

정책연구
2019-06

데이터 산업의 노동시장 분석

방형준 · 손연정 · 노세리

한국노동연구원

목 차

요 약	i
제1장 서 론 (방형준)	1
제1절 문제의식	1
제2절 연구의 구성	4
제2장 국내 데이터 산업 현황 (손연정)	6
제1절 데이터 산업 시장 현황 및 전망	6
1. 데이터 산업의 정의 및 세부 시장 분류	6
2. 데이터 산업 시장 현황 및 전망	7
제2절 데이터 산업 인력 현황	17
1. 데이터 산업 인력 현황	17
2. 전 산업 내 데이터직무 인력 현황	21
제3절 데이터 산업에서의 고용 동향 및 전망	28
1. 데이터 산업 필요인력 및 인력 부족률	28
2. 전 산업 내 데이터직무별 필요인력 및 인력 부족률	31
3. 국내 데이터직무 인력 수급 비교 및 현황	35
제4절 소 결	36
제3장 빅데이터 산업 (방형준)	38
제1절 빅데이터의 개념	38
1. 빅데이터의 정의	38
2. 빅데이터의 특성	41

제2절 빅데이터 현황 및 기술	44
1. 빅데이터 현황	44
2. 빅데이터 관련 기술	45
3. 빅데이터 기술의 활용	48
제3절 빅데이터 산업에서의 고용	50
1. 빅데이터 산업에서의 고용 전망	50
2. 빅데이터 산업의 주요 직종	50
제4절 데이터 현지화 조치의 함의와 노동시장 영향 분석	60
1. 데이터 현지화 조치의 개념 및 현황	60
2. 데이터 현지화 조치의 고용 영향	67
제5절 소 결	74
제4장 데이터 산업의 고용관계 이해	(노세리) 76
제1절 들어가는 말	76
제2절 설문조사 개요와 표본의 특성	78
1. 조사 방법 및 대상	78
2. 표본의 특성	79
제3절 데이터 산업 사업체 특성	81
1. 경영 체계	81
2. 하도급 거래 여부 및 기업 간 네트워크	83
3. 노사관계 및 고용안정 수준	85
제4절 데이터 산업 사업체 고용 현황	87
1. 고용 구성	87
2. 근로자 학력 수준, 근속기간 및 이직률	93
제5절 데이터 산업 사업체 고용관리	98
1. 인사관리 체계 및 방향	98
2. 인력 충원	100
3. 근로시간, 휴가제도 및 근무제도 운영	106

4. 교육훈련 및 경력개발	111
5. 보상관리	116
제6절 소 결	126
제5장 데이터 관련 기업 사례 조사	(방형준) 129
제1절 데이터 서비스 수요 기업 사례 조사	129
1. 금융권 기업 조사	129
2. 제조업 기업 조사(외국계 기업 C)	135
제2절 데이터 서비스 공급 기업 사례 조사	138
1. 빅데이터 공급 기업 조사	138
2. 일반 데이터 공급 기업 조사	141
제3절 소 결	146
제6장 결 론	(손연정·노세리·방형준) 148
참고문헌	151

표 목 차

<표 2- 1> 세계 주요국 데이터 산업 규모(2016~2017)	8
<표 2- 2> 데이터 산업 시장 규모(2011~2018)	10
<표 2- 3> 데이터 산업 시장 전망(2019~2022)	10
<표 2- 4> 데이터 산업 부문별 직접매출 시장 규모(2015~2017)	11
<표 2- 5> 데이터 솔루션 영역별 시장 규모(2010~2017)	12
<표 2- 6> 데이터 구축 및 컨설팅 시장규모(2010~2017)	13
<표 2- 7> 데이터 서비스 시장 규모(2013~2017)	14
<표 2- 8> ICT 시장 규모 중장기 전망(2016~2022)	15
<표 2- 9> 국내 소프트웨어 시장 현황 및 전망(2016~2021)	15
<표 2-10> 데이터 관련 사업 투자 비중(2015~2016)	16
<표 2-11> 데이터 산업 인력 현황(2015~2017)	17
<표 2-12> 부문별 데이터직무 인력 현황(2015~2017)	18
<표 2-13> 데이터직무별 인력 현황(2017)	19
<표 2-14> 데이터 산업 내 빅데이터 관련 인력 현황(2016~2017)	20
<표 2-15> 데이터 산업 부문별 빅데이터 관련 인력 현황 (2016~2017)	20
<표 2-16> 전 산업 내 데이터직무 인력 현황(2015~2017)	21
<표 2-17> 전 산업 내 데이터직무별 인력 현황(2016~2017)	22
<표 2-18> 전 산업 내 기술등급별 데이터직무 인력 현황(2017)	23
<표 2-19> 전 산업 내 성별 데이터직무 인력 현황(2017)	25
<표 2-20> 전 산업 내 데이터직무 중 빅데이터 관련 인력 현황 (2017)	26
<표 2-21> 전 산업 내 데이터직무 중 빅데이터 관련 인력 비중 (2017)	27
<표 2-22> 데이터 산업 내 데이터직무별 필요인력(2017)	29

<표 2-23> 데이터 산업 내 데이터직무 인력 부족률(2017)	30
<표 2-24> 데이터 산업 내 데이터직무 중 빅데이터 관련 필요인력 (2017)	30
<표 2-25> 데이터 산업 내 데이터직무 중 빅데이터 관련 인력 부족률(2017)	31
<표 2-26> 전 산업 내 데이터직무별 필요인력(2017)	32
<표 2-27> 전 산업 내 데이터직무별 인력 부족률(2017)	32
<표 2-28> 전 산업 내 데이터직무 중 빅데이터 관련 필요인력 (2017)	33
<표 2-29> 전 산업 내 빅데이터 관련 데이터 직무별 필요인력 (2017)	34
<표 2-30> 전 산업 내 빅데이터 인력 부족률(2017)	34
<표 3- 1> 미국과 한국의 빅데이터 관련 고용 및 직무 현황 비교표	57
<표 3- 2> 주요 국가의 항목별 데이터 현지화 조치	63
<표 3- 3> 한국의 산업별 취업유발계수	69
<표 4- 1> 데이터 산업 사업체 수 및 종사자 수	77
<표 4- 2> 업종별 표본 수	79
<표 4- 3> 지역별 표본 수	80
<표 4- 4> 기업 규모별 표본 수	80
<표 4- 5> 표본 사업체 업력	81
<표 4- 6> 사업체 경영 체계	82
<표 4- 7> 외국인 지분율	83
<표 4- 8> 하도급 거래	84
<표 4- 9> 원청 기업에 대한 매출 비중	84
<표 4-10> 노동조합 및 노사관계 평가	86
<표 4-11> 고용안정	86
<표 4-12> 상용직 근로자 구성	89

<표 4-13> 비정규직 근로자 구성	90
<표 4-14> 데이터 관련 직무 근로자 구성	92
<표 4-15> 데이터 관련 직무 수행 근로자 학력 수준	94
<표 4-16> 데이터 비관련 직무 수행 근로자 학력 수준	95
<표 4-17> 근로자 근속기간	97
<표 4-18> 근로자 이직률	97
<표 4-19> 인사계획 수립 단위	99
<표 4-20> 인사관리 특성	100
<표 4-21> 채용 시도	101
<표 4-22> 구인 수단	103
<표 4-23> 채용 시 애로사항	104
<표 4-24> 인턴십 제도	106
<표 4-25> 초과근로시간	107
<표 4-26> 연차휴가 소진율	108
<표 4-27> 유연근무제 시행 및 활용	110
<표 4-28> 교육훈련 프로그램	112
<표 4-29> 연평균 인당 공식 교육훈련 시간	113
<표 4-30> 교육훈련 비용	114
<표 4-31> 경력상당제도 시행	115
<표 4-32> 초임 수준	117
<표 4-33> 평균 임금	119
<표 4-34> 인사고과 방식	120
<표 4-35> 인사고과 구성 항목	121
<표 4-36> 인사점수에 따른 임금 인상 차등	122
<표 4-37> 기본급 지급 기준	123
<표 4-38> 상여금 또는 수당 항목	124
<표 4-39> 성과급 실시 현황	126

그림목차

[그림 3- 1] 미국의 연도별 데이터 생성량	44
[그림 3- 2] 빅데이터 시장 규모 전망	51
[그림 3- 3] 직군별 빅데이터 고용 규모	51
[그림 3- 4] 데이터 현지화 조치의 국가별 현황	62

요 약

1. 서 론

데이터 산업은 최근 4차 산업혁명 등과 맞물려 관심을 받고 있다. 데이터 산업은 독립된 산업으로서의 가치뿐만 아니라 기존에 존재하던 사회의 여러 분야와 접목되며 산업은 물론 사회 전반의 새로운 혁신을 낳기도 하고 있다. 그런데 지금까지 데이터 산업에 대한 대다수의 연구는 데이터 산업이 불러올 기술적 변화와 이에 따른 새로운 모습 혹은 사양되는 분야는 무엇인가에 주로 초점이 맞추어진 양상이다. 반면, 본 연구는 새롭게 등장하는 데이터 산업에서의 고용관계가 어떠한지, 데이터 산업과 접목된 기존의 제조업 및 서비스업에서는 어떤 고용관계의 변화 양상이 펼쳐지고 있는지를 살펴보고자 한다. 특히, 단순히 고용량의 변화만을 살펴보는 것이 아니라 고용의 질은 어떠한지, 노동 수요 측면에서 특정한 직무나 숙련, 학력 수준이 요구되는지 등 데이터 산업과 관련되어 기존의 행정통계로는 파악되지 않는 노동시장의 모습을 여러 각도에서 살펴보고자 한다. 이를 위해 본 연구에서는 행정통계를 사용한 데이터 산업에서의 노동 시장을 조망하고, 설문조사 등을 통해 노동의 질 측면을 살펴본 후 면접 및 방문조사를 통하여 실제 데이터 수요 및 공급 시장의 모습을 알아볼 것이다.

2. 국내 데이터 산업 현황

국내 데이터 산업 시장은 2010년 이후 연평균 약 7.5%의 비율로 꾸준히 성장하는 역동적인 분야로, 이후에도 5% 이상의 성장률을 기

록하면서 2022년 시장 규모가 18조 원까지 성장할 것으로 예상된다. 데이터 산업의 급속한 시장 확대는 인력 수요의 증가로 이어지고 있다. 2017년 기준 데이터 관련 기술 인력이 총 11만 명 정도로 추산되나 아직도 1만 3천여 명 정도의 기술 인력이 추가로 필요한 것으로 조사되어 현재 고용 규모의 10% 이상에 대한 추가 수요가 존재하고 있다.

하지만 데이터 산업의 고용 인력 다수가 데이터 축적, 저장 및 데이터베이스 구축 등 기초적인 분야에 주로 몰려 있는 반면, 앞으로 각광받을 데이터 활용 및 분석 관련 전문 인력은 여전히 시장 규모도 작고 인력 공급도 적은 것으로 나타났다. 특히나 데이터 관련 인력에 대한 기업 수요의 많은 부분이 실무 경험이나 활용 능력을 갖춘 인력에 주로 몰려 있는 반면 이러한 숙련이나 경력 및 능력을 갖춘 인력은 여전히 부족하여 노동시장에서 미스매치가 큰 것으로 조사되었다.

