의 수요측면 경제적 파급효과	68
〈표 3-14〉 데이터산업의 수요측면 생산유발효과	70
〈표 3-15〉 데이터산업의 수요측면 부가가치유발효과	72
〈표 3-16〉 데이터산업의 수요측면 취업유발효과	73
〈표 3-17〉 데이터산업의 수요측면 산업연관분석 결과의 대분류별 비교	76
〈표 3-18〉 데이터산업의 공급측면 생산유발계수	78
〈표 3-19〉 데이터산업 대분류별 공급측면 생산유발계수	80
〈표 4-1〉 미국의 빅데이터 R&D 전략	85
〈표 4-2〉 중국의 중앙정부 및 지방정부 주요 빅데이터 정책	88
〈표 4-3〉 국내 데이터 경제 정책	92
〈표 4-4〉 표본 설계 및 조사 개요	94
〈표 4-5〉 데이터서비스 기업 대상 정책수요 조사 항목	95
〈표 4-6〉 데이터 경제 활성화를 위해 필요한 정책	96

그림 목 차

[그림 1-1]	글로벌 빅데이터 및 분석 시장 규모	1
[그림 2-1]	데이터산업과 타 산업 간의 연관관계	8
[그림 2-2]	데이터 경제: 데이터산업과 타 산업 파급효과	9
[그림 2-3]	데이터산업의 직접효과, 후방연쇄효과, 전방연쇄효과	10
[그림 2-4]	EU 회원국별 데이터 업체 비중	14
[그림 2-5]	중국 빅데이터 산업규모와 성장률	21
[그림 2-6]	중국 빅데이터 세부 분야 구성	22
[그림 2-7]	2020년 중국 빅데이터 산업 구성	22
[그림 2-8]	국내 데이터산업 분야별 시장 비중 추이	27
[그림 2-9]	국내 데이터산업 직접매출 규모 추이	28
[그림 2-10]	국내 데이터산업 직접매출 기준 대분류별 규모 및 전체대비 비중	28
[그림 2-11]	국내 데이터산업 인력현황	29
[그림 2-12]	국내 데이터산업 내 데이터 직무 별 인력현황	30
[그림 2-13]	국내 데이터서비스 부문 내 데이터 거래 현황	32
[그림 2-14]	분류별 데이터 구매·판매 비중	33
[그림 3-1]	산업연관표의 기본구조	42
[그림 3-2]	10대 후방산업	71
[그림 3-3]	10대 전방산업	80
[그림 4-1]	데이터 경제 활성화를 위해 시급한 정책	95
[그림 4-2]	데이터 경제 활성화를 위한 정책수요	97

요 약 문

1. 제 목

데이터 경제 진전에 따른 산업별 파급효과 분석과 정책적 활용방안 연구

2. 연구목적 및 필요성

지능정보화시대와 4차 산업혁명 시대로 진입하며 데이터의 중요성이 부각되고 있다. 데이터의 중요성이 부각되며 전 세계에서는 데이터 경제로의 진전이 가속화되고 있다. 따라서 지속적이고 안정적인 신성장 동력 창출을 위해서는 데이터 경제 활성화의 경제적 파급효과를 고려한 정책 방안 제시가 요구된다. 하지만 국내 데이터 경제를 체계적으로 분석하고 산업별 파급효과를 바탕으로 지속적인 발전을 위한 정책적 방향성을 제시하는 연구는 많지 않다. 이에 본 연구에서는 데이터 경제를 데이터산업과 데이터산업의 경제적 파급효과 두 축으로 나누어 살펴보고 국내외 데이터 경제 활성화 관련 정책들을 바탕으로 데이터 경제 활성화를 위한 정책적 시사점을 제공하고자 한다.

3. 연구의 구성과 범위

데이터의 활용을 통한 경쟁력 제고가 강조되고, 데이터산업의 전 세계적 성장세를 고려할 때, 우리나라 역시 데이터 경제 육성에 힘을 쏟아야 할 것으로 판단된다. 이에 본 연구에서는 우선 2장에서 데이터 경제의 정의를 내리고, 이에 따라 국내외 데이터 경제 현황을 살펴보고자 한다. 3장에서는 우리나라의 데이터 경제에 주목하여, 데이터산업이 타 산업에 미치는 경제적 파급효과를 도출하였다. 그리고 4장에서는 해외 주요국 및 우리나라의 데이터 경제 활성화 정책을 개괄하고, 국내 데이터 서비스 기업들을 대상으로 실시한 정책수요에 대해 살펴보았다. 마지막으로 5장에서는 경제적 파급효과와 개괄한 정책들을 바

탕으로 데이터 경제 활성화를 위한 정책적 제언을 제시하고 한다.

4. 연구 내용과 결과

2장에서는 데이터 경제의 정의와 국내의 현황에 대해 살펴보았다. 먼저 데이터 경제를 어떠한 관점에서 바라보아야하는지에 대한 논의를 바탕으로 데이터 경제를 데이터산업과 데이터산업의 경제적 파급효과 두 축으로 이루어진 것으로 정의하였다. 이러한 정의를 바탕으로 유럽, 미국 등 해외 주요국의 데이터 경제 현황을 데이터산업과 데이터산업의 경제적 파급효과로 나누어서 살펴보았다. 우리나라의 데이터 경제 현황의 경우 3장에서 다룰 데이터산업의 경제적 파급효과를 제외하고 데이터산업에 집중하여 우리나라의 데이터산업의 규모, 고용현황, R&D 투자, 데이터 유통·거래 현황 등에 대해 살펴보았다.

3장에서는 산업연관분석을 바탕으로 우리나라 데이터산업의 경제적 파급효과를 도출하였다. 먼저 국내외 문헌에서 나타난 데이터산업에 대한 정의를 살펴본 뒤 데이터산업의 경제적 파급효과 도출에 가장 용이한 ICT 통합분류체계에 따른 데이터산업 정의를 활용하여 산업연관분석을 진행하였다. 기존 산업연관표에 데이터산업을 재구성하기 위해 산업연관표와 함께 제공되는 품목별·부문별 공급액표를 활용하여 산업연관표 내에서 데이터산업을 재분류하고 하나의 산업으로 추가하였다.

데이터산업이 재분류·추가된 산업연관표를 바탕으로 데이터산업의 전·후방연쇄효과를 도출하였는데, 먼저 후방연쇄효과 도출을 위해 데이터산업을 외생화하여 수요측면 산업연관 분석을 수행하였다. 그 결과 데이터산업의 최종수요 또는 매출액이 1원 증가함에 따라 후방산업 전체에서 약 0.652원의 생산이 유발되는 것으로 나타났고, 여기에 데이터산업에서 최종수요 1원 증가를 포함한 총 생산유발효과는 1.652원으로 추정되었다. 또한 생산유발효과와 함께 부가가치유발효과와 취업유발효과 역시 도출하였는데, 후방산업 전체에서 부가가치유발효과는 약 0.251원, 데이터산업에서 추가적인 생산으로 인해 발생하는 부가가치는 0.368원으로 데이터산업의 총 부가가치유발효과는 0.619원으로 나타났다. 데이터산업의 최종수요 또는 매출액이 10억 원 증가함에 따라 후방산업 전체에서 증가하는 고용은 약 4.509명, 데이터산업 내에서 증가하는 고용은 3.203명으로 총 취업유발효과는 7.712명으로 추정되었다.

〈데이터산업의 수요측면 경제적 파급효과〉

	생산유발 효과	부가가치 유발효과	취업유발효과
타 산업 전체 유발효과	0.652	0.251	4.509
자기산업의 직접적인 연관효과	1.000	0.368	3.203
합계	1.652	0.619	7.712

데이터산업의 직접적인 자기연관효과를 제외하고 데이터산업의 대분류별로 각각 생산유발효과, 부가가치유발효과, 취업유발효과를 살펴본 결과, 먼저 데이터 서비스와 데이터 인프라의 생산유발효과는 각각 0.920과 0.784로 데이터산업 전체의 평균을 상회하였다. 부가가치 유발효과의 경우에는 데이터 서비스, 데이터 인프라, 데이터 구축/컨설팅 분야에서 모두 데이터산업 평균을 상회하는 것으로 나타났다. 마지막으로 취업유발효과에서는 데이터 서비스, 데이터구축/컨설팅, 데이터 솔루션이 데이터산업 전체 평균을 상회하였다.

〈데이터산업 대분류별 수요측면 경제적 파급효과〉

	생산유발 효과	부가가치 유발효과	취업유발효과
데이터산업 전체	0.652	0.251	4.509
데이터 솔루션	0.491	0.222	4.827
데이터 구축/ 컨설팅	0.647	0.263	5.258
데이터 서비스	0.920	0.382	6.282
데이터 인프라	0.784	0.277	4.457

전방연쇄효과는 공급측면에서의 산업연관분석을 통해 도출되었다. 공급지장효과 또는 공급측면의 생산유발효과는 데이터산업의 최종수요 또는 매출액이 1원 증가함에 따라 타 산업에서 이 1원 어치의 생산물을 중간재로 활용하면서 증가하는 생산액의 가치로 이해할 수 있는데, 데이터산업 전체의 공급측면 생산유발효과는 약 0.892원으로 도출되었다. 데이

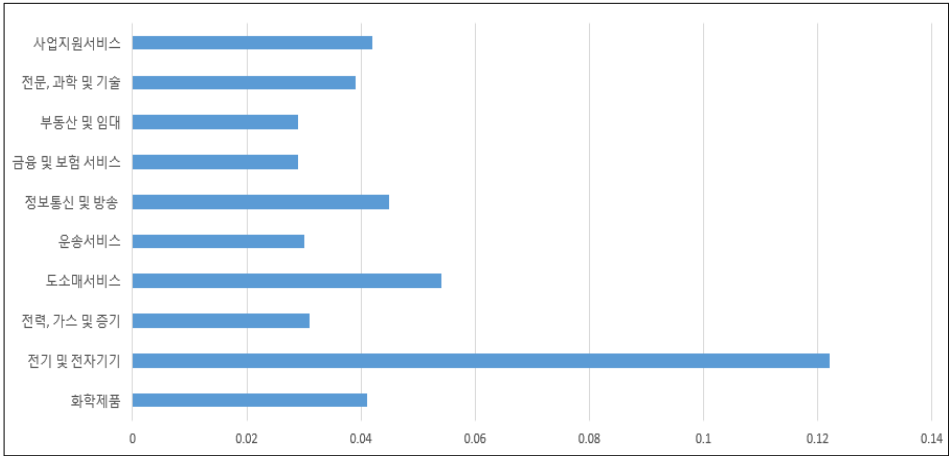
터산업의 대분류별로 살펴본 공급측면 생산유발계수는 데이터 서비스가 1.395원으로 가장 높았고, 데이터 구축/컨설팅 1.308로 데이터 서비스 부문의 뒤를 이었다. 데이터 서비스와 데이터 구축/컨설팅 모두 데이터산업 전체보다 높은 전방연쇄효과를 갖는 것으로 나타난 반면, 데이터 솔루션과 데이터 인프라 부문은 각각 0.608과 0.742의 생산유발효과를 가져 데이터산업 전체의 전방연쇄효과에 미치지 못하는 것으로 나타났다.

〈데이터산업 대분류별 공급측면 생산유발계수〉

분류	공급측면 생산유발효과	
	값	순위
데이터산업 전체	0.892	-
데이터 솔루션	0.608	4
데이터구축/ 컨설팅	1.308	2
데이터 서비스	1.395	1
데이터 인프라	0.742	3

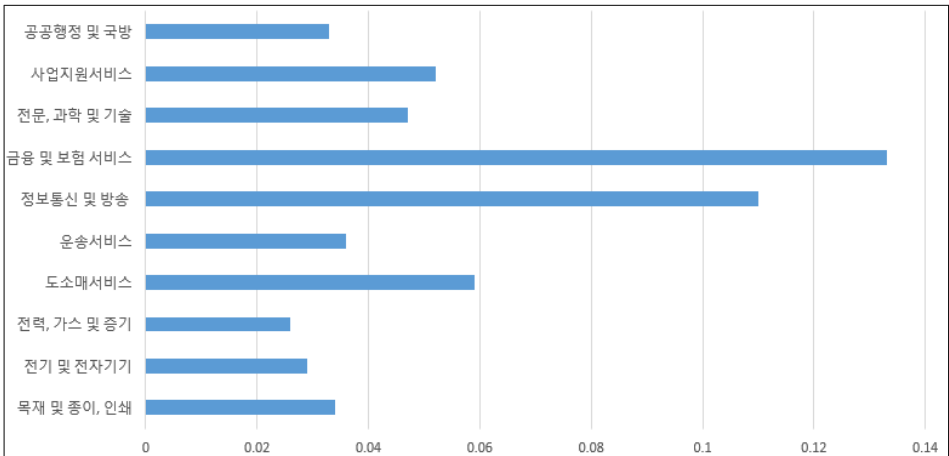
위에서 살펴본 수요측면과 공급측면의 생산유발계수를 바탕으로 10대 전방산업과 10대 후방산업을 도출할 수 있는데, 먼저 데이터산업의 후방연쇄효과가 높은 상위 10개 후방산업으로는 전기 및 전자기기, 도소매서비스, 사업지원 서비스, 화학제품, 전문, 과학 및 기술, 정보통신 및 방송 산업 등이 꼽혔다.

[10대 후방산업]



다음으로 데이터산업의 전방연쇄효과가 높은 상위 10개 전방산업으로는 금융 및 보험 서비스, 정보통신 및 방송, 도소매서비스, 사업지원 서비스, 운송서비스 등이 꼽혔는데, 이들 중 특히 금융 및 보험서비스, 도소매서비스, 운송서비스는 데이터 경제 활성화에 따른 효율성 제고 등의 사례에서 주로 나타나는 산업인 점을 주목할 만하다.

[10대 전방산업]



4장에서는 국내외 데이터 경제 활성화 관련 정책을 개괄하고, 국내 데이터서비스기업들을 대상으로 정책수요에 대한 설문조사 결과에 대한 분석을 진행하였다. 해외 주요국들의 데이터 경제 활성화 정책의 패러다임은 자칫 서로 상충되어 보이는 개인정보의 보호와 데이터의 적극적인 활용을 두 축으로 구성되는 추세이다. 최근 발표된 우리나라의 데이터 산업 활성화 전략 역시 전 세계적인 추세에 맞게 개인정보의 보호와 데이터의 적극적인 활용을 모두 강조하고 있다. 정책수요 설문조사 결과 시급성과 필요성 모두에서 높은 응답을 받은 정책은 개인정보보호법 개정(가명·익명 정보개념 정립, 개인정보의 범위 명확화 등), 데이터 소유권의 명확한 지정, 데이터 활용 촉진 사업(기존 중소기업에 빅데이터 분석 서비스 제공 기업을 매칭하고 해당 서비스를 활용할 수 있는 자금 또는 바우처 지원) 등으로 나타났다. 이러한 정책적 수요는 최근 발표된 데이터산업 활성화 전략에 모두 포함되어있어 우리나라의 데이터 경제 활성화 정책의 방향성은 전 세계적 추세와 국내 정책 수요를 모두 포괄하고 있다고 평가할 수 있다. 하지만 개인정보보호법 개정과 데이터 소유권은 단지 현재 시급하고 필요하다고 단기간에 성과를 거두기 위해 성급히 진행해야 하는 문제가 아니라고 판단된다. 물론 데이터 경제가 하루 빨리 활성화되기를 바라는 산업계의 요구를 무시하지는 말아야하나 개인정보보호법과 데이터 소유권에 대한 문제는 사회적 합의를 바탕으로 진행되지 않을 경우 계속하여 제자리 걸음을 할 수 밖에 없기 때문에 사회적 합의 도출을 위한 정책적 노력이 장기적인 관점으로 지속되어야 할 것이다.

5. 정책적 활용 내용

본 연구결과는 데이터 경제 관련 정책효과 분석을 위한 기초자료로 활용될 수 있다. 즉, 데이터산업의 산업 별 파급효과를 살펴봄으로써 데이터산업 및 데이터 경제 활성화 정책의 타 산업에의 영향 및 국민 경제 전체 영향 등 정책평가에 활용이 가능하다. 이에 더하여 데이터 경제의 산업 별 파급효과를 바탕으로 정책입안 시 영향평가를 위한 참고자료로서도 활용가능하다. 마지막으로 국내외 데이터 경제 활성화 관련 정책을 개괄 및 비교 분석하고 산업계의 정책 수요를 파악함으로써 향후 데이터 경제 활성화를 위한 정책 발굴 및 입안에 참고자료로 활용될 수 있을 것으로 기대된다.

6. 기대효과

본 연구에서 제시된 데이터산업의 파급효과는 데이터산업의 중요성과 그로 인한 국민경제에의 함의를 평가하는데 기초자료로서 활용될 수 있다. 따라서 데이터산업 더 나아가 데이터 경제 활성화 정책을 객관적으로 평가하여 데이터 경제 활성화 정책 발굴 및 입안의 효율성을 증대시킬 수 있을 것이다. 데이터 경제 활성화 정책 전반의 효율성 증대는 데이터 경제로의 진전을 촉진시켜 우리나라 경제의 신성장 동력을 확보하는 한편 경제 체제의 선진화를 통해 경제 발전 및 일자리 창출에 기여할 수 있을 것으로 기대된다.

SUMMARY

1. Title

A Study on the Economic Impact of Data Economy and its Policy Implications

2. Purpose and Necessity

As we enter the age of intelligent information and the era of the Fourth Industrial Revolution, data has been taking on increasing significance. As the importance of data grows, the world has been accelerating its transition into a data economy. Therefore, in order to create new engines for continuous and stable growth, it is required to propose policy plans, taking into consideration the economic ripple effects of an invigorating data economy. However, there have not been many studies that systematically analyze the Korean data economy and propose policy orientations to ensure continuous development based on its ripple effects on individual industries. Hence, this study intends to look into the data economy from two perspectives – the data industry and its economic ripple effects – and then derive policy implications for the vitalization of the data economy, based on relevant policies in Korea and other countries.

3. Organization and Scope

Given the emphasis placed on enhancing competitiveness through data utilization as well as the worldwide growth of the data industry, it is considered necessary that Korea should also concentrate its strength on nurturing a data economy. In this regard, Chapter

2 offers a definition of data economy and looks into the current states of both Korean and other countries' data economies. Chapter 3 identifies the economic ripple effects of the data industry on other industries, focusing on the Korean data economy. Chapter 4 outlines the policies of Korea and other major countries to invigorate the data economy and examines policy demands in a survey with Korean data service providers. Lastly, Chapter 5 offers policy proposals to vitalize the data economy based on the economic ripple effects and policies outlined in the previous chapters.

4. Contents and Results

Chapter 2 defines data economy and looks into its current status at home and abroad. Based on discussions on what perspective we should take to look at the data economy, the data economy is defined to be built on two pillars, data industry and its economic ripple effects. On the basis of this definition, we look into the current states of major data economies, including Europe and the U.S., in terms of data industry and its economic ripple effects. As for the Korean data economy, the economic ripple effects of the data industry are excluded from this chapter as they will be addressed in Chapter 3. Instead, we focus on the data industry, discussing its scope, employment, R&D investments, and data distribution and trades.

Chapter 3 identifies the economic ripple effects of Korea's data industry. We first look into the definitions of data economy in various Korean and foreign literature and then proceed to perform an input-output analysis using the definition of data economy by the Harmonized Classification System for ICT, as it is the easiest to figure out the economic ripple effects. To incorporate the data industry into the existing Input-Output Table, we reclassified it within the Input-Output Table using the Supply Amount Table by Item/by Sector, which is provided with the Input-Output Table, and added it as a new industry.

We identified the forward and backward linkage effects of the data industry based on the revised Input-Output Table where the data industry is reclassified and added. First, to identify backward linkage effects, we performed a demand-side input-output analysis by developing exogenous specifications for the data industry. The results show that, for every one won increase in the final demand or revenue of the data industry, production increases by about 0.652 won across all the upstream industries. The total production-inducing effects of the data industry, which include the increase of one won in final demand within the data industry, are estimated to reach 1.652 won. We also estimated value-added creation effects and employment-inducing effects: Total value-added creation effects are calculated to be 0.619 won, with value-added created across the entire spectrum of upstream industries estimated at 0.368 won and value-added generated by additional production within the data industry at about 0.251 won. On the other hand, total job-inducing effects are estimated to be 7.712, with employment in all the upstream industries increasing by about 4.509 persons and employment within the data industry growing by 3.203 for each 1 billion won growth in the final demand or revenue of the data industry.

〈Demand-side economic ripple effects of the data industry〉

	Production-inducing effects	Value-added creation effects	Employment-inducing effects
Other industries - total	0.652	0.251	4.509
Data industry - total	1.000	0.368	3.203
Grand total	1.652	0.619	7.712

Forward linkage effects were identified through a supply-side input-output analysis. Supply shortage effects or supply-side production-inducing effects can be understood as the production(value) that every increase of one won in the final demand or revenue of the data industry causes to increase in other industries, which use that one won's worth

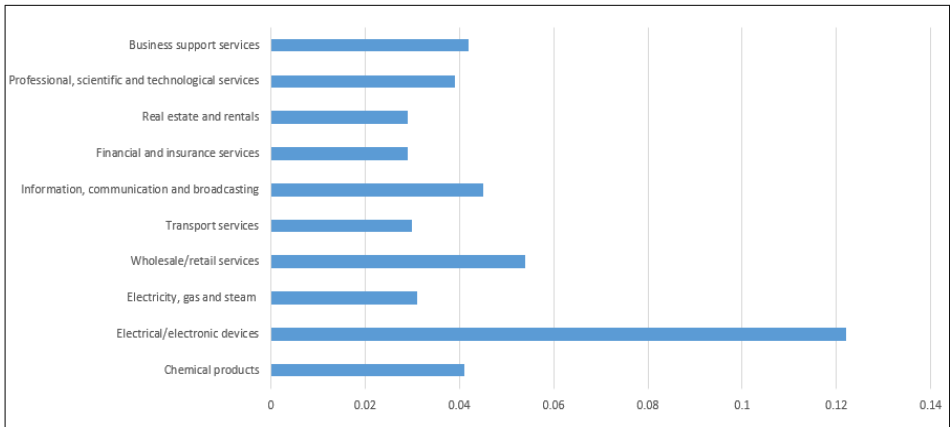
of product from the data industry as intermediate goods. The supply-side production-inducing effects of the whole data industry are determined to be about 0.892 won. In terms of production-inducing coefficients for each high-level category of the data industry, Data Services is the highest at 1.395 won, followed by Data Construction and Consulting at 1.308 won. This shows that both Data Services and Data Construction/Consulting have higher forward linkage effects than those of the data industry as a whole, while Data Solutions and Data Infrastructure are at 0.608 and 0.742, respectively, falling short of the forward linkage effects of the entire industry.

〈Supply-side production-inducing effects, by high-level data industry categories〉

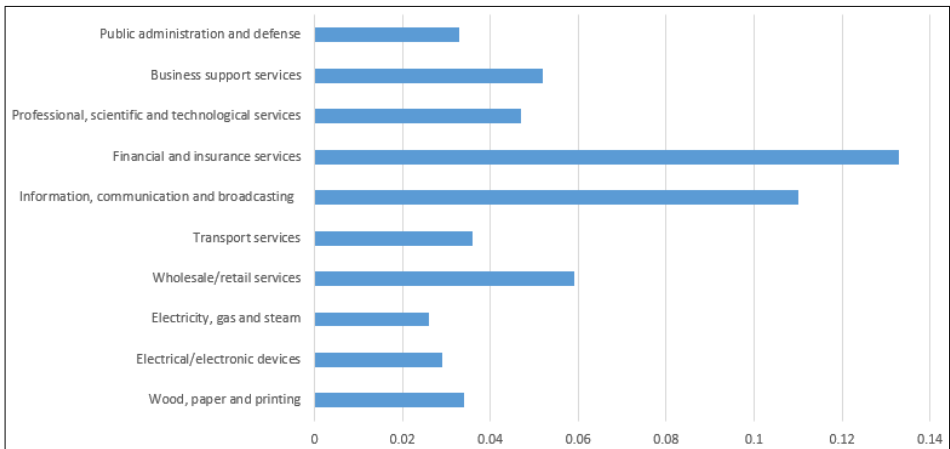
Category	Supply-side production-inducing effects	
	Value	Ranking
Data industry as a whole	0.892	-
Data solutions	0.608	4
Data construction/consulting	1.308	2
Data services	1.395	1
Data infrastructure	0.742	3

Top 10 upstream industries and top 10 downstream industries can be identified based on the demand-side and supply-side production-inducing effects described above. Top 10 upstream industries for which the data industry has high backward linkage effects include electric and electronic devices; wholesale and retail services; business support services; chemical products; professional, scientific, and technical Services; and information, communications and broadcasting.

[Top 10 upstream industries]



[Top 10 downstream industries]



Next, top 10 downstream industries for which the data industry has high forward linkage effects include financial and insurance services; information, communications and broadcasting; wholesale and retail services; business support services; and transport services. It is noteworthy that financial and insurance, wholesale/retail, and transport services in particular are cited frequently in case studies of how invigorating the data

economy can boost efficiency.

Chapter 4 gives an overview of policies to invigorate data economies in and out of Korea and analyzes the results of a survey on policy demands, which was conducted with Korean data service providers. The paradigms of major countries' policies tend to be forming on two pillars that appear to be mutually conflicting: privacy protection and aggressive use of data. Korea's data industry invigoration strategy that has recently been announced also emphasizes both personal information protection and active use of data in keeping with the global trends. According to a survey on policy demands, the revision of the Personal Information Protection Act(to establish the concepts of anonymity and pseudonymity, clarify the scope of personal information, etc.), the clear designation of data ownership, and projects to facilitate data uses(matching big data analysis service providers to existing SMEs and providing financial or voucher support required to use such services) received high responses in terms of both urgency and necessity. All these policy demands are included in the recently-announced strategy. Therefore, it can be said that the orientation of Korea's policy to invigorate its data economy reflects both the global trends and domestic policy demands. However, we believe that the revisions of the Act and data ownership are not something to be rushed just because they are urgent and necessary now and quick results are wanted. Of course, demands of the industrial circles for having the data economy invigorated as soon as possible should not be downplayed, but the issues of privacy protection and data ownership are bound to go around in circles unless they are addressed based on social consensus. Therefore, policy efforts to reach social consensus should continue from a long-term perspective.

5. Uses in Policy Making

Results of this study can be used as basic data to analyze the effects of data economy policies. That is, they can be used to evaluate policies including the impact of the data industry and policies to vitalize the data economy on other industries and the national

economy by examining the ripple effects of the data industry on other industries. In addition to this, the study results, particularly the ripple effects of the data economy on each industry, can be used as reference data for impact assessment when drafting policies. Lastly, this study provides an overview and comparative analysis of domestic and overseas policies related to the invigoration of the data economy and identifies the policy demands of the industrial circles. In this respect, the study results are also expected to be used as references for discovering and drafting future policies to vitalize the data economy.

6. Benefits Expected

The ripple effects of the data industry presented in this study can be used as basic data to evaluate the importance of the data industry and its implications on the national economy. Consequently, the data industry as well as policies to invigorate the data economy can be evaluated objectively, which will in turn increase efficiency in discovering and drafting policies to invigorate the data economy. Increases in efficiency across the board of the invigoration policies are expected to advance the economic system further and thus contribute to economic development and job creation, while securing new growth engines for our economy by facilitating advances into the data economy.

제1장 서론

제1절 연구의 배경

지능정보화혁명이라고도 불리는 4차 산업혁명이 화두로 떠오르며 그 핵심자원인 '데이터'는 '원유(原油)'에 비견되며 데이터에 대한 관심과 중요성 역시 급증하고 있다. 이에 따라 전 세계의 데이터산업 또는 데이터시장의 규모 역시 빠른 속도로 성장하고 있다. 전 세계 데이터시장은 2015년 1,200억 달러(약 135조 원)에서 2017년에는 1,508억 달러(약 170조 원)에 이를 것으로 전망되며, 이러한 성장세는 지속되어 2020년에는 약 2,100억 달러(약 236조 원)에 이를 것으로 예상되고 있다.¹⁾ 우리나라의 데이터산업 역시 지속적으로 성

[그림 1-1] 글로벌 빅데이터 및 분석 시장 규모



주: *는 전망치임

자료: IDC(2017a), 박선우(2018) 재구성

1) IDC(2017a), 박선우(2018) 재인용.

장하고 있다. 한국데이터진흥원(2018)에 따르면 우리나라의 데이터산업은 2013년 약 11조원에서 지속적으로 성장하여 2017년에는 약 14조원으로 2016년에 비해 약 4.0% 성장할 것으로 전망된다.

또한 데이터산업의 전 세계적인 성장세에 더하여 데이터의 수집과 분석을 통한 부가가치 창출과 데이터 활용의 중요성은 2017년 3월 기준으로 데이터 기반 혁신을 원동력으로 각각 시가총액 2위, 4위, 8위에 오른 알파벳(구글의 모회사), 아마존, 페이스북의 경우에서도 잘 나타난다.

데이터산업은 원시데이터를 저장·가공·처리·활용·분석하여 새로운 유형의 정보를 생산하는데 이러한 새로운 유형의 정보는 국가 경제에 주입되어 신산업을 창출²⁾하거나 기존산업에서 활용된다. 특히 기존 산업에서는 새로운 정보를 활용함으로써 다양한 영역에서 의사결정 합리화, 새로운 서비스·제품 생산 및 공급할 수 있다. 생산관리·위험관리 등의 영역에서 의사결정 합리화는 매출액 증대, 비용 절감으로 이어져 기존 산업의 수익성과 경쟁력 강화하고, 다양한 소비자 행태 분석을 활용한 예측배송·맞춤형 마케팅 등 새로운 서비스 창출되어 공급된다.

〈표 1-1〉 산업별 데이터 활용 주요사례

산업분류	기업	데이터 활용 내용
제조업	볼보	자사가 생산한 자동차의 엔진 및 브레이크 등에 센서를 부착하여 차량운행 데이터를 수집·분석하여 작동오류 방지에 활용
	오르론*	생산 공정 데이터 수집·분석하여 불량원인 분석에 필요한 노동자의 수를 6명에서 1명으로 축소
유통/소매업	아마존	소비자의 구매행태를 분석하여 소비자가 구매결정을 내리기 전 수요상품을 예측하여 배송하는 예측배송에 대한 특허를 취득
	Otto Versand*	소비자행태 분석을 바탕으로 약 2백만 건의 반품량을 줄임

2) 신산업 창출의 대표적인 예로서 Uber는 해당 서비스 소비자의 위치 정보와 서비스 공급자의 위치 정보를 활용하여 새로운 형태의 산업을 개척하였다.

산업분류	기업	데이터 활용 내용
금융	테스코	물류 및 재고 데이터 분석을 활용하여 재고량 최적화를 통해 약 6백만 유로의 재고비용 절감
	Progressive*	피보험자의 운행정보를 활용하여 적절한 수준의 보험료를 산정
	신한카드	소비자의 소비패턴을 분석하여 개인별 맞춤형 카드 제공하여 경쟁력 제고
건강/보건	IBM	왓슨(Watson)의 비정형데이터 분석을 활용하여 암진단 비용 절감
	동아ST	약물의 부작용 사례 데이터 분석을 통해 신약개발 비용 절감

주: 오픈마켓은 일본의 산업용 제어기기 및 헬스케어 제품 생산업체, Otto Versand는 독일의 온라인 쇼핑 플랫폼, Progressive는 미국의 자동차 보험 업체
 자료: 최재경(2016), 금융위원회(2018), IDC(2017b) 재구성

이처럼 데이터 활용의 중요성이 전 산업에 걸쳐 강조되며 전통적인 산업들에서 데이터가 핵심적인 자원으로 활용되며 새로운 부가가치를 창출하고 데이터가 모든 경제활동의 근간이 되는 경제 체제인 ‘데이터 경제’가 또 다른 화두로 부상하고 있다. 이에 따라 해외 주요국과 우리나라는 데이터 경제로의 이행을 촉진시키는 정책을 추진하고 있다. EC는 데이터의 잠재력을 활용하여 사회·경제를 이롭게 하는 것을 목적으로 하는 ‘유럽 데이터 경제 구축(Building a European Data Economy)’에서 다양한 종류의 데이터 활용과 유럽 내 국가 간 데이터의 자유로운 흐름을 원활하게 할 것을 천명하였고, 일본 역시 최근 데이터 유통 추진 협의회의를 설립하고 데이터 유통을 촉진시켜 데이터 활용을 활성화하기 위한 기술적·제도적 환경을 정비하고 있다. 우리나라 역시 최근 관계부처합동으로 ‘데이터산업 활성화 전략’을 발표하고 데이터산업 육성과 데이터의 안전한 활용을 위한 기반조성에 힘을 쏟고 있다.

이처럼 데이터의 활용을 통한 경제 전체의 혁신동력 제고가 화두로 부상하고, 데이터산업의 전 세계적 성장세를 고려할 때, 우리나라 역시 ‘데이터산업’ 뿐만 아니라 ‘데이터 경제’ 육성에 힘을 쏟아야 할 것으로 판단된다. 이에 본 보고서에서는 먼저 데이터 경제를 정의하고, 이에 따라 국내외 데이터 경제 현황을 살펴본 뒤 데이터산업의 경제적 파급효과에 대해 살펴보고자 한다. 또한 해외 주요국 및 우리나라의 데이터 경제 활성화 정책을

개괄한 뒤, 경제적 파급효과와 개괄한 정책들을 바탕으로 국내 데이터 경제 활성화를 위한 정책적 제언을 제시하고 한다.

제 2 절 연구의 목표 및 주요내용

본 연구는 크게 두 가지를 목표로 한다. 첫 번째로, 데이터산업 현황을 파악하고 산업연관 분석을 통해 데이터산업의 산업별 파급효과를 도출하고자 한다. 먼저 2장에서는 국내외 문헌에 나타난 데이터 경제의 정의를 바탕으로 데이터 경제를 어떠한 관점으로 바라볼 것인지에 대해 살펴보고자 한다. 데이터 경제는 여러 문헌에서 언급되어 왔으나, 그 정의에 대한 합의가 이루어지지 않은 상태이다. 본 연구에서는 데이터 경제를 데이터산업과 그 파급효과로 정의하고 국내외 데이터 경제 현황에 대해 파악하였다. 유럽 및 미국 등 해외 주요국의 데이터 경제 현황을 데이터산업과 그 파급효과로 나누어서 살펴본 뒤, 우리나라의 경우 본 연구에서 도출될 데이터산업의 경제적 파급효과를 제외한 데이터산업 현황을 데이터 경제의 구성원 별 민간·공공부문의 R&D 투자현황 파악, 공공데이터 활용 현황 파악, 민간부문 데이터 유통·활용 등으로 나누어 파악하고자 한다. 다음으로 3장에서는 산업연관분석을 위해 데이터산업을 산업연관표 내에서 재구성·재정의하여 데이터산업이 국내 산업들에 미치는 경제적 파급효과를 도출하였다. 경제적 파급효과 도출을 위해 사용된 방법론은 산업연관분석(Input-output analysis)로 산업연관표를 바탕으로 데이터산업이 전·후방산업에 미치는 파급효과를 수요측면 및 공급측면 분석을 통해 각각 도출하였다. 후방연쇄효과란 데이터산업의 최종수요 또는 최종 매출액이 1원이 증가할 경우 이를 위해서 추가적으로 투입되어야 하는 중간재들의 수요가 증가하며 발생하는 생산유발 효과이다. 반면 전방연쇄효과란 데이터산업의 최종수요 또는 최종 매출액이 1원 증가하면서 이 증가된 생산물을 중간재로 활용하는 산업에서 유발되는 생산량을 뜻한다.

본 연구의 두 번째 목표는 국내외 데이터 경제 활성화 정책을 개괄하고 이를 비교·분석함으로써 데이터 경제 활성화를 위한 정책적 시사점을 도출하는 것이다. 먼저 4장에서는 해외주요국의 데이터 경제 활성화 정책들에 대해 살펴본 뒤, 우리나라의 데이터 경제 활성화 정책을 개괄하고 한다. 다음으로 데이터산업 구성원으로부터 데이터 경제 활성화 관

런 정책수요를 위해 실시한 설문조사 결과를 살펴보았다. 국내 데이터산업 구성원들의 정책수요와 국내외 데이터 경제 활성화 정책의 비교·분석을 바탕으로 데이터 경제 활성화를 위한 정책적 시사점을 도출하였다.

마지막으로 5장에서는 3장에서 도출된 데이터산업의 경제적 과급효과를 정책입안 또는 정책 평가 과정에서 활용하는 방안과 주의사항에 대해 논의한 뒤, 4장에서 도출된 정책적 시사점과 연계하여 데이터 경제 활성화를 위한 정책의 방향성을 제시하였다.

제2장 국내외 데이터 경제 현황

제1절 데이터 경제의 정의

데이터 경제가 화두로 떠오르고 있지만 아직까지 데이터 경제의 정의에 대한 전 세계적 인 합의는 이루어지고 있지 않다. 우리나라에서도 최근 들어 데이터 경제에 대한 여러 연구들이 생산되고 있으나, 데이터 경제에 대한 정의에 대한 합의는 아직 이루어지지 않은 것으로 보인다. 본 절에서는 국내외 문헌에 나타난 데이터 경제에 대한 정의를 개괄하고 이를 바탕으로 본 연구에서 사용될 데이터 경제의 정의를 내리고자 한다.

국내의 문헌에서 데이터 경제라는 단어가 쓰이기 시작한 것은 최근의 일로 먼저 ‘데이터 기반 경제(Data-driven Economy)라는 단어가 EC에서 사용되기 시작하며 주목을 받기 시작하였다. 현재까지 국내외 문헌에서 나타난 ‘데이터 경제란 광의로는 ‘데이터 중심의 경제 체제를 좁은 의미로는 ‘데이터산업 생태계 또는 ‘데이터산업을 의미한다. 먼저 ‘데이터 중심의 경제 체제(Data-driven Economy)란 전통적인 산업들에서 데이터가 핵심적인 자원으로 활용되며 새로운 부가가치를 창출하고, 데이터가 모든 경제활동의 근간이 되는 경제 체제 정도로 정의될 수 있다. IDC(2017b)의 연구에서는 데이터 경제(Data Economy)를 데이터의 생성·축적·가공·거래·활용을 다루는 구성원들이 만들어 내는 생태계와 그로 인한 경제적 파급효과로 정의하였다. 데이터의 생성·축적·가공·거래·활용을 다루는 구성원들이 만들어 내는 생태계³⁾는 한국데이터진흥원의 데이터산업 현황조사에서 정의된 데이터산업과 유사하다.⁴⁾ 또한 그 경제적 파급효과는 ‘데이터 생태계의 경제적 파급효과로 전·후방 간접효과(forward and backward indirect effect)와 유도효과(induced effect)라고 정

3) 이하에서는 혼동이 되지 않는 경우 데이터의 생성·축적·가공·거래·활용을 다루는 구성원들이 만들어 내는 생태계를 ‘데이터 생태계’를 지칭한다.

4) 한국데이터진흥원은 매년 발간하는 데이터산업 현황조사에서 데이터산업을 ‘데이터의 생산·수집·처리·분석·유통·활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업으로 정의한다.

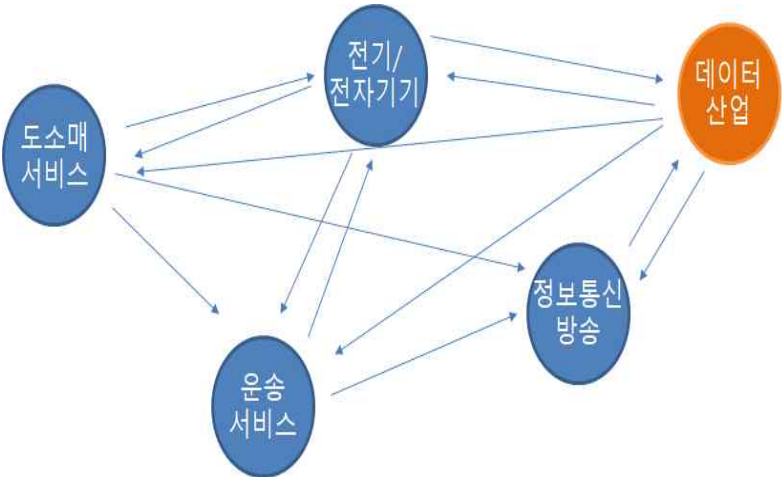
의하고 있다. 먼저 전방간접효과는 해당 산업이 생산에서 활용하는 중간재를 생산하는 산업들에서 일어나는 경제적 파급효과이다. 가령 데이터생태계에서 생산물이 한 단위 증가하였을 때 이 추가적인 생산을 위해서 더 많은 중간재가 투입되어야한다. 따라서 이 추가적인 수요를 충족하기 위해서 중간재를 생산하는 산업에서 역시 생산이 증가하여야한다. 데이터 생태계에서의 생산 증가가 타 산업, 특히 데이터산업에 중간재를 공급하는 산업에서의 증가로 이어지는 경제적 파급효과를 전방간접효과라고 정의하고 있다. 반면 후방 간접효과라 함은 데이터 생태계에서 생산된 생산물을 중간재로 활용하는 산업들에서 일어나는 경제적 파급효과이다. 예를 들어 데이터 생태계에서 데이터 분석 리포트가 한 단위 생산되었다고 할 때, 운송서비스에서 이 분석 리포트를 중간재로 활용하여 서비스를 생산해 낼 수 있다. 이 경우 데이터 생태계에서의 생산량 증가가 운송서비스에서의 추가적 생산으로 이어지는 경제적 파급효과를 후방간접효과라고 할 수 있다. 마지막으로 유도효과(Induced effect)는 이러한 전·후방간접효과를 제외한 나머지 경제적 파급효과로서, 예를 들어, 데이터 생태계에서의 생산량 증가로 인해 데이터 생태계에 종사하는 인구의 임금이 증가하고 이로 인해 국내 소비가 진작되어 다시 국내 생산이 증가하는 효과 등을 포괄한다.

경제(Economy)라는 단어의 사전적 의미는 '인간의 생활에 필요한 재화나 용역을 생산·분배·소비하는 모든 활동. 또는 그것을 통하여 이루어지는 사회적 관계'⁵⁾이다. 따라서 경제는 비단 한 산업을 지칭하기보단 경제체제 전체를 포괄하는 의미이므로, 본 연구에서는 데이터 경제(Data Economy)를 광의로 해석하고 정의하고자 한다. 경제체제의 의미를 갖는 데이터 경제를 체계적으로 바라보기 위한 하나의 프레임워크(framework)는 데이터 경제를 '데이터산업과 '데이터산업 외 타 산업으로 분류하여 바라보는 것이다. 즉, 데이터 경제를 데이터산업 자체에서의 생산 활동과 이를 제외한 타 산업의 생산 활동의 합으로 정의하는 것이다. 하지만 이 경우 단순히 우리나라의 경제체제를 재분류하여 바라보는 것에 그치게 된다. 따라서 단순한 재분류에 그치지 않기 위해서는 데이터기반경제(Data-driven economy)의 개념처럼 '데이터산업에 경제체제 전체의 기반산업으로서의 성격을 부여하고, '데이터산업 외 타 산업'이 기반산업인 데이터산업에 의해 받는 영향으로 나누어 보는 것

5) 네이버 국어사전 웹페이지 (<https://ko.dict.naver.com/#/search?query=%EA%B2%BD%EC%A0%9C&range=all>, 2018. 11. 2. 검색).

이 타당하다. 이를 위해서는 먼저 데이터산업과 데이터 외 타 산업이 어떻게 서로 영향을 주고받는지에 대한 논의가 필요하다. 모든 경제체제에서 어떠한 산업이 타 산업과의 연관관계 없이 작동하는 경우는 거의 없다고 해도 무방하다. 데이터 경제에서도 역시 기반산업인 데이터산업과 데이터산업 외 타 산업은 계속해서 아래 그림처럼 서로가 서로에게 영향을 주고받게 된다.

[그림 2-1] 데이터산업과 타 산업 간의 연관관계

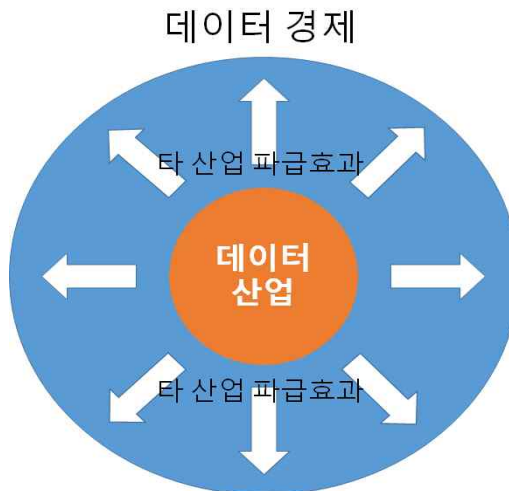


예를 들어 데이터산업은 최종재인 데이터분석서비스를 생산하기 위해서 전기/전자기기 산업에서 컴퓨터 등 하드웨어를 구입해야하고, 대용량 데이터를 다운로드하기 위해서는 정보통신방송산업의 최종생산물인 네트워크 서비스를 이용해야한다. 이 경우 전기/전자기기산업과 정보통신방송 산업은 데이터산업에서 최종재 생산을 위한 중간재를 공급하는 산업으로서 데이터산업과 영향을 주고받는다 할 수 있다. 반면에 데이터산업의 최종재인 데이터분석서비스가 타 산업에서 다시 중간재로 사용되기도 한다. 예를 들어 운송서비스나 도소매서비스 산업에서는 데이터산업에서 생산된 데이터분석서비스를 중간재로 활용하여 운송 또는 도소매서비스라는 최종재를 생산하는 경우를 생각해 볼 수 있다. 이 경우 데이터산업은 운송, 도소매서비스 산업에 중간재를 공급하는 산업으로서 이들 산업에 영

향을 주고받는다. 즉, 데이터 경제에서 데이터산업과 데이터산업 외 타 산업은 각자의 생산물을 중간재·최종재로 활용하며 상호작용한다.

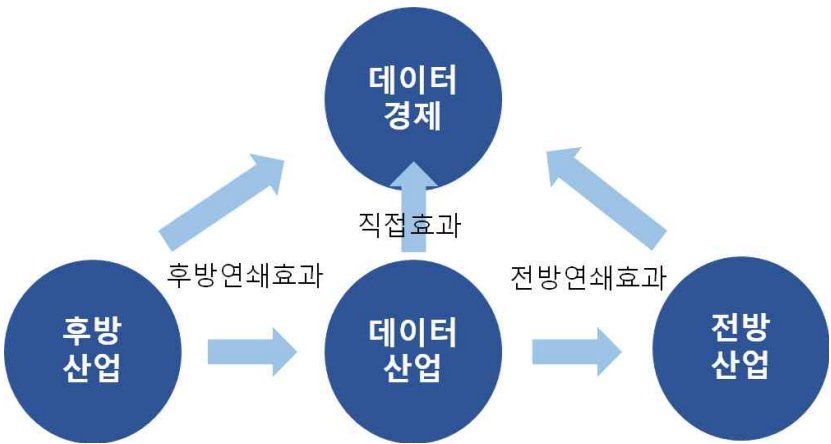
이상의 논의를 바탕으로 본 연구에서는 데이터 경제를 다음과 같이 정의하고자 한다. 먼저 데이터 경제를 데이터산업과 데이터산업의 타 산업 파급효과로 나누어 보고자 한다. 이러한 프레임워크(framework)에서 데이터산업은 데이터 경제 전반에 '데이터'라는 자원을 데이터 경제에 생산·공급하는 기반산업으로서의 의미를 가진다. 기반산업인 데이터산업의 타 산업 파급효과는 크게 직접효과와 간접효과 두 가지로 나누어 볼 수 있다. 먼저 직접효과는 데이터산업의 생산 활동 증가로 인해 직접적으로 증가하는 생산물의 가치로 데이터산업의 매출액 등으로 그 규모를 추정할 수 있다. 간접효과는 다시 전방연쇄효과와 후방연쇄효과로 나누어 볼 수 있는데, 먼저 후방연쇄효과는 기반산업인 데이터산업에 중간재를 공급하는 후방산업에 미치는 영향으로 데이터산업의 생산 활동 증가로 인해 이를 중간재로 활용하는 타 산업에서의 생산물 가치로 그 효과를 추정할 수 있다. 전방연쇄효과는 기반산업인 데이터산업의 최종재를 중간재로 활용하는 전방산업에 미치는 경제적 영향으로 데이터산업의 생산 활동 증가에 필요한 중간재를 타 산업에서 추가적으로 구매하여 증가하는 타 산업 생산물 가치로 추정할 수 있다.

[그림 2-2] 데이터 경제: 데이터산업과 타 산업 파급효과



데이터 경제는 데이터산업이 생산한 최종재의 가치, 이를 생산하기 위해 추가적으로 필요한 중간재들이 후방산업에서 생산되면서 증가한 생산물의 가치, 데이터산업의 최종재를 중간재로 활용하여 타 산업에서 생산되는 생산물의 가치로 구성되어 있다고 할 수 있다. 이는 앞서 살펴본 IDC(2017b)가 데이터 경제 규모를 추정할 때 사용된 정의와 동일한 것인데, IDC(2017b)의 연구에서는 이들의 합을 데이터 경제의 규모로 정의하고 있다.⁶⁾ 하지만 본 연구에서는 직접효과, 전방연쇄효과, 후방연쇄효과를 단순하게 합하기에는 한계⁷⁾가 있는 방법론을 사용함으로 데이터 경제의 구성요소를 기반산업인 데이터산업과 데이터산업의 전·후방연쇄효과로 정의하고 이를 각각 살펴보도록 하겠다.

[그림 2-3] 데이터산업의 직접효과, 후방연쇄효과, 전방연쇄효과



6) IDC(2017b)는 이에 더하여 앞서 언급하였듯이 데이터산업에 종사하는 임금노동자의 임금상승으로 인한 수요 증가 등의 유도효과 역시 데이터 경제에 포함하고 있다.
 7) 뒤에 상세히 설명되겠지만 본 연구에서 사용한 산업연관분석에서 도출된 전방연쇄효과와 후방연쇄효과는 동일한 정보로부터 추출되는데 이에 따라 두 효과가 서로 완벽히 상호배타적(mutually exclusive)하지 않아 이를 단순하게 합할 경우 경제적 파급효과 과대 추정될 수 있다.

제 2 절 국내외 데이터 경제 현황

본 절에서는 앞서 살펴본 데이터 경제의 정의에 따라 국내외 데이터 경제 현황을 데이터 산업과 데이터산업의 경제적 파급효과로 나누어 살펴보고자 한다. 먼저 해외 주요국인 유럽, 미국, 일본, 중국의 데이터 경제 현황을 살펴보고, 국내의 경우 제3절에서 다른 데이터산업의 경제적 파급효과를 제외하고 데이터산업 현황을 중점적으로 살펴보도록 하겠다.

1. 해외 주요국의 데이터 경제 현황

유럽, 미국, 일본의 데이터 경제 현황은 유럽위원회의 의뢰로 유럽, 미국, 일본의 데이터 경제 규모를 추정한 IDC의 보고서를 바탕으로 재구성하였다. 우선 데이터 관련 인력, 데이터 기업, 데이터산업으로 나누어 살펴본 뒤, 데이터산업의 경제적 파급효과에 대해 논의하도록 하겠다. 다음으로 중국의 경우 빅데이터산업 중심으로 데이터 경제 현황을 살펴보도록 한다.

가. 유럽연합(EU)⁸⁾

유럽의 데이터 경제는 계속해서 성장세를 이어오고 있다. IDC는 2017년과 2018년 보고서에서 2013년부터 2017년까지의 데이터 경제의 규모를 추정하고 2020년에 규모를 전망하였다. 먼저 데이터산업⁹⁾의 규모는 2017년 약 650억 유로에서 2020년 774억 유로로 성장할 것으로 전망하였다. 이에 따라 데이터 관련 인력 및 데이터 기업 역시 계속하여 증가할 것으로 전망하였다. 데이터 관련 인력은 2017년 669만 명에서 2020년 831만 명으로, 데이터 기업¹⁰⁾의 경우 2017년 27만 6천개에서 2020년 30만 6천개 개로 증가할 것으로 전망하였다.

유럽의 데이터산업 규모에 대해 조금 더 자세히 살펴보면 영국의 데이터산업 규모를 포

8) 이후에는 혼란의 여지가 없을 경우 유럽연합을 유럽이라고 지칭하도록 한다.

9) IDC에서는 데이터산업 대신 데이터 시장(Data Market)이라는 용어를 사용하였는데 데이터 시장의 정의는 “디지털 데이터가 가공되어 ‘상품’ 또는 ‘서비스’로서 교환되는 시장”으로 데이터산업과 정의 상 큰 차이가 없다.

10) 데이터 기업의 상세한 정의는 “디지털 데이터의 생산과 분배를 주된 활동으로 하는 데이터 공급 조직”이다.

합할 경우 2013년 약 470억 유로에서 꾸준히 증가하여 2017년에는 약 650억 유로로 조사되었고, 약 6.0%의 평균성장률로 성장이 지속될 경우 2020년에는 약 774억 유로에 달할 것으로 전망된다. 영국의 데이터산업을 포함하지 않더라도 2013년 약 380억 유로의 규모에서 2017년 약 505억 유로로 성장하여, 약 6.1%의 연평균 성장률로 성장할 경우 2020년에 약 602억 유로에 이를 것으로 전망된다.

〈표 2-1〉 유럽 데이터산업 규모 추이

(단위: 백만유로)

구분	2013	2014	2015	2016	2017	2020(E*)	2017-2020 연평균성장률(%)
데이터산업 규모 (EU28)	47,419	50,888	54,351	59,496	65,038	77,407	6.0
데이터산업 규모 (EU27)	38,022	41,012	42,586	46,183	50,483	60,254	6.1

주: 1) EU28은 Brexit 후에도 영국이 EU와 종전과 같은 수준으로 EU역내에서 경제활동을 할 것이라는 가정을 하고 영국의 데이터산업 규모가 포함된 수치이고(EU27)은 영국을 포함하지 않은 수치임

2) 2013년부터 2015년까지 추정 규모는 IDC(2017b), 2016년부터 2017년 추정규모 및 2020년 전망규모는 IDC(2018)

자료: IDC(2017b), IDC(2018)

IDC(2018)에서는 데이터 경제 규모를 직접효과인 데이터산업의 규모와 데이터산업의 전후방 간접 파급효과 및 유도효과의 합으로 정의하고 추산하였는데, 이에 따르면 유럽의 데이터 경제 규모는 2017년 약 3,356억 유로에서 연평균 10.4% 성장해 2020년 약 4,522억 유로에 이를 것으로 전망하고 있다. 이는 영국의 데이터 경제 규모를 포함한 수치로, 영국을 제외한 유럽의 데이터 경제 규모는 2017년 약 2,432억 유로에서 연평균 11%씩 증가하여 2020년 약 3,658억 유로에 달할 것으로 예상하고 있다.

〈표 2-2〉 유럽 데이터 경제 규모 추이 및 전망

(단위: 백만유로)

구분		2016	2017	2020	2016~2017 성장률(%)	2017~2020 연평균성장률 (%)
데이터 경제 규모 (직접효과 및 전후방과급효과, 유도효과 포함)	EU28	305,977	335,618	452,190	9.7	10.4
	EU27	243,205	267,006	365,761	9.8	11.1

주: EU27은 영국을 포함하지 않은 수치임

자료: IDC(2018)

다음으로 데이터산업 내의 고용현황은 영국을 포함할 경우 2017년 까지 약 6백만 명이 고용되어 있으며, 2020년에는 2017년 이후 연평균 7.5% 증가해 약 8백만 명의 인원이 데이터산업에 고용될 것으로 추산된다. 영국을 제외할 경우 2017년 약 5백만 명이 데이터산업에 종사하고 있으며, 2020년에는 약 6백만명에 이를 것으로 전망된다.

〈표 2-3〉 유럽 데이터산업 고용 인원 추이

(단위: 천명)

구분	2014	2015	2016	2017	2020(E*)	2017~2020 연평균성장률 (%)
데이터 관련 인력(EU28)	5,818	6,005	6,187	6,685	8,309	7.5
데이터 관련 인력 (EU27)	4,707	4,730	4,875	5,273	6,619	7.9

주: 1) EU27은 영국을 포함하지 않은 수치임

2) 2014년부터 2015년까지 추정 규모는 IDC(2017b), 2016년부터 2017년 추정규모 및 2020년 전망규모는 IDC(2018)

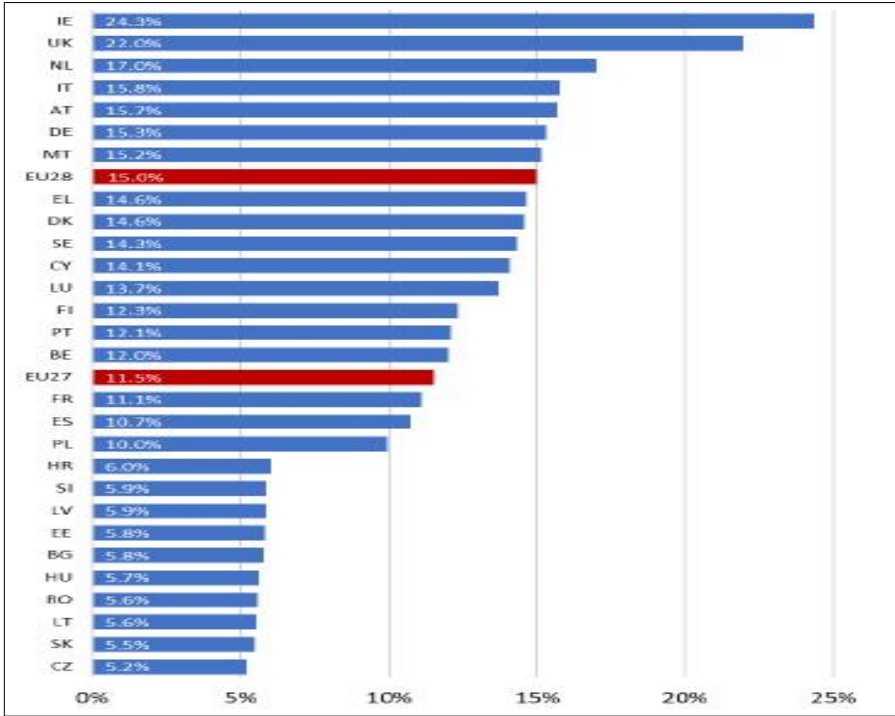
자료: IDC(2017b), IDC(2018)

다음으로 데이터 관련 제품, 서비스, 기술을 제공하는 기업으로 정의된 데이터 기업 혹은 데이터 공급기업 수의 변화추이를 살펴보면 2017년 기준 데이터산업을 구성하고 있는 데이터공급업체인 데이터 기업의 수는 약 27만 6천개로 이 역시 꾸준히 증가하여 2020년에는 약 30만 6천개에 이를 것으로 추산된다. 주목할 점은 유럽연합 회원국 별 데이터 기

업의 수는 2017년 기준 영국이 13만 4,550개로 가장 많지만, 영국 데이터 기업이 영국 시장에서 차지하는 비중은 22% 수준으로 영국의 데이터 기업들은 주로 중소기업임을 추론할 수 있다. 반면 아일랜드의 경우 데이터 기업 수가 타 국가에 비해 높지 않은 수준임에도 유럽 데이터산업에서 차지하는 비중은 24.3%로 나타나 아일랜드에 속한 데이터 기업들의 규모가 상대적으로 큰 것을 알 수 있다.

[그림 2-4] EU 회원국별 데이터 업체 비중

(단위: %)



주: ICT 및 전문서비스분야의 전체 기업에서 차지하는 비중
 자료: IDC(2018)

유럽 전체의 데이터 기업을 규모별로 살펴볼 경우 약 99%가 중소기업으로 고용인원이 250명 미만인 기업이 대부분으로 나타나 유럽 전반에 걸쳐 데이터산업이 중소기업 위주로 구성되어 있음을 알 수 있다.

〈표 2-4〉 EU 회원국 별 데이터 기업 수 및 시장비중 현황

(단위: 개, %)

연도	249명 미만 고용 업체	250명 이상 고용업체	합계 (EU28)	중소기업 비중(%)
2013	237,128	2,717	239,845	98.9
2014	240,840	2,760	243,600	98.9
2015	246,313	2,787	249,100	98.9
2016	258,650	2,800	261,450	98.9
2017	273,250	3,200	276,450	98.8

주: 1) 영국 포함 EU28 기준임

2) 2013년부터 2015년까지 추정 규모는 IDC(2017b), 2016년부터 2017년 추정규모는 IDC(2018)

자료: IDC(2017b), IDC(2018)

나. 미국

미국의 데이터산업은 두 자릿수 성장률을 기록할 정도로 가파른 성장세에 있다. IDC(2018)에 따르면 미국 데이터산업의 규모는 유럽 데이터산업 규모의 약 2배로 2017년 기준 약 1,455억 4,600만 유로(전년 대비 12.7% 성장)로 추정된다.¹¹⁾ 또한, 데이터산업의 직접효과¹²⁾와 전후방연과효과¹³⁾는 2017년 기준 각각 1,136억 7,700만 유로(전년 대비 4.8% 증가), 77억 6,600만 유로(전년대비 6.8% 증가)로 추정된다. 이를 2017년 EU 28개국의 데이터 경제 규모와 비교해보면, EU의 직접효과 규모(650억 3,800만 유로)와 간접 후방과급효과규모(33억 800만 유로)를 각각 74.8%, 134.8% 초과한 수준이다. 또한, 데이터산업 고용인원 규모도 2017년 전년 대비 10.1% 증가하여 EU 28개국의 669만 명을 2배 이상 상회하는 1,401만 2,000명이다.

11) 민대홍 외(2018).

12) 데이터 공급기업의 데이터 상품과 서비스 매출을 기반으로 추정된 직접적인 효과이다.

13) 데이터공급기업의 간접 후방과급효과 규모이다.

〈표 2-5〉 미국 데이터산업 및 데이터 경제 규모(2016~2017년)

구분	2016년	2017년	2016-2017 성장률(%)
데이터산업 규모(백만 유로)	129,173	145,546	12.7
데이터 경제 규모(직접효과: 백만 유로)	108,521	113,677	4.8
데이터 경제 규모(간접후방과급효과: 백만 유로)	7,270	7,766	6.8
데이터산업 고용 인원 규모(천 명)	12,732	14,012	10.1

자료: IDC(2018)

미국의 데이터산업이 이처럼 가장 큰 규모로 성장할 수 있었던 주요기반은 데이터 브로커 시장으로 9개 주요 업체들의 총매출액은 2012년 기준 4억 2,600만 달러로 조사되었다.¹⁴⁾ 이미 20세기 중반부터 활성화된 미국의 데이터 브로커는 소비자의 개인 정보를 수집, 가공하여 판매하는 서비스를 제공하는데, 이들은 데이터베이스를 활용해 마케팅, 신용조회, 사기탐지 등에 활용하고 있다. 데이터 브로커의 M&A도 활발하게 진행되고 있어 FTC (2014)의 시장규모 추정에 포함된 9개 업체 중 5개 업체¹⁵⁾가 데이터 관련 솔루션회사, 신용 정보회사 등에 인수된 상태이다.¹⁶⁾

〈표 2-6〉 미국 데이터 브로커 업체 현황

회사 명	설립연도	사업분야
Acxiom	1969	• 부정사용 탐지를 위한 고객데이터 분석 및 마케팅캠페인 서비스 제공
Corelogic	1991	• 산업계와 정부에 재무정보와 부동산정보에 기초한 분석서비스 제공
Datalogix	2002	• 소비자 대상 마케팅 데이터 제공 • 2014년 12월 Oracle에 인수됨

14) FTC(2014).

15) Datalogix, eBureau, ID Analytics, Intelius, Rappleaf 5개 업체가 현재 각각 Oracle, Transunion, Symantec, H.I.G. Capital, TowerData에 인수되었다(민대홍 외, 2018).

16) 민대홍 외(2018).

회사 명	설립연도	사업분야
eBureau	2004	<ul style="list-style-type: none"> • 마케터와 재무 관련 회사, 온라인유통업체에 수익성이 높은 잠재 고객과 부정 거래 예측 서비스 제공 • 2017년 10월 개인신용정보회사인 Transunion에 인수됨
ID Analytics	2002	<ul style="list-style-type: none"> • 2002년 설립, 특정인 확인, 부정 거래 확인 서비스 제공 • 2017년 2월 Symantec에 인수됨
Intelius	2003	<ul style="list-style-type: none"> • 신원 조회와 공문서 정보 제공 • 2015년 7월 투자회사 H.I.G. Capital에 인수됨
PeekYou	2006	<ul style="list-style-type: none"> • 소셜미디어사이트, 홈페이지, 블로그의 콘텐츠를 분석해 작성자 확인 서비스 제공
Rapleaf	2005	<ul style="list-style-type: none"> • 이메일 주소, 이메일 주소 소유자의 연령, 성, 우편번호, 소득, 결혼 여부, 자녀 여부와 취미, 구매 유형 등 정보 제공 • 2013년 12월 TowerData에 인수됨
Recorded Future	2009	<ul style="list-style-type: none"> • 소비자와 기업의 과거 이력 데이터 분석을 통해 미래 행동 예측 정보 제공

자료: FTC(2014), 정용찬(2015), 민대홍 외(2018) 재인용

위에 언급된 데이터 브로커 업체를 제외하고도 많은 데이터 브로커들이 존재하는데, 이들은 소비자 관련 정보(소비자들은 거주 지역 또는 소득 수준 등)를 포함한 다양한 종류의 데이터(부동산, 자연재해, 교육, 신용 등)를 여러 정보원에서 수집·제공하고 있다. 데이터 브로커들은 정부의 공공데이터를 활용하거나 특정 데이터를 구입하는 방법으로 데이터를 수집한다.

〈표 2-7〉 데이터 브로커의 분야별 제공 데이터 및 업체

분야	데이터	제공업체
소비자	<ul style="list-style-type: none"> • 개인신상정보인 이름, 생일, 전화번호, 이메일, 신용카드번호, 사회보장번호 등을 바탕으로 계층, 소득, 지역, 가족 수, 미디어이용행태 등 그룹형태로 제공 	<ul style="list-style-type: none"> • Acxiom, comScore, Harte Hanks, Nielsen, Tru Optik, US Data Corporation
상업	<ul style="list-style-type: none"> • 회사정보인 기업 명, 주소, 전화번호, 재무정보, 인력, 물류, 투자 등 	<ul style="list-style-type: none"> • Alacra, Bloomberg, CoreLogic, Dow Jones, FactSet, Lexis-Nexis, Risk Solutions, Pitch-Book

분야	데이터	제공업체
신용	• 소비자 및 기업의 신용점수, 재무이력 등 신용 정보	• Dun & Bradstreet, Equifax, Experian, TransUnion
정부	• 정부에서 제공하는 교통, 도로, 도시개발, 법, 세제, 구매, 판결, 행정서비스 등의 데이터	• BidNet, NTIS, Data.gov, Onvia
과학 및 기술	• 자연재해, 날씨, 우주, 해양, 특허, 혁신, 지식재산권, 생명과학, 약제조, 소비자인구통계, 의약품연구, 정부허가, 질병연구 등	• Apervita, IMS Health, Meteogroup, Monsanto, The Weather Company(IBM)
교육 및 훈련	• 교육과정, 교재, 입학, 졸업 등	• Jstor, ProQuest
부동산	• 부동산정보로 주택, 땅, 농장, 상가, 빌딩 등의 주소, 소유권, 세금, 판매 및 렌트 가격 등	• MLS.com, Redfin, Zillow
제품 및 서비스	• 제품, 서비스, 브랜드, 리뷰, 기술정보, 가격, 구입처, 뉴스, 온오프라인 발간물 내용 등	• BrightPlanet, Gracenote, HG Data, Sportvision, Zagat
지리	• 전 세계 도로, 교통상황, POI 등	• Esri, Factual, Inrix, MapLarge
위험관리	• 사이버테러, 범죄, 전쟁 등	• Recorded Future, Verisk Analytics

자료: Gartner(2017), 민대홍 외(2018) 재인용

데이터 브로커 서비스의 주요 구매자는 다양한 산업의 기업들과 정부와 같은 공공기관 또는 대학교와 같은 비영리조직 등인데, 이들은 직접적으로 소비자로부터 정보를 수집하지 않고 주로 공개정보를 활용하거나 정부 또는 민간으로부터 데이터를 수집하여 서비스를 제공하고 있다. 이들이 수집하는 공개정보를 구체적으로 살펴보면 먼저 연방정부 기관 중 공개정보를 취득하는 기관은 인구조사국(U.S. Census Bureau)으로 특정 지역(도시나 구역)의 정보, 예를 들어, 인종분포, 연령분포, 인구의 교육수준, 그들의 가족구성, 수입 분포, 직업의 종류, 통근 시간 등 해당 기관에서 수집·공개된 정보들을 활용하고 있다. 또한 사회보장국으로부터는 사망자 이름, 사회보장번호, 사망일 등의 정보를 활용하고 있는데, 사회보장번호는 우리나라의 주민등록번호와 같은 개념이다. 그 밖에도 우편서비스, 연방법원으로부터 주소변경 정보와 과산정보를 수집하여 활용하고 있다. 주정부로부터는 전문직 면허 정보, 사냥·낚시 면허와 같은 세부정보를 수집하거나 부동산 관련 정보, 유권자 등록정보, 자동차 관련 정보, 법원과 같은 정보를 추가적으로 수집하여 활용한다.

이 밖에도 인터넷 상에 공개된 개인신상정보 - 예를 들어 우리나라의 사람인과 같은 역

할을 하는 취업플랫폼인 링크드인(Linked In)에 공개된 학력, 직장, 연봉수준, 전문분야, 인적네트워크 등을 활용하기도 하며, 우리나라의 대형식료품 매장이나 전자상거래 플랫폼으로부터 소비자들의 소비 패턴 관련 정보를 매입한다. 또한, 금융회사로부터는 개인들의 금융상품 거래 정보를 구입하고, 포털로부터는 개인의 이메일, 우편번호, 성명 등의 정보를 구입하여 활용하기도 한다. 이에 더하여 다른 데이터 브로커들이 보유한 정보들을 서로 거래하여 활용하기도 한다.

〈표 2-8〉 미국 데이터 브로커의 분야별 수집 데이터

분야		데이터
정부	연방 정부	<ul style="list-style-type: none"> • 인구조사국: 특정 도시나 구역의 인종, 연령, 교육수준, 가족 구성, 수입, 직업, 통근 시간, 도로, 주소, 행정구역, 선거구, 학군 정보 • 사회보장국: 사망자 이름, 사회보장번호, 사망일 • 우편서비스의 주소 변경 이력 정보 • 연방법원: 파산 정보
	주정부 지방정부	<ul style="list-style-type: none"> • 전문직 면허(의사, 변호사, 조종사, 건축사) 정보와 사냥, 낚시 면허 정보 • 부동산: 세무, 주택소유권과 담보, 소유자 정보, 부동산 세무 내역 정보 (욕실 수, 침실 수, 수영장 유무 등) • 유권자 등록정보: 이름, 주소, 생일, 가입 정당 • 자동차등록정보와 운전 기록 • 법원: 범죄 기록, 민사 소송, 출생, 결혼, 이혼, 사망 기록
공개정보		<ul style="list-style-type: none"> • 전화번호부, 보도자료, 블로그와 소셜미디어사이트(링크드인과 같은 공개형) 등 인터넷에 올린 개인정보
상업정보		<ul style="list-style-type: none"> • 소매업, 카탈로그회사의 거래내역, 품목(예를 들어 고가의 구두, 친환 경식품 등), 구매일자, 결제 방법 데이터 • 잡지사의 구독자 데이터 • 전자상거래, 뉴스, 여행 사이트의 고객 이름, 이메일, 우편번호 정보 • 금융서비스회사의 거래 데이터 • 다른 데이터 브로커가 보유한 데이터

자료: FTC(2014), 정용찬(2015), 민대홍 외(2018) 재인용

다. 일본

최근 일본의 데이터산업 및 데이터 경제는 ICT 투자가 재개되고 데이터 경제가 자리 잡아가면서 2016년 대비 견조한 성장세를 실현한 것으로 추정된다.¹⁷⁾ IDC(2018)에 따르면, 2017년 일본 데이터시장 규모는 전년 대비 8.7% 성장한 277억 2,300만 유로로 추정되며, 데이터산업 고용 인원 규모는 2016년의 374만 명에 비해 8.0% 증가한 404만 명으로 조사되었다. 이에 더하여 데이터산업 자체에서 생산된 가치를 의미하는 데이터 경제의 직접영향 규모는 2017년 기준 전년대비 9.3% 증가한 299억 4,900만 유로로, 데이터산업의 간접후방과급효과 규모는 2016년 대비 6.7% 증가한 12억 6,900만 유로로 추정된다.¹⁸⁾

〈표 2-9〉 일본 데이터산업 및 데이터 경제 규모(2016-2017년)

구분	2016년	2017년	2016~2017 성장률(%)
데이터시장 규모(백만 유로)	25,513	27,723	8.7
데이터 경제 규모(직접효과: 백만 유로)	27,394	29,949	9.3
데이터 경제 규모(간접후방과급효과: 백만 유로)	1,189	1,269	6.7
데이터산업 고용 인원 규모(천 명)	3,740	4,040	8.0

자료: IDC(2018), 민대홍 외(2018) 재인용

〈표 2-10〉 일본과 미국의 데이터산업 및 데이터 경제 규모 비교(2017년)

구분	규모		미국 대비 비중(%)
	일본	미국	
데이터시장 규모(백만 유로)	27,723	145,546	19.0
데이터 경제 규모(직접효과: 백만 유로)	29,949	113,677	26.3
데이터 경제 규모(간접후방과급효과: 백만 유로)	1,269	7,766	16.3
데이터산업 고용 인원 규모(천 명)	4,040	14,012	28.8
데이터 기업 수	104,664	302,810	34.6

자료: IDC(2018) 재구성, 민대홍 외(2018) 재인용

17) IDC(2018).

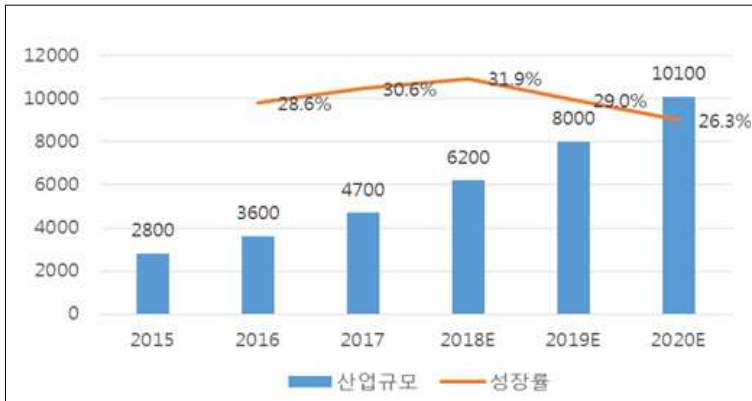
18) 민대홍 외(2018).

그러나 일본의 데이터산업 및 데이터 경제 규모는 급속한 성장에도 불구하고, 미국과 비교하면 크지 않은 편이다. 2017년 기준 일본의 데이터시장 규모, 직접효과규모, 간접후방과급효과 규모, 데이터산업 고용인원 규모, 공급업체 규모는 각각 미국의 19.0%, 26.3%, 16.3%, 28.8%, 34.6%으로 조사되었다.

라. 중국

중국의 경우 전 산업에 걸쳐 빅데이터 활용이 확산되면서 빅데이터 산업이 빠른 속도로 성장 중이다. 중국정보통신연구원(2018.4)에 따르면, 2017년 중국의 빅데이터 산업 규모는 4,700억 위안으로 전년 동기대비 30.6% 증가한 것으로 추정된다. 폭발적인 성장세는 향후에도 이어져 2017년부터 2020년까지 연평균 29.0%씩 중국의 빅데이터 산업규모가 확대될 것으로 전망된다.