3. 빅데이터 산업

빅데이터 산업은 데이터 산업 내에서도 최근 가장 주목받고 있는 분야이다. 빅데이터는 고용량, 고속의 이동성, 그리고 다양성을 갖춘 데이터로 정의할 수 있는데, 기존에 없던 새로운 분야이기에 빅데이터 관련 조직이나 기업에서는 새로운 직무와 직종이 등장하고 있다. 대체적으로 빅데이터 관련 직종에서는 연봉 수준 및 학력 수준이 기존의 데이터 산업 종사자보다 높으며, 노동 수요 역시 폭발적이라 할 정도로 빠르게 증가하고 있다. 하지만 한국은 미국과 비교하면 빅데이터 산업의 고도화 수준이 높지 않아서 앞의 ‘2. 국내 데이터 산업 현황’ 요약 부분에서도 확인했듯이 실제 데이터를 분석하고 활용하는 직군이 낮다는 문제점도 안고 있다.

한편 빅데이터분야는 그 특성상 국경 간에 많은 데이터 이동이 발생할 가능성을 필연적으로 내포하고 있다. 이에 대해서 각국은 개인

정보 및 정보에 대한 국가의 통제권 강화 등을 목적으로 데이터 현지화 조치를 취하고 있다. 이러한 데이터 현지화 조치는 데이터 산업 후발 주자들에게 자국의 데이터 산업을 보호하고 육성할 수 있는 기회와 명분을 주기도 하는데, 이로 인해 자국 데이터 산업이 커짐에 따라 해당 분야의 직접 고용이 늘어날 수 있다. 하지만 이러한 비관 세 무역 장벽 및 보호주의 정책은 평균적인 데이터 서비스의 질을 낮추고 가격을 인상하여 데이터 서비스를 활용하여 나타날 수 있는 여러 간접 고용 효과를 낮추는 역할도 할 수 있다. 본 장에서는 이러한 직접 고용 효과와 간접 고용 효과를 비교하였는데, 이에 따르면 직접 고용에서의 증가 효과가 간접 고용에서의 감소 효과보다 크지 않을 가능성이 높다는 점도 확인하였다.

4. 데이터 산업의 고용관계 이해

데이터 산업에서의 노동시장을 자세히 들여다보기 위해 실시한 설문조사 결과 데이터 산업에 종사하는 사업체 다수는 소유 경영 체제로 운영되며 외국인이 지분을 가진 경우가 거의 없는 영세한 업체들이 다수를 이루는 것으로 파악되었다. 한편 성별로는 남성 근로자가 다수를 차지하고 연령대로는 20~30대의 젊은 층이 많은 것으로 조사되었다. 하지만 기업 규모가 영세하기 때문에 이직이 잦고 근로자들의 평균 근속 연수도 매우 낮은 편이라는 점도 확인되었다.

한편 기업 규모가 작기 때문에 인력을 채용 공고를 통해서 충원하기보다는 상시 채용 및 지인을 통한 소개 등으로 충당하는 경우가 많았으며, 채용 과정에서 기업들이 겪는 가장 큰 어려움은 적합한 능력이나 경력을 가진 사람을 찾지 못하는 점을 꼽았다.

하지만 여타 산업과 비교하면 비교적 젊은 층으로 구성되어 있으며 새로운 기술을 다루기 때문에 1인당 교육 및 훈련 시간과 비용은 높은 편으로 나타났다. 한편 초임은 기업 규모에 따라 다르지만 대략 300만 원 정도로 낮지 않은 편이며, 직무급제의 도입 비율이 높은 것

으로 조사되었다. 또한 다수의 기업에서 보상과 평가를 연동하는 형태의 보수 체계를 가지고 있다고 응답하였다.

종합해보자면, 데이터 산업의 다수 기업들은 데이터 산업체로서 독특한 특성을 일부 보이기도 하지만 여타 산업 분야의 영세한 업체들과 유사한 특징도 공유하고 있다. 따라서 새롭게 떠오르는 데이터 산업을 육성하기 위해서는 정부가 영세한 기업들의 고용관계에 대한 정책적 지원을 할 필요가 있으며, 아울러 높은 이직률을 낮추고 근로자들의 평균적인 숙련 수준을 높이기 위해서 영세업체들이 고용을 유지할 수 있는 지원 방안에도 대해서도 논의해 봐야 할 것이다.

5. 데이터 관련 기업 사례 조사

본 장에서는 데이터 서비스 수요 기업과 공급 기업에 대해서 조사하였는데, 데이터 서비스 수요 기업은 금융권 서비스업과 제조업으로 나누었으며, 공급 기업의 경우 중소 빅데이터 기업 한 개소, 중견 데이터 기업 한 개소와 중소 데이터 기업 한 개소를 조사하였다.

데이터 서비스 기업들의 경우, 컴퓨터 프로그래머 등 직접적인 분석 및 빅데이터 기술을 가진 인력보다는 상경계열 및 통계직군 등 응용 분야에 집중하여 수요가 발생하고 있다. 하지만 이러한 응용 능력을 갖춘 경력직 및 신입들의 숫자는 많지 않기 때문에 대기업에서는 주로 사내 직업훈련 프로그램 등을 통하여 기존 인력의 전직 등으로 부족한 인력을 충당하나 이러한 프로그램을 운영하기 어려운 소규모 영세 사업체들에서는 인력 부족이 상대적으로 큰 문제로 작용하고 있다. 또한 데이터 서비스 공급 기업들 다수가 앞서 4장에서 살펴본 것처럼 영세하기 때문에 프로그래머들의 이직률도 높은 편이며, 이들의 연봉 수준도 응용 개발자들보다 낮기 때문에 장기적으로 빅데이터 및 데이터 분석 관련 기초 소양을 갖춘 인력의 공급 부족 문제가 야기될 수 있어 이에 대한 대책이 필요하다 하겠다.

아울러 사내 재교육 프로그램 등으로 새로운 분석 및 데이터 관리

기술을 익히고자 하는 인력들이 많은 것이 확인되었으나, 중소·중견 기업에서는 해당 프로그램을 운영하기 힘든 여건이다. 하지만 이에 대한 대비책으로 정부가 교육훈련 프로그램을 운영하기에는 교육 내용 수요 측면에서의 미스매치 및 장시간 근로로 인한 교육 시간 확보 부족 문제 등이 있기 때문에 시장의 수요에 맞는 프로그램이 운영될 수 있는 방안을 강구할 필요가 있다.

6. 결론

본 연구는 기술적인 측면에서 접근한 데이터 산업에 대한 여타의 연구와는 달리 데이터 산업의 노동시장에 대해서 다각도로 살펴보았다. 데이터 산업은 그 특성상 사회의 여러 기존 분야와 접목되기 쉽기 때문에 데이터 산업 자체에서의 인력 수요뿐만 아니라 다양한 기존 산업에서의 데이터 관련 인력 수요 역시 크게 증가하고 있으나, 시장에서 요구하는 학력 및 경력을 갖춘 인력이 원활하게 공급되고 있지는 않다. 특히나 일반 기업체에서는 빅데이터 관련 분석 능력을 갖춘 응용 분석가에 대한 수요가 많이 있으나 시장에서는 이를 모두 충족시키지는 못하는 것으로 조사되었다. 다만 이러한 응용 분야에 대한 인력 수요가 많고 해당 분야의 보수와 처우가 좋다는 점은 장기적으로 기초 기술을 학습하는 인력의 공급이 줄어들 가능성이 있어 해당 분야에 대한 육성책이 필요하다 하겠다.

아울러 빅데이터와 관련하여, 빅데이터의 핵심은 많은 양의 데이터를 수집하고 구축하는 것에 더해 이를 효과적이고 빠르게 분석하는 것도 포함되나, 한국은 이러한 빅데이터 관련 분석 기술이 선진국에 비해 낮은 것으로 조사되어 이미 구축된 데이터를 효과적으로 활용하지 못할 가능성이 있다. 따라서 빅데이터 관련 기술의 고도화를 추구하여 단순한 데이터의 수집 및 분석에만 머무르지 않고 고급 기술을 갖춘 인재를 육성할 필요가 있을 것이다.

한편 기존에 데이터 관련 직무를 수행하지 않은 근로자들도 기회

가 있다면 빅데이터 혹은 데이터 관련 기술을 학습하고 싶은 수요가 많을 가능성이 대기업의 사례 조사를 통해 확인되었다. 하지만 중소기업 및 중견 기업의 경우 이러한 교육훈련 프로그램을 운영하기가 어렵다. 이로 인해 영세한 업체들의 경우 핵심 인력의 유지 및 기존 근로자들의 기술 적용도 향상 등에 어려움을 겪고 있어 영세한 업체들에 대한 다양한 지원 방안을 강구할 필요가 있다 하겠다.

제1장 서론

제1절 문제의식

본 연구는 최근 발전하고 있는 데이터 산업에서의 노동시장을 분석하는 것이 목적이다. 데이터 산업은 특히나 빅데이터를 중심으로 최근 떠오르고 있으며, 각종 사물인터넷(IoT) 관련 제품 및 기술의 확산과 5G의 보급 등으로 많은 데이터를 축적, 보관, 전송, 활용하는 것이 용이해지면서 더욱 각광을 받고 있다.

그러나 이러한 세간의 주목하는 바 가운데 다수는 새롭게 등장하는 데이터 기술과 빅데이터 기법에 초점이 맞추어져 있을 뿐 이로 인해 발생하는 여타의 다양한 사회경제적 문제들로부터는 약간 빗겨나 있다. 그러나 과거 증기기관의 등장이 단순한 기술의 확산과 새로운 산업의 등장에만 그친 것이 아니라 사회 전반의 경제 구조, 환경, 노동시장, 국제 무역 등에 영향을 미친 것을 감안하면, 최근 새롭게 등장하는 데이터 산업의 실체와 미래를 조망하는 것 못지않게 데이터 산업의 부상(浮想)으로 인해 발생할 다양한 사회적 변화의 모습 역시 살펴볼 필요가 있다 하겠다.

이와 관련되어 등장하는 하나의 주장은 데이터 산업의 확산과 기술의 발전이 장기적으로 고용을 줄일 것이라는 우려이다. 이는 작업 현장의 자동화 및 무인화와 맞물려 앞으로 고용이 줄어들 것이라는 공포감으로 인

2 데이터 산업의 노동시장 분석

식되고 있기도 하다. 하지만 산업혁명 시기에 등장한 다양한 기술과 기계가 궁극적으로 고용을 증대시켰으며, 산업혁명에 맞추어 사회경제적 대응을 펼친 국가와 그렇지 못하고 변화에 적응하지 못한 국가의 명암이 뚜렷했기 때문에 반드시 데이터 산업의 확장과 기술 발전에 대해서 걱정하고 막으려 할 필요는 없을 것이다. 오히려 이럴 때일수록 데이터 산업의 현재와 미래를 살펴봄과 아울러 그로 인해 초래될 다양한 사회적 경제적 변화를 조망해 보는 것이 필요할 것이다.