[그림 2-5] 중국 빅데이터 산업규모와 성장률
(단위: 억 위안)

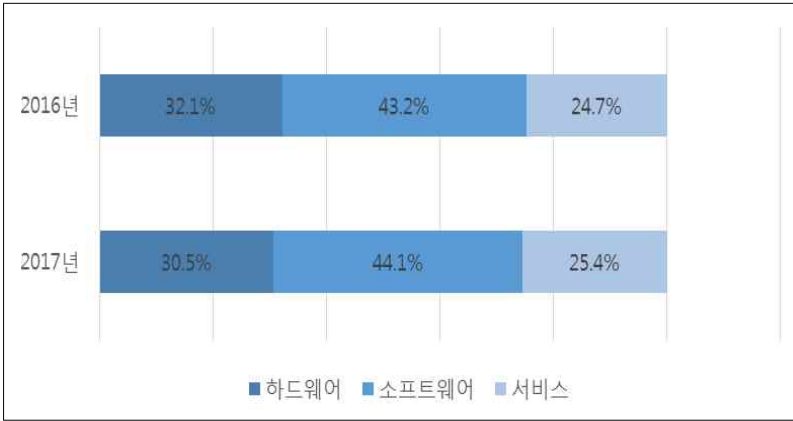


자료: 中国信息通信研究院(2018. 4), 민대홍 외(2018) 재인용

2017년 기준 세부 분야별 중 비중이 가장 큰 분야는 SW로 전년 대비 0.9% 증가한 104억 위안의 시장규모로 전체 빅데이터 산업의 44.1%를 차지한 것으로 추정된다. SW 다음으로 비중이 높은 HW는 2017년 전년 대비 1.6% 감소한 72억 위안으로 전체 산업에서 차지하는 비중은 30.5%에 달한다.¹⁹⁾

19) 中国信息通信研究院(2018. 4).

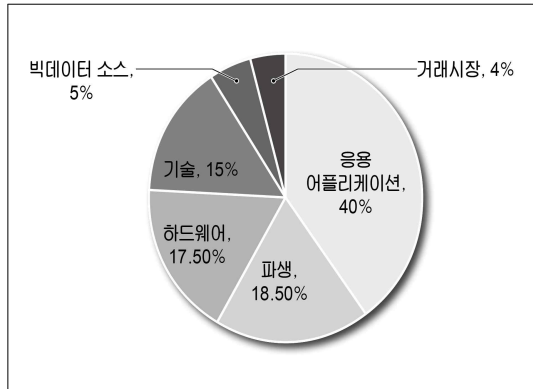
[그림 2-6] 중국 빅데이터 세부 분야 구성



자료: 中国信息通信研究院(2018. 4)

빅데이터의 분야별 구성을 살펴보면 응용어플리케이션산업이 가장 큰 부분을 차지하며, 2020년 기준 빅데이터 산업 규모의 40%에 이를 것으로 기대된다. 반면, 가장 낮은 비중을 차지하는 분야는 빅데이터거래시장으로 빅데이터 산업의 4%에 불과한 것으로 추정되었다.²⁰⁾

[그림 2-7] 2020년 중국 빅데이터 산업 구성



자료: 贵阳大数据交易(2016. 5)

20) 贵阳大数据交易(2016. 5).

중국의 빅데이터산업 생태계는 빅데이터를 생산 및 유통 관련 데이터구축 및 솔루션 포함 빅데이터 인프라를 담당하는 기술 분야와 데이터 수집, 처리, 분석, 거래 등 데이터 서비스, 그리고 금융, 의료 등 빅데이터를 활용하는 분야 등 크게 세 분야로 구성되어 있다. 특히, 빅데이터 활용 분야는 데이터 전문 활용기업 뿐만 아니라 물류기업 차이나오, 차량 공유기업 디디추싱, 인터넷포탈기업 바이두처럼 타 산업에서의 주요 기업들이 데이터 활용을 통한 부가가치 창출에 적극적으로 나서고 있다.

〈표 2-11〉 중국 빅데이터 산업 분야별 주요 기업

구분	서비스 유형	주요 기업
빅데이터활용	정부빅데이터	
	금융빅데이터	
	공업빅데이터	
	여행빅데이터	
	안보빅데이터	
	소셜미디어빅데이터	
	의료빅데이터	
	농업빅데이터	
	통신빅데이터	
	교통빅데이터	

구분	서비스 유형	주요 기업
데이터서비스	데이터거래서비스	
	데이터 수집 및 처리 서비스	
	데이터 분석 및 가시화 서비스	
	데이터 안보 서비스	
기술서비스	데이터 저장 관리	
	데이터 분석 소프트웨어	
	데이터 플랫폼 구축 솔루션	
	빅데이터 관련 하드웨어	
	서버 및 시스템 관리	
	빅데이터 처리 도구	

자료: 中国电子信息产业发展研究院(2017)

〈표 2-12〉 중국 주요 기업의 빅데이터 활용 사례

주요 기업	세부 내용
 차이나오(菜鸟)	<ul style="list-style-type: none"> 물류 빅데이터와 창고 빅데이터를 통합하여 물류 기초설비 플랫폼을 구축, 물류기업을 대상으로 스마트한 수요공급체인 솔루션을 제공
 바이두(百度)	<ul style="list-style-type: none"> 사용자의 검색과 열람 등 사용기록 데이터를 통합하여 편리한 검색 서비스를 제공 중, 동시에 광고기업을 대상으로 이용자의 맞춤형 광고 서비스를 제공
 디디추싱(滴滴出行)	<ul style="list-style-type: none"> 사용자의 사용기록과 교통 데이터를 통합하여 맞춤형 이용 서비스 제공
 차이나텔레콤(电信)	<ul style="list-style-type: none"> 사용자의 모바일 데이터를 통합하여 맞춤형 모바일 데이터를 구축, 이를 통해 마케팅, 마켓 포지셔닝, 신용분석 등의 서비스를 제공
 앤트파이낸셜(蚂蚁金服)	<ul style="list-style-type: none"> 사용자의 소비데이터를 통합하여 신용평가 체계를 구축, 이를 통해 온라인 결제, 예금, 신용대출 등 온라인 금융 체계를 완성
 텐센트(腾讯)	<ul style="list-style-type: none"> 웨이신 등 SNS 상 사용자의 이용 데이터를 통해 사용자의 소비행동을 분석, 이를 통해 전자상거래 등 기업을 대상으로 맞춤형 광고 서비스 등을 제공

자료: 中国电子信息产业发展研究院(2017)

2. 우리나라의 데이터산업 현황

우리나라의 경우 데이터산업의 경제적 파급효과는 향후 자세히 논의될 것이기에 데이터 산업에 집중하여 현황을 살펴보도록 하겠다. 우리나라의 경우 한국데이터진흥원에서 매해 실시하는 ‘데이터산업시장 현황(2018)’을 통해 2016년과 2017년 예측치를 살펴볼 수 있다. 데이터산업시장 현황(2018)에서는 ‘데이터산업시장’을 ‘데이터의 생산·수집·처리·분석·유통·활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업’으로 정의하고 있는데 이는 IDC(2017b)의 데이터 경제 정의 중 데이터산업으로 해석할 수 있는 ‘데이터의 생성·축적·가공·거래·활용을 다루는 구성원들이 만들어 내는 생태계’와 매우 유사한 정의임을 알 수 있다. IDC(2017b)가 데이터산업을 데이터관련 인력, 데이터 기업, 데이터산업

의 전체 규모로 조사하였다면, 데이터산업시장 현황(2018)에서는 데이터산업을 크게 데이터솔루션, 데이터 구축/컨설팅, 데이터서비스로 나누어 조사하였다. 먼저 데이터솔루션 부문에 속하는 업체는 데이터 모델링, 분석/시각화, 검색엔진, 품질 등 관련 솔루션 제품으로 사업을 영위하는 사업체로 지적재산권, 유지·보수, 개발에서 매출 발생하는 업체로 주로 소프트웨어를 생산하는 업체이다. 데이터 구축/컨설팅 부문에 속하는 업체들은 데이터베이스 구축, 문서/음성/영상 등을 데이터로 변환/저장하는 데이터 처리, 데이터 외부 제공을 위한 API, LOD 구축 등의 사업을 영위하는 사업체이거나 데이터 품질, 설계 및 데이터 활용 등에 대한 컨설팅 사업을 영위하는 사업체로서 일반 컨설팅기업은 제외되어있다. 마지막으로 데이터서비스 부문에 속하는 업체들은 데이터 판매·중개 서비스, 원천데이터 가공을 통해 정보를 제공 또는 데이터 분석서비스를 제공함으로써 발생하는 데이터 이용료/수수료 등으로 사업을 영위하는 사업체로서 데이터 경제의 맥락에서 의미하는 ‘데이터 기업’과 가장 근접한 개념의 사업체들이다.²¹⁾

〈표 2-13〉 우리나라 데이터산업 중분류 및 정의

구분	정의
데이터솔루션	<ul style="list-style-type: none"> • 데이터 모델링, 분석/시각화, 검색엔진, 품질 등 관련 솔루션 제품으로 사업을 영위하는 사업체로 지적재산권, 유지·보수, 개발에서 매출 발생
데이터 구축 컨설팅	<ul style="list-style-type: none"> • 데이터베이스 구축, 문서/음성/영상 등을 데이터로 변환/저장하는 데이터 처리, 데이터 외부 제공을 위한 API, LOD 구축 등의 사업을 영위하는 사업체 • 데이터 품질, 설계 및 데이터 활용 등에 대한 컨설팅 사업을 영위하는 사업체
데이터서비스	<ul style="list-style-type: none"> • 데이터 판매·중개 서비스, 원천데이터 가공을 통해 정보를 제공 또는 데이터 분석서비스를 제공함으로써 발생하는 데이터 이용료/수수료 등으로 사업을 영위하는 사업체

자료: 한국데이터진흥원(2018)

먼저 국내 데이터산업 시장 규모는 2017년 기준 14조 3,047억 원으로 2016년 대비 4.0% 성장할 것으로 예상되며, 2010년 이후 연평균 7.5%의 증가율로 매년 꾸준한 성장세를 유

21) 한국데이터진흥원(2018).

지하고 있다

〈표 2-14〉 우리나라 데이터산업 규모 추이

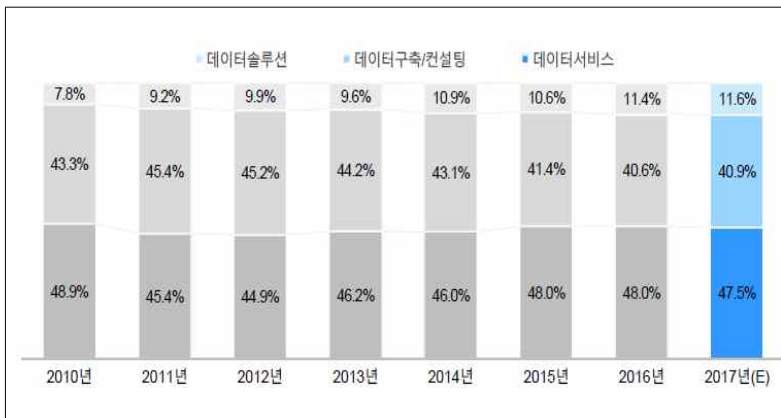
(단위: 억원)

구분	2013	2014	2015	2016	2017(E)	증감률 2016~2017(%)
데이터 솔루션	10,789	13,619	14,124	15,720	16,536	5.5
데이터 구축/ 컨설팅	49,985	53,730	55,280	55,850	58,565	4.9
데이터 서비스	52,258	57,329	64,151	65,977	67,946	3.0
전체	113,032	124,678	133,555	137,547	143,047	4.0

자료: 한국데이터진흥원(2018)

데이터산업의 분야별 비중을 살펴보면 데이터 솔루션이 약 10%, 데이터 구축 및 컨설팅이 약 40%, 데이터 서비스가 약 47%로 데이터 구축 및 컨설팅과 데이터 서비스가 큰 비중을 차지하고 있다. 이와 같은 추세는 지난 7년간 큰 변동 없이 유지되고 있으며, 이를 통해 우리나라의 데이터산업의 핵심 부문이 데이터서비스 부문인 것을 알 수 있다.

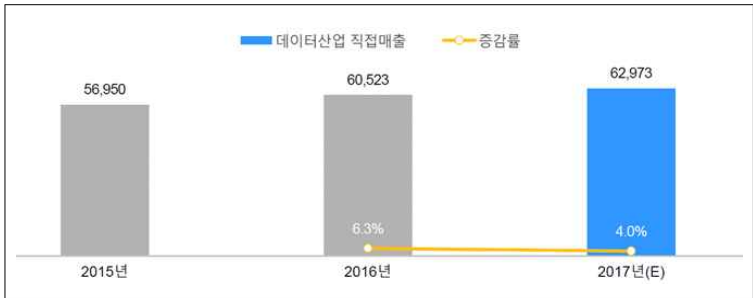
〔그림 2-8〕 국내 데이터산업 분야별 시장 비중 추이



자료: 한국데이터진흥원(2018)

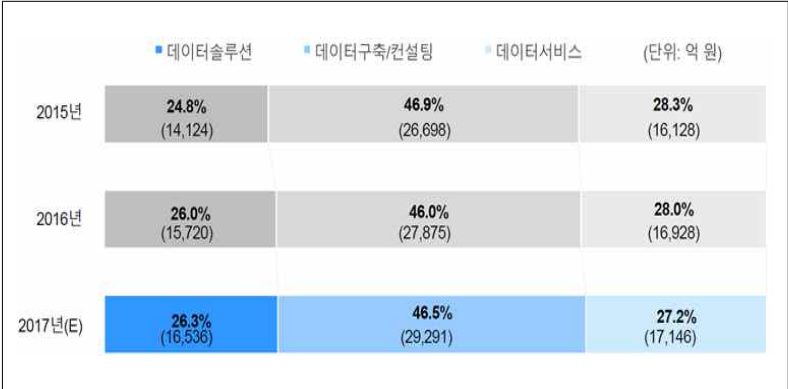
앞서 살펴본 데이터산업 시장 규모는 데이터를 활용한 광고 매출, 데이터베이스 시스템 구축 용역 등 데이터베이스 관련 간접 매출까지 포함한 광의의 시장 규모로 이를 제외한 직접매출 시장 규모는 2016년 기준 6조 523억 원에서 2017년 6조 2,973억으로 성장할 것으로 예상된다. 데이터산업 대분류별로 직접매출 시장 규모 비중을 살펴보면 데이터 솔루션이 약 26%, 데이터구축/컨설팅이 약 46%, 데이터서비스가 약 28%로 나타났다. 위의 간접매출을 포함한 산업규모와 비교해보면, 데이터서비스 부문의 경우 간접매출 포함 시 전체 산업에서 차지하는 비중이 약 47%였으나 간접매출을 제외할 경우 그 비중이 약 28%으로 데이터서비스 부문 내 간접매출 비중이 높은 것을 확인할 수 있다.

[그림 2-9] 국내 데이터산업 직접매출 규모 추이 (단위: 억 원)



자료: 한국데이터진흥원(2018)

[그림 2-10] 국내 데이터산업 직접매출 기준 대분류별 규모 및 전체대비 비중

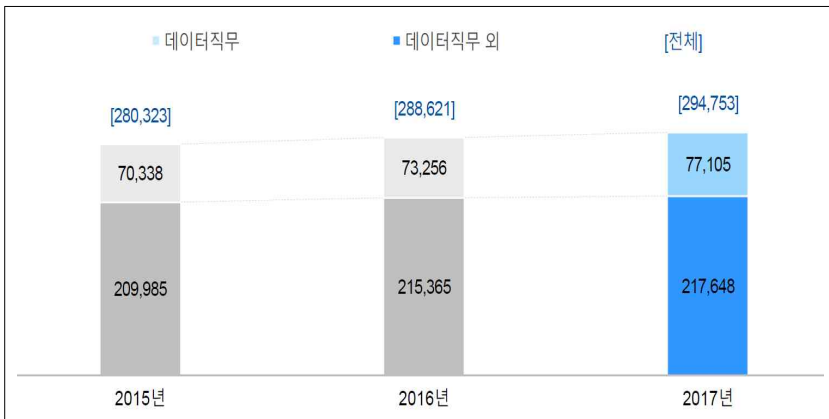


자료: 한국데이터진흥원(2018)

다음으로 국내 데이터산업의 고용현황을 살펴보면, 2017년에 데이터산업이 고용하고 있는 인력은 총 29만 4,753명으로 2016년 대비 약 2% 증가하여 계속하여 성장세에 있는 것으로 나타났다. 데이터산업에 고용된 인력은 데이터관련 직무 인력과 데이터 외 직무인력으로 나누어 볼 수 있는데, 먼저 데이터관련 직무 인력은 DA, 데이터개발자, 데이터 엔지니어, 데이터 분석가, DBA, 데이터 사이언티스트, 데이터 컨설턴트, 데이터 기획·마케터로서 고용된 자로 정의된다. 반면 데이터 직무 외 인력은 위에서 언급된 직무 이외의 인력으로 경영지원, 회계 등의 직무에 종사하는 인력이다. 데이터관련 직무의 인력의 경우 2016년 대비 약 5% 증가한 반면 데이터직무외 인력은 약 1% 증가한 것으로 나타났다.

[그림 2-11] 국내 데이터산업 인력현황

(단위: 명)

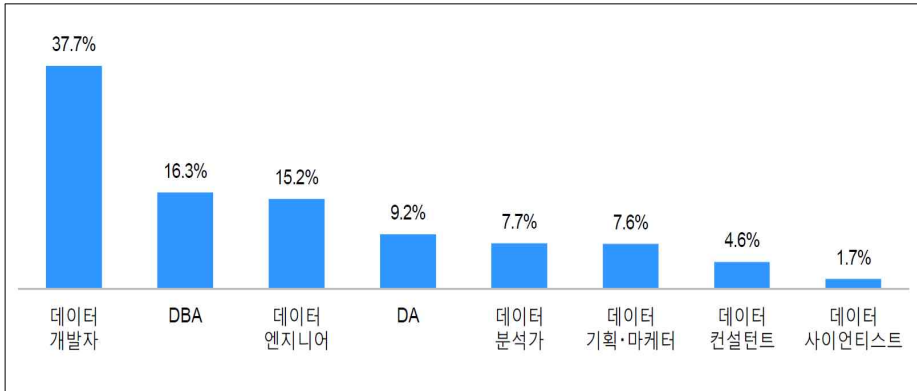


자료: 한국데이터진흥원(2018)

다음으로 데이터산업 내 데이터직무별 구성 비율을 살펴보면, 데이터 개발자가 약 40%, DBA와 데이터엔지니어가 각각 약 15%의 비중을 차지하는 것으로 나타났고, 데이터 분석가와 데이터 사이언티스트는 각각 약 8%와 약 2%의 비중을 차지하는 것으로 나타났다.

[그림 2-12] 국내 데이터산업 내 데이터 직무 별 인력현황

(단위: %)



자료: 한국데이터진흥원(2018)

데이터산업 내 R&D 등 투자는 큰 변동 없이 유지되고 있는데, 데이터산업 내 기업들의 매출액 대비 사업 투자 비중은 2016년 20.4%로 2017년에도 비슷한 수준을 유지할 것으로 조사되었다. 데이터산업 대분류 중에서 매출액 대비 투자액 비중이 가장 높은 분야는 데이터 솔루션으로 2016년의 경우 매출액의 약 24%를 투자에 지출하여 전 산업 평균을 크게 상회하는 것으로 나타났다.

<표 2-15> 국내 데이터산업 대분류별 매출액 대비 투자액 비중 현황

(단위: %)

구분	2016년	2017년(E)
데이터산업 전체	20.4	20.5
데이터 솔루션	23.5	23.2
데이터 구축	22.1	22.3
데이터 컨설팅	17.4	18.7
데이터 서비스	17.2	17.4

자료: 한국데이터진흥원(2018)

데이터산업 내 투자액 중 가장 높은 비중을 차지하는 것은 R&D(평균 투자액의 약 68%)로 대분류 중 투자액에서 R&D가 가장 높은 비중을 차지하는 분야는 데이터구축으로 나타났다. 전체 투자액으로 살펴본 대분류별 순위에서 1위를 차지하던 데이터서비스의 경우

전체 투자액 대비 R&D 투자 비중이 가장 낮은 것으로 나타났다.

〈표 2-16〉 국내 데이터산업 대분류 별 투자분야 비중 현황

(단위: %)

구분	설비 투자	R&D 투자	라이선스 지출	마케팅 /홍보	해외 투자	기술 교육투자	기타
데이터산업 전체	8.1	65.7	4.4	10.8	0.6	4.0	6.4
데이터솔루션	6.6	71.4	5.1	6.4	0.6	6.1	3.8
데이터구축	8.4	73.6	2.7	8.5	0.3	3.0	3.5
데이터컨설팅	8.1	67.8	4.2	6.2	0.0	6.5	7.2
데이터서비스	9.0	54.1	5.3	17.7	0.8	2.3	10.8

자료: 한국데이터진흥원(2018)

데이터산업 내 기업들이 사업을 영위하는 대상을 살펴보면, 약 60%가 국내 기업을 대상으로 사업을 진행하고 있으며, 약 18%와 약 19%의 기업들은 각각 개인과 정부/공공부문을 대상으로 사업을 영위하고 있다. 이를 통해 국내 데이터산업 내 기업들의 주요 수요는 국내에서 창출되며 해외 시장에서 수요를 창출하거나 사업을 영위하는 기업은 거의 전무한 것으로 추론해 볼 수 있다.

〈표 2-17〉 국내 데이터산업 대분류 별 주요 사업 대상 현황

(단위: %)

구분	국내			해외		
	기업	개인	정부/공공	기업	개인	정부/공공
데이터산업 전체	58.1	18.3	18.9	0.1	0.1	0.1
데이터솔루션	70.4	4.4	20.1	0.0	0.0	0.4
데이터구축	58.0	2.7	35.1	0.4	0.4	0.0
데이터컨설팅	68.6	3.5	25.6	0.0	0.0	0.0
데이터서비스	46.9	42.4	5.5	0.0	0.0	0.0

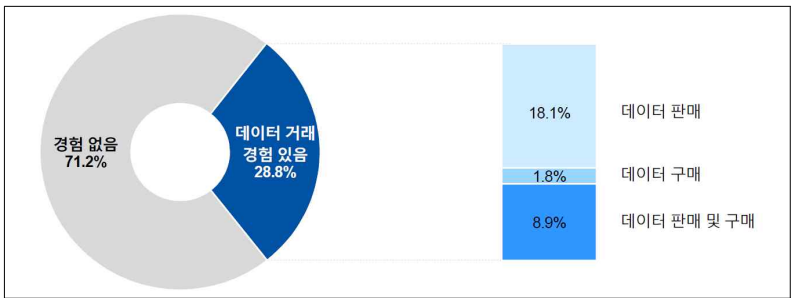
주: *무응답 부분을 포함하지 않아 각 행의 합이 100이 되지 않을 수 있음

자료: 한국데이터진흥원(2018)

미국이 세계에서 가장 큰 규모의 데이터산업을 보유할 수 있었던 것은 데이터브로커 산업을 기반으로 하여 데이터의 활발한 유통과 활용이 이루어졌기 때문이다. 우리나라에서

도 데이터 거래·유통을 촉진하려고 하는 시도가 있어왔으나 현재 데이터 거래 시장의 활성화는 아직 걸음마 단계이다. 이러한 우리나라의 데이터 거래 현황을 살펴보면 다음과 같다. 한국데이터진흥원이 데이터서비스부문의 하위분류인 데이터 거래, 정보제공, 데이터 분석 제공 업체를 대상으로 한 2017년 설문²²⁾에 따르면 데이터 거래 경험이 있는 기업이 28.8%로 대부분의 기업들이 데이터 거래 경험이 없는 것으로 나타났다. 여기에서 데이터 거래 기업은 데이터 거래 플랫폼 참여 기업 및 데이터 융합·가공 업체, 정보제공 기업은 포털, 신용/재무, 통계 등 원시 데이터 보유 업체, 데이터 분석 제공 기업은 데이터 분석을 통해 유의미한 정보를 제공하는 업체를 각각 의미한다.

[그림 2-13] 국내 데이터서비스 부문 내 데이터 거래 현황



자료: 한국데이터진흥원(2018)

다시 데이터 거래 경험이 있는 업체들을 대상으로 데이터 거래 유형을 조사한 결과 데이터 거래 기업들의 약 97%, 정보제공 업체들의 약 88%가 데이터 판매 경험이 있는 것으로

22) 한국데이터진흥원이 데이터 거래의 경우 데이터 서비스부문에 집중하여 설문을 실시한 것은 데이터산업 내 타 부문, 데이터솔루션, 데이터구축/컨설팅의 경우에는 데이터 거래에 대한 수요가 없을 가능성이 높기 때문이다. 데이터솔루션은 데이터 관련 소프트웨어 개발 업체가 주를 이루고 있기 때문에 실제로 데이터를 활용한 서비스를 제공하지 않고 있다. 또한 데이터구축/컨설팅의 경우에도 이미 사업대상 업체의 데이터를 효율적으로 관리하기 위한 체계를 구축하거나 이미 보유하고 있는 데이터를 어떻게 활용해야 할지에 대한 컨설팅이 주를 이루고 있기 때문에 데이터 거래에 대한 수요가 없다. 반면 데이터 서비스 부문의 경우 정보제공, 데이터 거래, 분석서비스의 세 하위부분으로 나뉘는데 이들은 모두 데이터 거래를 통해 분석서비스를 제공하거나 데이터 거래 그 자체를 사업으로 영위하기에 데이터 거래에 대한 수요가 가장 확실한 부문이다.

로 나타나 이 두 분류에 속한 업체들은 데이터 거래 시장에서 주로 공급자 역할을 담당하는 것으로 나타났다. 반면 데이터분석제공 기업들은 모두 데이터 구매 경험만 존재하는 것으로 나타나 이 분류에 속한 업체들은 데이터 거래 시장에서 주로 수요자 역할을 담당하고 있다.

[그림 2-14] 분류별 데이터 구매·판매 비중



자료: 한국데이터진흥원(2018) 자료 재구성, 민대홍 외(2018) 재인용

마지막으로 데이터산업에서 또 다른 한 축인 공공데이터와 관련하여 ‘공공데이터 개방·활용 현황’을 살펴보면, 2013년 ‘공공데이터의 제공 및 이용 활성화에 관한 법률’이 시행된 이후 공공데이터 개방·활용에 대한 노력에 힘입어 개방 및 활용 모두에서 성과를 거두고 있는 것으로 나타났다. 데이터 개방 목록 개수는 2017년 23,804개로 2013년 대비 4.3배의 큰 폭으로 증가하였다.

<표 2-18> 국내 공공데이터 개방 목록 개수 추이

(단위: 개)

구분	2013년	2014년	2015년	2016년	2017년 (10월말 누적)
파일데이터	4,718	11,796	14,115	19,184	20,629
오픈API	554	1,361	1,790	2,143	2,397
표준데이터	-	-	7	31	58
합계	5,272	13,157	15,912	21,358	23,084

자료: 공공데이터포털(2017)

양적인 지표인 공공데이터 개방 건수 뿐만 아니라 질적인 지표인 공공 데이터 다운로드 신청 건수 역시 2017년 기준 총 3,505,731건으로 2013년 대비 약 252배 증가하였다.

〈표 2-19〉 국내 공공데이터 개방 다운로드 신청 건수 추이

(단위: 개)

구 분	총 계	2013년	2014년	2015년	2016년	2017년 (10월말 누적)
파일데이터	3,184,984	4,108	115,525	589,783	978,104	1,497,464
오픈API	287,247	9,815	23,872	40,159	108,766	104,635
표준데이터	33,500	-	-	511	13,607	19,382
합 계	3,505,731	13,923	139,397	630,453	1,100,477	1,621,481

자료: 공공데이터포털(2017)

또 다른 질적 지표인 공공데이터 활용 웹 또는 앱 서비스 개발사례 역시 증가하여 2017년 기준 총 1,401개로 2014년 대비 약 33배 증가하여 공공데이터 개방·활용 역시 활발하게 진행되고 있는 것으로 나타났다. 하지만 국내 데이터산업계에서 공공데이터의 질적 문제에 대한 이슈가 계속하여 제기되는 것으로 보아 위와 같은 지표로만 공공데이터 관련 현황을 평가하기보다 산업계와의 면밀한 협의를 통해 새로운 지표를 개발하고 산업계의 의견을 반영한 공공데이터 개방·활용 수준에 대한 조사가 실시되어야 할 것으로 보인다.

〈표 2-20〉 국내 공공데이터 활용 웹/앱 서비스 개발 사례

(단위: 개)

구 분	2013년	2014년	2015년	2016년	2017년 (10월말 누적)
개발사례 건수	42	395	695	1,056	1,401

자료: 공공데이터포털(2017)

제3장 데이터산업의 경제적 파급효과

제1절 데이터산업의 정의

이 절에서는 데이터산업의 경제적 파급효과를 살펴보기 이전에 데이터산업의 정의에 대해 살펴보고자 한다. 먼저 국내외 선행연구에서 데이터산업이 어떻게 정의되어 왔는지에 대해 살펴본 후 데이터산업 파급효과 산출을 위해 본 연구에서 사용된 정의에 대해 알아보겠다.

1. 국내외 문헌에서 데이터산업의 정의

국내외 선행연구에 따르면 데이터산업 또는 데이터산업 생태계를 바라보는 시각에 따라 데이터산업에 대한 다양한 정의들이 존재한다. 먼저 한국정보화진흥원(2012)에서는 빅데이터산업 생태계를 서비스 공급자, 빅데이터 사용자, 어플리케이션 공급자 간의 상호작용이 일어나는 공간으로 정의하였는데, 서비스 공급자, 서비스 사용자, 어플리케이션 공급자는 각각 빅데이터 서비스를 제공하는 주체, 빅데이터 서비스를 구매하는 주체, 공급자와 사용자를 연결하는 주체로 정의하고 이들 간의 상호작용이 일어나는 공간을 빅데이터산업 생태계로 보았다.

김사혁(2013)은 빅데이터산업 생태계를 인프라, 소프트웨어, 서비스로 이루어진 생태계라고 정의하고 있다. 이들 세부구성요소 중 분석인프라는 저장 공간 등 하드웨어 제공 주체이며 소프트웨어는 데이터 가공, 통합 등 서비스 제공 주체, 서비스는 데이터 분석을 통한 의미 있는 정보를 제공하는 주체로 정의하였다.

다음으로 한국데이터진흥원(2018)은 앞 장에서 살펴본바와 같이 데이터산업을 데이터 솔루션, 데이터구축 및 컨설팅, 데이터서비스로 구성된 산업으로 정의하였는데, 데이터솔루션은 데이터 시각화, 검색엔진 등의 업체로, 데이터 구축/컨설팅은 데이터베이스 구축, 데이터 변화/저장 관련 업체로, 데이터서비스는 데이터 분석 서비스 제공 업체로 구성되어 있다.

해외 문헌인 IDC(2017b)에서는 앞서 살펴보았듯이 데이터산업을 데이터기업, 데이터 시장, 데이터 관련 인력으로 구분하여 분석하였다. 데이터관련 인력은 데이터 관련 직무에 종사하는 자이며, 데이터 기업은 데이터의 생산과 분배에 관여하는 데이터 공급 조직, 데이터 시장은 디지털 데이터가 가공되어 상품으로서 교환되는 시장으로 각각 정의된다.

방동희(2018)은 데이터산업 생태계를 데이터의 생애주기²³⁾에 따라 생산자, 소비자, 거래자, 가공자의 종합 활동 체계로 정의하고 있는데, 앞서 살펴본 한국정보화진흥원(2012), 김사혁(2013), 한국데이터진흥원(2018)의 산업분류에 기반한 정의에서 벗어나 데이터의 생애주기를 중심으로 데이터산업을 정의하였다는 점에서 그 의의가 있다.

〈표 3-1〉 데이터산업 생태계 분류 또는 정의

개별 연구	데이터산업 생태계 분류 체계 또는 정의
한국정보화진흥원 (2012)/ 김신곤 외(2016)	<ul style="list-style-type: none"> • 데이터 서비스 공급자(빅데이터 관련 서비스 공급) • 어플리케이션 서비스 공급자(서비스 공급자와 사용자를 중계하는 역할) • 데이터 서비스 사용자(빅데이터 관련 서비스 사용)
김사혁(2013)	<ul style="list-style-type: none"> • 인프라(저장 공간 등 하드웨어 제공) • 소프트웨어(데이터 가공, 통합 등의 서비스 제공) • 서비스(데이터 분석을 통한 의미 있는 정보 제공)
한국데이터진흥원 (2018)	<ul style="list-style-type: none"> • 데이터 솔루션(데이터 시각화, 검색엔진 등의 업체) • 데이터 구축·컨설팅(데이터베이스 구축, 데이터 변환/저장 관련 업체) • 데이터서비스(데이터 분석 서비스)
IDC(2017b)	<ul style="list-style-type: none"> • 데이터 인력(데이터산업 종사 인력) • 데이터 기업(데이터산업 내 데이터 생산·분배 조직) • 데이터 시장(데이터 서비스 또는 가공된 데이터가 거래되는 시장)
방동희(2018)	<ul style="list-style-type: none"> • 데이터의 생산, 저장, 수집, 유통, 거래, 가공, 처리, 이용, 폐기 등의 일련의 데이터 생애주기에 따른 생산자, 소비자, 거래자, 가공자의 종합 활동체계
이영환 외(2015)	<ul style="list-style-type: none"> • 원시데이터의 수집, 취합, 가공 및 정제를 거쳐 해석 및 분석을 통해 소비되는 복잡한 다단계로 연결

자료: 민대홍 외(2018) 재인용

23) 데이터의 생산에서 저장, 수집, 유통, 거래, 가공, 처리, 이용, 폐기에 이르기까지의 전 생애주기를 지칭한다.

이영환 외 (2015)에서도 방동희(2018)과 비슷한 관점으로 데이터산업 생태계를 분석하였는데, 해당 연구에서 데이터산업 생태계란 원시데이터의 수집, 취합, 가공 및 정제, 분석 단계를 거쳐 최종 소비되는 복잡다단한 체계로 정의하였다.

2. 데이터산업의 정의

본 연구에서는 데이터산업을 ICT 통합분류체계에 따른 정의를 활용하여 살펴보도록 하겠다. 위에서 살펴본 여러 정의들 중 ICT 통합분류체계에 따른 정의를 채택한 이유는 이러한 정의가 산업연관분석과 연계하기 가장 적합하기 때문이다. 또한 이번 연구에서 데이터산업의 파급효과의 추정치를 활용하기 위해서는 데이터산업의 규모가 필요한데, 현재 우리나라에서 데이터산업의 규모를 추정하는 작업은 한국데이터진흥원의 '데이터산업 현황조사'가 유일하다. 2018년도부터 데이터산업 현황조사에서 활용되는 데이터산업의 정의 역시 ICT 통합분류체계로 향후 연구와의 정합성 및 활용도 측면에서도 데이터산업을 ICT 통합분류체계에 따른 정의를 사용하는 것이 적절하다.

먼저 ICT 통합분류체계에 따르면 데이터산업은 데이터솔루션, 데이터구축/컨설팅, 데이터서비스, 데이터인프라의 4개 대분류로 구성되어있다. 데이터솔루션 업체는 데이터 모델링, 분석/시각화, 검색엔진, 품질 등 관련 솔루션 제품으로 사업을 영위하는 사업체로 지적재산권, 유지·보수, 개발에서 매출 발생하는 사업체들이다. 데이터구축/컨설팅 업체는 데이터베이스 구축, 문서/음성/영상 등을 데이터로 변환/저장하는 데이터 처리, 데이터 외부제공을 위한 API, LOD 구축 등의 사업을 영위하는 사업체, 데이터 품질, 설계 및 데이터 활용 등에 대한 컨설팅 사업을 영위하는 사업체이며, 데이터서비스 업체는 데이터 판매·중개 서비스, 원천데이터 가공을 통해 정보를 제공 또는 데이터 분석서비스를 제공함으로써 발생하는 데이터 이용료/수수료 등으로 사업을 영위하는 사업체로 정의될 수 있다. 마지막으로 데이터인프라 업체는 위 세 분류에 속하는 업체들이 사업을 영위하기 위해 필요한 인프라를 제공하는 업체로 서버, 스토리지, 네트워크, 클라우드 구축 및 서비스 제공 등과 관련된 사업을 영위하는 사업체이다.

각각의 대분류 데이터 솔루션, 데이터구축/컨설팅, 데이터서비스, 데이터인프라에 속하는 세부 분야들을 살펴보면, 먼저 데이터 솔루션 부문은 다시 데이터 수집, 데이터 설계,

DBMS, 데이터관리, 데이터 품질, 데이터 분석, 데이터 플랫폼의 7개 중분류로 구성되어있다. 이들 업체의 한국표준산업분류(KSIC9) 상에서 응용 소프트웨어 개발 및 공급업으로 분류할 수 있다. ICT 통합분류체계 상 산업편에서는 KSIC9과 동일하게 소프트웨어 개발 및 공급업으로 분류되며 품목편에서는 시스템소프트웨어와 응용소프트웨어로 분류된다.

데이터 구축/컨설팅 부문은 구축과 컨설팅 부문 2개의 중분류로 구성되어있는데, 한국표준산업분류(KSIC9) 상에서 시스템 소프트웨어 개발 및 공급업, 기타 정보기술 및 컴퓨터 운영 관련 서비스업, 자료처리업, 컴퓨터시스템 통합 자문 및 구축 서비스업에 해당된다. ICT 통합분류체계 산업편에서는 시스템 소프트웨어 개발 및 공급업, 기타 정보기술 및 컴퓨터 운영 관련 서비스업, 정보인프라서비스업, 컴퓨터시스템 통합 및 관리업에 해당하며, 품목편에서 IT컨설팅 및 시스템 개발 및 IT시스템 관리 및 지원서비스에 해당된다.

데이터서비스 부문은 데이터 거래, 데이터 신디케이션, 정보제공, 데이터분석·제공, 데이터 인프라 서비스의 5개 중분류로 구성되어있다. 한국표준산업분류(KSIC9) 상에서는 포털 및 기타 인터넷 정보매개 서비스업, 뉴스제공업, 데이터 베이스 및 온라인정보 제공업 및 그 외 기타정보서비스업에 해당되며, ICT 통합분류체계 산업편에서 정보매개서비스업과 정보 제공 서비스업, 품목편에서는 자료처리서비스, 온라인 정보처리, 광고제공서비스, 인터넷 정보제공서비스, 기타 정보제공 서비스에 해당된다.

마지막으로 데이터인프라 부문은 온프레미스와 클라우드 2개의 중분류로 구성되어 있는데, 한국표준산업분류(KSIC9) 상에서는 컴퓨터제조업, 기억장치 제조업, 유선 통신장비 제조업, 기타 무선통신장비 제조업에 해당된다. ICT 통합분류체계 산업편에서 컴퓨터기기업, 컴퓨터 주변기기업, 유선 통신장비업, 방송 및 무선통신장비업에, 품목편에서 기타 컴퓨터 주변기기, 네트워크장비, 호스팅 서비스, 부가 네트워크서비스, 인터넷 관리 서비스, 기타 정보인프라서비스, 스토리지에 해당된다.

〈표 3-2〉 ICT 통합분류체계에서 데이터산업 분류

대분류	중분류	소분류	KSIC9	ICT통합분류체계	
				[산업편]	[품목편]
데이터 솔루션	데이터 수집	데이터 검색	58222 응용소프트 웨어 개발 및 공급업	응용 소프트웨어 개발 및 공급업	3110.0000 시스템 소프트웨어 (3113.0000 스토리지 제외) 3120.0000 응용소프트웨어
		로그데이터 수집			
		웹데이터 수집			
	데이터 설계	데이터 아키텍처			
		데이터 모델링			
	DBMS	디스크 DBMS			
		인메모리 DBMS			
		DBMS 어플 라이언스			
		기타 DBMS			
	데이터 관리	DB 운영 관리			
		DB 성능 관리			
		DB 보안 관리			
		데이터 분산 처리			
		데이터 통합· 흐름 관리			
	데이터 품질	데이터 품질 관리			
메타데이터 관리					
마스터데이터 관리					
데이터 분석	정형 데이터 분석				
	비정형 데이터 분석				
	실시간 데이터 분석				
	데이터 시각화 분석				
	데이터 플랫폼	데이터 플랫폼			

대분류	중분류	소분류	KSIC9	ICT통합분류체계		
				[산업편]	[품목편]	
데이터 구축/ 컨설팅	구축	DB설계·구축	58221 시스템 소프트웨어 개발 및 공급업	시스템 소프트웨어 개발 및 공급업 기타 정보기술 및 컴퓨터운영 관련 서비스업 정보인프라 서비스업 컴퓨터시스템 통합 및 관리업	3310.0000 IT컨설팅 및 시스템 개발 3320.0000 IT시스템 관리 및 지원서비스	
		데이터 이행·처리				
		기계처리형 데이터구축				
	컨설팅	데이터 설계	62090 기타 정보기술 및 컴퓨터운영 관련 서비스업 63111 자료처리업 62021 컴퓨터시스템 통합 자문 및 구축 서비스업			
		데이터 품질				
		데이터 관리 성능개선				
		데이터 거버넌스 체계				
	데이터 분석					
데이터 서비스	데이터 거래	데이터 마켓	63120 포털 및 기타 인터넷 정보매개 서비스업 63910 뉴스제공업 63991 데이터베이스 및 온라인 정보 제공업 63999 그 외 기타 정보서비스업	정보매개 서비스업 정보제공 서비스업	2311.0000 자료처리 서비스 2311.1000 온라인 정보처리 2321.0000 광고제공서비스 (웹서치포털) 2331.0000 인터넷 정보제공서비스 2339.0000 기타 정보제공서비스	
	데이터 신디케이션					
	정보 제공	교육/훈련				63120 포털 및 기타 인터넷 정보매개 서비스업 63910 뉴스제공업 63991 데이터베이스 및 온라인 정보 제공업 63999 그 외 기타 정보서비스업
		신용/재무				
		통계				
		경영/ 비즈니스				
		취업/창업				
		뉴스				
		포털				
		행정/법률				
		학술				
	문화/예술 생활					
	데이터 분석 제공	환경분석 서비스				63999 그 외 기타 정보서비스업
		예측분석 서비스				
심사분석 서비스						
기타 데이터분석 서비스						
데이터 인프라 서비스	운영					
	컨설팅					

대분류	중분류	소분류	KSIC9	ICT통합분류체계		
				[산업편]	[품목편]	
데이터 인프라	온프레미스	서버	26310 컴퓨터 제조업 26321 기억장치 제조업 26410 유선 통신장비 제조업 26429 기타 무선 통신장비 제조업	컴퓨터기업 컴퓨터 주변기기업 유선 통신장비업 방송 및 무선통신장비업	1225.9000 기타 컴퓨터주변기기 1311.5000 네트워크장비	
		스토리지			2312.0000 호스팅 서비스	
		네트워크			2313.0000 부가 네트워크서비스	
	클라우드	클라우드 서비스 제공			클라우드 구축	2314.0000 인터넷 관리 서비스
						2319.0000 기타 정보 인프라서비스 3113.0000 스토리지

제 2 절 연구방법론— 산업연관 분석

이 절에서는 데이터산업의 경제적 파급효과 도출에 사용될 산업연관분석을 개괄적으로 살펴보고자한다. 즉, 산업연관분석에서 활용되는 산업연관표의 구조와 공급측과 수요측 분석을 통해 도출되는 생산유발효과, 부가가치유발효과, 취업유발효과, 공급지장효과, 물가파급효과가 어떻게 도출되는지에 대해 살펴보도록 하겠다.

산업연관분석은 산업연관표를 바탕으로 하여 진행되는데 산업연관분석을 이해하기 위해서는 먼저 산업연관표에 대한 이해가 필수적이다. 산업연관표는 국내 여러 산업의 상호의존관계를 일람표 형태로 나타낸 표이다. 산업연관표의 기본 구조는 내생부분과 외생부분으로 나누어 볼 수 있는데 그 기본구조는 [그림 3-1]과 같다.

로 국내 총산출액을 기록할 때 제하고 기록하게 된다. 예를 들어 국내에서 총 생산된 우유의 양이 100 단위라고 할 때, 이 중 50 단위는 치즈생산에 사용되고 70 단위는 최종수요에서 소비된다고 할 때, 이를 합하면 국내에서 최종재 또는 중간재로서 소비된 우유의 양은 총 120 단위가 된다. 즉, 국내에서 산출된 우유의 양 100 단위 보다 20 단위가 더 사용되었다는 뜻인데 이는 수입을 통해서 국내에서 소비된 것이다. 따라서 산업연관표 상 위의 그림에서 내생부문에 50, 최종수요 부문에 70이 기입되고, 나머지 20은 수입부문에 기입된다. 마지막으로 50과 70의 합에서 수입부문 20을 제하여 총 산출액 100이 기입되게 되는 것이다.

산업연관표의 행이 어떤 산업의 최종생산물이 중간재 또는 최종재로서 어떻게 사용되었는지에 대한 정보를 포함하고 있다면, 열은 어떤 산업의 최종생산물의 생산을 위해서 타 산업들의 최종재가 얼마만큼 투입되었는지에 대한 정보와 해당 산업의 최종생산물 생산에서 얼마만큼의 부가가치가 창출되었는지에 대한 정보를 포함하고 있다. 먼저 내생부문은 타 산업들의 최종재가 어떤 산업에서 얼마만큼 중간재로 사용되었는지에 대한 정보가 기록되어있다. 우유산업을 예로 들면 우유생산에는 젖소, 젖소들의 먹이, 기타 장비, 인터넷 서비스, 운송서비스 등이 활용되는데 이들이 얼마만큼 우유생산에 기여했는지에 대한 정보를 담고 있다. 부가가치 부문에서는 우유생산에 소요된 노동력이 임금으로 기입되어 최종가치에서 중간재로 투입된 생산물의 가치에서 생산된 최종생산물에 얼마만큼의 부가가치를 추가하였는지에 대한 정보가 담겨있다. 예를 들어 우유 생산에 있어서 젖소들의 먹이가 80 단위 정도로 소비되었고, 최종생산물인 우유의 가치가 100 단위이라고 한다면, 먼저 열 부분의 내생부문에 80 단위가 기입되고, 부가가치 부문에 20 단위가 기입되게 되어 우유산업에서 총 투입액은 100 단위가 된다.

마지막으로 산업연관표는 총산출액과 총 투입액이 같도록 고안되었는데 이는 국내에서 수입되거나 생산되어 소비된 것이 국내에서 생산된 것과 같아야한다는 기본적인 가정이 투영된 것으로 볼 수 있다.

이처럼 산업연관표는 내생부문과 외생부문으로 나뉘어 국내에서 생산되고 소비되는 재화를 체계적으로 기입하여 어떤 산업에서의 최종생산물 생산에 어떤 산업의 최종생산물이 얼마만큼 중간재로 사용되었는지 또 부가가치는 얼마나 창출되었으며 수입되어 소비된 양을 얼마인지에 대해 일목요연하게 정리한 표로, 이를 통해서 국민 경제에서 여러 재화들

이 어떻게 배분되어 활용되었는지에 대한 정보를 담고 있다.

특히 특정 산업의 경제적 파급효과 도출을 위해 중요한 부문은 내생부문이다. 위에서 살펴본 대로 내생부문은 산업 간 중간수요에 대한 정보를 담고 있다. 즉, 이러한 내생부문을 행으로 보느냐 열로 보느냐에 따라 어떤 산업의 최종생산물이 타 산업의 중간재로 어떻게 활용되는지 또는 타 산업들의 최종재들이 어떤 산업의 생산에서 어떻게 중간재로 활용되는지에 대해 알 수 있다. 기본적인 산업연관표 상에서 내생부문은 총 30개 부문의 산업들로 이루어져있다. 기본적인 산업연관표 상의 30대 산업명은 아래 표에 나타나있다.

〈표 3-3〉 산업연관표 내 30대 산업

1	2	3	4	5
농림수산물	광산품	음식료품	섬유 및 가죽제품	목재 및 종이, 인쇄
6	7	8	9	10
석탄 및 석유제품	화학제품	비금속 광물제품	1차 금속제품	금속제품
11	12	13	14	15
기계 및 장비	전기 및 전자기기	정밀기기	운송장비	기타 제조업 제품 및 임가공
16	17	18	19	20
전력, 가스 및 증기	수도 및 폐기물	건설	도소매서비스	운송서비스
21	22	23	24	25
음식점 및 숙박서비스	정보통신 및 방송	금융 및 보험 서비스	부동산 및 임대	전문, 과학 및 기술
26	27	28	29	30
사업지원 서비스	공공행정 및 국방	교육서비스	보건 및 사회복지	문화 및 기타 서비스

위의 30대 산업을 산업연관표의 내생부문에 맞게 재구성하면 행과 열 각각 30개 부문의 산업으로 구성된 30×30 행렬을 구할 수 있다.

〈표 3-4〉에서는 내생부문에 대한 이해를 돕기 위해서 외생부문을 제외한 형태의 산업연관표를 상정하였다. 즉, 위의 표에서는 부가가치창출이나 최종수요 또는 수입 부문이 존재하지 않는 가상의 경제체제를 상정하였다. 위 표에서 첫 행을 살펴보면 제 1부문이 농림수산물에서 생산된 총 800원어치의 최종재는 농림수산물 그 자체에서 500원 만큼 중간재로 사용하고, 나머지 300원어치는 광산품에서 100원, 문화 및 기타서비스에서 200원 만큼 중간재로 사용된 것으로 기입되어있다. 반대로 첫 열을 살펴보면 농림수산물에서 어떠한 산업의 최종재를 중간재로 사용하였는지에 대한 정보를 알 수 있다. 즉, 위 표의 첫 열이 담고 있는 정보는 800원 만큼의 농림수산품을 생산하는데 농림수산품의 최종재 중 500원만큼이 다시 중간재로 활용되었고, 광산품 부문의 200원 만큼의 최종재, 문화 및 기타서비스에서 100원 만큼의 최종재가 투입되어 800원 만큼의 최종재 생산이 이루어진 것을 알 수 있다.

〈표 3-4〉 산업연관표 내의 내생부문 예시

		1	2	...	30	총산출액
		농림수산물	광산품	문화 및 기타서비스	
1	농림수산물	500	100	200	800
2	광산품	200	0	100	300
.
.
.
.
30	문화 및 기타서비스	100	200	50	350
총 투입액		800	300	350	1450

이 중간수요 행렬을 이용하여 수요측면에서 생산유발계수, 부가가치유발계수, 취업유발계수를 도출할 수 있는데, 이를 위해서는 먼저 위와 같은 계수의 도출을 위해 30×30 행렬을 투입요소비율의 형태로 재구성하여야 한다. 내생부문의 정보를 투입계수표로 전환하는 방식은 아래와 같다.

〈표 3-5〉 산업연관표 내의 내생부문의 투입계수표로의 전환

		중간수요				
		1	2	n
중간 투입	1	a_{11}	a_{12}	a_{1n}
	2	a_{21}	a_{22}	a_{2n}

	n	a_{n1}	a_{n2}	a_{nn}

위의 각 셀에 위치한 문자들은 $a_{ij} = x_{ij} / X_j$ 라는 산식을 이용하여 도출된 숫자들로 a_{11} 은 1번 산업의 총 생산물 중에서 1번 산업 중간재가 차지하는 비중이다. 또 다른 예로 a_{12} 의 경우 2번 산업의 총 생산물 중에서 1번 산업 중간재가 차지하는 비중을 뜻하며, a_{2n} 은 n 번째 산업의 총 생산물 중에서 2번 산업 중간재가 차지하는 비중을 뜻한다. 마지막으로 산업연관분석에서 가장 중요한 가정 중 하나인 국가전체의 총수요와 총투입액이 동일하다는 것을 이용할 수 있다. 위와 같이 도출된 30×30 행렬과 총수요액과 총투입액이 동일해야 한다는 가정을 사용하면 아래와 같은 수식이 도출된다.

$$A^d X + Y = X$$

위 식에서 A^d 는 수출입을 제외한 국내부문에서 중간투입계수표이고 X 는 국내 전 산업에서 생산된 생산액, 그리고 Y 는 국내 총수요이다. 위의 행렬식에서 $A^d X$ 을 우변으로 이항한 뒤 다시 X 를 공통항으로 인수분해하면 아래와 같은 수식을 도출할 수 있다.

$$X = (I - A^d)^{-1} Y$$

위 식에서 $(I - A^d)^{-1}$ 가 핵심이 되는 부문으로 이를 활용하여 생산유발계수, 부가가치유발계수, 취업유발계수를 계산할 수 있다. 먼저 위 식을 통해서 최종수요 Y 가 한 단위 증가할 때 이를 위해 전 산업부문에 걸쳐 증가해야 하는 생산물을 가늠할 수 있다. 이를 이해하기 위해서 $(I - A^d)^{-1}$ 에 대한 이해가 필요하다. $(I - A^d)^{-1}$ 는 $(I - A^d)$ 의 역행렬로서 이 역시 30×30 행렬의 형태를 가지며 아래와 같이 나타낼 수 있다.

$$(I - A^d)^{-1} = \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{pmatrix}$$

먼저 $(I - A^d)^{-1}$ 의 i 번째 열은 i 번째 산업의 최종수요가 1단위 증가하였을 때 각 산업에서 추가적으로 생산되어야 하는 생산물 단위로 구성되어있다. 반면 $(I - A^d)^{-1}$ 의 j 번째 행은 각 열의 산업들에서 최종수요가 한 단위 증가하였을 때 j 번째 산업에서 추가적으로 생산해야하는 생산단위를 의미한다.

예를 들어, 한 국가의 산업이 3개 밖에 존재하지 않는 경우를 상정할 경우 다음과 같은 3×3 행렬을 생각해 볼 수 있다.

$$(I - A^d)^{-1} = \begin{pmatrix} 1 & 0.2 & 0.4 \\ 0.2 & 1 & 0.3 \\ 0.5 & 0.7 & 1 \end{pmatrix}$$

첫 열을 보면 1, 0.2 0.5로 구성되어 있는데, 이는 1번째 산업의 최종수요가 1단위 증가하였을 때 1번째, 2번째, 3번째 산업에서 추가적으로 생산되어야 하는 생산물이 각각 1단위, 0.2단위, 0.5 단위인 것을 의미한다.

다른 예로 2행을 살펴보면 0.2, 1, 0.3으로 구성되어 있는데, 이는 1번째 산업에서 최종수요가 1단위 증가하였을 때 2번째 산업에서 0.2 단위만큼의 생산물이 추가적으로 생산되어야 함을 의미하고 있다.

산업연관분석에서 생산유발계수란 각 열의 세로 합으로 해당산업의 최종수요가 한 단위 증가할 때 국민경제 전체에서 부가적으로 생산되어야 하는 생산물 단위의 합을 의미한다. 따라서 위의 3 산업 경제의 예에서 제 1 산업의 생산유발 계수는 1.7, 제 2 산업의 생산유발계수는 1.9, 제 3 산업의 생산유발계수는 1.7로 산출됨을 알 수 있다.

생산유발계수는 후방연쇄효과와 밀접한 관련이 있는데 어떠한 산업의 후방연쇄효과란 해당산업의 최종수요가 한 단위 증가하였을 때 이를 생산하기 위해 전 산업부문에서 중간재로 활용되기 위해 재화나 서비스가 얼마나 추가적으로 생산되어야 하는지에 대한 정보를 포괄하고 있다.

후방연쇄효과를 측정하기 위해 주로 사용되는 척도는 영향력계수로, 영향력계수란 어떠한

한 산업의 최종수요가 한 단위 증가하였을 때 전 산업에 미치는 영향을 전 산업 평균에 대한 상대적 크기로 나타낸 계수로 해당 산업의 생산유발계수의 열합계를 전 산업 평균으로 나누어 산출한다.²⁴⁾

위의 3 산업 경제에서 전 산업의 생산유발효과 평균은(1.7+1.9+1.7)/3=1.77이며 제 1 산업의 영향력계수는 제 1 산업의 영향력 계수인 1.7을 평균치 1.77로 나눈 값(1.77/1.77)인 1로 산출되어 전 산업평균 후방연쇄효과와 동일하다. 이는 전 산업의 후방연쇄효과 평균치와 제 1 산업의 후방연쇄효과가 같다는 의미로 국민경제 전반에서 후방연쇄효과 측면을 고려할 때 제 1 산업이 평균적 위치를 차지함을 의미한다.

다음으로 부가가치유발계수 역시 $(I - A^d)^{-1}$ 를 활용하여 산출 가능한데, 이를 위해서는 먼저 부가가치계수행렬이 먼저 구성되어야 한다. 부가가치계수행렬은 부가가치율이 정방행렬의 대각선에 위치하고 나머지는 모두 0으로 이루어진 행렬로서 부가가치율 또는 부가가치계수는 총 투입액으로 부가가치를 나눈 값으로 $v_j = \frac{a_j^v}{X_j}$ 의 산식으로부터 도출될 수 있다. 여기에서 a_j^v 는 j 산업의 부가가치로 v_j 는 j 산업에서 발생한 부가가치가 총 투입액 X_j 에서 차지하는 비중이다. 이들 $a_1^v, a_2^v, \dots, a_n^v$ 를 $n \times n$ 행렬의 대각에 위치시키고 나머지 부분은 모두 0으로 채운 행렬이 부가가치계수행렬로 A^v 로 정의할 수 있다.

위의 3 산업 경제의 예에서 A^v 는 아래와 같이 나타낼 수 있다.²⁵⁾

$$A^v = \begin{pmatrix} a_1^v & 0 & 0 \\ 0 & a_2^v & 0 \\ 0 & 0 & a_3^v \end{pmatrix} = \begin{pmatrix} 0.3 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}$$

이 행렬을 $(I - A^d)^{-1}$ 에 곱한 행렬인 $A^v(I - A^d)^{-1}$ 을 부가가치유발계수행렬로 해당 행렬의 각 열의 세로합이 산업 j 의 부가가치유발계수이다. 위의 3 산업 경제의 예에서 부가가치유발계수행렬과 부가가치유발계수는 아래와 같이 계산 가능하다.

24) 한국은행(2014b).

25) 이 예에서는 내생부문만이 존재한다고 가정하고 논의를 진행하고 있기 때문에 해당되는 부가가치율을 직접 계산할 수 없으므로, 임의의 부가가치를 상정하고 만들어낸 값이다.

$$A^v(I-A^d)^{-1} = \begin{pmatrix} 0.3 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 0.5 \end{pmatrix} \begin{pmatrix} 1 & 0.2 & 0.4 \\ 0.2 & 1 & 0.3 \\ 0.5 & 0.7 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0.3 & 0.06 & 0.12 \\ 0.14 & 0.7 & 0.21 \\ 0.25 & 0.35 & 0.5 \end{pmatrix}$$

1행 1열의 0.3은 1 산업에서 최종수요 한 단위가 증가함에 따라 1 산업에서 창출되는 부가가치, 1행 2열의 0.14는 1 산업에서 최종수요 한 단위가 증가함에 따라 2 산업에서 생산물이 0.2 단위 생산되며 창출되는 부가가치, 마지막으로 1행 3열의 0.25는 1 산업에서 최종수요 한 단위 증가함에 따라 3 산업에서 생산물이 0.5 단위 생산되며 창출되는 부가가치를 뜻한다. 따라서 산업별 부가가치유발계수는 1 산업에서 0.55, 2 산업에서는 1.11, 3 산업에서는 0.83으로 산출할 수 있다.

마지막으로 $(I-A^d)^{-1}$ 를 활용하여 취업유발계수 역시 산출할 수 있는데 부가가치유발계수의 경우처럼 먼저 취업계수행렬을 구성해야 한다. 취업계수행렬은 부가가치계수행렬처럼 취업계수들을 $n \times n$ 행렬의 대각선상에 위치시키고 나머지는 0으로 채움으로써 구성할 수 있다. 취업계수는 취업자수를 산출액(10억 원 기준)으로 나눈 뒤 100을 곱해서 얻을 수 있는데, 산업 j 의 취업계수는 산업 j 의 취업자수를 총투입액 X_j 로 나누고 100을 곱해서 구할 수 있다.

위의 3 산업 경제에서 각 산업별 취업자수를 10명, 30명, 5명으로 가정하고, 각 산업의 산출액(10억 원 기준)을 1000, 5000, 1500으로 가정할 경우 각 산업의 취업계수는 아래와 같이 도출된다.

$$l_1 = 1, l_2 = 0.6, l_3 = 0.33$$

이를 바탕으로 취업계수행렬을 구성하면 아래와 같은 결과를 얻을 수 있다.

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.33 \end{pmatrix}$$

이 취업계수행렬을 다시 $(I-A^d)^{-1}$ 에 곱함으로써 취업유발계수를 산출할 수 있는데, 위의 3 산업 예에서 각 산업의 취업유발계수는 아래와 같이 계산할 수 있다.

$$L(I-A^d)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.33 \end{pmatrix} \begin{pmatrix} 1 & 0.2 & 0.4 \\ 0.2 & 1 & 0.3 \\ 0.5 & 0.7 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0.2 & 0.4 \\ 0.12 & 0.6 & 0.18 \\ 0.16 & 0.231 & 0.33 \end{pmatrix}$$

이들 각 열의 세로합이 각 산업의 취업유발계수로 각 산업의 취업유발계수는 제 1 산업이 1.28, 제 2 산업이 1.031, 제 3 산업이 0.91로 산출된다. 이들 취업유발계수는 각 산업들의 최종재 수요가 10억 원 만큼 증가할 때 타 산업에서 추가적으로 생산물을 공급하기 위해 유발되는 취업자수를 뜻한다.

이상 살펴본 것은 수요측면의 산업연관분석으로 주로 후방연쇄효과와 깊은 관련이 있는 계수들 산출에 용이하다. 이상의 분석이 수요측면의 분석이라고 불리는 이유는 가장 기본이 되는 아래 식을 활용하면서 최종수요 Y 가 한 단위 증가함에 따라 타 산업에서 얼마만큼의 추가적인 중간재 수요가 창출되는가가 분석대상이기 때문이다.

$$X = (I - A^d)^{-1} Y$$

반면 전방연관효과와 밀접한 관련이 있는 산업연관분석은 공급측면의 분석이다. 수요측면의 분석은 어떤 산업의 최종생산물의 최종수요가 한 단위 증가할 때 이에 수반되는 중간재 수요에 집중하였다면, 공급측면 분석은 어떤 산업의 최종생산물이 한 단위 더 공급되어 중간재로 활용될 때의 경제적 파급효과를 분석한다. 어떤 산업의 최종생산물이 국민경제 전체에 중간재로서 공급되는 측면에서 분석을 진행한다는 점에서 공급측면 분석이라 불린다.

공급측면의 분석에서도 기본적인 계수도출과정은 수요측면에서와 비슷하나 가장 큰 차이점은 수요측면 분석모형이 산업연관표의 열에 집중하였다면 공급측면 분석 모형은 행에 집중하였다는 점이다. 앞에서 간략하게 언급하였듯이 내생부문을 따로 떼어내 구성한 30×30 행렬의 열은 어떤 산업의 최종생산물 생산과정에서 타 산업의 최종생산물이 얼마만큼 중간재로 활용되었는지에 대한 정보를 담고 있는 반면, 해당 행렬의 행에는 어떤 산업의 최종생산물이 타 산업의 최종생산물 생산과정에서 얼마만큼 중간재로 활용되었는지에 대한 정보를 담고 있기 때문에 전방효과분석을 위한 공급측면 분석모형에서는 행에 집

중하여 분석을 진행한다.

앞서 살펴보았듯이 산업연관표의 j 번째 열은 j 번째 산업의 생산에 투입되는 타 산업의 최종재들을 열거한 것인 반면 산업연관표의 i 번째 행은 i 번째 산업에서 생산된 최종재가 타 산업들에서 중간재로 사용된 것을 열거한 것이다.

따라서 수요측면 분석에서는 j 번째 산업의 총투입액으로 각 산업에서 j 번째 산업에 투입된 생산액으로 나누어 구성한 투입계수행렬을 사용하는 반면 공급측면 분석에서는 i 번째 산업의 총산출액을 각 부문별 중간수요액으로 나누어 구성한 산출계수행렬 또는 배분계수행렬을 사용하여 분석을 진행한다.

예를 들어 조금 더 간단한 2 산업 경제를 가정하고 1 산업과 2 산업의 중간수요액 - 즉, 내생부문만을 따로 떼어낸 행렬 - 이 다음과 같다고 가정해 보자.

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

이 경우 제 1산업의 최종산출액(1행의 합) X_1 은 3, 제 2산업의 최종산출액(2행의 합) X_2 는 5이고, 제 1산업의 최종투입액(1열의 합) X_1 은 3, 제 2산업의 최종투입액(2열의 합) X_2 는 5이다.

위와 같은 2 산업 경제에서 중간투입계수행렬과 배분계수행렬은 다음과 같이 도출할 수 있다. 먼저 중간투입계수행렬은 앞서 논의되었듯 각 열을 열의 합계로 나눈 행렬이므로 $a_{ij} = x_{ij} / X_j$ 식에 의해 아래와 같이 도출된다.

$$A = \begin{pmatrix} 1/3 & 2/5 \\ 2/3 & 3/5 \end{pmatrix}$$

반면 배분계수행렬은 각 행을 행의 합으로 나눈 행렬로 수식으로 표현하면 $\bar{a}_{ij} = x_{ij} / X_i$ 이며 아래와 같이 도출된다.

$$B = \begin{pmatrix} 1/3 & 2/3 \\ 2/5 & 3/5 \end{pmatrix}$$

산업연관표의 구조 상 행렬 A 에 $[X_1 \ X_2]$ 를 곱할 경우 최종수요와 합하여 각 최종투입액과 같은 등식이 아래와 같이 도출된다.

$$\begin{pmatrix} 1/3 & 2/5 \\ 2/3 & 3/5 \end{pmatrix} \times \begin{pmatrix} 3 \\ 5 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

위의 식은 앞서 이미 수요측면 분석의 핵심이 되었던 $AX+Y=X$ 의 형태이다. 반면에 행렬 B 에 $[X_1 \ X_2]$ 를 곱할 경우 산업별 부가가치와 합하여 각 최종 투입액과 같은 등식이 도출된다.

$$(35) \begin{pmatrix} 1/3 & 2/3 \\ 2/5 & 3/5 \end{pmatrix} + (w_1 \ w_2) = (3 \ 5)$$

여기에서 w_1 과 w_2 는 각각 제 1 산업과 제 2 산업의 부가가치로 위의 2 산업 경제에서 따로 상정하지 않았으나, 산업연관표 내에 어떠한 값으로 기입되어있다고 생각하고 논의를 진행하겠다. 위의 관계를 행렬식으로 정리하면 다음과 같은 행렬식을 도출할 수 있다.

$$X'B+W=X'$$

먼저 위의 식을 $AX+Y=X$ 와 비교함으로써 수요측면 분석과 공급측면 분석의 차이점을 더욱 잘 이해할 수 있다. $AX+Y=X$ 의 경우 내생부문만을 따로 떼어낸 행렬을 재구성하여 총산출물의 값을 곱하여 얻은 부분이 AX 로 이는 결국 내생부문을 따로 떼어낸 행렬과 동일한 값을 지니게 된다. 또한 외생부문인 최종수요 Y 를 더하게 될 경우 이는 결국 산업연관표의 각 행의 합을 더하게 되어 반드시 최종산출물인 X 과 같아지는 결과를 얻게 된다. 위의 그림을 참고하여 보면 결국 $AX+Y=X$ 은 산업연관표의 행의 합으로 보아도 무방하다.

반면 $X'B+W=X'$ 을 살펴보면 이는 결국 산업연관표의 열의 합인 것을 알 수 있다. 내생부문을 따로 떼어내 재구성한 B 행렬에 총 투입액을 곱한 행렬인 XB 는 내생부문을 따로 떼어낸 행렬과 같다. 이러한 내생부문에 부가가치 값을 더하게 되면 이는 결국 산업연관표 상 각 열을 합한 것과 같아져 결국 $X'B+W$ 은 총투입액 X' 과 같은 값을 가지게 된다. 결국은 열의 합으로 귀결되는지 또는 행의 합으로 귀결되는지에 따라 공급측면 분석 또는 수요측면의 분석이 되는 것이다.

다시 공급측면 분석의 논의를 이어가자면, 먼저 배분계수행렬을 B^d 로 정의할 경우 위에서 살펴본 것과 같이 아래와 같은 등식이 반드시 성립되어야 한다.

$$X' B^d + W = X'$$

위 식을 수요측면 분석에서와 마찬가지로 XB^d 를 우측으로 이항하고 X 를 공통항으로 인수분해 하여 정리하면 아래와 같은 등식이 도출된다.

$$X' = W(I - B^d)^{-1}$$

수요측면 분석에서 $(I - A^d)^{-1}$ 이 경제적 파급효과 도출을 위해 핵심이 되는 행렬이었듯이 공급측면 분석에서는 $(I - B^d)^{-1}$ 이 생산유발효과 등에 대한 모든 정보를 요약하고 있는 행렬이다. 이 $X' = W(I - B^d)^{-1}$ 을 활용하여 어떠한 산업의 부가가치가 한 단위 증가할 때 이로 인해 해당산업의 생산물을 중간재로 활용하는 타 산업에서의 생산이 얼마만큼 증가할 수 있는지에 대한 정보를 알 수 있다.

예를 들어 2 산업 경제에서 $(I - B^d)^{-1}$ 가 아래와 같은 형태를 가진다고 가정하자.

$$(I - B^d)^{-1} = \begin{pmatrix} 0.2 & 0.5 \\ 0.4 & 0.7 \end{pmatrix}$$

위의 예시에서 $(I - B^d)^{-1}$ 의 의미를 살펴보면, 제 1산업에서 부가가치가 한 단위 증가할 경우 이로 인해 제 1산업에서 추가적으로 0.2 단위만큼, 제 2산업에서 0.5 만큼 생산량이 증가해야 하며, 제 2 산업에서 부가가치가 한 단위 증가할 경우 이로 인해 제 1산업에서 0.4 단위, 제 2산업에서 0.7 단위만큼 생산량이 증가해야 한다는 것을 의미한다.

이를 통해 각 산업의 공급측면에서의 총산출이 증가하는 승수²⁶⁾를 구할 수 있는데 이는 간단히 각 행의 가로 합과 같다. 따라서 제 1산업의 공급측면 총산출 증가는 0.7, 제 2산업의 공급측면 총산출 증가는 1.1로 도출된다. 앞서 살펴보았듯이 이렇게 도출된 공급측면 총산출 증가는 전방연쇄효과와 밀접한 관련이 있는 총산출 증가 효과라고 할 수 있다.

26) 공급모형의 $(I - B^d)^{-1}$ 를 output inverse라고 하며, output inverse의 행합은 i 산업의 부가가치 1원의 증가가 경제 모든 부문을 통해 총산출에 미치는 영향을 나타내며 투입승수(input multiplier) 또는 공급승수(supply multiplier)라고 부른다. 반면, 수요모형의 $(I - A^d)^{-1}$ 를 input inverse 또는 Leontief inverse라고 하며, input inverse의 열합은 j 산업의 최종수요 1원의 증가가 경제 모든 부문을 통해 총산출에 미치는 영향을 나타내며, 산출승수(output multiplier) 또는 수요승수(demand multiplier)라고 부른다.

공급유도형 모형은 공급량(투입량)의 변화에 따라 총생산량이 변화하는 것을 전제로 한다. 그러므로 공급유도형 모형을 이용하면, 산업의 부가가치가 늘어나는 것을 가정하면 중간재로의 공급(투입)이 증가하기 때문에 총산출이 증가하는 효과를 측정할 수 있으며, 반대로 산업의 부가가치가 줄어드는 것을 가정하면 공급(투입)에 지장이 생기는 때문에 총산출이 감소하는 효과를 측정할 수 있다. 일반적으로 후자의 경우를 공급지장효과(supply shortage effect)라고 지칭한다.

제 3절 데이터산업의 산업별 파급효과

이 절에서는 산업연관분석을 활용하여 도출한 데이터산업의 산업별 파급효과에 대해 살펴보도록 하겠다. 먼저 제2절에서 살펴본 산업연관분석을 활용하기 위해서는 기존 산업연관표 내에서 데이터산업을 따로 분류해 내어 정의하는 과정이 필요하다. 따라서 먼저 어떻게 데이터산업을 산업연관표 내에서 분류해 낼 수 있는지에 대해 살펴본 뒤 이를 통해 재구성한 산업연관표를 바탕으로 도출된 데이터산업의 전·후방연쇄효과에 대해 살펴보도록 하겠다.

데이터산업의 산업별 파급효과 측정을 위한 산업연관분석을 위해서는 산업연관표 상에서 데이터산업을 정의하고 재분류할 필요가 있다. 이는 한국은행에서 제공되는 산업연관표에 데이터산업이 따로 분류되어있지 않기 때문인데, 본 연구에서는 앞서 살펴본 데이터산업의 정의에 따라 데이터산업을 산업연관표에서 재분류하였다. 기존 연구들에서보다 조금 더 데이터산업에 가깝게 재분류하기 위해서²⁷⁾ ‘품목별·부문별 공급액표’를 활용하였는데, 품목별·부문별 공급액표는 총 384개의 기본부문과 총 3,136개의 기초부문으로 구성된 표로 기초부문들은 각 기본부문의 세부부문으로 해당 표의 국산산출액 총합은 국내 총 산출량과 일치한다.

27) 본 연구 이외에 데이터산업의 경제적 파급효과를 산업연관분석을 활용하여 도출한 연구로는 민서현 외(2017)가 있는데, 이 연구에서는 데이터산업을 384개의 기본부문을 활용하여 재정의하였는데, 본 연구에서는 기본부문의 하위 개념인 3,136개 기초부문을 활용하여 재정의한다.

〈표 3-6〉 품목별·부문별 공급액표

(단위: 백만 원)

기본 부문	기초 부문	부문명칭	국산산출액	수입액	잔폐물	총공급액
001	****a	벼	6,970,742	0	0	6,970,742
001	0000	벼	0	0	0	0
001	0100	벼	6,463,795	0	0	6,463,795
001	9900	부산물	506,947	0	0	506,947
002	****a	맥류 및 잡곡	193,108	3,567,460	0	3,760,568
002	0000	맥류 및 잡곡	0	0	0	0
002	0101	쌀보리(나맥)	44,313	0	0	44,313
002	0102	겉보리(대맥)	14,691	8,296	0	22,987
.
.
.
383	****a	가사서비스	1,695,173	0	0	1,695,173
383	0000	가사서비스	1,695,173	0	0	1,695,173
384	****a	기타 개인서비스	3,246,566	31	0	3,246,597
384	0000	기타 개인서비스	238,545	31	0	238,576
384	0101	장례식장	691,936	0	0	691,936
384	0102	화장 및 공원묘지 관리	290,072	0	0	290,072
384	0200	예식장	1,116,424	0	0	1,116,424
384	8801	점술 서비스	213,264	0	0	213,264
384	8802	개인 간병인	441,598	0	0	441,598
384	8803	결혼 상담	111,214	0	0	111,214
384	8804	애완동물 관련 서비스	11,832	0	0	11,832
384	8888	기타	131,681	0	0	131,681
		소계	3,048,067,463	571,494,231	9,443,844	3,629,005,538
		해외직접구매	0	20,183,485	0	20,183,485
		잔폐물발생액	0	0	-9,443,844	-9,443,844
		합계	3,048,067,463	591,677,716	0	3,639,745,179

자료: 한국은행 2010년 산업연관표

현재 활용 가능한 품목별·부문별 공급액표는 한국은행에서 2010년에 발표한 자료로 현재 2014년 자료는 활용이 불가능한 상황이다. 따라서 본 연구에서는 2010년 자료를 활용하여 데이터산업을 재분류하였다. 품목별·부문별 공급액표는 위에서 살펴본 바와 같이 30개 산업을 세분류한 384개 기본부문을 다시 세분류하여 약 3,136개의 기초부문들의 총산출액에 대한 정보를 담고 있다. 384개 기본부문의 하위개념인 기초부문들의 정의를 살펴보고 이를 위에서 정의한 데이터산업의 각 부문별 정의와 맞는 기초부문들을 분류하여 데이터산업을 재구성했다. 먼저 기본부문에 속하는 기초부문 중 데이터산업과 연계된 한국 표준산업분류체계 또는 ICT통합분류체계에 해당하는 기초부문을 확정하였다. 다음으로 확정된 기초부문이 기본부문 총 국산산출액에서 얼마의 비중을 차지하는지를 산출하여 산업연관표 내에서 데이터산업을 재구성하였다. 예를 들어 기초부문 1의 총 산출액이 90만원이고 해당 기초부문 1이 3개의 기초부문으로 구성되어 있으며, 각 기초 부문이 동일하게 30만원의 산출량을 가지고 있다고 가정할 때, 3개의 기초부문 중 2개의 기초부문이 데이터산업으로 확정될 경우 기초부문 1의 60%가 데이터산업에 해당하는 것으로 볼 수 있다.