이를 위해 본 연구에서는 데이터 산업의 노동시장을 분석할 것이다. 하지만 데이터 산업 그 자체는 데이터를 분석하거나, 저장 및 보관을 위한 각종 설비 및 시스템을 보급하는 데 한정되지만, 빅데이터 산업에서도 데이터 관련 서비스를 수요하거나, 다양한 형태로 자체적인 생산과 소비를 하고 있는 실정이다. 따라서 본고에서는 데이터 산업 자체에서의 인력 현황을 살펴보는 것에만 그치지 않고, 데이터 서비스를 생산하고 수요하는 모든 산업을 함께 살펴보려 할 것이다. 이를 위해서 실제 데이터 서비스를 수요하는 데이터 및 제조 업체에 대해서도 조사를 진행할 것이다.

아울러 본 연구에서는 고용에 대한 양적 분석뿐만 아니라 질적인 측면에서의 변화도 함께 살펴볼 것이다. 데이터 산업의 확산과 데이터 서비스의 보급에 따라 새롭게 등장하는 직업이나 직무도 있을 것이고, 이들에 대해서 기존과는 다른 형태의 새로운 고용관계나 인력 관리 방식이 나타날 수도 있다. 또한 데이터 기업들에서의 총고용만이 아니라 고용의 질도 아울러 살펴보고자 한다. 특히나 고용의 질은 행정통계를 통해서 쉽게 조사되거나 각종 자료를 통해 바로 파악되는 것이 아니므로 이를 위해서는 별도의 심층적인 설문조사나 인터뷰를 수행할 필요가 있다.

최근 데이터 산업이 사회적으로 주목을 받는 가장 큰 이유는 아마도 빅데이터 때문일 것이다. 빅데이터는 용어를 사용하는 주체마다 서로 다른 정의를 내리고 있을 뿐만 아니라 이를 사용하는 많은 사람들도 해당 용어의 정의나 범위에 대해서 명확하게 인지하지 못하고 쓰기도 한다. 문제는 빅데이터에 대한 용어의 혼선이 빅데이터 관련 정책을 수립하거나 향후 인재를 양성하는 데 있어 큰 방해가 될 수 있으므로 이에 대해서 일반 대중도 이해할 수 있는 개념에 대한 정의가 필요하다는 사실이다. 본

연구에서는 이러한 시도를 해 볼 것이나 전문용어를 완벽히 배제하고 난해한 기술을 쉽게 소개하기란 여전히 쉽지 않다.

한편 인터넷 속도가 빨라짐에 따라 국경 간 데이터 이동 역시 중요한 문제로 부각되고 있다. 이에 대해서 어떤 국가는 데이터의 자유로운 이동을 더욱 허용하려 하기도 하며, 일부 국가는 데이터의 이동에 대한 국가의 통제를 강화하려 하기도 한다. 이러한 국경 간 데이터 이동 문제가 개인 정보 보호 및 국내총생산의 증감에 미치는 영향에 대해서는 이미 많은 분석이 있었다. 하지만 데이터의 이동을 방해하는 비관세 장벽이 노동시장에 미치는 영향에 대해서 다룬 연구는 많지 않다. 일반적으로 관측하거나 측정하기 어려운 비관세 장벽의 고용 효과를 살펴본 연구가 적은 실정이나, 국경 간 데이터 이동을 규제하는 데이터 현지화 관련 규정은 여타 비관세 장벽과 비교하여 상대적으로 규정하고 측정하기 쉬운 편이다. 따라서 이러한 데이터 현지화 조치의 고용 효과를 간단하게 살펴보고, 해당 정책이 단순히 국내총생산에 미치는 영향에만 주목할 것이 아니라 노동시장에 미치는 영향도 분석하여 정책 수립 및 집행 시 참고하는 것이 바람직할 것이다.

제2절 연구의 구성

본 연구는 데이터 서비스를 공급하는 데이터 산업과 수요하는 연관 산업들에서 노동시장의 여러 특성을 살펴보고자 한다. 데이터 산업 자체에서는 고용량의 변화, 고용의 질에서의 변화 및 현재의 인력 현황을 중심으로 살펴보고, 빅데이터와 관련된 이슈에 대해서도 알아볼 것이다. 아울러 데이터 서비스를 수요하는 기업들이 직면한 노동시장의 현실과 그에 따른 문제를 살펴봄으로써 이후 데이터 서비스가 더욱 확산됨에 따라 우리 사회에서 예상되는 문제와 그에 대한 대응책도 아울러 모색해 볼 필요가 있겠다.

우선 2장에서는 데이터 산업의 고용 현황을 주로 살펴볼 것이다. 현재의 데이터 산업 자체의 규모 및 현황과 전문가들의 전망을 살펴본 후, 이에 기초하여 데이터 산업의 인력 현황과 고용 동향 및 전망을 조사함으로써 이후 전개될 장에서 활용할 기초 자료로 활용할 것이다.

제3장에서는 국내 데이터 산업의 현황에 대해서 실태조사를 벌이고 그 결과를 소개할 것이다. 특히 고용의 질이나 고용관계, 인력 관리 방식 등은 노동시장에서 무척 중요함에도 통계상으로 관찰하기 힘들어서 자칫 간과하기 쉬운 이슈이다. 하지만 고용량이 늘어남에도 고용의 질이 악화된다면 이를 바람직하다 할 수 없기 때문에 데이터 산업에서 고용의 질이 어떠한지 실태조사를 통해 살펴보는 것은 고용량에 대한 분석만큼이나 의미있다 할 것이다. 아울러 새롭게 나타나는 고용관계나 인력 관리 방식이 있는지 살펴봄으로써 이후 데이터 산업 관련 정책을 수립하거나 제고할 때 참고할 수 있는 자료가 될 것이다.

제4장에서는 데이터 산업 중 최근 주목을 받고 있는 빅데이터 산업만을 별도로 떼어내서 살펴볼 것이다. 특히 빅데이터의 경우 다양한 개념이 존재하며, 이에 따라 사람들이 그리는 해당 기술의 형태나 미래의 활용 가능성에 대한 그림도 제각각인데, 본 장에서는 빅데이터의 개념과 현황 및 여러 관련 기술을 언급하여 빅데이터에 대한 이해를 높이고자 한다.

아울러, 전송 기술의 발전은 빅데이터를 가능케 하는 한 요소임과 동시에 데이터의 자유로운 이동에 따른 여러 문제점 발생의 원인이기도 한데, 후자의 일부가 개인정보 보호 및 국경 간 데이터 이동에서 나타나고 있다. 본 장에서는 각국이 빅데이터 구축에 따라 쉽게 유출될 수 있는 개인 정보 보호를 위해서 어떠한 정책을 펼치고 있는지 살펴보고, 아울러 데이터의 국경 간 이동을 규제하기 위한 데이터 현지화 조치는 무엇이고 각국은 어떤 입장을 취하고 있는지 알아보려고 한다. 이를 통해 데이터 현지화 조치가 노동시장에 어떠한 영향을 미칠지도 아울러 분석하여 데이터 현지화 조치를 고려할 때 발생할 수 있는 문제의 다양한 측면도 살펴보고자 한다.

5장에서는 데이터 서비스 관련 수요 업체와 공급 업체들을 방문하거나 인터뷰한 결과를 정리하여 현재 기업들이 체감하고 있는 데이터 서비스에 대한 노동시장의 모습을 그려보고자 한다. 특히, 시장에서는 어떠한 데이터 관련 인재를 수요하며, 현재 어떠한 인재들이 공급되고 있는지, 또 이들은 어떻게 관리되고 있는지에 이르기까지 종합적으로 살펴볼 것이다. 그리고 인력 수급 불균형이 발생하는 부분은 어디인지 살펴보고 이를 해결하기 위한 방안도 모색해 볼 수 있는 기회를 가질 것이다.

마지막으로 6장에서는 상술한 논의들을 종합하여 데이터 산업 노동시장의 모습을 종합적으로 살펴보고, 향후의 양상을 전망해보고자 한다. 본 연구의 마무리는 데이터 산업 및 서비스와 관련된 다양한 모습을 살펴봄으로써, 이후 데이터 관련 노동시장의 문제에 대응하고 보다 양질의 일자리를 창출할 수 있는 방안은 무엇인지 고민하는 내용으로 채워질 것이다.

제 2 장

국내 데이터 산업 현황

제1절 데이터 산업 시장 현황 및 전망

1. 데이터 산업의 정의 및 세부 시장 분류

데이터는 기업과 국가의 경쟁력의 원천으로서 다양한 데이터 구축과 활용은 산업 전반의 패러다임을 변화시킬 수 있는 원동력이 되고 있다. 이른바 4차 산업혁명이 진전됨에 따라 IoT 및 센서 등에서 생산되는 빅데이터가 데이터 기반 산업과 경제 활성화를 위한 핵심 자원으로 부상하고 있으며, 산업의 발전과 새로운 가치 창출에 데이터가 필수가 되는 ‘데이터 경제(Data Economy)¹⁾’라는 새로운 패러다임이 구축되고 있는 것이다. 이에 유럽과 미국 등 세계 주요국을 중심으로 데이터 경제가 경제성장과 일자리의 동인으로 새로운 시대의 경쟁우위를 주도할 것으로 인식되면서 범정부 차원에서 주요 데이터를 확보하고, 산업·사회에 데이터 활용을 확대하는 한편, 데이터 분석을 담당할 인재를 양성하고 안전하게 데이터를 활용할 수 있도록 제도를 정비하는 등 다양한 정책이 마련되고 있

1) 2011년 David Newman의 Gartner 보고서 “How to Plan, Participate and Prosper in the Data Economy”에서 처음 등장한 개념으로, 유럽 집행위원회가 2014년부터 디지털 싱글 마켓 전략을 채택하면서 그 일환으로 ‘Data Economy’/‘Data-driven Economy’ 개념을 도입하면서 집중 조명되었다.

는 상황이다.

본 장에서는 데이터 산업 관련 각종 데이터와 자료를 이용하여 국내 데이터 산업의 전반적인 시장 및 인력 현황을 살펴보고자 한다. 본문에 주로 인용된 자료는 과학기술정보통신부와 한국데이터진흥원의 「2017년 데이터 산업 현황조사」 결과보고서의 내용으로, 데이터와 관련된 상품과 서비스를 생산 및 제공하는 전국의 1인 이상 사업체를 대상으로 한 조사를 바탕으로 한다.

본 장에서는 「2017년 데이터 산업 현황조사」에서 사용된 정의에 따라 데이터 산업은 ‘데이터의 수집·생산 및 처리·분석, 데이터 활용·유통 등을 통해 새로운 가치를 창출하는 상품 또는 서비스를 생산하고 제공하는 산업’으로 정의하였다. 이는 데이터로부터 가치가 창출되는 일련의 과정들, 즉, 데이터의 생명주기상에 나타난 데이터 생산, 수집, 처리, 분석, 유통, 활용 등의 활동들을 모두 포함한다. 이러한 정의에 따라 데이터 산업의 세부 유형은 크게 관련 제품을 판매하거나 데이터 관련 기술을 제공하는 데이터 솔루션, 데이터 구축 및 데이터 컨설팅 분야와 데이터와 관련된 서비스를 제공하는 데이터 서비스 분야로 구분된다(서동혁, 2018: 47). 본 장에서는 국내 데이터 산업 전반에 대한 현황 및 전망은 물론 이와 같이 구분된 데이터 산업 각각의 세부 산업 시장 현황과 인력 현황을 살펴봄으로써 현재 국내 데이터 산업의 위치를 알아보하고자 한다.