먼저²⁸⁾ 데이터산업의 대분류 부문 중 하나인 데이터솔루션과 관련된 기본부문은 329번째 소프트웨어 개발 공급 부문으로 이 중 데이터솔루션의 세부분류와 일치하는 기초부문은 시스템소프트웨어와 응용소프트웨어 부문이다. 따라서 329번째 기본부문의 국내총산출액에서 해당 기본부문에 속하는 두 기초부문인 시스템소프트웨어와 응용소프트웨어의 국내산출액의 비중을 산출하였다. 산출결과 329번째 기본부문 중 데이터산업으로 분류될 수 있는 산출액은 해당 기본부문의 총 산출액의 16.5%로 나타났다. 시스템소프트웨어와 응용소프트웨어 부문의 일부분만이 데이터솔루션에 속하지만 이보다 더 상세한 분류는 불가능하여 둘 모두가 데이터솔루션 부문에 해당한다고 가정하고 데이터솔루션 부문을 확정하였다.

데이터산업의 두 번째 대분류인 데이터구축/건설링 부문과 관련된 기본부문은 330번째 컴퓨터관리서비스 부문으로 해당 부문의 모든 기초부문이 데이터 구축/건설링 부문으로 분류 가능하다. 따라서 330번째 기본부문의 국내총산출액 전체를 데이터구축/건설링 부

28) 아래에 기술한 데이터산업의 재분류 과정을 담은 품목별·부문별 공급액표는 부록에 수록되어있다.

문으로 분류하고 330번 째 기본부문의 국내총산출량 모두가 데이터구축/건설팅에 해당한다고 볼 수 있다.

다음으로 세 번째 대분류인 데이터서비스 부문과 관련된 기본부문은 328번 째 정보서비스 부문으로 이중 하나의 기초부문인 자료처리를 제외한 나머지 기초부문이 모두 데이터서비스 부문으로 분류 가능하다. 따라서 328번 째 기본부문의 국내총산출량에서 해당 기본부문에 속하는 기초부문들 중 자료처리를 제외한 나머지 기초부문의 국내총산출량이 328번 째 기본부문의 국내총산출량에서 차지하는 비중을 산출하였다. 산출결과 328번 째 기본부문 중 데이터산업으로 분류될 수 있는 산출량은 해당 기본부문의 총 산출량의 96.5%로 나타났다.

데이터산업의 마지막 대분류인 데이터인프라 부문과 관련된 기본부문은 컴퓨터기억장치(232), 유선통신기기(234), 기타 무선통신장비 및 방송장비(236), 유선통신서비스(323), 기타전기통신서비스(325), 정보서비스(328)이다. 각 기본부문과 ICT통합분류체계를 활용하여 연결하면 컴퓨터주변기기는 232, 네트워크장비는 234과 236, 호스팅서비스는 328, 부가네트워크서비스는 323, 인터넷 관리 서비스 및 기타 정보인프라서비스는 325번, 호스팅서비스는 328번 째 부문과 각각 연결 가능하다. 각 기본부문에 속한 기초부문 중 위에서 연결된 부문에 해당하는 기초부문들을 확정한 해당 기본부문에서 데이터인프라산업이 차지하는 비중을 산출하였다. 컴퓨터기억장치(232), 유선통신기기(234), 기타 무선통신장비 및 방송장비(236)의 모든 기초부문이 데이터인프라에 속하여 이들 기본부문의 국내총산출량 모두가 데이터산업의 산출량으로 분류될 수 있다. 유선통신서비스(323)의 경우 부가 네트워크서비스와 연결 가능한 부가네트워크서비스를 제외한 타 기초부문은 데이터인프라 부문과 상관없는 산출량으로 확정한 결과, 323번 째 기본부문의 국내총산출량의 4.7%만이 데이터인프라로서 데이터산업의 산출량으로 확정할 수 있는 것으로 나타났다. 기타전기통신서비스(325)의 기초부문 중 부가통신응용 및 중개서비스를 제외한 나머지 부문은 데이터산업과 무관한 것으로 판단하여 제외한 결과 데이터인프라 부문으로서 데이터산업의 산출량으로 분류할 수 있는 국내총산출량은 기타전기통신서비스(325)의 국내총산출량의 60.6%를 차지한다. 마지막으로 정보서비스(328)의 기초부문 중 호스팅서비스와 연결가능한 부가통신응용 및 중개서비스를 제외한 나머지 부문은 데이터산업과 무관한 것으로 판단하여 제외한 결과 328번 째 기본부문의 국내총산출량 중 데이터인프라로서 데이터산업의 산출

량 비중은 3.1%에 불과한 것으로 나타났다.

위와 같은 방법으로 제정되던 데이터산업은 아래 <표 3-7>에 정리되어있다. 먼저 데이터솔루션 부문은 329번 제 기본부문인 소프트웨어 개발 공급 부문과 연결될 수 있으며, 329번 제 부문의 기초부문 중 데이터솔루션 부문으로 분류할 수 있는 매출액은 329번 제 기본부문의 약 16.5%를 차지하는 것으로 산출되었다. 데이터 구축/컨설팅 부문은 330번 제 기본 부문과 동일하게 연결되어 330번 제 부문의 매출액 100%가 모두 데이터 구축/컨설팅 부문으로 분류되었다. 데이터서비스 부문은 328번 제 부문의 전체 매출액의 약 97%를 차지하는 것으로 산출되었으며, 마지막으로 데이터 인프라 부문은 232, 234, 236번 제 부문과 동일하게 연결되어 이들 부문의 전체 매출액을 포괄하고, 328 번째 부문 매출액의 약 3.1%, 323번 제 부문 매출액의 약 4.7%, 325번 제 부문 매출액의 약 60.6%까지 포함하는 것으로 나타났다.

<표 3-7> 품목별·부문별 공급액표에서 데이터산업의 비중

(단위: 백만원)

구분	기본부문	기본부문 총공급액 (A)	데이터산업 매출액 (B)	데이터 산업 비중 (B/A)	
데이터 솔루션	329	24,643,411	4,066,547	16.5%	
데이터 구축/컨설팅	330	9,521,487	9,521,487	100.0%	
데이터 서비스	328	7,398,392	7,169,552	96.9%	
데이터 인프라	기타 컴퓨터주변기기	232	6,299,798	6,299,798	100.0%
	네트워크장비	234	5,801,994	5,801,994	100.0%
		236	9,671,972	9,671,972	100.0%
	호스팅 서비스	328	7,398,392	228,840	3.1%
	부가 네트워크서비스	323	17,089,067	804,858	4.7%
	인터넷 관리 서비스, 기타 정보인프라서비스	325	3,239,488	1,962,151	60.6%

위와 같이 품목별·부문별 공급액표를 바탕으로 산출한 각 부문의 비중을 바탕으로 2014년 산업연관표를 활용하여 산업연관분석을 진행하였다. 앞서 언급되었듯이 품목별·부문별 공급액표는 2010년 자료만이 활용가능하기 때문에 산업연관분석에서는 가장 최근 자료인 2014년의 자료를 이용하되, 데이터산업을 재분류에서는 2010년 자료가 활용되었다. 2014년 자료를 활용한 산업연관분석에서는 2010년 품목별·부문별 공급액표에서 나타난 데이터산업 대분류가 차지하는 각 기본부문의 비중이 변하지 않았다고 가정하고 위에서 구한 비중을 그대로 2014년 산업연관표에 적용하여 데이터산업을 재정의한 후 분석을 진행하였음을 밝힌다. 즉, 산업을 분리하고 부분합을 적용하는 부분에서만 2010년 산업연관표 공급액표를 기준으로 정의한 후 나머지 기본적인 산업연관분석은 2014년도 산업연관표를 이용하여 분석하였다.

이를 위해서 먼저 산업연관표의 내생부문인(투입요소행렬)인 30×30 행렬에 재정의된 데이터산업을 구성함으로써 본 연구에서 사용할 산업연관표를 재구성할 수 있다. 데이터산업이 기존 30개 부문 산업에 추가적으로 재정의되면서 데이터산업이 포함된 산업연관표는 아래 예시처럼 31×31 행렬로 구성된다.

〈표 3-8〉 데이터산업 포함 이전의 산업연관표 예시

		1	2	...	30	총산출액
		농림수산물	광산품	문화 및 기타서비스	
1	농림수산물	500	100	200	800
2	광산품	200	0	100	300
.
.
.
.
30	문화 및 기타서비스	100	200	50	350
총 투입액		800	300	350	1450

〈표 3-9〉 데이터산업을 추가한 산업연관표 예시²⁹⁾

		1	2	...	30	31	총산출액
		농림수산물	광산품	문화 및 기타 서비스	데이터 산업	
1	농림수산물	500	100	150	50	800
2	광산품	200	0	70	30	300
.
.
30	문화 및 기타서비스	90	150	25	0	265
31	데이터산업	10	50	20	5	85
총 투입액		800	300	265	85	1450

먼저 〈표 3-8〉과 〈표 3-9〉를 살펴봄으로써 데이터산업을 기존 산업연관표 상에서 어떻게 추가하였는지를 이해할 수 있다. 〈표 3-8〉과 〈표 3-9〉는 이러한 과정의 이해를 돕기 위해 예시로 작성된 표로 내생부문만을 고려하였고 표 안의 수치들은 실제 수치와는 무관함을 밝힌다.

먼저 〈표 3-8〉을 살펴보면 30개 산업과 내생부문으로만 구성된 30×30 행렬의 형태를 갖고 있다. 이 가상의 산업연관표를 살펴보면, 농림수산품의 총산출액과 총투입액은 800 단위로 동일하고, 광산품 역시 300 단위, 문화 및 기타서비스 역시 350 단위로 총산출액과 총투입액이 동일한 것을 확인할 수 있다. 다음으로 문화 및 기타서비스의 최종생산물이 어떻게 활용되고 어떻게 생산되었는지를 살펴보면, 먼저 마지막 행을 통해 문화 및 기타서비스의 최종생산물이 농림수산물 산업에서 100 단위, 광산품에서 200 단위가 각각 산업에서 최종재를 생산하며 중간재로 활용되었으며, 50 단위는 문화 및 기타서비스 산업 내에서 자체적으로 활용되었음을 확인할 수 있다. 다음으로 30번째 열을 살펴보면 문화 및 기타서비스 산업에서 최종재를 생산하는 과정에서 농림수산물 200 단위, 광산품 100 단위,

29) 해당 표는 데이터산업을 추가하는 과정을 개념화한 것으로 실제 수행된 과정과는 차이가 있을 수 있다.

그리고 문화 및 기타서비스 50 단위가 투입되어 총투입액이 350 단위가 되는 것을 확인할 수 있다. 문화 및 기타서비스는 384개 기본부문 중 몇 개의 기본부문의 합으로 구성되는데, 위에서 살펴보았듯이 품목별·부문별 공급액표를 활용하여 이들 기본부문 중에서 데이터산업으로 분류될 수 있는 매출액 비중을 산출하였다고 가정하자. 그 결과를 바탕으로 문화 및 기타서비스 전체 총산출량 중 85 단위가 데이터산업 매출로 재분류될 수 있는 매출액을 85로 산출할 수 있다.

<표 3-9>는 이를 바탕으로 데이터산업을 <표 3-8>에 추가한 형태로 해당 표를 살펴보면 국내 총산출액과 총투입액은 그대로인 상태로 문화 및 기타서비스의 산출량이 총 350 단위에서 265 단위로 줄어들고 줄어든 총산출액 85 단위가 데이터산업으로 귀속되었음을 알 수 있다. <표 3-9>에 새로 데이터산업이 추가되면서 기존 30×30 행렬의 형태가 31×31로 바뀌었음을 알 수 있다. <표 3-9>의 31번째 행을 살펴보면 데이터산업으로 재분류된 산출액들이 어떠한 산업에서 중간재로 얼마만큼 투입되었는지를 알 수 있는데, 먼저 <표 3-8>에서는 문화 및 기타서비스 산업의 최종산출물 100 단위가 농림수산물에서 중간재로 활용되었다면 <표 3-9>에서는 문화 및 기타서비스 산업의 최종산출물 90 단위가 농림수산물에서 중간재로 활용되어 <표 3-8>에서의 100 단위보다 10 단위 감소한 것을 확인할 수 있다. 이 감소한 10 단위는 데이터산업 계정에 새로 귀속되어 데이터산업의 총산출량 85 단위 중 10 단위가 농림수산물 산업에서 중간재로 활용되었음을 확인할 수 있다. 이러한 방식으로 기존 <표 3-8>에서 문화 및 기타서비스 산업의 최종산출물 중 200 단위가 광산품 산업에서 최종생산물 생산에 사용되었다면 <표 3-9>에서는 문화 및 기타서비스의 최종산출물 중 150 단위가 광산품에서 중간재로 활용되고 <표 3-8>과의 차이인 50 단위는 데이터산업의 최종산출물이 광산품에서 중간재로 활용된 것으로 재기입 되었다. 그 다음으로 <표 3-8>에서는 문화 및 기타서비스의 최종산출물이 다시 자기산업 내에서 중간재로 활용된 양이 50으로 나타나었지만, <표 3-9>에서는 자기산업에서 다시 중간재로 활용된 양이 25 단위로 줄고 데이터산업의 최종산출물 20 단위가 문화 및 기타서비스 산업에서 중간재로 활용되었음을 볼 수 있다. 마지막으로 데이터산업에서의 최종생산물 중 자기산업에서 중간재로 다시 활용된 양이 5 단위로 기입되며 데이터산업의 최종산출액 85 단위와 일치하는 것을 볼 수 있다.

이러한 변동은 31 번째 열에서도 대칭적으로 반영되어 나타나는데 기존 <표 3-8>에서

는 농림수산품의 최종산출물 200 단위가 문화 및 기타서비스 산업의 중간재로 활용되었다면 <표 3-9>에서는 농림수산품의 최종산출물 150 단위가 문화 및 기타서비스 산업에서 나머지 50 단위는 데이터산업에서 중간재로 활용되는 것으로 변동되었음을 확인할 수 있다. <표 3-8>에서 광산품의 최종산출물 100 단위가 문화 및 기타서비스 산업에서 중간재로 활용되었다면 <표 3-9>에서는 기존 100 단위가 문화 및 기타서비스 산업에서 70, 데이터산업에서 30으로 나뉘어 중간재로 활용된 형태를 갖게 된다.

이러한 방식으로 재구성된 산업연관표를 바탕으로 먼저 데이터산업이 국민 경제 전체에서 차지하는 위상을 파악해 볼 수 있는데, 먼저 31개 부문으로 구성된 국민 경제 전체에서 31 개 산업이 가지고 있는 전·후방연쇄효과를 측정한 결과는 아래와 같다.

<표 3-10> 31개 산업 전·후방연쇄효과

30대 산업명	전방효과		후방효과	
	값	순위	값	순위
농림수산품	0.955	13	0.966	19
광산품	0.585	30	0.912	22
음식료품	1.092	9	1.230	4
섬유 및 가죽제품	0.871	20	1.059	13
목재 및 종이, 인쇄	0.971	12	1.105	9
석탄 및 석유제품	1.260	7	0.695	31
화학제품	1.972	2	1.107	8
비금속광물제품	0.756	23	1.126	7
1차 금속제품	2.007	1	1.282	1
금속제품	1.081	10	1.243	3
기계 및 장비	0.922	18	1.221	5
전기 및 전자기기	1.280	6	1.003	15
정밀기기	0.645	26	1.093	11
운송장비	0.936	15	1.279	2
기타 제조업 제품 및 임가공	1.009	11	1.076	12
전력, 가스 및 증기	1.366	4	0.767	27
수도, 폐기물	0.731	24	1.009	14
건설	0.603	27	1.177	6

30대 산업명	전방효과		후방효과	
	값	순위	값	순위
도소매서비스	1.673	3	0.954	20
운송서비스	1.300	5	0.877	24
음식점 및 숙박서비스	0.884	19	1.104	10
정보통신 및 방송	0.940	14	0.969	18
금융 및 보험 서비스	1.221	8	0.893	23
부동산 및 임대	0.928	17	0.748	29
전문, 과학 및 기술	0.929	16	0.877	25
사업지원서비스	0.858	21	0.812	26
공공행정 및 국방	0.597	28	0.749	28
보건 및 사회복지	0.589	29	0.937	21
문화 및 기타 서비스	0.721	25	0.997	16
데이터산업	0.778	22	0.987	17

먼저 전방효과는 해당산업의 생산물이 증가하였을 때 이를 중간투입물로 활용하여 전 산업에서 발생하는 생산물 증가의 상대적인 크기를 의미하고, 후방효과계수는 해당산업의 생산물이 증가하였을 때 해당 산업 생산증가에 필요한 중간재나 원료를 공급하기 위해 전 산업에서 발생하는 생산물 증가의 상대적인 크기를 의미한다. 즉, 전방효과계수는 해당산업이 생산한 최종재가 타 산업에서 중간재로 사용되어 유발되는 생산유발효과의 척도로 공급측면의 공급승수(supply multiplier)와 관련이 깊고, 후방효과계수는 해당산업의 최종재 생산을 위해 타 산업으로부터 생산된 최종재를 중간재로 사용됨에 따라 유발되는 생산유발효과의 척도로 수요측면의 수요승수(demand multiplier)와 관련이 있다.

위의 표에서 도출된 전·후방 연쇄효과는 영향력계수와 감응도 계수를 활용하여 계산되었다. 영향력계수와 감응도 계수는 생산유발계수를 이용하여 각 산업 간의 상호의존관계의 정도를 전 산업의 평균치를 기준으로한 상대적 크기로 표시한 것인데, '영향력계수는 어떤 산업부문의 생산물에 대한 최종수요가 한 단위 증가하였을 때 전 산업부문에 미치는 영향, 즉 후방연쇄효과의 정도를 전 산업 평균에 대한 상대적 크기로 나타낸 계수로서 당해 산업의 생산유발계수의 열합계를 전 산업 평균³⁰⁾으로 나누어 구한다. '감응도계수는 모

30) 한국은행(2014b).

든 산업부문의 생산물에 대한 최종수요가 각각 한 단위씩 증가하였을 때 어떤 산업이 받는 영향, 즉 전방연쇄효과가 어느 정도인가를 전 산업 평균에 대한 상대적 크기로 나타낸 계수로서 그 산업의 생산유발계수의 행 합계를 전 산업 평균³¹⁾으로 나누어 구한다.

이를 종합해 보면 후방연쇄효과 도출을 위해 사용하는 영향력계수는 수요측면 분석과정에서 사용되는 $(I-A^d)^{-1}$ 의 열합계를 사용하고, 전방연쇄효과 도출을 위해 사용하는 감응도계수는 공급측면 분석과정에서 사용되는 $(I-B^d)^{-1}$ 의 행 합계를 사용하는 것이 적절하다. 마지막으로 어떤 산업의 전·후방연쇄효과가 1이라는 것은 전 산업 평균과 해당산업의 전·후방연쇄효과가 동일하다는 것으로 위의 표에서 전·후방연쇄효과가 1보다 낮다는 것은 전 산업평균보다 낮은 것을, 1보다 높다는 것은 전 산업평균보다 높다는 것을 의미한다.

먼저 전방연쇄효과가 높은 상위 10개 산업을 살펴보면 1위부터 10위 순으로, 1차 금속제품, 화학제품, 도소매서비스, 전력 및 가스 및 증기, 운송서비스, 전기 및 전자기기, 석탄 및 석유제품, 금융 및 보험서비스, 음식료품, 금속제품으로 나타났다. 후방연쇄효과가 높은 상위 10개 산업을 살펴보면 1위부터 10위 순으로, 1차 금속제품, 운송장비, 금속제품, 음식료품, 기계 및 장비, 건설, 비금속광물제품, 화학제품, 목재 및 종이, 인쇄, 음식점 및 숙박서비스로 나타났다. 이상에서 알 수 있듯이 전·후방연쇄효과가 높은 산업들은 주로 제조업 군에 속하거나 제조업과 관련이 깊은 산업들로 제조업 중심 우리나라 경제체제에서 당연한 결과로 해석할 수 있다.

데이터산업의 전방연쇄효과는 0.778, 후방연쇄효과는 0.957로 모두 평균인 1보다 낮은 것을 알 수 있다. 전방연쇄효과의 경우 총 31개 산업 중에서 22위 후방연쇄효과의 경우 총 31개 산업 중에서 17위로 나타나 전방연쇄효과 보다 후방연쇄효과가 더 높은 산업으로 추정되었다. 이러한 결과는 데이터산업이 데이터 상품을 전 산업에 공급하여 효율성을 높이는 산업이라는 일반적인 관념에 반하는 것이다. 여기에는 데이터산업의 전방연쇄효과가 신산업 창출 등을 포괄하지 못하기 때문으로 보인다. 산업연관분석은 이미 작성된 산업 간의 연관관계를 활용하는 분석이므로 데이터산업의 활성화 또는 데이터산업의 최종 생산물을 활용에 따라 창출되는 신산업으로 인한 파급효과 등을 포괄하지 못한다.

데이터산업의 경제적 파급효과가 총 31개 산업과 비교하여 평균에 미치지 못하지만 이

31) 한국은행(2014b).

는 데이터산업이 서비스업의 성격을 강하게 갖는 것에 기인한다. 타 서비스업들과 비교해 볼 경우 금융 및 보험서비스, 도소매 서비스, 사업서비스를 제외한 모든 서비스산업들보다 높은 전·후방연쇄효과를 갖는 것으로 미루어 볼 때, 데이터산업이 국민 경제에서 차지하는 위상은 동 산업 군 - 서비스산업 -에 비해 높은 것으로 생각해 볼 수 있다.

다음으로 데이터산업의 대분류별 전·후방연쇄효과를 비교해보면, 먼저 전방연쇄효과와의 경우 데이터구축/컨설팅이 0.704로 가장 높은 전방연쇄효과를 갖는 것으로 나타났다. 그 다음으로 데이터인프라, 데이터서비스, 데이터솔루션 순으로 높은 전방연쇄효과를 갖는데, 데이터 구축/컨설팅의 경우 여러 산업에서 이미 중간재로서 활용되는 경우가 많기 때문에 가장 높은 전방연쇄효과를 갖는 것으로 보이며, 반면 데이터분석 등을 포함하는 데이터서비스의 경우 신산업 창출 효과를 포괄하지 못하고, 또한 데이터를 활용이 사회 전반적으로 확산되지 못한 2014년도의 자료를 바탕으로 파급효과가 도출된 한계점으로 인해 전방연쇄효과가 비교적 낮게 나타난 것으로 보인다. 반면 후방연쇄효과와의 경우 1순위부터 4순위 순으로 데이터서비스, 데이터구축/컨설팅, 데이터인프라, 데이터솔루션 부문이 높은 후방연쇄효과를 갖는 것으로 나타났는데, 전·후방연쇄효과 모두에서 데이터솔루션이 가장 낮은 경제적 파급효과를 갖는 것으로 확인되었다.

〈표 3-11〉 데이터산업 대분류별 전·후방연쇄효과 비교

데이터산업 대분류	전방효과		후방효과	
	값	순위	값	순위
데이터산업 전체	0.778	-	0.987	-
데이터 솔루션	0.544	4	0.797	4
데이터구축/ 컨설팅	0.704	1	1.032	2
데이터 서비스	0.591	3	1.035	1
데이터 인프라	0.609	2	0.989	3

위의 전·후방연쇄효과는 데이터산업을 외생화하지 않고 도출된 수치이다. 데이터산업의 외생화에 대해서는 아래에서 더 자세하게 논의될 것이지만, 간략하게 데이터산업의 외생화란 데이터산업을 내생부문, 즉, 기존 주요 30개 주요 산업에 추가하여 31번째 산업으로

로써 포함시키지 않고, 데이터산업을 외생부문에 삽입한 뒤 데이터산업이 30개 내생부문에 미치는 영향을 보기 위한 분석 방법이다. 외생화 기법을 적용할 경우에는 데이터산업의 직접적인 자기연관효과³²⁾가 제외되어 경제적 파급효과 과소 추정될 수 있다. 또한 위와 같이 데이터산업을 31번째 산업으로 취급하여 국민경제 전체에서 데이터산업이 차지하는 위상을 파악하는데 한계가 있을 수 있다. 반면 외생화 기법을 적용하지 않을 경우 데이터산업의 직접적인 자기연관효과가 제외되지 않아 데이터산업에서 증가된 지출이 다시 자기산업에 대한 중간재 수요로 이어져 경제적 파급효과가 과대 추정될 가능성이 있다.

이 후의 분석에서는 데이터산업이 타 산업에 어떠한 영향을 미치는지에 대해 집중하여 논의를 진행하기 때문에, 이 후에는 데이터산업을 국내 경제 30대 산업 중 하나로 취급하여 분석하기보다, 데이터산업을 외생부문에 추가하여 데이터산업이 타 산업에 미치는 순경제적 파급효과를 살펴보고자한다.

데이터산업 외생화의 의미는 데이터산업을 내생부문에서 분리해내 외생부문³³⁾에 위치시키는 것을 의미한다. 데이터산업을 외생화시키지 않은 경우에는 내생부문에서 데이터산업을 분리해내어 하나의 산업으로서 30×30 행렬에 추가하여 분석을 진행하였는데, 이 경우 내생부문의 행렬이 31×31로 확장된 산업연관표를 활용하여 산업연관분석을 진행하였다. 데이터산업을 외생화시킬 경우 기존 내생부문의 산출량에서 데이터산업에 해당하는 부문의 열을 따로 떼어내어 외생부문에 위치시킴으로써 산업연관표의 내생부문이 30×30 행렬로 변형되고 내생부문의 산출량 행렬 내의 숫자들이 변화하게 된다.

32) 내생화된 경우의 자기산업의 최종수요가 1단위 증가했을 경우 자기산업의 생산유발효과는 1단위 보다 반드시 높게 나타나게 되어 있다. [부록 4] 2014년 산업연관표 생산유발계수(통합대분류)에 대각행렬은 자기산업의 유발효과에 직접적인 자기연관효과가 1이 포함되어 항상 1보다 높게 나타남을 확인할 수 있다. 외생화하는 경우 외생화 하는 산업은 이러한 직접인 효과를 제외한다.

33) 외생부문은 크게 부가가치, 최종수요, 그리고 수입으로 이루어져있는데, 부가가치는 노동이나 자본투입으로, 최종수요는 민간부문소비지출, 정부부문소비지출, 민간 및 정부의 고정자본형성 등으로 구성되어 있다.

〈표 3-12〉 데이터산업의 외생화 예시³⁴⁾

		1	2	...	30	내생부문 총계	데이터 산업	최종 수요	총산출액
		농림 수산물	광산품	문화 및 기타 서비스				
1	농림수산물	500	100	150	750	50	100	900
2	광산품	200	0	70	270	30	200	500
.
.
30	문화 및 기타서비스	90	150	25	265	-	300	565
내생부문총계		790	250	...	245	1285	-	-	-
부가가치		110	250	...	320	685	-	-	-
총 투입액		900	500	565	600	-	-	-

〈표 3-12〉는 데이터산업을 외생화하는 과정의 예를 보여주는데 〈표 3-12〉는 앞서 살펴본 〈표 3-8〉과 〈표 3-9〉에서 사용한 예시를 그대로 사용하였으며, 차이점은 외생부문을 임의로 추가한 것이다. 〈표 3-9〉와 〈표 3-12〉를 비교해보면 데이터산업의 외생화 과정을 이해할 수 있는데 외생화를 한 후에도 국민경제의 총산출액과 총투입액은 변하지 않는 것을 확인할 수 있다. 또한 데이터산업을 재분류하여 재정의하면서 데이터산업 계정으로 추가된 산출액 등을 해당 산출액이 속해있었던 기존 산업에서 제하고 데이터산업계정으로 옮겨 기입한 것을 확인할 수 있다. 〈표 3-12〉에서 내생부문은 30×30으로 유지되나 〈표 3-9〉에서 데이터산업을 추가하면서 변화된 산출액 내역은 그대로 유지된다. 다만 새로 추가된 데이터산업이 외생부문에 삽입됨으로써 〈표 3-9〉의 내생부문에서 차감된 매출액들을 보존하여 국내 총산출은 변화하지 않는다.

위와 같은 외생화 과정을 거친 후 먼저 데이터산업의 경제적파급효과를 수요측면에서

34) 해당 표는 외생화 과정을 개념적으로 보여주기 위한 표로서 실제 외생화 과정과는 차이가 있을 수 있다. 정확한 논의는 송준혁(2017)에 잘 나타나있다.

분석한 결과는 다음과 같다. 이후의 결과는 데이터산업이 외생부문으로 옮겨진 산업연관표 상에서 분석하므로 경제적 파급효과는 자기산업의 직접적인 연관효과를 제외한 타 산업에 대한 파급효과만을 포함한 계수가 산출되었다.

데이터산업을 외생화한 후에 분석한 수요측면의 산업연관분석 결과 데이터산업의 타 산업 생산유발효과는 0.652원, 부가가치 유발효과는 0.251원, 취업유발효과는 4.509명으로 나타났다. 데이터산업의 산출량이 1원 증가할 때 데이터산업을 제외한 국민경제 전체에서 유발되는 생산액은 0.652원, 유발되는 부가가치는 0.251원이며, 데이터산업의 산출량이 10억 원 증가할 때 데이터산업을 제외한 국민경제 전체에서 유발되는 취업인원은 약 4.5명으로 추정되었다.

〈표 3-13〉 데이터산업의 수요측면 경제적 파급효과

	생산유발 효과	부가가치 유발효과	취업유발효과
타 산업 전체 유발효과	0.652	0.251	4.509
자기산업의 직접적인 연관효과	1.000	0.368	3.203
합계	1.652	0.619	7.712

수요측면 분석에서 생산유발효과는 데이터산업의 최종생산물을 1원 만큼 더 생산할 경우 이를 위해서 필요한 다른 30개 부문 산업의 생산물 역시 더 생산되어야 하는데 생산유발효과 0.652원은 데이터산업의 최종생산물이 1원 더 생산될 때 국민경제를 구성하는 30개 산업에서 총 0.652원의 추가적인 생산물이 생산되어야 하는 것을 의미한다. 총 0.652원의 추가생산에 데이터산업에서 1원의 추가생산액을 합하면 데이터산업의 수요측면에서 총 생산유발효과인 1.652원이 도출된다. 부가가치유발효과 역시 비슷한 맥락에서 이해할 수 있는데, 부가가치유발효과는 데이터산업의 최종생산물을 1원 만큼 더 생산할 경우 이를 위해서 필요한 다른 30개 부문 산업의 생산물이 더 생산될 때 유발되는 부가가치이다. 예를 들어 데이터산업의 최종생산물을 1원 만큼 더 생산할 경우 총 0.251원의 부가가치가 다른 30개 부문에서 생산의 추가적인 생산과정에서 유발되는 것을 의미한다. 위에서 살펴본 생산유발효과와 마찬가지로 데이터산업에서 1원의 추가적 생산이 일어나며 데이터산업

내에서 발생하는 부가가치인 0.368을 더하면 데이터산업의 수요측면에서 총 부가가치유발효과 0.619원이 도출된다. 마지막으로 취업유발효과는 데이터산업의 최종생산물이 10억 원 만큼 더 생산될 경우 이를 위해서 필요한 다른 30개 부문 산업의 생산물이 더 생산되면서 이들 30개 산업에서 증가되어야 하는 취업인원수이다. 위의 결과에 따르면, 데이터산업에서 10억 원만큼 더 생산되면서 추가적인 중간재 수요를 충족하기 위해 다른 30개 산업에서 생산을 늘리며 총 4.509명의 인원이 추가적으로 고용되는 것을 뜻한다. 또한 데이터산업에서 10억 원만큼의 추가적 생산에 따라 데이터산업 내에서 추가적으로 고용해야하는 인원은 3.203명으로, 타 산업에서 추가적인 고용인원인 4.509명을 합하여 데이터산업의 수요측면에서 총 취업유발효과인 7.712명이 도출될 수 있다.

다음으로 수요측면 분석에서 도출된 생산유발효과, 부가가치유발효과, 취업유발효과를 30대 산업별로 살펴보도록 하겠다.

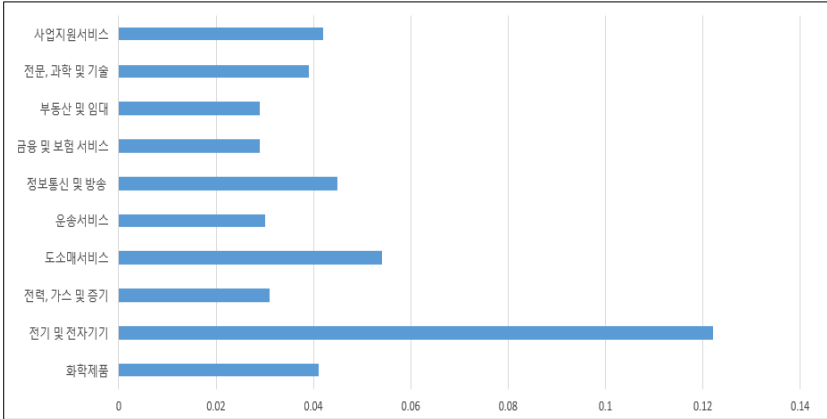
먼저 수요측면에서의 생산유발효과는 외생화된 산업연관표를 활용하여 도출되었는데, 데이터산업의 외생화 이후에도 앞서 연구방법론에서 살펴본 생산유발효과의 도출과정이 변형되어 그대로 사용될 수 있다. <표 3-14>에 나타난 30대 산업별 생산유발효과는 데이터산업을 산업연관표 상에서 재정의하고 외생화 한 후의 산업연관표에서 계산된 $(I-A^d)^{-1}$ 을 바탕으로 도출된 값이다. 데이터산업의 재정의 및 외생화 과정을 거친 후 도출된 생산유발효과는 $(I-A^e)^{-1}A_K^e$ 로 나타나며³⁵⁾, 30×30 행렬의 형태를 갖는데 이 행렬에서 각 열은 각 산업의 최종생산물이 타 산업에서 중간재로 얼마나 활용되는지에 대한 정보를 담고 있다. 따라서 각 열의 열 합계는 각 산업의 최종생산물이 타 산업에서 중간재로 활용되는 정도로 볼 수 있는데, <표 3-14>에 나타난 산업별 생산유발효과는 외생화된 데이터산업에서 1원의 추가적인 수요가 일어날 때 각 산업들이 이를 위해 추가적으로 생산해야하는 생산물의 가치를 정규화(normalized)하여 나타난 값들이다.

35) $(I-A^e)^{-1}$ 는 투입계수행렬 A^d 에서 외생화 K 산업의 행과 열을 제외한 레온티에프 역행렬이며, A_K^e 는 투입계수행렬 A^d 의 외생화 K 부문 열벡터에서 K 부문 원소를 제외한 열벡터이다.

〈표 3-14〉 데이터산업의 수요측면 생산유발효과

30대 산업명	생산유발효과	
	값	순위
농림수산물	0.004	24
광산물	0.000	30
음식료품	0.009	18
섬유 및 가죽제품	0.006	22
목재 및 종이, 인쇄	0.016	16
석탄 및 석유제품	0.019	13
화학제품	0.041	5
비금속광물제품	0.005	23
1차 금속제품	0.026	11
금속제품	0.018	14
기계 및 장비	0.011	17
전기 및 전자기기	0.122	1
정밀기기	0.007	21
운송장비	0.007	20
기타 제조업 제품 및 임가공	0.017	15
전력, 가스 및 증기	0.031	7
수도, 폐기물	0.004	25
건설	0.003	26
도소매서비스	0.054	2
운송서비스	0.030	8
음식점 및 숙박서비스	0.021	12
정보통신 및 방송	0.045	3
금융 및 보험 서비스	0.029	10
부동산 및 임대	0.029	9
전문, 과학 및 기술	0.039	6
사업지원서비스	0.042	4
공공행정 및 국방	0.003	27
교육서비스	0.001	29
보건 및 사회복지	0.003	28
문화 및 기타 서비스	0.009	19
타산업 합계	0.652	-

[그림 3-2] 10대 후방산업



따라서 수요측면의 생산유발효과는 후방연쇄효과로 생각해 볼 수 있는데, 데이터산업 전체의 후방연쇄효과는 총 0.652원으로 데이터산업에서 수요 1원 증가로 인해 전체 30개 산업에서 추가적으로 중간재로 사용되는 0.652원 어치의 생산이 유발된다고 볼 수 있다. 이 0.652원 이라는 수치는 각 30대 산업에서 생산유발계수의 총 합이다. 따라서 생산유발효과를 30대 산업별로 살펴볼 수 있는데, 가장 큰 후방연쇄효과를 갖는 산업은 전기 및 전자기기 산업으로 나타났다. 그 뒤를 이어 도소매서비스, 정보통신 및 방송, 사업지원서비스, 화학제품, 전문 과학 및 기술, 전력 가스 및 전기, 운송서비스, 부동산 및 임대, 금융 및 보험서비스 산업 순으로 높은 후방연쇄효과를 갖는 것으로 나타났다.

다음으로 수요측면에서의 부가가치유발효과 역시 외생화된 산업연관표를 활용하여 도출되었는데, 데이터산업을 산업연관표 상에서 재정의하고 외생한 후의 산업연관표에서 계산된 부가가치유발효과는 $A_v^e(I-A^e)^{-1}A_K^e$ 이며³⁶⁾, 이를 바탕으로 도출한 결과는 다음과 같다.

36) A_v^e 는 부가가치 대각행렬에서 외생화 K산업의 행과 열을 제외한 행렬이다.

〈표 3-15〉 데이터산업의 수요측면 부가가치유발효과

30대 산업명	부가가치유발효과	
	값	순위
농림수산물	0.002	18
광산품	0.000	30
음식료품	0.001	24
섬유 및 가죽제품	0.001	26
목재 및 종이, 인쇄	0.004	15
석탄 및 석유제품	0.001	27
화학제품	0.008	10
비금속광물제품	0.001	25
1차 금속제품	0.004	16
금속제품	0.005	13
기계 및 장비	0.003	17
전기 및 전자기기	0.033	1
정밀기기	0.002	20
운송장비	0.002	22
기타 제조업 제품 및 임가공	0.007	12
전력, 가스 및 증기	0.008	9
수도, 폐기물 및	0.002	21
건설	0.001	28
도소매서비스	0.028	3
운송서비스	0.011	8
음식점 및 숙박서비스	0.008	11
정보통신 및 방송	0.020	6
금융 및 보험 서비스	0.015	7
부동산 및 임대	0.022	5
전문, 과학 및 기술	0.022	4
사업지원서비스	0.028	2
공공행정 및 국방	0.002	19
교육서비스	0.000	29
보건 및 사회복지	0.002	23
문화 및 기타 서비스	0.005	14
타산업 합계	0.251	-

따라서 부가가치유발효과인 0.251 역시 각 30대 산업의 부가가치유발효과를 합한 것이고, 생산유발효과에서 살펴본 것처럼 30대 산업별로 그 정도를 살펴볼 수 있다. 먼저 데이터산업의 최종수요 한 단위가 증가할 때 가장 큰 부가가치 유발효과를 갖는 산업은 생산유발효과와 동일한 전기 및 전자기기산업으로 나타났다. 그 뒤를 이어 사업지원서비스, 도소매서비스, 전문, 과학 및 기술, 부동산 및 임대, 정보통신 및 방송, 금융 및 보험 서비스, 운송서비스, 전력, 가스 및 증기, 화학제품 순으로 부가가치유발효과가 큰 것으로 나타났다. 부가가치유발효과가 큰 상위 10개 산업들은 앞서 살펴본 10대 후방산업들과 대동소이하나 그 순위에서 차이가 있다. 이러한 차이는 10대 후방산업들에서 부가가치 비중이 얼마나 큰가에 따라 재조정된 것으로 보인다.

마지막으로 수요측면에서 분석한 취업유발효과는 데이터산업 제정의 및 외생화 과정을 완료한 산업연관표를 바탕으로 도출되었는데 외생화 후 수요측면의 취업유발효과는 $L^e(I-A^e)^{-1}A_K^e$ 이다³⁷⁾. 데이터산업 전체의 취업유발효과 4.509는 데이터산업의 최종수요가 10억 원 증가함에 따라 추가적으로 30대 산업에서 중간재 공급을 위한 생산을 증가시킬 때 필요한 추가적 고용인원 수로, 4.509는 30대 산업 전체의 합으로 구해진 값이다. 취업유발효과 역시 30대 산업별로 살펴볼 수 있는데, 가장 높은 취업유발효과를 갖는 전방산업은 사업지원서비스로 나타났다. 그 뒤를 이어 2순위부터 10순위 순으로 도소매서비스, 전문, 과학 및 기술, 음식점 및 숙박서비스, 운송서비스, 정보통신 및 방송, 전기 및 전자기기, 문화 및 기타 서비스, 금융 및 보험 서비스, 기타 제조업 제품 및 임가공 산업이 취업유발효과가 높은 것으로 나타났다.

〈표 3-16〉 데이터산업의 수요측면 취업유발효과

30대 산업명	취업유발효과	
	값	순위
농림수산물	0.109	11
광산품	0.002	30
음식료품	0.028	18

37) L^e 는 취업계수 대각행렬에서 외생화 K 산업의 행과 열을 제외한 행렬이다.

30대 산업명	취업유발효과	
	값	순위
섬유 및 가죽제품	0.027	19
목재 및 종이, 인쇄	0.076	13
석탄 및 석유제품	0.002	29
화학제품	0.057	14
비금속광물제품	0.013	27
1차 금속제품	0.022	24
금속제품	0.046	15
기계 및 장비	0.036	17
전기 및 전자기기	0.201	7
운송장비	0.014	26
정밀기기	0.024	21
기타 제조업 제품 및 임가공	0.123	10
전력, 가스 및 증기	0.023	23
수도, 폐기물 및 건설	0.019	25
건설	0.026	20
도소매서비스	0.763	2
운송서비스	0.327	5
음식점 및 숙박서비스	0.350	4
정보통신 및 방송	0.269	6
금융 및 보험 서비스	0.157	9
부동산 및 임대	0.107	12
전문, 과학 및 기술	0.421	3
사업지원서비스	1.022	1
공공행정 및 국방	0.024	22
교육서비스	0.008	28
보건 및 사회복지	0.043	16
문화 및 기타 서비스	0.173	8
타산업 합계	4.509	-

지금까지는 데이터산업 전체를 기준으로 데이터산업의 수요측면 산업연관분석은 결과를 살펴보았다. 앞서 데이터산업을 산업연관표에서 재구성할 때 데이터산업의 대분류별로

나누어 재정의하였기 때문에 본 연구에서는 데이터산업의 대분류 - 즉, 데이터솔루션, 데이터 구축/컨설팅, 데이터 서비스, 데이터 인프라 부문 - 별로 수요측면 분석이 가능하다. 데이터산업 전체의 수요측면 산업연관분석 결과 - 생산유발효과, 부가가치유발효과, 취업유발효과는 데이터산업의 대분류 별 생산유발효과, 부가가치유발효과, 취업유발효과와의 평균과 비슷한 성격을 지닌다. 데이터산업의 대분류별 생산유발효과, 부가가치유발효과, 취업유발효과를 살펴보면 다음과 같다. 먼저 생산유발효과와 취업유발효과와 비교하면 데이터서비스가 0.920원으로 가장 높게 나타났고, 그 다음이 데이터 인프라, 데이터 구축/컨설팅, 데이터 솔루션 순으로 높게 나타났다. 특히 데이터 서비스, 데이터 인프라의 경우 데이터산업 전체의 생산유발효과인 0.652를 상회하는 것으로 나타났다. 하지만 데이터 구축/컨설팅의 경우 데이터산업 전체의 생산유발효과에 조금 못 미치는 0.647로, 데이터 솔루션의 경우 데이터산업 전체 생산유발계수인 0.652에 크게 미치지 못하는 0.491로 나타났다.

부가가치유발계수는 다른 양상을 보이는데 데이터산업의 대분류 중에서 자기산업 내 부가가치유발효과를 포함하여 가장 높은 부가가치 유발계수를 보이는 것은 데이터 솔루션으로 약 0.825원의 부가가치유발효과를 갖는 것으로 나타났다. 그 뒤를 이어 데이터서비스, 데이터 구축/컨설팅, 데이터 인프라 순으로 높은 부가가치유발효과를 갖는 것으로 나타났다. 특히 데이터솔루션, 데이터 구축 컨설팅, 데이터 서비스 부문은 모두 데이터산업 전체보다 높은 부가가치유발효과를 갖는 것으로 나타났다.

데이터산업 대분류별로 살펴본 취업유발계수 역시 비슷한 양상을 보이는데, 자기산업 고용효과를 포함하여 데이터솔루션이 14.651명으로 가장 높은 취업유발계수를 가지는 것으로 나타났고, 그 다음으로 데이터 서비스, 데이터 구축/컨설팅, 데이터 인프라 순으로 높은 취업유발계수를 갖는 것으로 나타나 부가가치유발계수에서의 대분류별 순위와 동일한 것으로 나타났다.

부가가치유발계수와 취업유발계수는 모두 외생부문으로부터 도출된 정보인데 부가가치유발계수의 경우 부가가치계수 행렬, 취업유발계수의 경우 취업계수 행렬을 활용하여 도출된다. 이들 부가가치계수 행렬과 취업계수 행렬은 모두 외생부문에 포함된 정보로서 산업연관표 내에서 이미 기입된 값들의 비율과 값들로 볼 수 있다. 이들 정보는 데이터산업의 대분류별 성격에 따라 정해지는데, 데이터 인프라와 같이 제조업의 특성이 강한 부문의 경우 외생부문에서 부가가치나 고용과 관련된 비중이 낮을 수밖에 없다. 따라서 부가

가치유발효과나 취업유발계수가 타 대분류 부문에 비해 낮게 산출될 수밖에 없다. 데이터 구축/컨설팅의 경우에도 역시 제조업의 성격이 포함되어 부가가치유발계수 및 취업유발계수가 대분류 별 순위에서 3 순위로 산출된 것으로 보인다. 데이터솔루션의 경우 소프트웨어 개발 및 이와 관련된 지적재산권으로 매출이 발생하므로 부가가치나 임금의 투입비중이 높은 부문임을 생각해 볼 수 있다. 데이터 서비스 역시 데이터 솔루션 보다 기타 장비 등을 더 많이 활용하므로 제조업의 특성이 조금 더 강하기는 하지만, 데이터 구축/컨설팅 또는 데이터 인프라에 비해 인적자본의 투입이 높고, 부가가치 창출이 용이하기 때문에 부가가치유발효과 및 취업유발효과에서 2 순위를 차지한 것으로 보인다.

〈표 3-17〉 데이터산업의 수요측면 산업연관분석 결과의 대분류별 비교

		전체 데이터 산업	데이터 솔루션	데이터 구축/ 컨설팅	데이터 서비스	데이터 인프라
생산 유발 계수	타산업	0.652	0.491	0.647	0.920	0.784
	자기산업	1.000	1.000	1.000	1.000	1.000
	합계	1.652	1.491	1.647	1.920	1.784
	순위	-	4	3	1	2
부가 가치 유발 계수	타산업	0.251	0.222	0.263	0.382	0.277
	자기산업	0.368	0.602	0.374	0.414	0.284
	합계	0.619	0.825	0.637	0.796	0.560
	순위	-	1	3	2	4
취업 유발 계수	타산업	4.509	4.827	5.258	6.282	4.457
	자기산업	3.203	9.824	2.486	4.532	1.310
	합계	7.712	14.651	7.744	10.814	5.767
	순위	-	1	3	2	4

지금까지 살펴본 외생화 이후 수요측면 산업연관분석은 전체 30대 산업을 후방산업 - 즉, 데이터산업에 중간재를 공급하는 산업으로 보고 경제적과급효과를 도출하기 위한 분석이었다. 반면 지금부터 살펴볼 공급측면 산업연관분석은 데이터산업의 최종생산물을 타 30대 산업에서 중간재로 활용하는 것에 초점을 맞춘 분석으로 30대 산업을 전방산업으로

보고 도출한 전방연쇄효과와 깊은 관련이 있는 분석이다.

공급측면 분석은 앞서 연구방법론에서 살펴본 것처럼 내생부문의 행 - 어떠한 산업의 최종생산물이 어느 산업에서 얼마만큼 중간재로 사용되었는지에 대한 정보를 포함 -에 초점을 맞춘 분석이다. 다음에 살펴볼 공급지장효과 역시 이러한 맥락에서 이해할 수 있는데 공급지장효과란 어떤 산업에서의 최종생산물이 1원만큼 덜 생산됨으로써 타 산업에서 중간재로서 사용되는 생산물의 양이 줄어들어 생산에 차질을 빚는 효과이다. 외생화 후에 살펴본 데이터산업의 공급지장 효과란 데이터산업에서 최종수요가 1원만큼 줄어들음으로써 타 30개 부문의 산업에서 중간재로 활용되는 데이터산업 최종생산물이 감소하여 발생하는 생산액 감소분으로 이해할 수 있는데, 이를 생산 증가의 측면에서 해석하면 데이터산업의 최종수요가 1원 증가할 때 타 30개 부문의 산업에서 이 증가된 중간재를 활용한다는 가정하에 얼마만큼의 생산이 증가했는지를 가늠해 볼 수 있다.

이러한 공급지장효과는 내생부문으로 구성된 30×30 행렬을 이용하여 도출되는데 앞서 연구방법론에서 살펴본 것처럼 핵심이 되는 정보는 $(I - B^d)^{-1}$ 에 모두 포함되어 있다. 앞서 설명한 바와 같이 공급측면에서 공급지장효과는 아래 식과 같이 나타낼 수 있다.

$$X' = W(I - B^d)^{-1}$$

위와 같은 행렬식에서 데이터산업을 외생화하면 아래와 같은 수식을 도출할 수 있다.

$$X^{e'} = (W^e + B_K^e X_K)(I - B^e)^{-1}$$

위의 수식에서 외생화된 부가가치, 즉 W^e 의 변동이 없다고 가정하고 데이터산업의 최종수요 X_K 가 증가하였을 경우 타 산업에서의 산출량 변동분 ΔX^e 를 도출하면 아래 수식과 같이 나타낼 수 있다.

$$\Delta X^{e'} = B_K^e \Delta X_K (I - B^e)^{-1}$$

즉, 외생화된 데이터산업의 최종수요가 한 단위 증가하였을 때 타 30개 부문의 산업의 산출량 변동분은 $B_K^e (I - B^e)^{-1}$ 을 통해 산출³⁸⁾된다. 위와 같은 과정을 거쳐서 산출된 데이

38) $(I - B^e)^{-1}$ 는 산출계수행렬 B^d 에서 외생화 K 산업의 행과 열을 제외한 산출 역행렬이며 ,

터산업 전체의 공급지장효과(또는 공급측면에서 도출된 생산유발효과)는 0.892원으로 데이터산업에서 최종수요 1원이 증가하였을 때, 이를 중간재로 활용하여 타 30개 부문에서 추가적으로 생산되는 최종재의 가치가 총 0.892원임을 뜻한다. 공급측면에서 도출된 생산유발효과 역시 30대 산업별로 살펴볼 수 있는데, 이는 아래 <표 3-18>에 정리되어 있다.

<표 3-18> 데이터산업의 공급측면 생산유발계수

30대 산업명	공급측면 생산유발계수	
	값	순위
농림수산물	0.011	28
광산물	0.017	20
음식료품	0.012	26
섬유 및 가죽제품	0.011	29
목재 및 종이, 인쇄	0.034	7
석탄 및 석유제품	0.019	17
화학제품	0.021	15
비금속광물제품	0.019	18
1차 금속제품	0.021	14
금속제품	0.016	22
기계 및 장비	0.013	25
전기 및 전자기기	0.029	9
정밀기기	0.013	24
운송장비	0.012	27
기타 제조업 제품 및 임가공	0.020	16
전력, 가스 및 증기	0.026	10
수도, 폐기물 및	0.023	12
건설	0.015	23
도소매서비스	0.059	3
운송서비스	0.036	6
음식점 및 숙박서비스	0.024	11

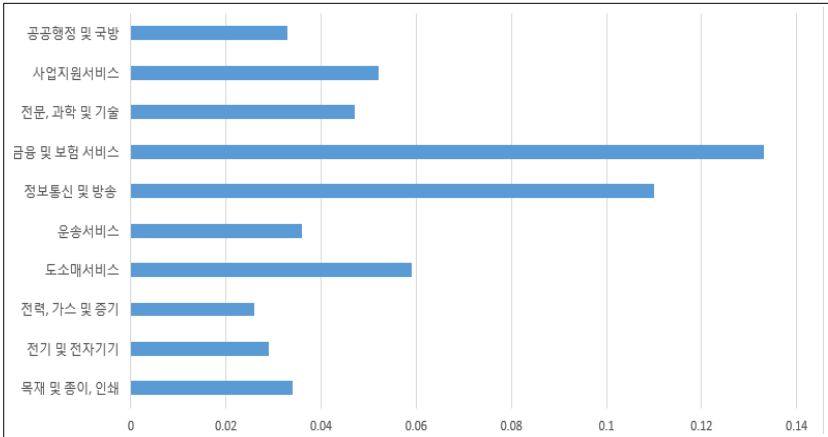
B_K^c 는 산출계수행렬 B^d 에서 외생화 K 부문의 행벡터에서 K 부문 원소를 제외한 행벡터이다.

30대 산업명	공급측면 생산유발계수	
	값	순위
정보통신 및 방송	0.110	2
금융 및 보험 서비스	0.133	1
부동산 및 임대	0.017	21
전문, 과학 및 기술	0.047	5
사업지원서비스	0.052	4
공공행정 및 국방	0.033	8
교육서비스	0.018	19
보건 및 사회복지	0.008	30
문화 및 기타 서비스	0.022	13
타산업 합계	0.892	-

먼저 가장 높은 공급측면 생산유발계수를 갖는 산업은 금융 및 보험 서비스로 0.133원의 추가적인 생산이 일어나는 것으로 나타났다. 그 뒤를 이어 정보통신 및 방송, 도소매서비스, 사업지원서비스, 전문, 과학 및 기술, 운송서비스, 목재 및 종이, 인쇄, 공공행정 및 국방, 전기 및 전자기기, 전력, 가스 및 증기 산업 순으로 높은 전방연쇄효과를 갖는 것으로 나타났다.

10대 전방산업을 살펴보면 금융 및 보험서비스와 정보통신 및 방송 산업의 전방효과가 가장 크게 나타났는데, 데이터산업의 최종재를 중간재로서 사용하는 산업들을 개념적으로 생각해 볼 때 타당한 결과로 볼 수 있다. 데이터산업의 최종재인 데이터분석 및 데이터 구축/컨설팅을 중간재로 활용하는 산업들이 어떠한 산업인지에 대해 생각해 보면 30대 산업 중 금융 및 보험서비스, 도소매서비스, 운송서비스가 대표적이다. 금융 및 보험서비스의 경우 최근 핀테크 열풍을 차치하더라도 개인의 신용정보 또는 보험료 산출을 위해서는 데이터산업에서의 최종생산물이 중간재로 활용되어야 한다. 또한 미국의 아마존의 예에서처럼 개인 선호에 대한 데이터를 바탕으로 어떠한 상품을 누구에게 마케팅할지를 결정하기 위해서는 데이터산업의 최종생산물이 중간재로 활용되어야 한다. 마지막으로 운송서비스의 경우 최적의 운송경로를 산출할 때는 물론이고, 기본적인 운송서비스를 제공하기 위해서는 어느 곳에 언제 운송서비스를 제공해야하는지에 대한 정보가 필요하기 때문에 데이터산업의 최종생산물이 높은 비중으로 사용될 수밖에 없다.

[그림 3-3] 10대 전방산업



다음으로 데이터산업을 외생화한 후에 분석한 공급측면의 산업연관분석 결과를 데이터 산업의 각 부문별로 비교해 볼 경우 경제 전체에 가장 큰 전방연쇄효과를 갖는 부문은 데이터서비스 부문이 1.395의 생산유발효과를 나타내는 것으로 나타났다. 이 뒤를 이어 데이터 구축/건설팅이 1.308로 2순위를 차지하였으며, 데이터 인프라 및 데이터 솔루션은 데이터산업 전체의 생산유발효과인 0.892를 하회하는 생산유발효과를 가지며 각각 3순위와 4순위를 차지하였다.

<표 3-19> 데이터산업 대분류별 공급측면 생산유발계수

분류	공급측면 생산유발효과	
	값	순위
데이터산업 전체	0.892	-
데이터 솔루션	0.608	4
데이터구축/ 건설팅	1.308	2
데이터 서비스	1.395	1
데이터 인프라	0.742	3

제 4 장 데이터 경제 활성화 정책 방안

이 장에서는 데이터 경제 활성화 정책 방안에 대한 논의를 다룬다. 먼저 해외 주요국의 데이터 경제 활성화 정책을 개괄하고, 우리나라의 데이터 경제 활성화 정책의 역사와 최근 동향을 살펴보도록 하겠다. 다음으로 우리나라 200개의 데이터서비스 기업을 대상으로 실시한 정책수요 조사를 살펴보고, 이들을 바탕으로 데이터 경제 활성화 정책 방안에 대해 논의하도록 하겠다.

제 1 절 해외 주요국의 데이터 경제 활성화 정책

EU, 영국, 미국, 일본, 중국 등 주요 국가들은 데이터 경제를 생산성 증진을 통해 경제성장과 고용창출을 유도할 수 있는 수단으로 인식하고 혁신전략 및 정책 수립을 지속적으로 해오고 있다. 이 절에서는 먼저 EU, 영국, 미국, 일본, 중국 각국의 데이터 경제 활성화 정책에 대해 각각 살펴보도록 하겠다.

1. 유럽

유럽은 데이터 경제 진전을 위해 우선 데이터 경제를 측정하여 현황을 파악하고 데이터 R&D 투자, 인력양성 등 인프라 구축 및 데이터 개방, 공유, 보호 등 데이터 활용을 촉진하기 위한 정책을 지속적으로 추진해 오고 있다. 먼저 EU 데이터 경제를 측정하는 European Data Market 연구의 일환으로 <Data Landscape 프로젝트>를 2014년부터 시작하여 현재까지 이어오고 있는데, EU 데이터 시장 상황을 모니터링하는 도구 개발을 목표로, 데이터 시장 및 경제 활동 주체의 정의, 데이터 가치사슬 설계, 측정 지표 개발을 통해 데이터 경제 규모(현황, 전망)를 산출하고, 정책을 제시하고 있다. 해당 프로젝트는 비단 유럽의 데이터 경제뿐만 아니라 전 세계 데이터 시장의 현황 및 전망, 그리고 미국, 일본 등 세계 각국의 현황을 총망라하고 이를 유럽의 현황과 비교·분석하여 정책적 제언을 도출하는 것을

목적으로 하고 있다.

다음으로 데이터 기반 경제 기반 마련을 위해 2014년 유럽위원회는 민관의 협업을 목적으로 유수의 ICT기업들³⁹⁾, 그리고 빅데이터협회(Big Data Value Association)와 함께 총 2016년부터 총 4년에 걸쳐 에너지, 제조, 헬스 등 분야의 빅데이터 연구 및 혁신활동에 25억 유로를 투입하기로 결정하였다.⁴⁰⁾

이와 아울러 데이터를 보호하면서 데이터 소유권 및 이동권 촉진을 위한 정책 추진하고 있다. 2016년 'Digital Single Market' 전략에서 '디지털 싱글 마켓 내 유럽의 자유로운 데이터 흐름'이라는 영향 평가 보고서에서 이동권 등은 데이터 활동 시 필수조건으로 명시하였고, 2017년에는 유럽 데이터 경제 육성(Building a European Data Economy)계획을 통해 데이터 소유권(data sovereignty) 및 이동권(free flow of data)을 강화함으로써 새로운 비즈니스 모델을 개발 하는 것을 목표로 설정하였다.

데이터 소유권 및 이동권에 대한 정책을 추진함과 동시에 안전한 데이터 활용과 데이터 유통에 대한 신뢰와 책임을 강조하기 위해 개인정보보호규정 수립 역시 동시에 추진하고 있다. 1995년 EU 개인정보보호지침(Directive 95/46/EC)은 개인정보보호라는 목적에도 불구하고 새로운 기술의 도입, 국가별로 상이한 제도기반과 법적 환경 등에 따라 EU 회원국 간 자유로운 정보의 이동이 제한되고 있다는 한계점을 파악하고, 2016년 EU 개인정보보호지침을 대체하기 위해 개인정보보호법(GDPR: General Data Protection Regulation)' 제정하였다. GDPR은 정보 이동권, 프로파일링 권리 강화, 데이터 삭제권에 대한 원칙을 확립하고 GDPR의 테두리 내에서 개인정보를 포함한 데이터가 자유롭게 유통될 수 있는 체계를 구성함으로써 개인정보 유통에 대한 신뢰를 제고하는 것이 목적이며, 현재 GDPR은 2018년 5월 발효되어 전 세계 개인정보보호법의 기준으로서의 위상을 차지하고 있다.

유럽의 여러 국가 중 데이터 경제가 가장 활성화된 국가로 영국을 꼽을 수 있는데, 영국은 공공 데이터의 공유와 활용을 활성화하는 오픈데이터 정책 및 데이터 인력 양성, R&D 체계 수립 등을 포괄하는 데이터 역량강화 정책 추진해오고 있다.

먼저 방대한 공공데이터 활용 촉진을 위해 2010년 '정보공개법(Freedom of Information

39) ATOS, Nokia Solutions and Networks, Orange, SAP, SIEMENS 등이 포함되어 있다.

40) 민대홍 외(2018).

Act 2000)’ 으로 오픈데이터 활용이 용이하게 됨에 따라 ‘데이터 공개 사이트인 ‘Data. Gov.uk’를 개설하고, 2012년에는 기업혁신숙련부 등 17개 주요 부처가 ‘오픈데이터 전략’을 마련하여, 의료, 교육, 세금, 고용, 기상데이터 등을 순차적으로 개방하기로 결정하였다. 뒤 이어 2014년에는 데이터의 공유와 활용 활성화를 위한 전략으로서 데이터 공개 의무화, 데이터양과 질 모두에서 수준 향상, 모두가 이용할 수 있는 데이터, 데이터 개방을 통한 정부역량 강화, 데이터 개방을 통한 혁신의 5대 원칙을 제시하였다. 2015년 오픈데이터 전략의 추진을 위한 정책으로 오픈데이터 전략을 감독·추진하는 최고데이터 책임자를 임명함으로써 데이터 공개와 관련된 자금 지원, 더 많은 데이터 공개, 이미 공개된 오픈데이터의 활용촉진을 위한 정부와 민간에 대한 교육 지원, 오픈데이터를 활용한 연구개발 강화를 추진해오고 있다. 또한 2017년에는 데이터 경제법을 입법하고, 공공기관 간 개인정보 공유 및 외부 공개를 확대·추진하기로 하였다. 해당 법안에서는 공공기관이 연구, 사기방지 등의 목적으로 개인정보를 활용할 경우 타 공공기관과 개인정보를 교환할 수 있으며, 비식별화되어 개인식별이 불가능할 경우 개인정보를 공개할 수 있음을 명시하였다.⁴¹⁾

또 다른 한편으로는 데이터 관리 및 분석에 관한 기술 확보 및 전문성을 갖춘 인력 양성 등을 위해 2013년 ‘데이터역량강화전략(A strategy for UK data capability)’ 발표하였는데, 위 전략은 데이터 분석 및 연구를 위한 인프라 및 연구개발 추진체계 구축하는 것을 목표로 하고 있다. 해당 전략 중 데이터분석 및 연구 촉진을 위한 세부 전략을 살펴보면 먼저 (1) 고성능 컴퓨팅, 클라우드 컴퓨팅 등을 통합해 대량의 데이터 수집·저장·분석·시각화 등 빅데이터의 전 과정을 지원할 수 있는 인프라인 ‘e-인프라 시스템’을 구축하고, (2) ‘빅데이터 분석 국가 네트워크’를 바탕으로 민·관·학·연의 R&D 협력 체계 지원하여 빅데이터 R&D를 촉진하며, (3) 영국 정부 산하 연구위원회들로 구성된 TF 주도로 공공데이터 기반 대규모 분석 및 연구를 위한 ‘공공데이터 연구 네트워크’를 구성하는 것을 골자로 하고 있다. 인력양성 부분에서는 데이터 전문 인력(Human Capital) 확보를 위해 교육체계 개선 내용을 구체적으로 제시하였는데, (1) 초중등 교육을 개혁하여 2014년 3/4분기부터 5-16세 학생들을 대상으로 컴퓨팅 과정을 도입하여, 데이터 분석에서 필요한 우수 능력을 지닌 학생을 양성하는 계획을 추진하는 동시에, (2) 대학 졸업자들이 실제 현장에서 필요로 하

41) 한국정보화진흥원(2017), 민대홍 외(2018) 재인용.

는 전문성을 습득할 수 있도록 전문 기업과의 제휴를 통해 빅데이터 기술 교육 프로그램 개발 및 대학 졸업자들이 박사과정 트레이닝 센터(Centres for Doctoral Training)에서 4년 이상의 빅데이터 관련 전문 교육을 추가로 이수할 수 있도록 지원하는 정책을 추진하고 있다.

2. 미국

미국은 데이터 활용을 위한 오픈데이터 정책 및 연구개발을 통한 핵심기술 확보에 중점을 둔 정책을 추진하고 있다.

먼저 개인정보를 보호하면서 데이터 활용으로 경제적, 사회적 편익을 증대하는 방향으로 데이터를 개방하고 공개된 데이터의 표준화 마련하였는데, 2012년 '디지털 전략'에서 수치 데이터 이외에 문서 정보 등의 비정형 데이터의 공개를 추진하였고, 2013년 오픈데이터 정책에서 오픈데이터를 최종사용자가 완전한 형태로 입수·사용 가능한 공개 데이터로 정의하고 오픈데이터의 활용을 위한 정보관리 원칙을 제시하였다. 오픈데이터 정책에서 제시한 정보관리 원칙은 데이터표준과 메타데이터 표준 준수, 정보시스템 구축을 통한 상호운용성과 정보접근성의 담보, 데이터 관리와 개방 강화를 위한 데이터 저장소 구축, 사용자 친화적 API 방식 적용, 개인정보 보호, 기관별 정보 자원관리 전략으로 요약할 수 있다. 2014년에는 재무데이터에 관한 정부 표준을 확립하는 것을 목적으로 제정된 '데이터법'을 근거로 연방정부의 재무데이터를 공개하는 웹사이트 개설함으로써 연방정부 기금의 투명성을 제고하고 정부 지출 등에 관한 데이터 활용을 촉진하고자 하였다. 또한 2014년 "빅데이터: 기회 포착과 가치보호"에서 오픈데이터 정책에 의한 데이터 개방으로 야기되는 개인정보 보호 이슈에 대한 접근, 큰 틀에서 빅데이터 정책의 방향, 민관의 데이터 관리를 언급하였다. 구체적으로는 개인정보보호와 관련되어 전 지구적 관점의 프라이버시 보호, 빅데이터로 활용으로 인해 발생가능한 차별 방지에 대해 논의하는 한편, 데이터관리와 관련하여 공공재로서의 빅데이터 이용 강화, 빅데이터 혁명 촉진을 위한 연구·기술 개발, 빅데이터를 통한 학습 기회 제공, 빅데이터를 이용한 법집행과 국가안보 영역에서 책임 있는 사용을 강조하였다. 2016년 '알고리즘과 인권보호'에서는 특정 알고리즘을 활용하여 개인정보 데이터를 분석하는 과정에서 야기될 수 있는 특정 인종 또는 계층의 인권침해를

각 분야별(신용, 교육, 고용, 법률) 사례를 언급하며, 추후에 발생할 수 있는 윤리적 문제에 대한 논의도 진행하였다.

위와 같은 개인정보보호와 데이터 활용을 강조한 정책에 발맞추어 데이터 과학 핵심기술 및 인력 확보에 중점을 둔 R&D 정책 역시 동시에 추진하고 있다. 대표적인 예로 2012년에는 대통령실 내 과학기술정책국(OSTP)은 빅데이터 기술 R&D, 빅데이터 전문 인력 양성을 목표로 '빅데이터 R&D 이니셔티브(Big Data R&D Initiative)'를 발표하였다. 2015년에는 이니셔티브 일환으로 데이터 과학 혁신을 촉진하는 거점으로서 전국 네 곳에 빅데이터 지역 혁신 허브(Big Data Regional Innovation Hub; BD Hubs)를 설립해 정부부처, 민간기업, 대학 등 205개 이상의 조직과 컨소시엄을 구성하였다. 뒤이어 2016년 연방정부 네트워크

〈표 4-1〉 미국의 빅데이터 R&D 전략

7대 전략	세부 과제
미래 빅데이터 특성을 반영한 기술 개발로 차세대 능력 함양	<ul style="list-style-type: none"> • 데이터의 크기, 전달·처리 속도, 복잡성에 보조를 갖춘 기술 개발 • 미래에 요구되는 새로운 빅데이터 기술의 방법론 개발
데이터의 신뢰성 및 더 나은 빅데이터 기반 의사결정을 위한 R&D 지원	<ul style="list-style-type: none"> • 데이터의 신뢰성과 타당성을 제고시켜 더 나은 결과 도출 • 데이터 기반 의사결정을 지원하는 도구 개발
빅데이터 혁신을 가능하게 하는 사이버 인프라 구축 및 강화	<ul style="list-style-type: none"> • 국가 데이터 인프라 강화 • 빅데이터에 대한 응용과학 사이버 인프라 역량 강화
데이터 공유 및 관리를 촉진하는 정책을 통해 데이터 활용 가치 향상	<ul style="list-style-type: none"> • 데이터 투명성과 효용성을 증가시키는 메타데이터의 모범사례 개발 • 데이터 자산에 효율적이고 지속적이며 안전한 접근 제공
개인정보보호, 보안 및 빅데이터의 수집·공유·활용의 윤리적 측면 이해	<ul style="list-style-type: none"> • 올바른 개인정보보호 • 안전한 빅데이터 사이버공간 구축 • 데이터 거버넌스를 위한 정보윤리 이해
국가의 빅데이터 교육 및 훈련 환경 개선, 폭 넓은 인력 확충	<ul style="list-style-type: none"> • 데이터 과학자 양성 • 데이터 영역 전문가 커뮤니티 확장 • 데이터 사용이 가능한 인력 확충 • 공공데이터 활용 역량 개선
정부기관, 대학, 기업, 비영리 단체와의 협력에 의한 빅데이터 혁신 생태계 지원	<ul style="list-style-type: none"> • 기관 간 빅데이터 협력 장려 • 빠른 대응과 영향력 측정이 가능한 정책과 정책추진 프레임워크 구축

자료: 한국데이터산업진흥원(2017) 재구성

킹 IT R&D 프로그램(NITRD) 산하 빅데이터 협의체(BDIWG)에서 '빅데이터 R&D 전략 계획(The Federal Big Data R&D Strategic Plan)'를 발표하고 추진해오고 있다.

3. 일본

일본은 사회현안 해결을 위한 데이터개방과 활용을 촉진하는 정책 수립에 집중하고 있다. 2011년 동일본 대지진을 계기로 사회 전반의 협력을 위한 데이터의 중요성을 인식하고, 오픈데이터 유통 환경 정비 본격화하기 시작하였다. 먼저 2012년 공공데이터 활용 촉진을 위한 기본전략으로 '전자행정 오픈데이터 전략'을 제정하였는데, '정부가 직접 공공데이터를 공개, 기계판독이 가능한 형식 사용, 영리/비영리 목적을 불문한 데이터 활용 촉진, 그리고 공공데이터부터 신속하게 공개하는 것을 기본원칙으로 하며 내각관방장관의 총괄 하에 총무성, 경제산업성 등 여러 부처가 각 부처별 전략을 수립하는 것을 골자로 하고 있다. 또한 같은해인 2012년 경제산업성은 보유한 데이터를 기업이 활용할 수 있도록 개방하는 'DATA METI 구상' 역시 발표되었는데, 계량표준, 지질도, 시험·평가방법, 생물유전 정보, 표준물질, 제품사고, 첨단계측기술, 화학물질 관리 등 경제산업성이 보유한 데이터를 선행적으로 공개할 것을 골자로 하고 있다.

또한 일본은 데이터 활용을 촉진하는 방안을 포함하여 ICT 활성화 전략을 수립하고 추진해오고 있다. 대표적인 예로 2012년 총무성이 발표한 ICT 전략인 액티브 재팬 ICT 전략에 빅데이터 이용과 활용에 의한 사회·경제 성장의 액티브 데이터 전략 포함시켰다. 이 액티브 데이터 전략은 다양한 종류 및 대용량의 데이터를 실시간으로 수집·전송·분석하여 재난관리를 포함한 정책 과제의 해결에 연계함과 함께 수십 조 엔의 데이터 활용 시장이 창출되는 환경을 구축하여 빅데이터를 국가 자산화하여 성장 동력을 육성하겠다는 계획으로 요약될 수 있다. 2012년 일본 정부 산하 고도 정보통신 네트워크 사회추진 전략본부인 IT 종합전략 본부에서 정보통신 기술전략 공정표 개정안을 통해 빅데이터와 클라우드 추진 내용을 담은 국가 계획을 발표하였고, 뒤 이어 2013년 IT 종합전략 본부가 발표한 '세계 최첨단 IT 국가 창조 선언'에서 공공 데이터의 민간 개방 추진과 시장에 존재하는 빅데이터를 유기적으로 활용하는 것, 기업 활동과 소비자 행동에 이노베이션을 일으키는 사회를 실현하는 것을 목표로 설정하였다. 2016년에는 '일본재흥전략 2016'에서 4차 산업혁명의 핵심 요소 중 하나로 빅데이터 선정하고, 2017년 '미래투자전략 2017'에서는 새로운

사회의 인프라로서 데이터기반 플랫폼을 구축하고 데이터 활용을 촉진하기 위해 지적재산권이나 표준화에 관한 법적 고도화를 추진하겠다고 밝혔다.

또한 일본정부는 데이터 유통 촉진을 통해 데이터 활용을 확산하기 위한 제도 마련에도 힘쓰고 있다. 2016년 12월 ‘민관 데이터 활용 추진 기본법’을 제정을 시작으로, 2017년에는 ‘데이터 유통 플랫폼 간의 연계를 실현하기 위한 기본 사항’과 ‘데이터의 이용 권한에 대한 계약 가이드라인’을 발표했다. 전자는 데이터 이용자 측면에 중점을 둔 것인데, 이용자들의 편의성을 증대하기 위한 데이터 유통 플랫폼간의 상호 연계를 위한 필요한 사항(데이터 카탈로그, 카탈로그용 API)을 정리하였다. 후자는 데이터 거래 시 수집 또는 취득한 데이터 이용권한과 관련한 공정한 계약을 위한 가이드라인이다. 그 뒤를 이어 데이터 유통 추진 협의회의가 설립되었는데 해당 협의회의의 목적은 데이터 판매자나 데이터 구매자 간의 거래를 용이하기 위한 기술적·제도적 환경을 정비하는 것이다. 마지막으로 2018년에는 총무성의 총괄하에 2020년까지 공공부문과 민간부문의 데이터를 수집하여 공공·민간 통합 데이터 포털을 구축할 계획을 밝혔다.⁴²⁾

4. 중국

중국은 신산업성장동력으로 빅데이터산업을 중점적으로 육성하기 위한 정책을 추진하고 있다. 먼저 빅데이터를 국가 신산업으로 지정하고 정부 및 공공데이터의 개방과 빅데이터 수집, 관리 및 응용 촉진을 위한 기초 인프라, 시범기지를 건설하고 빅데이터 산업 발전을 위한 시장 환경 조성을 위해 노력해오고 있다. 2012년 국무원은 12·5 국가전략적 신흥산업발전규획을 통해 빅데이터 저장, 처리 기술의 연구개발 및 산업화를 천명한 후 2014년 3월 ‘2014년 정부 업무보고’에서 미래를 선도할 5G, 반도체, 첨단제조업, 신에너지, 신소재 등 신흥산업 중 하나로 ‘빅데이터’가 처음으로 등장하였다. 이후, 중앙정부의 적극적인 빅데이터 지원 정책을 통해 데이터 산업 관련 생태계가 형성되기 시작하였다.