2. 데이터 산업 시장 현황 및 전망

가. 데이터 산업 시장 규모(전체 시장)

IDC(International Data Corporation)와 The Lisbon Council에서 2018년 내놓은 “The European Data Market Monitoring Tool Report”에 따르면, 미국과 EU, 일본, 브라질 등 세계 여러 국가의 데이터 산업은 지속적으로 성장하고 있다. 이 보고서에서는 데이터 시장을 ‘데이터를 가공하여 제품 및 서비스로 재생산하는 디지털 데이터 시장’으로 정의하고 데이터 관련 기업의 매출액을 기반으로 데이터 시장 규모를 산출하였다. 미국의

8 데이터 산업의 노동시장 분석

데이터 산업은 전 세계에서 가장 규모가 크고 성장률 또한 가장 빠른 것으로 알려져 있는데, <표 2-1>에 제시된 바와 같이 미국은 지난 2017년 약 1,455억 유로의 시장 규모를 보였으며 이는 전년 대비 12.7% 증가한 수치이다. EU의 데이터 시장 역시 2016년과 2017년 사이 9% 전후의 빠른 성장세를 기록하며 2017년에 약 650억 유로에 이른 것을 확인할 수 있다. 일본은 ICT 산업에 대한 높은 투자에 힘입어 데이터 산업 역시 안정적인 성장세를 보이며 2017년 전년 대비 8.7% 성장하여 약 277억 유로로 도약한 것으로 나타났다. 라틴아메리카 국가 중 ICT 산업이 가장 발달한 브라질의 경우 데이터 산업이 빠른 속도로 성장하여 프랑스, 이탈리아 등 유럽의 주요 국가와도 경쟁할 수준으로 도약하고 있으며 2017년 시장 규모는 약 63억 유로 수준으로 나타났다.

아울러 <표 2-1>에서는 데이터 기업의 수를 국가별로 비교하여 제시하고 있다. 데이터 기업은 시장에서 데이터를 공급하는 기업으로, 데이터와 관련된 제품과 서비스, 기술을 공급하는 기업을 의미하며, 데이터 시장의 성장과 함께 데이터 기업의 수도 크게 성장해가고 있는 것을 볼 수 있다. 미국과 EU의 데이터 기업 수는 2017년에 각각 약 30만 개와 27만 개를 기록하였다. 일본의 데이터 기업 수는 2017년에 10만 개를 넘어선 것으로 집계되었으며 브라질은 3만 6천 개로 증가했다.

<표 2-1> 세계 주요국 데이터 산업 규모(2016~2017)

(단위: 백만 유로, 개, 천 명)

구분	미국		EU		일본		브라질	
	2016	2017	2016	2017	2016	2017	2016	2017
데이터 산업 규모	129,173	145,546	59,498	65,038	25,513	27,723	6,049	6,310
증감률(%)	12.67		9.30		8.66		4.31	
데이터 기업 수	289,556	302,810	261,450	276,450	101,612	104,664	35,979	36,387
증감률(%)	4.58		9.20		3.00		1.13	
데이터 산업 종사자 수	12,732	14,012	6,891	7,290	3,740	4,040	1,160	1,176
증감률(%)	10.05		8.00		8.03		1.37	

자료: IDC and Lisbon Council(2018) 재구성.

데이터 산업의 종사자 수 역시 데이터 기업의 수와 마찬가지로 데이터 시장의 성장과 함께 꾸준히 증가하고 있는 모습이다. 미국의 데이터 산업 종사자 수는 2017년 전년 대비 약 10% 증가하여 1천 4백만 명에 육박하는 것으로 추정되었으며, EU에서는 8% 증가하여 약 730만 명 규모로 나타났다. 일본의 데이터 산업 종사자 수 역시 4백만 명으로 전년 대비 8% 이상 증가하였으며, 브라질은 상대적으로 낮은 1.4%의 증가율을 보이며 2017년 약 118만 명으로 집계되었다.

국내 데이터 산업의 규모 역시 세계 시장과 마찬가지로 빠른 속도로 확대되고 있는데, 국내의 데이터 시장 규모는 2017년 6조 2,973억 원에 달했으며 2020년까지 7조 8,450억 원까지 성장할 것으로 전망되고 있어 연평균 7.6%의 성장률을 나타내고 있다.

<표 2-2>는 2011~2018년 기간 동안 국내 데이터 산업 시장 규모를 시장 전체 및 세부 유형에 따라 제시하고 있다. 이는 데이터와 관련한 간접매출까지 포괄하는 광의의 데이터 산업 시장 규모로 볼 수 있는데, 데이터 관련 비즈니스를 영위하는 사업체들에 있어 데이터와 직간접적으로 관련된 매출을 포괄하여 추정한 결과이다. 2018년 국내 데이터 산업의 시장 규모는 총 14조 8,884억 원으로 2017년 대비 4.0% 성장한 것으로 나타났다. 2011~2018년 연평균 증가율은 7.5%로 조사되었으며 매년 꾸준한 성장세를 유지하고 있는 것으로 보인다.

부문별로는 데이터 솔루션 시장이 2018년 기준으로 가장 큰 1조 7,117억 원 규모를 나타냈으며, 데이터 구축·컨설팅이 6조 1,343억 원, 데이터 서비스 시장이 7조 424억 원으로 나타났다. 2017년 대비 2018년 시장 규모 증가율은 데이터 솔루션 시장이 5.2%로 가장 크게 나타났으며, 2011~2018년 연평균 성장률 역시 데이터 솔루션 시장이 13.7%로 두드러진 성장세를 보였다.

데이터 산업은 <표 2-3>에 제시된 바와 같이, 2022년까지 연평균 5% 이상의 성장세를 이어나갈 것으로 보이며 2022년에는 시장 규모가 18조 원 대로 증가할 것으로 전망된다. 이는 4차 산업혁명의 진전으로 데이터의 부가가치가 더욱 커지고 있는 추세를 반영하고 있는 것으로 파악된다.

10 데이터 산업의 노동시장 분석

〈표 2-2〉 데이터 산업 시장 규모(2011~2018)

(단위: 억 원)

구분	데이터 솔루션	데이터 구축·컨설팅	데이터 서비스	전체
2011	8,717	43,180	43,218	95,115
2012	10,487	47,715	47,317	105,519
2013	10,789	49,985	52,258	113,032
2014	13,619	53,730	57,329	124,678
2015	14,124	55,280	64,151	133,555
2016	15,720	55,850	65,977	137,547
2017	16,536	58,565	67,946	143,047
2018	17,117	61,343	70,424	148,884
증감률(%)('17-'18)	5.2	4.9	3.0	4.0
연평균 증감률(%)	13.7	6.6	7.0	7.5

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-3〉 데이터 산업 시장 전망(2019~2022)

(단위: 억 원)

구분	데이터 솔루션	데이터 구축·컨설팅	데이터 서비스	전체
2019	18,010	65,391	75,869	159,270
2020	18,863	69,223	80,336	168,422
2021	20,234	73,052	83,832	177,118
2022	21,416	77,446	88,683	187,545
연평균 증가율(%)	5.3	5.6	5.1	5.3

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

데이터 관련 매출에는 데이터와 관련된 상품 및 서비스의 판매로부터 발생하는 직접매출 외에도 데이터를 매개로 발생하는 각종 간접매출이 있으며, 광고매출 등이 이에 포함된다. 이상에서 살펴본 데이터 산업 시장 규모 현황은 데이터 관련 간접매출까지 모두 포함하여 데이터 산업 시장 규모를 추정한 것이며, 데이터 직접매출을 중심으로 한 데이터 산업 시장

〈표 2-4〉 데이터 산업 부문별 직접매출 시장 규모(2015~2017)

(단위: 억 원)

구분	데이터 솔루션	데이터 구축·컨설팅	데이터 서비스	전체
2015	14,124	26,698	16,128	56,950
2016	15,720	27,875	16,928	60,523
2017	16,536	29,291	17,146	62,973
증감률(%)('16-'17)	5.2	5.1	1.3	4.0

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

규모를 살펴보면, 2017년 기준 전체 데이터 산업 시장 규모에서 직접매출이 차지하는 비중이 약 44%였으며, 부문별로는 데이터 솔루션의 직접매출이 전체 시장 규모의 100%, 데이터 구축·컨설팅은 50%, 데이터 서비스는 25%의 비중으로 나타났다.

2017년 데이터 산업 부문별 직접매출 시장의 규모를 부문별로 살펴보면, <표 2-4>에 제시된 바와 같이 데이터 솔루션 시장은 약 1조 7천억 원, 데이터 구축·컨설팅 시장은 2조 9천억 원, 데이터 서비스 시장은 1조 7천억 원으로 나타났다. 앞서 살펴본 광의의 데이터 산업 시장 규모와 비교 하였을 때, 간접매출은 대체로 데이터 구축·컨설팅 분야와 데이터 서비스 분야에서 발생하고 있음을 확인할 수 있다. 직접매출이 전체 매출에서 차지하는 비중은 데이터 구축·컨설팅은 약 50%, 데이터 서비스는 25% 수준인 것으로 나타났다.

나. 데이터 산업 시장 규모(세부 산업별)

과학기술정보통신부와 한국데이터진흥원의 「2017년 데이터 산업 현황조사」에서는 데이터 산업을 크게 데이터 솔루션, 데이터 구축·컨설팅, 데이터 서비스의 총 3개 대분류로 나누고 있으며 이는 다시 12개의 중분류와 47개의 소분류로 나누어진다. 데이터 산업 실태조사에서는 중분류 단위로 조사가 실시되고 있다.

3개의 데이터 산업 비즈니스 유형 중 먼저 데이터 솔루션은 DBMS

12 데이터 산업의 노동시장 분석

(DataBase Management System), DBMS 관리, 데이터 모델링, 데이터 분석·시각화, 검색엔진, 품질 등과 관련한 솔루션 제품으로 비즈니스를 영위하는 기업을 포함하며, 일반적으로 라이선스, 유지·보수, 커스터마이징(개발)을 통해 수익을 창출한다.

<표 2-5>에 제시된 바와 같이, 2017년 데이터 솔루션 시장 규모는 1조 6,536억 원으로 2016년 대비 5.2% 성장하였으며, 2010~2017년 간 연평균 성장률은 13.7%로 시장 규모가 급격히 성장해온 것을 확인할 수 있다.

데이터 솔루션 영역은 데이터 산업분류 중분류에 따라 다시 DBMS, 데이터 수집, 설계, 관리, 분석, 품질, 데이터 플랫폼 등 7개 영역으로 구분되는데, 2017년을 기준으로 이 중 DBMS가 37.5%로 가장 큰 비중을 차지하는 것으로 나타났으며, 데이터 관리와 데이터 플랫폼이 각각 26.7%, 11.0%로 그 뒤를 잇고 있다. 데이터 솔루션 영역에서는 데이터 관리가 연평균 19.9%로 가장 급속히 성장하고 있으며 데이터 수집도 17.5%의 성장률로 데이터 솔루션 시장의 발전에 상대적으로 큰 기여를 하고 있는 것을 확인할 수 있다.