2015년 국무원은 ‘빅데이터 발전 촉진을 위한 행동 요강’을 채택하였는데, 해당 행동요강에서는 정부정보시스템과 공공데이터의 개방 및 공유를 가속화하고, 통합관리 기능을 제고하여, 신흥 산업으로 빅데이터산업을 육성하기 위해 보안을 강화하고 관리 수준을 향

42) 민대홍 외(2018).

상시켜 안정적인 발전을 추구한다는 목표를 밝혔다. 같은해 10월에는 2016년부터 2020년까지 계획을 마련하는 18기 5중전회에서는 빅데이터를 국가전략 신흥산업 중 하나로 삼았다. 2016년 발표한 ‘제13차 5개년 계획’에서도 국가 빅데이터 전략을 명시하며 빅데이터산업 발전의 중요성을 강조하였고, 2017년 중국공업과정보화부에서 ‘빅데이터산업 발전계획(2016-2020년)’을 발표하여 빅데이터 기술과 상품 연구개발 강화, 공업 빅데이터 혁신응용 심화, 업계의 빅데이터 응용 발전 촉진, 빅데이터 산업 주체 육성, 빅데이터 표준 제정 및 확산 등을 주요 임무로 제시하였다.

또한, 중국 정부는 중앙정부 및 지방정부는 데이터 거래 활성화를 위해서도 노력하고 있는데, 중앙정부에서 데이터거래와 관련된 세부 정책을 빅데이터 발전계획에 포함시킨 후 지방정부에서도 이와 관련된 정책을 구상하여 발표하는 방식이다. 따라서 각 지방정부는 빅데이터 거래 시장을 조성하기 위해 빅데이터 종합연구소, 빅데이터 센터 등을 설립하고 있다.

〈표 4-2〉 중국의 중앙정부 및 지방정부 주요 빅데이터 정책

	발표 시기	발표 기관	관련 정책 및 회의	세부 내용
중앙 정부	2014년 3월	-	정부 업무보고	• 미래를 선도할 신흥산업 중 하나로 빅데이터를 포함(빅데이터, 5G, 반도체, 첨단제조업, 신에너지, 신소재)
	2015년 8월	국무원	빅데이터 발전 촉진 행동 강요	• 빅데이터 거래 시장 양성과 선도하기 위해 데이터 거래 시행 • 빅데이터 거래 탐색 및 전개로 생태계 각 분야 시장 주체의 데이터 교환과 거래를 독려 • 데이터 자원의 흐름을 촉진하고 빅데이터 거래 체계 및 가격결정 체계를 구축 • 데이터 거래 행위 규범화 • 빅데이터 시장거래 표준 체계 구축 가속화
	2015년 10월	-	18기 5중전회	• 13차5개년계획 기간(2016-2020년) 국가전략 신흥 산업 중 하나로 빅데이터를 포함
	2016년 1월	국가 발전 개혁 위원회	빅데이터 발전촉진을 위한 중대 공정 조직 및 실행에 관한 통지	• 국가 빅데이터 표준 체계 구축 및 개선을 통해 빅데이터 거래 플랫폼과 제도 구축

	발표 시기	발표 기관	관련 정책 및 회의	세부 내용
중앙 정부	2017년 1월	국가 정보 화부	빅데이터 산업 발전계획 (2016~2020년)	<ul style="list-style-type: none"> • 데이터 거래 기술을 발전시켜 데이터 거래의 기본 표준 연구체계 전개 • 데이터 거래 규정을 연구 및 제정하여 제 3자 데이터 거래 플랫폼 전개 및 데이터 거래 시범단지 구축
지방 정부	2015년 8월	샤먼 (厦門)	빅데이터 응용과 산업발전 계획(2015~2020년)	<ul style="list-style-type: none"> • 정부의 데이터 개방 플랫폼과 데이터 거래 시장 등 데이터 유통 플랫폼 구축 • 전국적인 '빅데이터 산업재산권 거래소' 구축 및 탐색 • 빅데이터 산업권 거래의 새로운 모델 탐색
	2015년 11월	쓰촨 (四川)	쓰촨성 빅데이터 발전 가속화를 위한 실행 의견	<ul style="list-style-type: none"> • 정부와 지방정부가 상호작용하는 빅데이터 수집 공유 체계 구축
	2015년 11월	쓰촨 (四川)	쓰촨성 빅데이터 발전 가속화를 위한 실행 의견	<ul style="list-style-type: none"> • 빅데이터 거래소 설립을 통해 전 방면의 빅데이터 거래 플랫폼 구축 • 부처, 업종별 데이터 순차적 개방 추진 • 빅데이터 가격결정, 거래, 규범화 사용 탐색 및 전개를 통한 데이터 자산화 촉진
	2016년 1월	구이저우 (貴州)	구이저우성 빅데이터 발전 응용 촉진 조례	<ul style="list-style-type: none"> • 빅데이터 거래 과정에서 나타나는 관련 요구사항을 제정하여 빅데이터 거래시장 양성 • 데이터 거래 행위를 규범화하여 법률적 근거 제공
	2016년 2월	광둥 (广东)	광둥성 빅데이터 발전 계획(2015~2020년)	<ul style="list-style-type: none"> • 데이터 평가체계와 거래 규칙을 구축 및 개선하여 민간자본의 참여 독려와 데이터 거래 플랫폼 구축
	2017년 2월	구이저우 (貴州)	구이저우 데이터 경제 발전 계획(2017~2020년)	<ul style="list-style-type: none"> • 데이터 거래 업무 양성 가속화 • 구이양 빅데이터 거래소 운영 모델 개선을 통한 국제 일류의 종합 빅데이터 거래 서비스 플랫폼 조성 • 데이터 상품 거래, 알고리즘 거래, 데이터 서비스 거래, 상업 데이터 관련 상품 거래 등 거래품종 발굴 • 데이터 자산 평가, 빅데이터 신용조회, 저당, 융자 등 패키지 업무 추진 • 빅데이터 거래 상품 체계 개선을 통한 데이터 거래 시장 형성

자료: 贵阳大数据交易(2016. 5), 민대홍 외(2018) 재인용

제 2 절 국내 데이터 경제 활성화의 정책 현황 및 정책 수요

1. 우리나라의 데이터 경제 활성화 정책

우리나라도 데이터의 중요성을 인식하고 공공데이터 개방 및 활용 촉진, 개인정보보호 가이드라인 수립, 전략산업으로 데이터산업육성 등 데이터 경제의 활성화를 위한 다양한 정책을 추진해 오고 있다.

2011년에 국가정보화전략위원회가 ‘빅데이터를 활용한 스마트 정부 구현(안)’을 발표한 후 2012년 방송통신위원회에서는 ‘빅데이터 서비스 활성화 방안’을 마련한데 이어 관계부처 합동으로 ‘스마트 국가 구현을 위한 빅데이터 마스터플랜’을 발표하였다. 뒤이어 2013년에는 관계부처가 합동으로 정부부문의 데이터 공개를 통한 공공데이터의 활용체계 구축을 목적으로 하는 ‘공공데이터의 제공 및 이용 활성화 기본계획(’13년~’17년)’을 발표하였다. 이에 따라 공공부문의 각 기관(중앙정부 및 지방자치단체)의 장에게 공공데이터의 공개 및 이용 활성화에 대한 계획 수립과 그 결과를 평가받는 것을 명문화한 ‘공공데이터의 제공 및 이용 활성화에 관한 법률(2013)’이 제정되었다.

공공데이터 활용 구축이 발표된 같은해인 2013년 관계부처 합동으로 창조경제 및 정부 3.0 지원을 위한 ‘빅데이터 산업발전전략’을 마련해 빅데이터 산업을 수요부문, 공급부문, 인프라 부문으로 나누고 각 부문에서 데이터 활용을 용이하게 하기 위한 정책 방향을 제시하였다. 각 부문의 정책방향은 빅데이터 시장창출을 위한 선도 시범사업 확대(수요부문), 산업육성기반 확충을 위한 핵심 기술 확보와 전문인력 양성(공급부문), 지속발전가능한 데이터 경제 생태계 조성을 위한 전략(인프라 부문)을 주요내용으로 하고 있다.

최근인 2016년에는 개인정보 관련 부처들(미래창조과학부, 방송통신위원회, 행정자치부, 보건복지부, 금융위원회)간의 협의 하에 “개인정보 비식별 조치 가이드라인”을 발표하였다. 해당 가이드라인은 비식별화 과정과 비식별화에 대한 정의를 구체적으로 제시하는 것을 목적으로 하였으나, 가이드라인의 성격으로 법적 구속력을 갖지 못하는 한계를 갖고 있다.

또한 같은해인 2016년에 정부기관, 민간기업 및 학계전문가로 이루어진 ‘민관 합동 빅데이터 TF’는 빅데이터 업무 추진현황의 종합적 파악과 공공과 민간 빅데이터 활성화 지원을 목적으로 출범하였는데, 해당 TF에서는 4차 산업혁명 촉진을 위한 부처별 빅데이터 관

런 정책 추진방향과 계획을 공유와 빅데이터 산업활성화를 위한 정책과 규제개선에 관한 논의가 이루어졌다.

2018년 5월에는 정부는 4차 산업혁명의 원동력인 데이터 자원의 효율적인 분배와 활용을 통해 전 산업의 혁신동력을 제고하고 신산업 창출을 목표로 하는 '데이터산업 활성화 전략'을 발표하였다. 해당전략은 '데이터를 가장 안전하게 잘 쓰는 나라'를 목표로 데이터 산업 활성화와 개인정보의 안전한 활용으로 요약할 수 있다. 전략의 구성은 데이터 이용제도 패러다임 전환, 데이터 가치사슬 전주기 혁신, 글로벌 데이터산업 육성 기반 조성의 3개 전략으로, 데이터 주권 찾기, 개인정보보호의 안전한 활용, 데이터 유통·분석·활용 활성화, 빅데이터 선도 기술 R&D, 데이터 인력 양성 등의 세부과제를 각 전략별로 추진하는 것을 목표로 하고 있다.

해당 활성화 전략의 정책의 목표는 데이터산업을 활성화를 통해 4차 산업혁명에 대응하고 기존 산업의 혁신 동력 확보 및 신산업을 창출하는 것으로, 데이터산업의 경제적 파급효과와 해외 주요국의 데이터 기반 산업 육성 전략을 고려할 때 적절하다고 생각된다. 이러한 방향성은 크게 개인정보보호에 관한 일반 규정과 유럽 데이터 경제 구축의 두 축으로 이루어져있는 EC의 4차 산업혁명 대응 전략인 'Digital Single Market'와 같다.

세부전략들을 차례로 살펴보면, 먼저 데이터 이용제도 패러다임 전환은 개인정보 활용에 대한 국민적 신뢰를 제고하는 것을 목표로 국민의 데이터 주권 찾기(Mydata)와 개인정보의 안전한 활용으로 구성되어 있다. 국민의 데이터 주권 찾기는 개인정보를 보유한 주체의 데이터 자기결정권 확립을 개인정보의 안전한 활용은 개인정보 활용 시 준수해야 할 필수사항을 제시하고 있다. 개인정보의 활용을 통해 데이터산업은 크게 비약할 수 있으나, 이를 위해서는 개인정보 활용에 대한 대국민 신뢰제도가 선제되어야한다. 해당 세부전략은 이를 위한 정책 방향성을 제시하고 있는데, 향후 데이터산업 활성화를 위한 기반 조성을 위해 장기적으로 추진되어야 할 것으로 판단된다.

데이터 가치사슬 전주기 혁신은 데이터 생애주기(구축·개방·유통·활용·분석)를 모두 고려한 산업 활성화 정책이다. 해당 전략은 구축, 개방, 유통, 활용, 분석 각 단계별로 AI 학습용 데이터 구축, 공공·민간데이터 개방, 데이터 거래기반 구축, 데이터 활용·분석을 통한 기존산업 활성화로 구성되어있는데, 우리나라의 데이터산업의 문제점이었던 활용 가능한 데이터 부족, 데이터 유통 기반 미비, 저조한 데이터 분석·활용을 해결하기 위한 정

책이다.

글로벌 데이터산업 육성 기반 조성은 빅데이터 관련 선도 기술(초연결 지능화, 인공지능 등) 수준 제고를 위한 R&D 지원과 빅데이터 전문인력 및 전문기업 양성으로 구성되어 있다. 세계적으로 빠르게 성장하고 있는 데이터산업에 비해 아직 걸음마 단계인 우리나라의 빅데이터산업을 고려할 때, 선도기술 확보를 위한 R&D, 전문인력 및 전문기업 양성은 산업활성화의 마중물 역할을 수행하는 정부가 마땅히 추진해야 할 정책으로 판단된다.⁴³⁾

데이터산업 활성화 전략은 데이터산업 활성화를 위한 선제조건, 세부 활성화 방향, 미래 대응 전략까지 포괄하며, 데이터산업 전반의 경쟁력 제고를 통한 기존 산업의 혁신성장동력 제고와 신산업 창출이라는 정책목표를 달성할 수 있을 것으로 판단된다. 데이터산업 활성화를 위한 선제조건인 개인정보보호 관련 이슈를 포괄하며, 데이터산업의 가치사슬에 대한 이해를 바탕으로 전주기적 활성화 전략을 제시하였다는 점에서 균형 잡힌 정책으로서 지금까지의 데이터 관련 정책에서 데이터 활용에만 치우쳐 개인정보보호 관련 이슈에 부딪혀 정책적 실효성이 낮았던 한계점을 극복한 것으로 보인다.

〈표 4-3〉 국내 데이터 경제 정책

년도	정책	내용	주무부처
2011	빅데이터를 활용한 스마트 정부 구현	<ul style="list-style-type: none"> • 빅데이터 활용 추진단 신설 • 사회적·기술적 주요 핵심기반 확보 	국가정보화 전략위원회
2012	스마트 국가구현을 위한 빅데이터 마스터플랜	<ul style="list-style-type: none"> • 데이터기반 정책수립 지원, 빅데이터 기반 신성장동력 창출을 목표로 4개 영역 12개 세부과제 선정 	관계부처 합동
2013	빅데이터 서비스 시범사업 추진	<ul style="list-style-type: none"> • 빅데이터 활용 스마트서비스 시범사업 6개 시행 	미래창조 과학부
2013	K-ICT 빅데이터센터 개소	<ul style="list-style-type: none"> • IoT, 빅데이터, 공공데이터(정부3.0 운영), 전자정부 사업 기획 및 추진 	미래창조 과학부
2013	빅데이터 산업 발전전략	<ul style="list-style-type: none"> • 창조경제 및 정부3.0 실현 	미래창조 과학부
2014	유능한 정부 구현을 위한 빅데이터 활용 확대방안	<ul style="list-style-type: none"> • 법제도 정비, 범정부적 지원체계 구축, 산업생태계 조성 지원 	관계부처 합동

43) 민대홍 외(2018).

년도	정책	내용	주무부처
2015	K-ICT 발전전략	• 9대 전략산업에 빅데이터 포함	미래창조과학부
2016	개인정보 비식별조치 가이드라인 발표	• 개인정보 보호 법령의 틀 내에서 빅데이터가 안전하게 활용될 수 있도록 개인정보의 비식별 조치 기준과 비식별 정보의 활용 범위 제시	관계부처합동
2016	민관 합동 빅데이터 TF	• 빅데이터 업무 추진현황을 종합적으로 파악하고, 공공과 민간 빅데이터 활성화를 지원	관계부처합동
2018	데이터산업 활성화 전략	• 데이터를 가장 안전하게 잘쓰는 나라를 목표로 개인정보통제권, 데이터산업 육성 기반 조성, 데이터 생애주기에 따른 산업생태계 구성을 추진	관계부처합동

자료: 박소영 외(2018), 관계부처합동(2018. 6), 민대홍 외(2018)

2. 데이터 경제 활성화 정책수요

여기에서는 지금까지 살펴본 데이터 경제 활성화 정책의 연장선상에서 데이터 경제 활성화를 위한 산업계 의견을 수렴한 정책수요 조사 결과를 살펴보도록 한다.

1) 조사 개요

지금까지 살펴본 데이터산업은 데이터 솔루션, 데이터 구축/컨설팅, 데이터 서비스, 데이터 인프라의 4개 대분류로 구성되어 데이터와 관련된 여러 분야의 산업들을 모두 포괄한다. 하지만 새로운 정보를 생산하고 해당 정보를 바탕으로 기존 산업의 효율화를 이루거나 신산업을 창출한다는 맥락에서 데이터 기반 경제 또는 데이터 경제를 생각해 볼 때 데이터산업은 정보제공, 데이터 분석 서비스 등을 포함하는 데이터 서비스 부문으로 좁게 바라볼 수 있다. 따라서 우리나라의 데이터 경제 활성화를 위한 정책 수요에 대해 데이터 산업 분류 중 ‘데이터 서비스’에 해당하는 전국 200개 기업을 대상으로 하였다. 조사방법으로는 구조화된 설문지를 활용한 온라인조사를 채택해 2018년 10월 1일부터 10월 17일 사이에 설문을 진행하였다.

〈표 4-4〉 표본 설계 및 조사 개요

구 분	내 용
조사 대상	<ul style="list-style-type: none"> • 전국에 소재한 데이터산업 중 한국데이터진흥원(2018)의 데이터산업현황 조사에서 “데이터서비스업체”로 분류된 2,326개 업체(데이터 거래, 정보제공, 데이터 분석제공) – 데이터 거래(데이터마켓, 데이터 신디케이션) – 정보제공(교육/훈련, 신용재무, 통계, 경영/비즈니스, 취업/창업, 뉴 스, 포탈, 행정/법률, 학술, 문화/예술, 생활) – 데이터 분석 제공(환경분석서비스, 예측분석서비스, 심사분석서비스, 기타 데이터분석서비스)
표본 크기	• 전체 200개 업체
조사 방법	• 구조화된 설문지 기반 온라인 조사
조사 기간	• 2018년 10월 1일 ~ 10월 16일
조사 내용	<ul style="list-style-type: none"> • 데이터 경제 활성화를 위한 세부 정책 필요도 • 데이터 경제 활성화를 위한 세부 정책 시급성

데이터 경제 활성화를 위한 정책을 파악하기 위해 먼저 전문가 자문과 우리나라의 데이터 정책 자료를 바탕으로 데이터 경제 활성화를 위한 총 9가지 정책 범주⁴⁴⁾를 설정하였다. 데이터 정책 범주는 <표 4-5>에 나타나 있듯이 데이터 소유권의 명확한 지정, 개인정보 보법 개정, 데이터 활용 촉진 사업, 공공부문 원시데이터 구축 및 개방, 민간·공공부문의 데이터 공유 및 거래를 위한 거래소 설립·운영, 데이터의 효율적인 저장·관리를 위한 클라우드 확산, 데이터 전문인력 양성, 빅데이터 선도 기술을 위한 R&D 지원, 데이터 전문 기업 맞춤형 지원의 9개 항목이다.

44) 본래 설문조사에서는 데이터안심존을 포함하여 10가지의 정책범주를 설정하고 설문조사를 진행하였으나, 설문결과 분석과정에서 전문가 인터뷰 결과 데이터안심존 관련 응답의 신뢰도가 낮아 본문에서는 9개 항목에 대해 다루었다. 설문조사 결과 중 데이터안심존 항목에 대한 자세한 논의는 부록에 수록하였다.

〈표 4-5〉 데이터서비스 기업 대상 정책수요 조사 항목

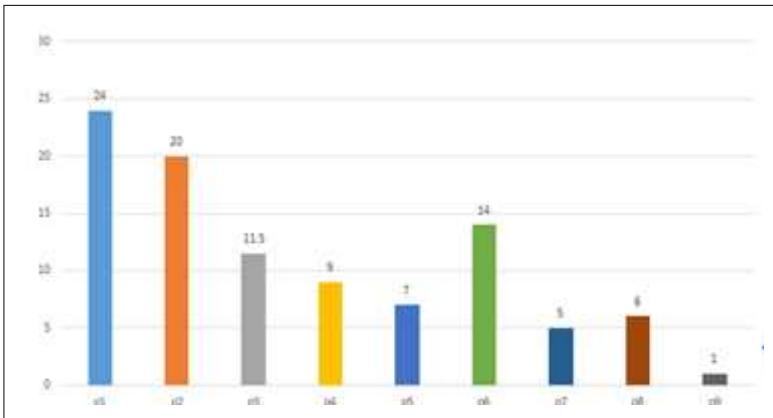
구분	정책 범주
p1	• 데이터 소유권의 명확한 지정
p2	• 개인정보보호법 개정 • (가명·익명 정보개념 정립, 개인정보의 범위 명확화 등)
p3	• 공공부문 원시데이터 구축 및 개방
p4	• 데이터의 효율적인 저장·관리를 위한 클라우드 확산
p5	• 민간·공공 부문의 데이터 공유 및 거래를 위한 거래소 설립·운영
p6	• 데이터 활용 촉진 사업 • (기존 중소기업에 빅데이터 분석 서비스 제공 기업을 매칭하고 해당 서비스를 활용할 수 있는 자금 또는 바우처 지원)
p7	• 빅데이터 선도 기술을 위한 R&D 지원
p8	• 데이터 전문인력 양성
p9	• 데이터 전문 기업 맞춤형 지원

2) 주요 조사 결과

먼저 9개 정책항목 중 데이터 경제 활성화를 위해 가장 시급한 정책은 무엇인지에 대해 설문한 결과 개인정보보호법 개정이 24%, 데이터 소유권의 명확한 지정이 20%, 데이터 활

〔그림 4-1〕 데이터 경제 활성화를 위해 시급한 정책

(단위: %)



주: n=200

용 촉진 사업이 14%로 나타났다. 그 뒤를 이어 공공부문 원시데이터 구축 개방, 클라우드 확산, 민간·공공 부문의 데이터 공유 및 거래를 위한 거래소 설립·운영, 데이터 전문인력 양성, 빅데이터 R&D 지원, 데이터 전문 기업 맞춤형 지원 정책 순으로 나타났다.

그 다음으로 위와 동일한 9개 정책 범주에 대해 데이터 경제 활성화를 위해 필요한 정책을 5점 척도로 설문하였는데, 그 결과는 다음과 같다. 데이터 경제 활성화를 위한 정책의 필요성은 개인정보보호법 개정(3.83/5), 데이터 소유권의 명확한 지정(3.77/5), 데이터 활용 촉진(3.71/5) 순으로 나타나 1순위부터 3순위까지 시급한 정책과 필요한 정책이 동일하게 나타났다. 그 뒤를 이어 공공부문 원시데이터 구축 및 개방, 빅데이터 기술 R&D, 데이터 거래소 구축 운영, 클라우드 확산, 데이터 전문인력 양성, 그리고 데이터 전문 기업 맞춤형 지원 순으로 나타나 4순위부터 10순위까지는 시급성과는 조금 다른 양상을 보였다.

〈표 4-6〉 데이터 경제 활성화를 위해 필요한 정책

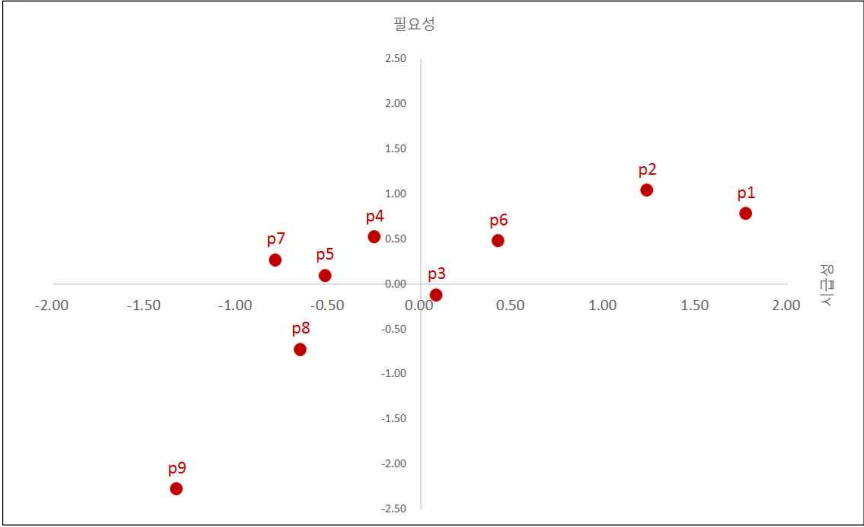
(n=200, 단위: %, 점/5점)

구분	전혀 필요하지 않다	필요하지 않다	보통	필요하다	매우 필요하다	5점 평균
p1	0.5	0.5	35.0	49.5	14.5	3.77
p2	0.5	5.5	29.0	40.5	24.5	3.83
p3	1.0	8.5	35.0	30.5	25.0	3.70
p4	0.5	7.0	42.5	36.0	14.0	3.56
p5	1.0	5.0	43.5	33.0	17.5	3.61
p6	0.5	1.5	41.0	41.0	16.0	3.71
p7	-	7.5	51.0	33.5	8.0	3.42
p8	-	6.5	43.5	29.0	21.0	3.65
p9	0.5	13.5	67.0	18.0	1.0	3.06

시급성과 필요성에 대한 설문조사 결과를 종합하여 살펴보기 위해 시급성을 가로축, 필요성을 세로축으로 설정한 뒤 분산형 그래프에 위의 정보를 나타내 보았다. 시급성은 순위의 척도로 필요성은 5점 척도로 조사되었기 때문에 용이한 비교를 위해 각 축을 평균을 바탕으로 정규화(normalized)하였다. 즉, 시급성의 경우 중위값을 갖는 정책을 기준으로 하여 정규화하였고, 필요성의 경우 모든 항목의 5점 척도 평균을 기준으로 하여 정규화하였다.

위와 같은 작업을 분산형 그래프로 나타낸 결과는 [그림 4-2]에 나타나있다. 시급성과 필요성이 모두 평균보다 높은 수준인 정책은 데이터 소유권의 명확한 지정(p1), 개인정보 보호법 개정(p2), 그리고 데이터 활용 촉진사업(p6)으로 나타났다. 그 다음으로 공공부문 원시데이터 구축 및 개방(p3)은 시급성 측면에서 평균을 약간 상회하나 필요성 측면에서는 평균이하로 나타났다.

[그림 4-2] 데이터 경제 활성화를 위한 정책수요



다음으로 데이터 경제 활성화를 위해 필요하지만 시급성은 상대적으로 낮은 정책들([그림 4-2]의 2사분면)로는 클라우드 확산(p4), 데이터 거래소 설립·운영(p5), 빅데이터 선도 기술을 위한 R&D 지원(p7)으로 조사되었다. 마지막으로 타 정책들에 비해 필요성과 시급성이 모두 낮은 정책들([그림 4-2]의 4사분면)은 데이터 전문인력 양성(p8)과 데이터 전문 기업 맞춤형 지원(p9)으로 나타났다.

제3절 데이터 경제 활성화를 위한 정책 제언

지금까지 해외 주요국의 데이터 경제 활성화 정책, 우리나라의 데이터 경제 활성화 정책 그리고 데이터산업 내 데이터 서비스 기업들의 정책 수요를 살펴보았다.

먼저 해외 주요국의 데이터 경제 활성화 정책들을 살펴보면 유럽 정책이 전 세계 데이터 경제 정책의 패러다임을 이끌어 나가고 있는 것으로 보인다. 유럽의 정책의 핵심은 개인정보보호법의 강화와 데이터 활용의 두 축으로 이루어져 있다. 미국의 경우 자칫 데이터 활용에만 무게를 둔 정책 방향으로 해석될 수 있지만, 그 세부 정책들을 살펴보면 활용 이전에 개인정보보호에 대한 안전장치를 언급하고 있는 것을 알 수 있다. 가장 공격적인 데이터 경제 활성화 정책을 추진하고 있는 중국의 경우에도 '보안'을 강조하며 개인정보보호에 대해 소홀하지 않음을 알 수 있다.

최근 우리나라의 데이터 경제 활성화 정책 역시 전 세계적인 흐름과 그 맥락을 같이 하고 있다. 최근 발표된 데이터산업 활성화 전략은 데이터산업 활성화 정책을 데이터의 생애주기 별로 제한하는 동시에 개인정보보호의 안전한 활용을 포괄하는 균형 잡힌 정책으로 평가받고 있다. 이는 데이터 활용에만 초점을 맞추어온 기존 데이터 관련 정책들의 정책적 실효성 측면에서 좋은 성과를 내지 못한 것을 타산지석 삼아 기존 정책들의 한계점을 극복하기 위한 방향성으로 해석된다. 특히 데이터 활용 촉진 정책의 애로사항이었던 개인정보보호 관련 이슈 해결을 위한 전략을 수립하여 기존 데이터 관련 정책의 한계점을 극복할 것으로 기대된다.

최근 발표된 데이터 경제 활성화 정책은 위에서 조사한 정책 수요를 고려할 때 옳은 방향성을 가지고 있는 것으로 해석된다. 정책수요 설문조사 결과, 시급성과 필요성 항목에서 모두 높은 응답을 받은 정책은 개인정보보호법 개정, 데이터의 소유권 확립, 그리고 데이터 활용 촉진 사업이다. 이들 모두 데이터산업 활성화 전략에 포함된 내용으로 먼저 2019년 내로 비식별화 조치 가이드라인에서 언급된 가명·익명 정보의 개념 정립 및 개인정보의 범위 명확화 등을 포함한 개인정보보호법 개정이 실시될 예정이다. 또한 데이터 소유권과 관련하여 Mydata 정책을 추진하여 개인정보의 안전한 활용을 추구하겠다는 전략 역시 데이터산업 활성화 전략에 나타나있다. 마지막으로 데이터 활용 촉진 사업 역시 '데이터 바우처 사업'으로 진행될 예정이다. 하지만 개인정보보호 관련 전략의 차질 없는 수행

을 위해 관계부처가 협력하여 사회적 합의를 도출하는 역할을 충실히 수행할 필요가 있을 것으로 판단된다. 유럽의 경우 1995년 EU 개인정보보호지침을 시작으로 2016년 GDPR에 이르기까지 약 11년의 시간을 가지고 개인정보보호법 개정을 추진해왔다. 정책수요의 설문조사 결과 개인정보보호법 개정에 대한 시급성과 필요성이 모두 높게 나타났지만 장기적인 관점에서 데이터 경제 활성화의 근간이 되는 개인정보보호법 개정은 긴 시간을 두고 사회적 합의를 도출하는데 먼저 힘을 쏟아야할 것이다. Mydata 사업 역시 시급성이 높게 나타나는 정책으로 해석될 수 있는데 데이터 소유권의 확립 역시 사회적 합의에 기반하지 않으면 사상누각이 될 여지가 있다. 따라서 데이터 소유권 확립과 관련한 정책을 추진함에 있어도 사회적 합의를 도출하는 것이 선제되어야 할 것이다.

또한 데이터산업 활성화는 기본적으로 기존 산업의 혁신동력을 제고하고 신산업을 창출하지만 데이터를 분석·활용에 대한 이해가 부족한 업체들도 존재할 것을 고려해야 한다. 이러한 고려가 반영된 정책으로 '데이터 바우처 사업'을 들 수 있는데, 이를 통해 데이터 분석·활용에 익숙하지 못한 업체의 데이터 경제로의 완만한 진입과 신산업 창출을 달성할 수 있도록 노력해야할 것이다.

마지막으로 데이터 경제 활성화로 인해 수익성이 악화될 가능성이 있는 산업의 완만한 퇴출을 위한 정책적 고민이 필요할 것으로 보인다. 어떠한 경제의 패러다임이 전환될 때 그로 인해 퇴출되는 산업들이 존재하게 된다. 데이터 경제로 우리나라 경제의 패러다임이 전환됨에 따라 데이터 활용을 통해 큰 부가가치를 창출할 수 없는 산업들은 그 위상이 낮아질 수밖에 없다. 따라서 이러한 산업들의 위상이 완만하게 낮아질 수 있는 정책을 고민하여 경제 전체에서의 충격을 최소화할 수 있는 방안에 대해 고민해야 할 것이다.

제5장 결론

4차 산업혁명 시대를 지나 바야흐로 데이터 경제로의 진전이 목격되고 있다. 본 연구에서는 데이터 경제를 데이터산업과 데이터산업의 파급효과로 정의하고 국내외 데이터 경제 현황에 대해 살펴보았다.

본 연구에서는 우리나라 데이터산업을 ICT 통합분류체계에 따라 데이터산업을 데이터 솔루션, 데이터 구축/컨설팅, 데이터 서비스, 데이터 인프라로 나누어 산업연관표와 연계하여 전·후방연쇄효과를 도출하였다.

먼저 후방연쇄효과 측정을 위해 수요측면의 산업연관분석을 수행한 결과, 데이터산업의 최종수요가 1원 증가(즉, 데이터산업에서 1원의 최종생산물 증가)에 따라 후방산업에서 추가적으로 유발되는 생산액은 약 0.652원으로, 추가적으로 유발되는 부가가치액은 0.251원으로 나타났다. 한편 데이터산업의 최종수요가 10억 원 증가함에 따라 후방산업에서 추가적으로 고용되는 인원은 약 4.5명으로 나타났다. 수요측면의 산업연관 분석 결과를 데이터산업의 대분류별로 살펴보면, 생산유발효과의 경우 데이터 서비스가 약 0.920원, 데이터 인프라가 0.784원으로 데이터산업 전체 생산유발효과를 상회하는 것으로 나타났으며, 데이터 구축/컨설팅, 데이터 솔루션이 각각 0.647원, 0.491원으로 그 뒤를 잇는다. 반면 부가가치 유발계수와 취업유발계수의 경우 데이터 솔루션, 데이터 서비스, 데이터 구축/컨설팅, 데이터 인프라 순으로 높게 나타났는데, 수요측면 분석에서 도출된 모든 유발효과를 고려할 때 데이터 서비스가 후방산업에 가장 지대한 경제적 파급효과를 갖는 것으로 나타났다. 한편 30대 산업들 중 데이터산업의 후방연쇄효과가 높은 산업들은 전지 및 전자기기, 사업지원 서비스, 도소매서비스, 정보통신 및 방송 산업 등으로 나타났다.

다음으로 전방연쇄효과 측정을 위해 공급측면의 산업연관분석을 수행하였다. 그 결과, 데이터산업 전체의 전방연쇄효과는 총 0.892원으로 데이터산업에서 1원의 추가수요가 발생할 때, 이를 중간재로 활용하여 타 산업에서 증가하는 최종생산물의 가치가 0.892원으로 해석할 수 있다. 이러한 전방연쇄효과 역시 데이터산업의 대분류별로 살펴보았는데, 데이터 서비스와 데이터 구축/컨설팅이 각각 1.395원 과 1.306원으로 데이터산업 전체의 전방

연쇄효과를 초과하는 수치를 보였다. 반면 데이터 인프라와 데이터 솔루션 부문의 전방연쇄효과는 각각 0.742원과 0.608원으로 데이터산업 전체의 전방연쇄효과보다 낮은 수치를 보였다. 한편 30대 산업들 중 데이터산업의 전방연쇄효과가 높은 주요 산업들은 금융 및 보험서비스, 도소매서비스, 운송서비스 등 데이터산업의 최종생산물인 데이터 - 정제되거나 새로운 정보 -를 활용하여 높은 부가가치를 창출할 수 있는 산업들이 꼽혔다.

본 연구에서 도출된 전방연쇄효과와 후방연쇄효과는 산업연관표라는 정보를 다른 관점에서 해석한 결과이다. 따라서 유럽 경우에서처럼 데이터 경제의 규모를 데이터산업의 규모와 전·후방연쇄효과의 단순 합으로 측정할 경우 경제적 파급효과에서 중복되어 측정되는 부분이 상당할 것으로 예측된다. 따라서 본 연구에서 도출된 전·후방 연쇄효과는 각각 따로 활용되어야 할 것이다.

여기에 더하여 정부정책으로 인한 경제적 파급효과 도출에 활용될 경우 후방연쇄효과에 주목해야 할 것이다. 후방연쇄효과의 경우 데이터산업의 최종수요 1원 증가(또는 데이터산업의 매출 1원 증가)에 따라 타 산업들에서 추가적으로 생산되어 데이터산업에서 중간재로 활용되는 생산물의 가치를 측정된 경제적 파급효과이다. 따라서 정부정책으로 인해 데이터산업에 투입되는 금액이 곧바로 경제적 파급효과로 환산될 수 있다. 하지만 본 연구에서 살펴본 전방연쇄효과의 경우 데이터산업의 최종수요 1원 증가(또는 데이터산업의 매출 1원 증가)에 따라 이 1원의 생산물이 전 산업에서 모두 중간재로 활용되어 생산될 생산물의 가치이다. 이 경우 데이터산업에서 1원의 추가생산이 이루진다고 할지라도 이 1원의 생산물이 반드시 모든 산업에서 중간재로 활용된다는 가정 없이는 정부정책에서 투입된 금액이 그대로 전방연쇄효과를 일으킨다고 할 수 없다. 따라서 본 연구에서 도출된 전방연쇄효과를 활용함에 있어 주의를 기울일 필요가 있다.