<표 2-5> 데이터 솔루션 영역별 시장 규모(2010~2017)

(단위:억 원)

구분	DBMS	데이터 수집	데이터 설계	데이터 관리	데이터 품질	데이터 분석	데이터 플랫폼	전체
2010	3,169	448	129	1,237	686	1,026	-	6,725
2011	4,272	570	136	1,516	949	1,274	-	8,717
2012	5,301	617	151	2,075	942	1,401	-	10,487
2013	5,488	622	166	2,489	855	1,169	-	10,789
2014	5,502	1,076	202	3,446	883	1,121	1,389	13,619
2015	5,727	1,115	207	3,574	918	1,157	1,426	14,124
2016	6,148	1,345	225	4,022	1,120	1,249	1,611	15,720
2017	6,192	1,382	269	4,417	1,127	1,328	1,821	16,536
증감률(%)('16-'17)	0.7	2.8	19.6	9.8	0.6	6.3	13.0	5.2
연평균 증가율(%)	10.0	17.5	11.1	19.9	7.3	3.3	6.4	13.7

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

데이터 구축·컨설팅 시장의 규모(표 2-6 참조)는 2017년에 5조 8,565억 원으로 2016년 대비 4.9% 증가하였다. 2010년 이후 데이터 구축·컨설팅 시장은 연평균 6.6%의 성장률을 보이고 있다. 데이터 구축 시장은 DB 설계 및 구축, 데이터 이행 및 처리, 기계 처리형 데이터 구축 분야를 포함하고 있으며, 데이터 컨설팅 시장은 데이터의 설계, 데이터 품질 및 관리 성능 개선 컨설팅, 데이터 활용 컨설팅, 데이터 거버넌스 체계 컨설팅 등의 영역을 포함하고 있다. 데이터 구축 시장은 2017년 시장 규모가 5조 6,893억 원으로 2016년 대비 4.3% 증가하였으며, 2010년 이후 연평균 6.5%의 성장률을 보이고 있다. 반면, 데이터 컨설팅 시장은 2017년 시장 규모가 1,672억 원으로 2016년에 비해 30% 이상 증가하였으며, 2010년 이후 연평균 11.2%의 높은 성장률을 보이고 있으나 전체 데이터 산업에서 차지하는 비중은 여전히 매우 낮은 것으로 보인다.

다음으로 <표 2-7>에 제시된 바와 같이 2017년 데이터 서비스 시장은 6조 7,946억 원 규모로, 연평균 6.8% 성장하고 있는 것을 확인할 수 있다. 데이터 서비스 시장은 크게 데이터를 판매 또는 중개하는 데이터 거래 서비스, 데이터나 DB를 이용해 정보를 제공하는 정보 제공 서비스, 데이

<표 2-6> 데이터 구축 및 컨설팅 시장 규모(2010~2017)

(단위:억 원)

구분	데이터 구축	데이터 컨설팅	전체
2010	36,610	797	37,409
2011	42,374	806	43,180
2012	46,865	850	47,715
2013	49,021	964	49,985
2014	54,142	1,138	53,730
2015	54,142	1,138	55,280
2016	54,571	1,279	55,850
2017	56,893	1,672	58,656
증감률(%)('16-'17)	4.3	30.7	4.9
연평균 증가율(%)	6.5	11.2	6.6

자료 : 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

14 데이터 산업의 노동시장 분석

터를 분석함으로써 새로운 가치를 창출하여 제공하는 데이터 분석 제공 서비스로 구분된다. 이 가운데 정보 제공 서비스가 데이터 서비스 시장의 대략 90%를 차지하고 있으며 연평균 6.3%로 지속적인 성장을 하고 있는 것으로 나타났다.

<표 2-8>은 전반적인 ICT 산업의 중장기 시장 규모 전망을 제시하고 있다. 국내 ICT 산업은 최근 세계적으로 경기가 회복세인 것과 더불어 이른바 4차 산업혁명과 맞물려 수요가 증가하고 있으며, 특히 데이터 산업과 관련이 깊은 소프트웨어 시장이 빅데이터, 클라우드, ICT 융합 관련 신산업에 대한 수요는 물론 공공부문의 패키지 소프트웨어 수요가 증가함에 따라 비교적 높은 성장률을 보이고 있다. 2018년부터 2022년까지 소프트웨어 산업 생산은 연평균 5%의 성장을 보여 2022년에 약 68조 원 규모를 형성할 것으로 전망된다.

<표 2-9>는 국내 소프트웨어 시장 현황 및 전망을 제시하고 있는데, IT 서비스의 경우 대형 IT 서비스 업체를 중심으로 빅데이터, 클라우드 컴퓨팅, IoT, AI 플랫폼 시장 확대로 2018~2022년까지 연평균 4.2% 성장이 전망된다. 특히 IT 서비스 생산은 2017년 9.0%, 2018년 5.8%의 높은 성장세를 보였는데, 이는 주요 대형 IT 서비스 업체들이 관계사를 중심으로 아웃소싱 서비스 및 산업 특화 솔루션 제공, 스마트 팩토리, 클라우드 등을

<표 2-7> 데이터 서비스 시장 규모(2013~2017)

(단위:억 원)

구분	데이터 거래	정보 제공	데이터 분석 제공	전체
2013	1,708	48,284	2,266	52,258
2014	2,019	51,985	3,325	57,329
2015	2,379	58,171	3,601	64,151
2016	2,594	59,854	3,529	65,977
2017	2,709	61,575	3,662	67,946
증감률(%)('16-'17)	4.4	2.9	3.8	3.0
연평균 증가율(%)	12.2	6.3	12.7	6.8

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

통해 고성장한 것에 기인한 것으로 보인다. 이에 따라 최근 IT 서비스 업체들은 클라우드 컴퓨팅 서비스 등에 대한 투자를 강화하고 관련 기업들과의 전략적 제휴 등을 통해 서비스를 확대하고 있는 추세다.

〈표 2-8〉 ICT 시장 규모 증장기 전망(2016~2022)

(단위:조 원)

구분	2016	2017	2018	2019	2020	2021	2022	CAGR ('18~'22)
합 계	428.2	463.2	478.3	487.6	497.1	504.1	513.9	1.8
정보통신방송기기	306.3	336.8	346.6	350.9	355.2	356.8	361.2	1.0
전자부품	175.4	208.4	214.8	215.7	216.2	214.4	215.7	0.1
컴퓨터·주변기기	8.2	9.3	10.7	11.7	12.7	13.6	14.5	8.0
통신·방송기기	54.9	47.4	47.5	47.0	47.0	47.2	47.6	0.1
영상·음향기기	11.6	10.8	10.8	10.7	11.0	11.1	10.5	-0.6
정보통신응용기반기기	56.2	61.0	63.0	65.8	68.3	70.5	72.9	3.7
정보통신방송서비스	71.8	73.5	75.6	77.6	79.8	82.2	84.7	2.9
통신서비스	37.3	37.4	37.5	37.7	38.0	38.4	38.9	1.0
방송서비스	17.2	17.9	18.7	19.4	20.0	20.5	21.1	3.0
정보서비스	17.3	18.2	19.4	20.6	21.9	23.2	24.7	6.2
소프트웨어	50.2	52.8	56.1	59.1	62.1	65.1	68.1	5.0

자료: 정혁 외(2018), 「ICT 산업 증장기 전망(2018~2022)」, 재구성.

〈표 2-9〉 국내 소프트웨어 시장 현황 및 전망(2016~2021)

(단위:조 원)

	2016	2017	2018	2019	2020	2021	2022	CAGR ('18~'22)
패키지SW	9.3	8.6	9.2	9.8	10.5	11.1	11.7	6.0%
게임SW	10.2	10.8	11.6	12.3	13.1	13.9	14.7	6.2%
IT서비스	30.6	33.3	35.3	36.9	38.6	40.1	41.7	4.2%
합 계	50.2	52.8	56.1	59.1	62.1	65.1	68.1	5.0%

자료: 정혁 외(2018), 「ICT 산업 증장기 전망(2018~2022)」, 재구성.

다. 데이터 산업의 투자 전망

데이터 산업은 기업 전체 매출액 대비 데이터 관련 사업에 대한 투자 비중이 2016년과 2017년에 각각 20.4%와 20.5%로 나타났다. 분야별로 특히 투자 비중이 높은 분야는 데이터 솔루션으로 2016년과 2017년에 각각 23.5%와 23.2%의 투자율을 보였다. 데이터 서비스 기업은 타 분야에 비해 투자 비중이 비교적 낮은 17%대로 나타났다.

〈표 2-10〉 데이터 관련 사업 투자 비중(2015~2016)

(단위: %, 개)

구분	2015	2016	2017	
	투자 비중	투자 비중	응답 수	투자 비중
전 체	21.3	20.4	1,004	20.5
데이터 솔루션	25.4	23.5	274	23.2
데이터 구축	22.3	22.1	262	22.3
데이터 컨설팅	23.1	17.4	86	18.7
데이터 서비스	14.6	17.2	382	17.4

주: 1) 데이터 솔루션, 구축, 컨설팅의 응답 수는 중복이 허용된 것임.

2) 투자 비중은 매출액 대비 투자 비중을 의미함.

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

제2절 데이터 산업 인력 현황

1. 데이터 산업 인력 현황

데이터 산업의 범위가 전통적인 데이터베이스 산업에서 빅데이터, 클라우드 및 사물인터넷 등 최근 각광받고 있는 신기술 시장 영역까지 확장되어 감에 따라, 데이터 전문 인력의 직무도 기존 데이터의 저장과 자료 축적 중심에서 데이터 분석 및 활용을 중심으로 변화하고 있다. 한국데이터진흥원에서 발간한 『2018 데이터 산업 백서』에서는 데이터직무를 ‘기업에서 필요로 하는 데이터 관련 기획, 분석, 개발, 운영 및 관리에 관한 직무를 수행하는 인력’으로 정의하고 있으며, DA(Data Architect), 데이터 개발자, 데이터 엔지니어, 데이터 분석가, DBA(DataBase Administrator), 데이터 사이언티스트, 데이터 컨설턴트, 데이터 기획·마케터 등 8개의 직무가 이에 포함된다.

2017년 국내 데이터 산업에 종사하고 있는 인력의 규모는 총 29만 4,753명으로 이 가운데 약 26%는 데이터직무 인력이 차지하고 있다. 국내 데이터 산업 인력은 2016년 대비 2017년에 2.1%가 증가하였으며, 특히 데이터직무 인력은 5.3%로 데이터직무 외 인력 증감률(1.1% 증가)에 비해 크게 늘어난 것으로 나타났다(표 2-11 참조).

〈표 2-11〉 데이터 산업 인력 현황(2015~2017)

(단위: 명, %)

구분	2015		2016		2017		증감률 (’16-’17)
	규모	비중	규모	비중	규모	비중	
전 체	280,323	100.0	288,621	100.0	294,753	100.0	2.1
데이터직무	70,338	25.1	73,256	25.4	77,105	26.2	5.3
데이터직무 외	209,985	74.9	215,365	74.6	217,648	73.8	1.1

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

데이터 산업의 부문별 인력 현황을 살펴보면, <표 2-12>에 제시된 바와 같이 데이터 솔루션 분야의 인력이 전체 데이터 산업 인력에서 차지하는 비중은 2017년 기준 13.3%로 가장 작지만 전년 대비 11.0%의 높은 성장률을 보이고 있다. 데이터 구축·컨설팅 및 데이터 서비스의 경우 각각 전체 인력 규모에서 차지하는 비중이 48.7%와 38.0%로 높게 나타났으며 데이터 구축·컨설팅 인력은 전년 대비 6.0%, 데이터 서비스의 경우 2.5% 증가했다.

<표 2-13>은 데이터 산업에 종사하는 인력을 데이터직무별로 구분하여 인력 현황을 제시하고 있다. 2017년 데이터 산업에서 데이터직무 인력은 총 7만 7,105명이었으며, 이 중 데이터 개발자의 비중이 전체의 38.0%로 가장 컸고, 그다음으로는 데이터 엔지니어가 17.0%, DBA가 15.4%로 비중이 높게 나타났다.

<표 2-14>는 데이터직무 중에서도 특히 빅데이터 관련 인력 현황을 보여주고 있는데, 여기서 빅데이터 관련 인력은 빅데이터 관련 기술²⁾을

<표 2-12> 부문별 데이터직무 인력 현황(2015~2017)

(단위: 명, %)

구분		2015		2016		2017		증감률 ('16-'17)
		규모	비중	규모	비중	규모	비중	
전체	소계	70,338	100.0	73,256	100.0	77,105	100.0	5.3
데이터 솔루션	소계	8,886	12.6	9,272	12.7	10,291	13.3	11.0
데이터 구축·컨설팅	소계	34,323	48.8	35,404	48.4	37,516	48.7	6.0
	구축	29,304	41.7	29,941	40.9	31,693	41.1	5.9
	컨설팅	5,019	7.1	5,463	7.5	5,823	7.6	6.6
데이터 서비스	소계	27,129	38.6	28,580	39.0	29,298	38.0	2.5

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

2) 데이터 산업 현황 조사에서 제시하고 있는 빅데이터 관련 기술에 대한 예시는 다음과 같음: 1) Hadoop, MapReduce, NoSQL, Spark 등 빅데이터 처리 또는 프로그램 개발을 위한 관련 기술 보유, 2) 두 개 이상의 데이터 포맷을 활용하는 빅데이터 분석 기술 보유, 3) 실시간으로 스트리밍되는 데이터의 분석 기술 보유, 4) 빅데이터 관련 제품 및 서비스에 대한 기획 능력 보유, 5) 빅데이터 분석을 통한 문제 해결, 의사결정 도출 및 제시 능력, 6) 기타 빅데이터 분석 및 처리, 개발 등에 필요한 관련 기술 보유 등이다.

보유한 인력을 의미한다. 2017년 데이터 산업 내 빅데이터 관련 인력은 총 6,324명으로 전체 데이터직무 인력의 8.2%를 차지하는 것으로 나타났다. <표 2-15>의 부문별 빅데이터 관련 인력의 분포를 살펴보면, 빅데이터 인력이 가장 많이 종사하고 있는 분야는 데이터 솔루션인 것으로 나타났다. 직무별 빅데이터 관련 인력 분포를 보면, 전체적으로 데이터 개발자가 가장 큰 비중을 차지하고 있으며 데이터 컨설턴트가 그 뒤를 잇고 있다.

<표 2-13> 데이터직무별 인력 현황(2017)

(단위: 명, %)

구분	항목	전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스
전체	규모	77,105	10,291	31,693	5,823	29,298
	구성비	100.0	100.0	100.0	100.0	100.0
DA	규모	5,025	618	3,392	526	489
	구성비	6.5	6.0	10.7	9.0	1.7
데이터 개발자	규모	29,267	4,429	13,702	2,188	8,948
	구성비	38.0	43.0	43.2	37.6	30.5
데이터 엔지니어	규모	13,127	1,468	6,529	696	4,434
	구성비	17.0	14.3	20.6	11.9	15.1
데이터 분석가	규모	4,792	629	1,638	335	2,190
	구성비	6.2	6.1	5.2	5.8	7.5
DBA	규모	11,851	702	3,732	388	7,029
	구성비	15.4	6.8	11.8	6.7	24.0
데이터 사이언티스트	규모	1,283	218	410	173	482
	구성비	1.7	2.1	1.3	3.0	1.7
데이터 컨설턴트	규모	4,263	1,530	1,376	1,231	126
	구성비	5.5	14.9	4.3	21.1	0.4
데이터 기획/마케터	규모	7,497	697	914	286	5,600
	구성비	9.7	6.8	2.9	4.9	19.1

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-14〉 데이터 산업 내 빅데이터 관련 인력 현황(2016~2017)

(단위: 명, %)

구분	2016			2017		
	빅데이터 관련 인력		데이터직무 인력 중 빅데이터 인력	빅데이터 관련 인력		데이터직무 인력 중 빅데이터 인력
	인력 규모	구성비	비중	인력 규모	구성비	비중
전 체	5,806	100.0	7.9	6,324	100.0	8.2
DA	-	-	-	-	-	-
데이터 개발자	1,790	29.8	6.3	1,866	29.5	6.4
데이터 엔지니어	1,006	17.3	7.9	1,081	17.1	8.2
데이터 분석가	341	5.9	7.6	381	6.0	8.0
DBA	-	-	-	-	-	-
데이터 사이언티스트	1,185	20.4	100.0	1,283	20.3	100.0
데이터 컨설턴트	1,274	21.9	31.6	1,369	21.7	32.1
데이터 기획/마케터	270	4.7	3.7	344	5.4	4.6

주: 1) 빅데이터 관련 인력: 데이터 산업에 종사하는 인력 중 빅데이터 관련 기술을 보유한 인력을 의미함.

2) DA와 DBA는 조사에서 제외되었으며, 데이터 사이언티스트는 100% 빅데이터 관련 인력으로 간주함.

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-15〉 데이터 산업 부문별 빅데이터 관련 인력 현황(2016~2017)

(단위: 명)

구분	2016					2017				
	전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스
전 체	5,806	2,314	1,465	1,260	767	6,324	2,515	1,600	1,354	855
데이터 개발자	1,730	800	404	332	194	1,866	896	409	350	211
데이터 엔지니어	1,006	787	115	49	55	1,081	829	137	57	58
데이터 분석가	341	28	121	187	5	381	37	134	196	14
데이터 사이언티스트	1,185	203	388	150	444	1,283	218	410	173	482
데이터 컨설턴트	1,274	264	437	531	42	1,369	271	484	562	52
데이터 기획/마케터	270	232	-	11	27	344	264	26	16	38

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

2. 전 산업 내 데이터직무 인력 현황

데이터는 데이터 산업을 초월하여 일반 기업에서도 새로운 자본으로 그 중요성이 커지고 있어, 데이터를 잘 생산해 축적하고 적재적소에 활용하는 것이 기업의 미래 가치를 평가하는 중요 지표가 되고 있으며 기업 경쟁력의 핵심으로 작용하고 있다. 최근에는 ICT 기업뿐 아니라 전통 기업들도 데이터를 활용해 제조 설비와 과정을 최적화하고 서비스를 효율화해 비용 절감과 기업 경쟁력 확보에 이용하고 있다. 따라서 일반 기업에서도 데이터직무 인력에 대한 수요가 점차 늘어날 것으로 전망된다.

지금까지 본 장에서는 데이터 산업에 국한하여 데이터직무 인력 현황을 살펴보았는데, 일반 산업을 포함한 전 산업으로 범위를 넓혀 데이터직무 인력 규모를 추산하면 <표 2-16>에 제시된 것과 같이 2017년에 총 10만 9,320명에 달하는 것으로 조사되었다. 그중 데이터 산업 내 데이터직무 인력은 전년 대비 5.3% 증가한 7만 7,105명으로 나타났으며, 일반 산업 내 데이터직무 인력은 전년 대비 10.6% 증가한 3만 2,215명으로 나타났다.

<표 2-16> 전 산업 내 데이터직무 인력 현황(2015~2017)

(단위: 명, %)

구분		전 산업	데이터 산업	일반 산업
2015	규모	100,440	70,338	30,102
	비중	100.0	70.0	30.0
2016	규모	102,375	73,256	29,119
	비중	100.0	71.6	28.4
2017	규모	109,320	77,105	32,215
	비중	100.0	70.5	29.5
증감률('16-'17)		6.8	5.3	10.6

주: 데이터 산업 현황 조사에서 일반 산업은 금융업, 제조업, 유통·서비스업, 공공, 통신·미디어, 의료 등 주요 6대 산업을 포함하고 있음.

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

<표 2-17>에서 보듯이 데이터직무별로는 데이터 개발자의 비중이 데이터 산업과 일반 산업 모두에서 가장 높게 나타났으며, 데이터 산업과 달리 일반 산업의 경우 DA 비중이 15.6%로 상대적으로 높게 나타났다. 이러한 결과는 기업의 인력 구성이 현재까지는 데이터를 활용하는 인력 보다는 자사에서 필요한 형태로 데이터를 가공하기 위한 프로그래밍 및 엔지니어링, 데이터 관리 부문에 더 집중되어 왔음을 의미한다.

<표 2-17> 전 산업 내 데이터직무별 인력 현황(2016~2017)

(단위: 명, %)

구분	항목	2016			2017			전 산업 전년 대비 증감률 ('16-'17)
		전 산업	데이터 산업	일반 산업	전 산업	데이터 산업	일반 산업	
전 체	규모	102,375	73,256	19,119	109,320	77,105	32,215	6.8
	비중	100.0	100.0	100.0	100.0	100.0	100.0	
DA	규모	9,267	4,736	4,531	10,071	5,025	5,046	8.7
	비중	9.1	6.5	15.6	9.2	6.5	15.7	
데이터 개발자	규모	38,948	27,379	11,569	41,254	29,267	11,987	5.9
	비중	38.0	37.4	39.7	37.7	38.0	37.2	
데이터 엔지니어	규모	15,670	12,664	3,006	16,634	13,127	3,507	6.2
	비중	15.3	17.3	10.3	15.2	17.0	10.9	
데이터 분석가	규모	7,339	4,477	2,862	8,398	4,792	3,606	14.4
	비중	7.2	6.1	9.8	7.7	6.2	11.2	
DBA	규모	17,116	11,488	5,628	17,863	11,851	6,012	4.4
	비중	16.7	15.7	19.3	16.3	15.4	18.6	
데이터 사이언티스트	규모	1,662	1,185	477	1,803	1,283	520	8.5
	비중	1.6	1.6	1.6	1.7	1.7	1.6	
데이터 컨설턴트	규모	4,513	4,028	485	5,004	4,263	741	10.9
	비중	4.4	5.5	1.7	4.6	5.5	2.3	
데이터 기획/마케터	규모	7,860	7,299	561	8,293	7,497	796	5.5
	비중	7.7	10.0	1.9	7.6	9.7	2.5	

자료 : 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

하지만 2016년 대비 2017년에 데이터 분석가의 증가율은 14.4%로 여타 직무에 비해 가장 크게 늘었고, 데이터 컨설턴트와 DA, 데이터 사이언티스트 역시 8~11%의 높은 증가율을 보이고 있다. 이러한 변화는 기업에서 데이터 활용에 특화된 전문 인력에 대한 수요가 점차 늘어나고 있음을 반영한다고 볼 수 있을 것이다.