데이터산업의 경제적 파급효과 도출과 함께 본 연구에서는 국내외 데이터 경제 활성화 정책을 살펴보고 국내 데이터서비스 기업들을 대상으로 정책수요 설문조사를 실시하였다. 최근 국내 데이터 경제 활성화 정책인 데이터산업 활성화 전략은 전 세계적인 추세인 개인정보보호와 데이터 활용을 두 축으로 한 균형 잡힌 전략으로 평가된다. 국내 데이터기업들이 가장 시급하고 필요하다고 응답한 정책 수요는 데이터 소유권의 명확한 지정, 개인정보보호법 개정, 그리고 데이터 활용 활성화 정책으로 나타났다. 데이터 소유권의 명확한 지정과 관련하여 현재 추진 중인 Mydata 사업을 확대 추진하여 소유권에 대한 명확한

지정을 정부가 주도하여 확실하게 해야 할 것이다. 또한 개인정보보호법 개정에 대한 요구에 즉각적으로 부응하되 사회적 합의의 도출에도 역시 크게 힘을 쏟아야할 것이다. 마지막으로 데이터 활용 활성화 사업의 일환인 데이터 바우처 사업을 지속적으로 추진하지만, 데이터 분석 및 활용에 적합하지 않아 데이터 경제 진전에 따라 퇴출이 가속화될 기업들의 완만한 퇴출을 위한 정책적 고민 역시 필요할 것으로 보인다.

참 고 문 헌

[국내 문헌]

공공데이터포털(2017), “공공데이터 개방 및 활용 주요 현황”.

금융위원회(2018), “금융분야 데이터활용 및 정보보호 종합방안”.보도자료.

김사혁(2013), “빅데이터 산업 생태계 분석 동향”, 정보통신정책동향 제25권 13호.

김신곤·이석준·김정곤(2016), “빅데이터 유통 생태계에 기반한 단계별 빅데이터 유통 모델 개발에 관한 연구”, Journal of Digital Convergence, 14(5).

관계부처합동(2018. 6), “데이터 산업 활성화 전략”.

민대홍·오정숙(2018), “ICT기반 신산업 발전을 위한 데이터 거래 활성화 방안”, 기본연구 18-02, 정보통신정책연구원.

민서현 외(2017), “데이터산업의 경제적 파급효과 분석”, 한국혁신학회지, 12(1), 25-50.

박소영·장현숙(2018), “빅데이터 거래의 한·중 비교: 기업 활용을 중심으로”, 한국무역협회.

박선우(2018), “빅데이터 시대와 데이터 융합”, 정보통신방송정책 제30권 1호.

방동희(2018), “데이터 경제 활성화를 위한 데이터 법제의 필요성과 그 정립방향에 관한 소고—4차 산업혁명과 지능정보사회에서 데이터 거래의 기반 확보를 중심으로”, 법학연구 59(1).

송준혁(2017), “외생화 및 내생화가 산업연관분석에 미치는 효과: 주거용 건물 건설업을 대상으로”, 한국부동산분석학회 2017년 정기학술대회.

이영환·전희주·윤정연(2015), “데이터 산업에서 창업 활성화를 위한 데이터 거래소 제안: 금융거래소형 데이터 거래소를 중심으로”, 한국창업학회지 제 10권 2호.

정용찬(2015), 『빅데이터 산업과 데이터 브로커』. KISDI Premium Report 2015-04. 정보통신정책연구원.

최재경(2016), “빅데이터 분석의 국내외 활용 현황과 시사점”, SK경영경제연구소.

한국데이터진흥원(2017), “2017 데이터 산업 백서”.

_____ (2018), “2017년 데이터 산업 현황 조사”.

한국은행(2011), “국민계정리뷰”, 2011년 1호.
 _____ (2014a), “국민계정리뷰”, 2014년 3호.
 _____ (2014b), “산업연관분석해설”, 2014년 산업연관표 해설편.
 한국정보화진흥원(2012), “빅데이터 시대: 에코시스템을 둘러싼 시장경쟁과 전략분석”.
 _____ (2017), “2017년 BIG DATA 시장현황 조사”.
 허재용 외(2008), “IT산업의 산업파급효과 분석 - RAS 기법의 응용을 중심으로”, 산업경제
 연구 21(2), 483-500.

[국외 문헌]

Federal Trade Commission(2014). “Data Brokers: A Call for Transparency and Accountability.”
 IDC(2017a). “Worldwide Semiannual Big Data and Analytics Spending Guide.”
 ____ (2017b). “European Data Market SMART 2013/0063 Final Report.”
 ____ (2018). “Updating European Data Market Monitoring Updating Tool.”
 IMF(2018. 2). “2017 ARTICLE IV CONSULTATION—PRESS RELEASE; STAFF REPORT;
 AND STATEMENT BY THE EXECUTIVE DIRECTOR FOR THE REPUBLIC OF
 KOREA.”
 贵阳大数据交易(2016. 5). “2016年中国大数据交易产业白皮书.”
 中国电子信息产业发展研究院(2017). “2017中国大数据产业发展白皮书.”
 中国信息通信研究院(2018. 4). “大数据白皮书(2018年).”

[기타 문헌]

네이버 국어사전 웹페이지.
<https://ko.dict.naver.com/#/search?query=%EA%B2%BD%EC%A0%9C&range=all>.

[부록 1] 외생화 이전 수요측면 산업연관 분석 결과

본문에서는 외생화 기법을 사용하여 수요측면 산업연관분석을 도출하였다. 하지만 산업연관분석 시 외생화 기법을 활용하는 것이 옳은지에 대한 학계에 일치된 의견은 존재하지 않는다. 따라서 본문에서 살펴본 외생화 기법을 활용한 분석자료 외에 외생화 기법을 적용하지 않은 분석 자료를 부록에 포함하였다.

데이터산업을 내생부문에 포함하여 분석한 수요측면의 산업연관분석 결과, 데이터산업의 생산유발효과는 1.865원, 부가가치 유발효과는 0.699원, 취업유발효과는 8.706명으로 나타나 외생화 기법을 적용할 경우보다 모다 높은 수치로 도출되었다.

30대 산업명	생산유발효과		부가가치 유발효과		취업유발효과	
	값	순위	값	순위	값	순위
농림수산물	1.825	19	0.809	8	31.247	1
광산물	1.723	22	0.803	10	7.356	25
음식료품	2.324	4	0.670	20	18.019	8
섬유 및 가죽제품	2.002	13	0.564	26	10.526	17
목재 및 종이, 인쇄	2.089	9	0.636	22	11.691	14
석탄 및 석유제품	1.313	31	0.166	31	1.866	31
화학제품	2.091	8	0.506	28	6.281	26
비금속광물제품	2.128	7	0.650	21	8.493	22
1차 금속제품	2.424	1	0.447	29	4.822	29
금속제품	2.349	3	0.675	18	8.059	23
기계 및 장비	2.307	5	0.670	19	9.133	19
전기 및 전자기기	1.895	15	0.547	27	5.357	28
정밀기기	2.066	11	0.622	24	8.634	21
운송장비	2.417	2	0.633	23	7.903	24
기타 제조업 제품 및 임가공	2.033	12	0.764	14	12.973	13
전력, 가스 및 증기	1.449	27	0.403	30	2.329	30
수도, 폐기물 및 건설	1.908	14	0.769	13	9.871	18
	2.225	6	0.733	16	13.922	11

30대 산업명	생산유발효과		부가가치 유발효과		취업유발효과	
	값	순위	값	순위	값	순위
도소매서비스	1.803	20	0.838	6	20.200	5
운송서비스	1.658	24	0.568	25	14.821	10
음식점 및 숙박서비스	2.086	10	0.762	15	25.871	3
정보통신 및 방송	1.832	18	0.805	9	13.165	12
금융 및 보험 서비스	1.688	23	0.843	5	11.091	15
부동산 및 임대	1.414	29	0.926	1	6.163	27
전문, 과학 및 기술	1.658	25	0.825	7	15.579	9
사업지원서비스	1.535	26	0.878	4	28.259	2
공공행정 및 국방	1.416	28	0.901	2	11.051	16
교육서비스	1.408	30	0.890	3	18.106	7
보건 및 사회복지	1.770	21	0.784	12	19.144	6
문화 및 기타 서비스	1.885	16	0.797	11	24.469	4
데이터산업	1.865	17	0.699	17	8.706	20

데이터산업의 대분류 부문별로 생산유발효과, 부가가치 유발효과, 취업유발효과를 살펴 보면 생산유발효과는 데이터서비스 부문이, 부가가치유발효과와 취업유발효과는 데이터 솔루션이 가장 큰 계수를 갖는 것으로 나타나 외생화 기업을 적용한 뒤와 비슷한 결과가 도출되었다.

데이터산업 대분류	생산유발효과		부가가치 유발효과		취업유발효과	
	값	순위	값	순위	값	순위
데이터산업 전체	1.865	-	0.699	-	8.706	-
데이터 솔루션	1.498	4	0.829	1	14.725	1
데이터구축/ 컨설팅	1.954	2	0.756	3	9.187	3
데이터 서비스	1.960	1	0.813	2	11.041	2
데이터 인프라	1.868	3	0.587	4	6.039	4

생산유발효과의 경우 데이터서비스, 데이터구축/건설링, 데이터인프라의 생산유발효과는 모두 데이터산업 전체의 생산유발효과를 상회하며, 데이터솔루션의 생산유발효과는 데이터산업 전체의 유발효과에 크게 미치지 못하는 것으로 나타났다. 가장 높은 부가가치유발효과를 갖는 부문은 데이터솔루션으로 해당 부문의 특성 상 인건비 비중이 높은 것이 주요요인으로 추정되며, 데이터인프라 부문의 경우 장비 판매 비중이 높아 데이터산업 전체의 부가가치유발효과에 비해 상당히 낮은 부가가치유발효과를 갖는 것으로 나타났다. 마지막으로 취업유발효과의 경우에도 인건비 비중이 높은 데이터솔루션 부문이 가장 높은 취업유발효과를 갖고 있으며, 부가가치유발효과와 전체적으로 비슷한 경향이 도출되었다.

[부록 2] 품목별 · 부문별 공급액표 상에서 데이터산업의 부분별 재분류 결과

기본 부문	기초 부문	부문명칭	국산 산출액	수입액	잔폐물	총공급액	재분류
232	****a	컴퓨터 기억장치	4,546,315	1,753,483	0	6,299,798	
232	0000	컴퓨터 기억장치	1,506,591	280,737	0	1,787,328	데이터 인프라 (기타 컴퓨터 주변기기)
232	0101	HDD	2,153,119	695,704	0	2,848,823	
232	0102	광디스크 드라이브	349,023	57,916	0	406,939	
232	0200	휴대용 저장장치	371,899	100,563	0	472,462	
232	6601	자기헤드	0	409,538	0	409,538	
232	6688	기타	137,181	0	0	137,181	
232	7700	수리	0	0	0	0	
232	8801	SSD	6,595	0	0	6,595	
232	8888	기타	21,907	209,025	0	230,932	
234	****a	유선통신기기	4,014,172	1,787,822	0	5,801,994	
234	0000	유선통신기기	158,395	18,933	0	177,328	데이터 인프라 (네트워크 장비)
234	0101	일반 전화기	6,279	36,075	0	42,354	
234	0102	무선 전화기	36,014	3,403	0	39,417	
234	0103	영상 전화기	29,150	1,993	0	31,143	
234	0188	기타	188,653	95,335	0	283,988	
234	0201	기간통신사업용 교환기	78,492	3,640	0	82,132	
234	0202	사설 교환기	95,389	5,170	0	100,559	
234	0288	기타	333	2,191	0	2,524	
234	0301	반송장치 (아날로그식)	85,012	0	0	85,012	
234	0302	반송장치(디지털)	317,225	8,098	0	325,323	
234	0303	아날로그 다중화장치	0	398	0	398	
234	0304	디지털 다중화장치	60,602	127,513	0	188,115	
234	0305	신호변환기	241,969	143,109	0	385,078	

기본 부문	기초 부문	부문명칭	국산 산출액	수입액	잔폐물	총공급액	재분류
234	0306	광전송 장비(다중계기, 분배기)	504,810	737,540	0	1,242,350	
234	0401	비디오오더폰	94,214	0	0	94,214	
234	0402	인터폰	28,387	0	0	28,387	
234	0403	홈네트워크 장비	160,026	0	0	160,026	
234	6601	유선 전화기 부품	26,816	17,790	0	44,606	
234	6602	교환기 부품	25,773	4,770	0	30,543	
234	6603	반송기기 부품	28,545	162,339	0	190,884	
234	6604	광통신기기 부품	402,120	55,802	0	457,922	
234	6688	기타	135,516	129,364	0	264,880	
234	7700	수리	7,055	0	0	7,055	
234	8801	중앙통제실 송신용 침입 및 화재경보 시스템	48,754	0	0	48,754	
234	8802	팩시밀리	692,969	7,020	0	699,989	
234	8803	라우터 및 게이트웨이	46,548	0	0	46,548	
234	8804	허브 및 네트워크관련 스위치	59,193	0	0	59,193	
234	8805	키폰세트 및 시스템	369,234	4,310	0	373,544	
234	8806	음성자동응답기	8,498	34	0	8,532	
234	8807	선로집선 장치	73	0	0	73	
234	8888	기타	78,128	222,995	0	301,123	
236	****a	기타 무선통신장비 및 방송장비	8,599,478	1,072,494	0	9,671,972	
236	0000	기타 무선통신장비 및 방송장비	230,748	2,979	0	233,727	
236	0101	주파수 공용통신 단말기	22,635	0	0	22,635	
236	0102	무선 호출기	13,307	0	0	13,307	
236	0103	무전기	142,792	82,908	0	225,700	
236	0188	기타	95,849	11,108	0	106,957	
236	0201	무선통신용 교환기	360,924	11,475	0	372,399	

기본 부문	기초 부문	부문명칭	국산 산출액	수입액	잔폐물	총공급액	재분류
236	0202	무선통신 기지국용 송수신기	278,888	13,861	0	292,749	
236	0203	무선통신용 중계기(릴레이)	1,116,249	4,349	0	1,120,598	
236	0204	안테나시스템	144,161	0	0	144,161	
236	0205	RFID(전자태그)	352,610	0	0	352,610	
236	0288	기타	1,459,391	1,751	0	1,461,142	
236	0301	방송용 증폭기	89,268	82,052	0	171,320	
236	0302	방송용 음향관련 기기	81,898	13,128	0	95,026	
236	0303	방송용 영상관련 기기	71,628	148,710	0	220,338	
236	0304	지상파방송 송수신기	1,211	19,846	0	21,057	
236	0305	유선방송 송수신기	8,481	0	0	8,481	
236	0306	기타 TV송수신기	51,136	0	0	51,136	
236	0307	라디오송수신기	51,351	6,009	0	57,360	
236	0308	DMB 송수신 기기	28,341	0	0	28,341	
236	0309	기타 송신관련 기기	8,806	28,968	0	37,774	
236	0401	케이블TV 수신자용 컨버터	11,053	0	0	11,053	
236	0402	케이블TV 수신자용 기기	8,018	0	0	8,018	
236	0403	위성방송 수신용 안테나	40,521	5,305	0	45,826	
236	0404	위성방송 수신기	1,724,438	1,173	0	1,725,611	
236	6601	방송국용 기기의 부분품	193,818	60,023	0	253,841	
236	6602	SAW Filter(표면탄성파 필터)	121,193	0	0	121,193	
236	6603	무선통신 단말기의 부품	664,758	37,407	0	702,165	
236	6604	무선통신기기의 부품	971,881	325,845	0	1,297,726	

기본 부문	기초 부문	부문명칭	국산 산출액	수입액	잔폐물	총공급액	재분류
236	7700	수리	46,656	0	0	46,656	
236	8801	기타 방송국용 기기	59,626	0	0	59,626	
236	8802	기타 무선통신기기	147,842	215,597	0	363,439	
236	8888	기타	0	0	0	0	
323	****a	유선통신서비스	16,584,882	504,185	0	17,089,067	
323	0000	유선통신서비스	0	504,185	0	504,185	
323	0100	전화서비스	6,267,378	0	0	6,267,378	
323	0200	유무선 설비접속 서비스 및 인터넷 백본서비스	2,988,604	0	0	2,988,604	
323	0300	전용회선서비스	2,146,525	0	0	2,146,525	
323	0400	초고속망서비스	4,328,344	0	0	4,328,344	
323	0500	부가네트워크 서비스	804,858	0	0	804,858	데이터 인프라 (부가 네트워크 서비스)
323	0600	전신 및 전보서비스	8,433	0	0	8,433	
323	8800	기타	40,740	0	0	40,740	
325	****a	기타 전기통신서비스	3,238,901	587	0	3,239,488	
325	0000	기타 전기통신서비스	0	587	0	587	
325	0100	부가통신응용 및 중개서비스	1,962,151	0	0	1,962,151	데이터 인프라 (인터넷 관리 서비스, 기타정보 인프라 서비스)
325	0200	회선설비 재판매	486,822	0	0	486,822	
325	0300	통신서비스 모집 및 중개서비스	789,928	0	0	789,928	
325	8800	기타	0	0	0	0	
328	****a	정보서비스	7,180,468	217,924	0	7,398,392	

기본 부문	기초 부문	부문명칭	국산 산출액	수입액	잔폐물	총공급액	재분류
328	0000	정보서비스	85,972	217,924	0	303,896	데이터 서비스
328	0100	자료처리	228,840	0	0	228,840	데이터인프라 (호스팅 서비스)
328	0200	인터넷포털서비스	3,923,209	0	0	3,923,209	데이터 서비스
328	0300	뉴스제공	219,434	0	0	219,434	
328	0400	데이터베이스 및 온라인 정보제공 서비스	1,623,240	0	0	1,623,240	
328	8800	기타	1,099,773	0	0	1,099,773	
329	****a	소프트웨어 개발 공급	22,670,780	1,972,631	0	24,643,411	
329	0000	소프트웨어 개발 공급	768,340	1,618,413	0	2,386,753	
329	0101	온라인 및 모바일 게임 소프트웨어	3,288,332	0	0	3,288,332	
329	0188	기타	341,745	0	0	341,745	
329	0201	시스템소프트웨어	1,560,380	0	0	1,560,380	데이터 솔루션
329	0202	응용소프트웨어	2,151,949	354,218	0	2,506,167	
329	0301	컴퓨터프로그래밍	451,528	0	0	451,528	
329	0302	시스템통합	13,686,257	0	0	13,686,257	
329	0400	자가개발 소프트웨어	0	0	0	0	
329	8800	기타	422,249	0	0	422,249	
330	****a	컴퓨터관리서비스	8,892,494	628,993	0	9,521,487	
330	0000	컴퓨터관리서비스	765,834	628,993	0	1,394,827	데이터 구축/ 컨설팅
330	0100	유지관리서비스	3,904,731	0	0	3,904,731	
330	8800	기타	4,221,929	0	0	4,221,929	

[부록 3] 수요측면 산업연관 분석 결과를 활용한 데이터 경제 규모의 국가 경제 대비 비중

1. 데이터산업 매출액 현황 및 전망

〈데이터산업 직·간접 매출액 규모 현황 및 전망〉

	데이터 솔루션	데이터 구축/컨설팅	데이터 서비스	합계	성장률
2010년	6,725	37,407	42,242	86,374	-
2011년	8,717	43,180	43,218	95,115	10.1%
2012년	10,487	47,715	47,317	105,519	10.9%
2013년	10,789	49,985	52,258	113,032	7.1%
2014년	13,619	53,730	57,329	124,678	10.3%
2015년	14,124	55,280	64,151	133,555	7.1%
2016년	15,720	55,850	65,977	137,547	3.0%
2017년(E)	16,536	58,565	67,946	143,047	4.0%
2018년(P)	17,117	61,343	70,424	148,884	4.1%
2019년(P)	18,010	65,391	75,869	159,270	7.0%
2020년(P)	18,863	69,223	80,336	168,422	5.7%
2021년(P)	20,234	73,052	83,832	177,118	5.2%
2022년(P)	21,416	77,446	88,683	187,545	5.9%

자료: 한국데이터진흥원(2018)

〈데이터산업 직접 매출액 규모 현황 및 전망〉

	데이터 솔루션	데이터 구축/컨설팅	데이터 서비스	합계	성장률
2010년	6,725	18,487	10,706	35,918	
2011년	8,717	21,340	10,954	41,010	14.2%
2012년	10,487	23,581	11,993	46,060	12.3%
2013년	10,789	24,703	13,245	48,737	5.8%
2014년	13,619	26,554	14,530	54,703	12.2%
2015년	14,124	26,698	16,128	56,950	4.1%
2016년	15,720	27,875	16,928	60,523	6.3%
2017년(E)	16,536	29,291	17,146	62,973	4.0%
2018년(P)	17,117	30,316	17,849	65,282	3.7%
2019년(P)	18,010	32,317	19,229	69,556	6.5%
2020년(P)	18,863	34,210	20,361	73,435	5.6%
2021년(P)	20,234	36,103	21,247	77,584	5.7%
2022년(P)	21,416	38,274	22,477	82,167	5.9%

주: 1) 데이터산업 직접매출 시장은 기존 부문별 시장 규모에서(특히 데이터구축과 데이터서비스 부문) 데이터를 매개로 하는 광고 매출, DB시스템 구축 용역 매출 등 DB 관련 간접매출을 제외

2) 직접매출 시장 규모는 2015~2017년도에만 실측되었으며, 나머지 기간에는 2015~2017년도 분야별 직접매출시장규모를 같은 기간 분야별 매출액으로 나눈 비중을 적용하여 추정함

자료: 한국데이터진흥원(2018)

2. 거시경제 현황 및 전망

〈거시경제 현황〉

(단위: 조 원, 만 명)

	2010	2011	2012	2013	2014	2015	2016	2017
매출액	3,856.3	4,335.3	4,449.1	4,429.2	4,462.9	4,697.3	4,930.5	5,196.7
부가가치	1,265.3	1,332.7	1,377.5	1,429.4	1,486.1	1,564.1	1,641.8	1,730.4
취업자	2,403.3	2,452.7	2,495.5	2,529.9	2,589.7	2,617.8	2,640.9	2,672.5

주: 2015년도 이후 기간의 매출액은 각 년도 부가가치에 2014년도의 매출액 대비 부가가치율의 역수를 곱하여 산정

출처: 매출액 - 한국은행 산업연관표 각 년도(2010~2014) 산업연관표 투입산출표 기초가격 총거래표 총공급계

부가가치 - 한국은행-10.1.1 연간지표-국내총생산(명목, 원화표시)

부가가치 - 한국은행-16.6.1 경제활동인구-취업자

〈거시경제 전망〉

(단위: 조 원, 만명)

	2018	2019	2020	2021	2022
매출액	5,352.6	5,507.8	5,662.0	5,814.9	5,966.1
부가가치	1,782.3	1,834.0	1,885.3	1,936.3	1,986.6
취업자	2,701.9	2,731.6	2,761.7	2,792.0	2,822.8

주: 1) 매출액과 부가가치 성장률은 IMF(2018)의 성장률⁴⁵⁾을 적용(2018: 3.0%, 2019: 2.9%, 2020: 2.8%, 2021: 2.7%, 2022: 2.6%)

2) 취업자 증가율은 2015년부터 2017년도의 평균 취업자 증가율 1.1%를 2018-2022년까지 적용

자료: 거시경제 현황 자료에 장기 성장률 적용

45) IMF(2018. 2).

3. 데이터 경제의 국가 전체 대비 비중

〈데이터 경제 규모 및 국가 경제 전체 대비 비중〉

(단위: 조원)

구분			'10년	'11년	'12년	'13년	'14년	'15년	'16년	'17년	'18년	'19년	'20년	'21년	'22년	
국가 총 매출액(A)			3,856.3	4,335.3	4,449.1	4,429.2	4,462.9	4,697.3	4,930.5	5,196.7	5,352.6	5,507.8	5,662.0	5,814.9	5,966.1	
데이터 경제 규모 (B)	해당 산업 매출액	데이터 솔루션	0.67	0.87	1.05	1.08	1.36	1.41	1.57	1.65	1.71	1.80	1.89	2.02	2.14	
		데이터 구축/컨설팅	1.85	2.13	2.36	2.47	2.66	2.67	2.79	2.93	3.03	3.23	3.42	3.61	3.83	
		데이터 서비스	1.07	1.10	1.20	1.32	1.45	1.61	1.69	1.71	1.78	1.92	2.04	2.12	2.25	
		합계	3.59	4.10	4.61	4.87	5.47	5.70	6.05	6.30	6.53	6.96	7.34	7.76	8.22	
	타산업 매출액 유발	데이터 솔루션	0.33	0.43	0.51	0.53	0.67	0.69	0.77	0.81	0.84	0.88	0.93	0.99	1.05	
		데이터 구축/컨설팅	1.20	1.38	1.53	1.60	1.72	1.73	1.80	1.90	1.96	2.09	2.21	2.34	2.48	
		데이터 서비스	0.98	1.01	1.10	1.22	1.34	1.48	1.56	1.58	1.64	1.77	1.87	1.95	2.07	
		합계	2.51	2.82	3.14	3.35	3.72	3.91	4.13	4.29	4.45	4.75	5.01	5.29	5.60	
	자기산업+타산업			6.10	6.92	7.75	8.22	9.19	9.60	10.19	10.58	10.97	11.70	12.36	13.04	13.81
	국가 대비 데이터 경제 비중 (B/A)	해당 산업 매출액	데이터 솔루션	0.02%	0.02%	0.02%	0.02%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.04%
데이터 구축/컨설팅			0.05%	0.05%	0.05%	0.06%	0.06%	0.06%	0.06%	0.06%	0.06%	0.06%	0.06%	0.06%	0.06%	
데이터 서비스			0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.04%	0.04%	0.04%	
합계			0.09%	0.09%	0.10%	0.11%	0.12%	0.12%	0.12%	0.12%	0.12%	0.13%	0.13%	0.13%	0.14%	
타산업 매출액 유발		데이터 솔루션	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	
		데이터 구축/컨설팅	0.03%	0.03%	0.03%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	
		데이터 서비스	0.03%	0.02%	0.02%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	
		합계	0.07%	0.06%	0.07%	0.08%	0.08%	0.08%	0.08%	0.08%	0.08%	0.09%	0.09%	0.09%	0.09%	
자기산업+타산업			0.16%	0.16%	0.17%	0.19%	0.21%	0.20%	0.21%	0.20%	0.21%	0.21%	0.22%	0.22%	0.23%	

주: 데이터산업의 매출액은 직접매출 시장 규모 기준

〈데이터 경제의 부가가치 및 국가 경제 전체 대비 비중〉

(단위: 조 원)

구분		'10년	'11년	'12년	'13년	'14년	'15년	'16년	'17년	'18년	'19년	'20년	'21년	'22년	
국가 총 매출액(A)		1,265.3	1,332.7	1,377.5	1,429.4	1,486.1	1,564.1	1,641.8	1,730.4	1,782.3	1,834.0	1,885.3	1,936.3	1,986.6	
데이터 경제 부가 가치 (B)	해당 산업 부가가치	데이터 솔루션	0.41	0.53	0.63	0.65	0.82	0.85	0.95	1.00	1.03	1.09	1.14	1.22	1.29
		데이터 구축/컨설팅	0.69	0.80	0.88	0.92	0.99	1.00	1.04	1.10	1.13	1.21	1.28	1.35	1.43
		데이터 서비스	0.44	0.45	0.50	0.55	0.60	0.67	0.70	0.71	0.74	0.80	0.84	0.88	0.93
		합계	1.54	1.78	2.01	2.12	2.42	2.52	2.69	2.80	2.90	3.09	3.26	3.45	3.65
	타산업 부가가치 유발	데이터 솔루션	0.07	0.10	0.11	0.12	0.15	0.15	0.17	0.18	0.19	0.20	0.21	0.22	0.23
		데이터 구축/컨설팅	0.32	0.36	0.40	0.42	0.45	0.46	0.48	0.50	0.52	0.55	0.58	0.62	0.65
		데이터 서비스	0.38	0.38	0.42	0.47	0.51	0.57	0.59	0.60	0.63	0.68	0.72	0.75	0.79
		합계	0.76	0.84	0.94	1.00	1.11	1.18	1.24	1.28	1.33	1.42	1.50	1.58	1.68
	자기산업+타산업		2.30	2.62	2.95	3.13	3.53	3.69	3.93	4.08	4.23	4.51	4.76	5.03	5.33
	국가 대비 데이터 경제 부가 가치 비중 (B/A)	해당 산업 부가가치	데이터 솔루션	0.03%	0.04%	0.05%	0.05%	0.06%	0.05%	0.06%	0.06%	0.06%	0.06%	0.06%	0.06%
데이터 구축/컨설팅			0.05%	0.06%	0.06%	0.06%	0.07%	0.06%	0.06%	0.06%	0.06%	0.07%	0.07%	0.07%	0.07%
데이터 서비스			0.04%	0.03%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.05%	0.05%
합계			0.12%	0.13%	0.15%	0.15%	0.16%	0.16%	0.16%	0.16%	0.16%	0.17%	0.17%	0.18%	0.18%
타산업 부가가치 유발		데이터 솔루션	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
		데이터 구축/컨설팅	0.02%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
		데이터 서비스	0.03%	0.03%	0.03%	0.03%	0.03%	0.04%	0.04%	0.03%	0.04%	0.04%	0.04%	0.04%	0.04%
		합계	0.06%	0.06%	0.07%	0.07%	0.07%	0.08%	0.08%	0.07%	0.07%	0.08%	0.08%	0.08%	0.08%
자기산업+타산업		0.18%	0.20%	0.21%	0.22%	0.24%	0.24%	0.24%	0.24%	0.24%	0.25%	0.25%	0.26%	0.27%	

〈데이터 경제 취업인구 및 국가 경제 취업인구 대비 비중〉

(단위: 만 명)

구분		'10년	'11년	'12년	'13년	'14년	'15년	'16년	'17년	'18년	'19년	'20년	'21년	'22년	
국가 총 취업자수(A)		2,403.3	2,452.7	2,495.5	2,529.9	2,589.7	2,617.8	2,640.9	2,672.5	2,701.9	2,731.6	2,761.7	2,792.0	2,822.8	
데이터 경제 취업인구 (B)	해당 산업 취업 인구	데이터 솔루션	0.66	0.86	1.03	1.06	1.34	1.39	1.54	1.62	1.68	1.77	1.85	1.99	2.10
		데이터 구축/컨설팅	0.46	0.53	0.59	0.61	0.66	0.66	0.69	0.73	0.75	0.80	0.85	0.90	0.95
		데이터 서비스	0.49	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		합계	1.61	1.53	1.62	1.67	2.00	2.05	2.24	2.35	2.44	2.57	2.70	2.89	3.06
	타산업 취업 유발	데이터 솔루션	0.16	0.21	0.25	0.26	0.32	0.33	0.37	0.39	0.41	0.43	0.45	0.48	0.51
		데이터 구축/컨설팅	0.63	0.87	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		데이터 서비스	0.62	0.63	0.69	0.77	0.84	0.93	0.98	0.99	1.03	1.11	1.18	1.23	1.30
		합계	1.41	1.71	1.62	1.02	1.16	1.27	1.35	1.38	1.44	1.54	1.62	1.71	1.81
	자기산업+타산업		3.01	3.24	3.24	2.69	3.16	3.32	3.59	3.74	3.87	4.11	4.33	4.59	4.86
	국가 대비 데이터 경제 취업인구 비중 (B/A)	해당 산업 취업 인구	데이터 솔루션	0.03%	0.03%	0.04%	0.04%	0.05%	0.05%	0.06%	0.06%	0.06%	0.06%	0.07%	0.07%
데이터 구축/컨설팅			0.02%	0.02%	0.02%	0.02%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
데이터 서비스			0.02%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
합계			0.07%	0.06%	0.06%	0.07%	0.08%	0.08%	0.08%	0.09%	0.09%	0.09%	0.10%	0.10%	0.11%
타산업 취업 유발		데이터 솔루션	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.02%	0.02%
		데이터 구축/컨설팅	0.03%	0.04%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		데이터 서비스	0.03%	0.03%	0.03%	0.03%	0.03%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.05%
		합계	0.06%	0.07%	0.06%	0.04%	0.04%	0.05%	0.05%	0.05%	0.05%	0.06%	0.06%	0.06%	0.06%
자기산업+타산업		0.13%	0.13%	0.13%	0.11%	0.12%	0.13%	0.14%	0.14%	0.14%	0.15%	0.16%	0.16%	0.17%	

[부록 4] 정책수요 설문조사 분석

본래 설문조사에서는 전문가 인터뷰 먼저 전문가 자문과 우리나라의 데이터 정책 자료를 바탕으로 데이터 경제 활성화를 위한 총 10가지 정책 범주 설정하였다.

〈데이터서비스 기업 대상 정책수요 조사 항목〉

구분	정책 범주
정책1	데이터 소유권의 명확한 지정
정책2	개인정보보호법 개정 (가명·익명 정보개념 정립, 개인정보의 범위 명확화 등)
정책3	데이터 안심존 구축 (데이터 반출은 금지되나 분석·활용할 수 있는 보안환경을 갖춘 가상 또는 실제적 공간)
정책4	공공부문 원시데이터 구축 및 개방
정책5	데이터의 효율적인 저장·관리를 위한 클라우드 확산
정책6	민간·공공 부문의 데이터 공유 및 거래를 위한 거래소 설립·운영
정책7	데이터 활용 촉진 사업 (기존 중소기업에 빅데이터 분석 서비스 제공 기업을 매칭하고 해당 서비스를 활용할 수 있는 자금 또는 바우처 지원)
정책8	빅데이터 선도 기술을 위한 R&D 지원
정책9	데이터 전문 인력 양성
정책10	데이터 전문 기업 맞춤형 지원

설문조사 결과 시급성과 필요성 측면 모두에서 가장 높은 응답을 받은 정책은 데이터안심존으로 나타났다. 전체 기업들 복수응답(1, 2, 3 순위 응답)을 포함하여 데이터안심존이 시급하다고 응답한 기업의 수는 약 101개로 전체기업 들 중 약 50.5%가 시급한 정책으로 꼽았다. 또한 5점 척도로 측정된 필요성에서도 데이터 안심존의 필요성 정도의 평균은 3.93으로 나타나 가장 높은 필요도를 보인 정책으로 꼽혔다. 설문조사 결과를 해석하기에 앞서 해당 결과를 산·학·연 전문가들과 논의한 결과 데이터 안심존에 대한 정책수요조사가 잘못되었을 가능성이 제기되었다. 먼저 설문조사 결과를 공개하기에 앞서 먼저 데이터 안심존 정책을 설명한 후 해당 정책의 실효성과 기대효과에 대해 논의한 결과 산·학·연

전문가들은 공통적으로 정책적 실효성과 기대효과가 모두 낮을 것이라는 의견이 수렴되었다. 뒤이어 설문조사 결과와 설문조사 내에서 데이터 안심존에 대한 설명(데이터 반출은 금지되나 분석·활용할 수 있는 보안환경을 갖춘 가상 또는 실제적 공간)을 공개한 뒤 이어진 논의에서는 산·학·연 전문가들은 공통적으로 설문조사 응답기업들이 데이터 안심존을 마치 데이터 관련 Sandbox로 해석하고 개인정보보호법 등 규제가 없이 작업이 가능한 공간으로 인식하였을 가능성이 높다는 견해를 보이며 데이터 안심존 관련 정책수요 결과 해석에 주의를 요한다는 의견이 수렴되었다.

따라서 설문조사 분석을 실시하기 전에 데이터 안심존을 가장 시급한 정책 1순위로 선택한 38개 기업들 중 2순위 정책에 대해 응답한 기업 33개를 살펴보았다. 그 결과 꼽은 2순위 정책을 살펴본 결과 이들 중 10개 기업(약 30%)이 개인정보보호법 개정이 시급하다고 응답한 것으로 나타나 데이터 안심존에 대한 정책수요와 개인정보보호법 개정에 대한 정책 수요 간에 어느 정도 연관관계가 있는 것으로 판단하였다.

이에 전문가들의 의견을 수용하여 본문에서 다룬 설문조사 분석에서는 데이터 안심존 항목을 제외하였다. 데이터 안심존을 설문조사 분석 결과에서 제외한 과정은 다음과 같다. 먼저 시급성과 관련된 설문조사 항목은 1, 2, 3순위까지 복수응답을 허용하였는데 1순위 응답에 초점을 맞추고 1순위로 데이터 안심존을 꼽은 기업들의 2순위 응답을 1순위로 가정하였다. 예를 들어 데이터 안심존을 1순위로 개인정보보호법 개정을 2순위로 꼽은 기업의 응답은 1순위로 개인정보보호법을 꼽은 응답으로 취급하여 설문조사 응답결과를 재구성하였다. 시급성에서 데이터 안심존을 1순위로 꼽은 38개 기업들 중에서 2순위 응답을 한 기업 33개의 2순위 응답 분포는 앞서 살펴보았듯이 개인정보보호법 개정이 10개로 가장 높았고, 그 다음으로 공공부문 원시데이터 개방이 7개, 클라우드확산이 5개로 나타났다. 그 이외의 응답들은 데이터소유권 확립이 4개, 거래소 설립·운영이 3개, 데이터 활용 촉진 사업이 2개, 데이터 전문인력 양성과 데이터 전문 기업 맞춤형 지원이 모두 1개로 나타났다. 이와 같은 과정을 거쳐 재구성한 설문조사 결과를 바탕으로 다시 재구성한 결과가 본문에서 살펴본 [그림 4-1]이다. 정책 필요성과 관련된 응답은 5점 척도로 구성되어 있어 데이터 안심존 항목만을 제거한 결과를 본문에서 다뤘다. 분산형 그래프를 구성할 때 진행한 정규화 작업에서 역시 데이터 안심존 항목을 제외한 전체 평균을 활용하였다.

● 저 자 소 개 ●

민 대 홍

- 성균관대학교 경제학 학사
- University of Arizona 경제학 박사
- 현 정보통신정책연구원 부연구위원

이 학 기

- 경희대학교 경제학 학사
- University of Pittsburgh 경제학 박사
- 현 정보통신정책연구원 부연구위원

오 정 숙

- 고려대학교 영어영문학 학사
- 고려대학교 경영학 석사
- 현 정보통신정책연구원 부연구위원

방송통신정책연구 2018-0-01686

데이터 경제 진전에 따른 산업별 과급효과
분석과 정책적 활용방안 연구

(A Study on the Economic Impact of Data
Economy and its Policy Implications)

2018년 12월 일 인쇄

2018년 12월 일 발행

발행인 과학기술정보통신부 장관

발행처 과학기술정보통신부

경기도 과천시 관문로 47

Homepage: www.msip.go.kr

인 쇄 인 성 문 화