<표 2-18>은 전 산업 내 데이터직무 인력 현황을 기술등급별로 구분하여 제시하고 있다. 데이터 기술등급은 크게 초급, 중급, 고급으로 구분되며, 초급은 정보처리기사, 정보처리산업기사 및 동급의 데이터 관련 자격 취득자, 전문학사 및 학사 학위 보유자 또는 고등학교 졸업 후 3년 이상의 관련 경력 보유자를 의미한다. 중급은 기사 자격이나 데이터 관련 자격 취득 후 경력이 3년 이상인 경우, 산업기사 자격 취득 후 경력이 7년 이상인 경우, 석사 학위 보유자의 경우 기사 자격 또는 데이터 관련 자격 취득 후 경력이 2년 이상인 경우를 의미한다. 마지막으로 고급은 컴퓨터 시스템응용기술사 및 정보관리기술사, 중급 이후 3년 이상의 경력 보유자, 박사 학위 보유자의 경우 기사 자격이나 데이터 관련 자격 취득자를 포함한다.

기술등급별 인력 현황을 구체적으로 살펴보면, 대체로 중급과 고급 인력이 많은 것으로 나타났으며, 특히 데이터 컨설턴트, DBA, 데이터 분석가 직무에서는 고급 인력의 비중이 높게 나타났다.

<표 2-18> 전 산업 내 기술등급별 데이터직무 인력 현황(2017)

(단위: 명)

구분	기술 등급별	2017					
		전 산업	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	일반 산업
전 체	계	109,320	10,291	31,693	5,823	29,298	32,215
	초급	20,734	2,247	7,654	1,084	3,608	6,141
	중급	46,769	4,087	12,335	2,344	13,159	14,844
	고급	41,817	3,957	11,704	2,395	12,531	11,230
DA	계	10,071	618	3,392	526	489	5,046
	초급	1,621	61	541	79	70	870

〈표 2-18〉의 계속

구분	기술 등급별	2017					
		전 산업	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	일반 산업
	중급	4,674	242	1,731	116	177	2,408
	고급	3,776	315	1,120	331	242	1,768
데이터 개발자	계	41,254	4,429	13,702	2,188	8,948	11,987
	초급	9,619	1,377	4,101	513	1,251	2,377
	중급	16,950	1,637	5,152	907	3,877	5,377
	고급	14,685	1,415	4,449	768	3,820	4,233
데이터 엔지니어	계	16,634	1,468	6,529	696	4,434	3,507
	초급	3,561	314	1,736	128	620	763
	중급	7,340	645	2,479	343	2,100	1,773
	고급	5,733	509	2,314	225	1,714	971
데이터 분석가	계	8,398	629	1,638	335	2,190	3,606
	초급	1,228	86	266	62	243	571
	중급	3,560	240	560	120	1,062	1,578
	고급	3,610	303	812	153	885	1,457
DBA	계	17,863	702	3,732	388	7,029	6,012
	초급	2,573	117	685	38	486	1,247
	중급	7,380	309	1,327	172	2,793	2,779
	고급	7,910	276	1,720	178	3,750	1,986
데이터 사이언티스트	계	1,803	218	410	173	482	520
	초급	228	19	41	63	34	71
	중급	839	138	92	2	333	274
	고급	736	61	277	108	115	175
데이터 컨설턴트	계	5,004	1,530	1,376	1,231	126	741
	초급	556	165	136	160	20	75
	중급	2,013	575	554	524	58	302
	고급	2,435	790	686	547	48	364
데이터 기획/마케터	계	8,293	697	914	286	5,600	796
	초급	1,348	108	148	41	884	167
	중급	4,013	301	440	160	2,759	353
	고급	2,932	288	326	85	1,957	276

자료 : 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

전 산업 내 데이터직무에 종사하는 남녀 인력 현황을 각각 살펴보면, <표 2-19>에 제시된 바와 같이 데이터직무에서는 여전히 남성 인력이 대부분을 차지하고 있는 것을 확인할 수 있다. 하지만 데이터 분석 및 데이터 기획/마케터 분야의 경우 타 분야에 비해 여성 인력의 비중이 상대적으로 높은 것으로 나타났다.

<표 2-19> 전 산업 내 성별 데이터직무 인력 현황(2017)

(단위: 명)

구분	성별	2017					
		전 산업	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	일반 산업
전 체	전체	109,320	10,291	31,693	5,823	29,298	32,215
	남성	89,641	8,607	26,506	4,848	22,358	27,322
	여성	19,679	1,684	5,187	975	6,940	4,893
DA	전체	10,071	618	3,392	526	489	5,046
	남성	8,733	564	3,050	426	339	4,354
	여성	1,338	54	342	100	150	692
데이터 개발자	전체	41,254	4,429	13,702	2,188	8,948	11,987
	남성	34,516	3,822	11,419	1,851	7,222	10,202
	여성	6,738	607	2,283	337	1,726	1,785
데이터 엔지니어	전체	16,634	1,468	6,529	696	4,434	3,507
	남성	14,390	1,254	5,641	607	3,715	3,173
	여성	2,244	214	888	89	719	334
데이터 분석가	전체	8,398	629	1,638	335	2,190	3,606
	남성	5,856	492	880	278	1,397	2,809
	여성	2,542	137	758	57	793	797
DBA	전체	17,863	702	3,732	388	7,029	6,012
	남성	15,131	542	3,383	349	5,711	5,146
	여성	2,732	160	349	39	1,318	866
데이터 사이언티스트	전체	1,803	218	410	173	482	520
	남성	1,587	198	369	104	479	437
	여성	216	20	41	69	3	83

〈표 2-19〉의 계속

구분	성별	2017					
		전 산업	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	일반 산업
데이터 컨설턴트	전체	5,004	1,530	1,376	1,231	126	741
	남성	4,018	1,228	1,105	1,026	58	601
	여성	986	302	271	205	68	140
데이터 기획/마케터	전체	8,293	697	914	286	5,600	796
	남성	5,410	507	659	207	3,437	600
	여성	2,883	190	255	79	2,163	196

자료 : 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-20〉 전 산업 내 데이터직무 중 빅데이터 관련 인력 현황(2017)

(단위: 명, %)

구분	2017				
	전 산업 내 데이터직무 인력 전체				전 산업 내 데이터직무 인력 중 빅데이터 인력
	소계		빅데이터 관련 인력		소계
	규모	구성비	규모	구성비	비중
전 체	109,320	100.0	9,955	100.0	9.1
DA	10,071	9.2	-	-	-
데이터 개발자	41,254	37.7	2,908	29.2	7.0
데이터 엔지니어	16,634	15.2	1,679	16.9	10.1
데이터 분석가	8,398	7.7	1,108	11.1	13.2
DBA	17,863	16.3	-	-	-
데이터 사이언티스트	1,803	1.7	1,803	18.1	100.0
데이터 컨설턴트	5,004	4.6	1,724	17.3	34.5
데이터 기획/마케터	8,293	7.6	733	7.4	8.8

자료 : 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-21〉 전 산업 내 데이터직무 중 빅데이터 관련 인력 비중(2017)

(단위: 명, %)

구분	전 산업		데이터 산업		일반 산업	
	규모	빅데이터 인력 비중	규모	빅데이터 인력 비중	규모	빅데이터 인력 비중
전 체	9,955	9.1	6,324	8.2	3,631	11.3
데이터 개발자	2,908	7.0	1,866	6.4	1,042	8.7
데이터 엔지니어	1,679	10.1	1,081	8.2	598	17.1
데이터 분석가	1,108	13.2	381	8.0	727	20.2
데이터 사이언티스트	1,803	100.0	1,283	100.0	520	100.0
데이터 컨설턴트	1,724	34.5	1,369	32.1	355	47.9
데이터 기획/마케터	733	8.8	344	4.6	389	48.9

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

2017년 일반 산업을 포함한 전 산업 내 데이터직무 인력 가운데 빅데이터 관련 인력은 총 9,995명으로 전체의 9.1%를 차지하고 있다(표 2-20 참조). 데이터 산업과 일반 산업에서 빅데이터 인력의 비중을 비교해 보면, <표 2-21>에 나타난 바와 같이 데이터 산업에서 빅데이터 인력 비중이 8.2%인 것에 비해 일반 산업 내 빅데이터 관련 인력의 비중은 11.3%로 더 높게 나타났다. 특히 데이터 컨설턴트 및 데이터 기획/마케터의 비중이 높은 것을 확인할 수 있다. 전 산업의 데이터직무 중 빅데이터 관련 인력 규모를 살펴보면, 데이터 개발자의 규모가 가장 크게 나타났으며, 다음으로 데이터 사이언티스트, 데이터 컨설턴트 순으로 높게 나타났다.

제3절 데이터 산업에서의 고용 동향 및 전망

1. 데이터 산업 필요인력 및 인력 부족률

본 절에서는 데이터 산업에서의 고용 동향 및 향후 인력 수요와 공급에 관한 전망을 살펴보고자 한다. 먼저, 데이터 산업 고용 동향을 파악하기 위하여 관련 필요인력 조사 자료를 보면(표 2-22 참조), 2017년 국내 산업에서 데이터직무에 추가로 필요할 것으로 예상되는 인력은 총 4,865명으로 나타났다. 또한, 전체 필요인력 가운데 비중이 가장 높은 직무는 데이터 개발자로, 전체 데이터 산업 내 데이터직무 인력의 52.7%에 해당하는 총 2,564명이 데이터 산업에 추가로 필요할 것으로 나타나고 있다. 다음으로는 데이터 엔지니어가 총 775명(15.9%)으로 데이터 산업 내에서 두 번째로 많은 인력 수요를 보이고 있다.

다음으로 현재 인력 대비 필요인력의 비중을 나타내는 데이터직무별 인력 부족률을 살펴보면(표 2-23 참조), 2017년 조사 자료에서 데이터 사이언티스트의 인력 부족률이 11.0%로 가장 높게 나타났으며 데이터 개발자(8.1%) 및 데이터 분석가(7.6%), DA(6.6%)도 비교적 부족률이 높은 직무에 해당하는 것으로 보인다.

데이터 산업 내 데이터직무에서 필요한 전체 인력 가운데 빅데이터 관련 필요인력은 총 1,380명으로 조사되었으며(표 2-24 참조), 이는 전체 필요인력의 약 28.4%를 차지하는 비중이다. 특히 빅데이터 관련 인력 중 향후 3년 이내 가장 수요가 높을 것으로 전망되는 직무는 데이터 개발자(47.8%), 데이터 분석가(16.5%), 데이터 사이언티스트(11.5%) 순으로 나타났다.

또한, 2017년 기준 빅데이터 관련 인력 부족률을 살펴보면, 데이터 분석가가 37.3%로 가장 크게 나타났으며, 다음으로 데이터 개발자(26.1%)와 데이터 기획·마케터(23.9%), 데이터 사이언티스트(11.0%) 순으로 나타났다(표 2-25 참조).

〈표 2-22〉 데이터 산업 내 데이터직무별 필요인력(2017)

(단위: 명, %)

구분	항목	전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스
전체	규모	4,865	1,038	1,802	789	1,236
	비중	100.0	100.0	100.0	100.0	100.0
DA	규모	354	140	117	44	53
	비중	7.3	13.5	6.5	5.6	4.3
데이터 개발자	규모	2,564	492	1,127	361	584
	비중	52.7	47.4	62.5	45.7	47.2
데이터 엔지니어	규모	775	163	358	128	126
	비중	15.9	15.7	19.9	16.2	10.2
데이터 분석가	규모	394	83	32	127	152
	비중	8.1	8.0	1.8	16.1	12.3
DBA	규모	145	15	41	19	70
	비중	3.0	1.4	2.3	2.4	5.7
데이터 사이언티스트	규모	158	68	50	29	11
	비중	3.2	6.6	2.8	3.7	0.9
데이터 컨설턴트	규모	164	35	19	56	54
	비중	3.4	3.4	1.0	7.1	4.4
데이터 기획/마케터	규모	311	42	58	25	186
	비중	6.4	4.0	3.2	3.2	15.0

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-23〉 데이터 산업 내 데이터직무 인력 부족률(2017)

(단위: %)

구분	전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스
전체	5.9	9.2	5.4	11.9	4.0
DA	6.6	18.5	3.3	7.7	9.8
데이터 개발자	8.1	10.0	7.6	14.2	6.1
데이터 엔지니어	5.6	10.0	5.2	15.5	2.8
데이터 분석가	7.6	11.7	1.9	27.5	6.5
DBA	1.2	2.1	1.1	4.7	1.0
데이터 사이언티스트	11.0	23.8	10.9	14.4	2.2
데이터 컨설턴트	3.7	2.2	1.4	4.4	30.0
데이터 기획·마케터	4.0	5.7	6.0	8.0	3.2

주: 인력 부족률: 필요인력/(현 인력+필요인력).

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-24〉 데이터 산업 내 데이터직무 중 빅데이터 관련 필요인력(2017)

(단위: 명, %)

구분	데이터직무 필요인력 전체				데이터직무 필요인력 중 빅데이터 필요인력
	소계		빅데이터 필요인력		소계
	인력 규모	구성비	인력 규모	구성비	비중
전체	4,865	100.0	1,380	100.0	28.4
DA	354	7.3	-	-	-
데이터 개발자	2,564	52.7	660	47.8	25.7
데이터 엔지니어	775	15.9	123	8.9	15.9
데이터 분석가	394	8.1	227	16.5	57.6
DBA	145	3.0	-	-	-
데이터 사이언티스트	158	3.2	158	11.5	100.0
데이터 컨설턴트	164	3.4	104	7.5	63.4
데이터 기획/마케터	311	6.4	108	7.8	34.7

주: 1) 빅데이터 관련 인력은 빅데이터 기술을 보유한 인력을 의미하며 데이터직무 인력 수에 포함됨.

2) DA와 DBA는 조사에서 제외되었으며, 데이터 사이언티스트는 100% 빅데이터 관련 인력으로 간주함.

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-25〉 데이터 산업 내 데이터직무 중 빅데이터 관련 인력 부족률(2017)

(단위: %)

구분	전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스
전체	17.9	13.8	15.7	13.6	35.3
데이터 개발자	26.1	14.4	23.3	3.6	63.8
데이터 엔지니어	10.2	3.6	31.2	12.3	27.5
데이터 분석가	37.3	68.1	11.3	37.4	50
데이터 사이언티스트	11.0	23.8	10.9	14.4	2.2
데이터 컨설턴트	7.1	5.9	3.6	6.3	37.3
데이터 기획/마케터	23.9	17.2	51.9	33.3	30.9

주: 1) 인력 부족률=필요인력/(현 인력+필요인력)

2) 2016년 데이터 구축의 경우 데이터 기획/마케터 인력이 없어 인력 부족률이 100%로 분석됨.

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

2. 전 산업 내 데이터직무별 필요인력 및 인력 부족률

2017년 일반 산업을 포함한 전 산업에서 총 1만 3,337명의 데이터직무 인력이 추가로 필요할 것으로 조사되었다. <표 2-26>에서 직무별로 살펴보면, 데이터 산업과 마찬가지로 일반 산업에서도 데이터직무 중 데이터 개발자에 대한 수요가 28.2%로 가장 높은 것으로 나타났으나, 데이터 산업에서는 데이터 엔지니어와 데이터 분석가가 그다음으로 높은 수요를 보인 반면 일반 산업에서는 DA가 17.0%, DBA가 16.8%로 상대적으로 높은 수요를 나타내고 있다. 이러한 데이터 개발자 및 데이터 분석가에 대한 높은 수요는 향후 데이터직무 분야 내 중·고급 인력 증가로 이어질 것으로 전망된다.

<표 2-27>에 제시된 바와 같이 전 산업 내 데이터직무 인력 부족률을 직무별로 나누어 살펴보면, 데이터 사이언티스트, 데이터아키텍트, 데이터 분석가가 상대적으로 부족률이 높은 것으로 나타났다.

2017년 현재 인력 대비 필요인력 비중인 인력 부족률의 경우, 전 산업에서 현재 인력의 약 10.9%의 데이터직무 인력이 추가로 필요한 상황이다. 특히 일반 산업에서는 현재 인력의 20.8%에 해당하는 인력이 더 필요한 것으로 파악되고 있다.

〈표 2-26〉 전 산업 내 데이터직무별 필요인력(2017)

(단위: 명, %)

구분	항목	전 산업	데이터 산업	일반 산업
전 체	규모	13,337	4,865	8,472
	비중	100.0	100.0	100.0
DA	규모	1,791	354	1,437
	비중	13.4	7.3	17.0
데이터 개발자	규모	4,953	2,564	2,389
	비중	37.2	52.7	28.2
데이터 엔지니어	규모	1,738	775	963
	비중	13.0	15.9	11.4
데이터 분석가	규모	1,371	394	977
	비중	10.3	8.1	11.5
DBA	규모	1,569	145	1,424
	비중	11.8	3.0	16.8
데이터 사이언티스트	규모	496	158	338
	비중	3.7	3.2	4.0
데이터 컨설턴트	규모	574	164	410
	비중	4.3	3.4	4.8
데이터 기획/마케터	규모	845	311	534
	비중	6.3	6.4	6.3

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-27〉 전 산업 내 데이터직무별 인력 부족률(2017)

(단위: %)

구분	전 산업	데이터 산업	일반 산업
전 체	10.9	5.9	20.8
DA	15.1	6.6	22.2
데이터 개발자	10.7	8.1	16.6
데이터 엔지니어	9.5	5.6	21.5
데이터 분석가	14.0	7.6	21.3
DBA	8.1	1.2	19.2
데이터 사이언티스트	21.6	11.0	39.4
데이터 컨설턴트	10.3	3.7	35.6
데이터 기획/마케터	9.2	4.0	40.2

주: 인력 부족률=필요인력/(현 인력+필요인력)

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

<표 2-28>에 제시된 바와 같이 전 산업에서 데이터직무 필요인력 가운데 빅데이터 관련 인력은 총 6,008명으로, 전체 데이터직무 필요인력의 약 45%를 차지하는 것으로 나타났다. 빅데이터 관련 필요인력 중 데이터직무 별로는 데이터 개발자가 39.1%로 수요가 가장 높고, 다음으로 데이터 분석가와 데이터 엔지니어가 각각 22.1%와 15.1%로 높은 수요를 보이고 있다.

빅데이터 관련 데이터직무 인력의 부족률 역시 <표 2-29>에 제시된 것처럼 빅데이터 관련 데이터 개발자와 데이터 분석가 인력이 가장 부족한 것으로 나타났다. 빅데이터 기획/마케터와 빅데이터 엔지니어도 그 뒤를 이어 높은 부족률을 보이고 있다. 빅데이터 관련 데이터직무 인력의 부족률은 특히 일반기업에서 크게 나타나고 있어 새롭게 빅데이터 활용을 준비하고 있는 기업이 많은 것으로 분석되며, 향후 빅데이터 도입 기업이 확산됨에 따라 빅데이터 인력에 대한 수요는 앞으로 더욱 증가할 것으로 예상된다.

<표 2-28> 전 산업 내 데이터직무 중 빅데이터 관련 필요인력(2017)

(단위: 명, %)

구분	데이터직무 필요인력 전체				데이터직무 필요인력 중 빅데이터 관련 필요인력
	소계		빅데이터 필요인력		소계
	규모	구성비	규모	구성비	비중
전 체	13,337	100.0	6,008	100.0	45.0
DA	1,791	13.4	-	-	-
데이터 개발자	4,953	37.2	2,350	39.1	47.4
데이터 엔지니어	1,738	13.0	907	15.1	52.2
데이터 분석가	1,371	10.3	1,331	22.1	97.1
DBA	1,569	11.8	-	-	-
데이터 사이언티스트	496	3.7	496	8.35	100.0
데이터 컨설턴트	574	4.3	509	8.5	88.7
데이터 기획/마케터	845	6.3	415	6.9	49.1

주: 1) 빅데이터 관련 인력은 빅데이터 기술을 보유한 인력을 의미하며 데이터직무 인력 수에 포함됨.

2) DA와 DBA는 조사에서 제외되었으며, 데이터 사이언티스트는 100% 빅데이터 관련 인력으로 간주함.

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-29〉 전 산업 내 빅데이터 관련 데이터직무별 필요인력(2017)

(단위: 명, %)

구분	항목	전 산업	데이터 산업	일반 산업
전 체	규모	6,008	1,380	4,628
	비중	100.0	100.0	100.0
데이터 개발자	규모	2,350	660	1,690
	비중	39.1	47.8	36.5
데이터 엔지니어	규모	907	123	784
	비중	15.1	8.9	16.9
데이터 분석가	규모	1,331	227	1,104
	비중	22.1	16.5	23.9
데이터 사이언티스트	규모	496	158	338
	비중	8.3	11.5	7.3
데이터 컨설턴트	규모	509	104	405
	비중	8.5	7.5	8.8
데이터 기획/마케터	규모	415	108	307
	비중	6.9	7.8	6.6

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

〈표 2-30〉 전 산업 내 빅데이터 인력 부족률(2017)

(단위: %)

구분	전 산업	데이터 산업	일반 산업
전 체	37.6	17.9	56.0
데이터 개발자	44.7	26.1	61.9
데이터 엔지니어	35.1	10.2	56.7
데이터 분석가	54.6	37.3	60.3
데이터 사이언티스트	21.6	11.0	39.4
데이터 컨설턴트	22.8	7.1	53.3
데이터 기획/마케터	36.1	23.9	44.1

주: 인력 부족률=필요인력/(현 인력+필요인력)

자료: 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」 재구성.

3. 국내 데이터직무 인력 수급 비교 및 현황

데이터직무 인력에 대한 수요가 점차 증가하면서 이에 대응하여 정부 산하기관, 지방자치단체, 대학 및 대학원, 민간 교육기관 등에서 데이터 전문인력 배출이 늘고 있다. 한국데이터진흥원 조사에 따르면, 국내 데이터 관련 학부 및 대학원 과정의 정원 기준으로 졸업생을 예측할 경우 매년 최소 400명 정도의 데이터 전문인력이 배출될 것으로 보인다. 아울러 한국데이터진흥원 및 고용노동부를 비롯해 정부 주도하에서 실시되고 있는 데이터 인력 양성과정과 지자체 주도로 추진되고 있는 프로그램을 통해 배출될 것으로 예상되는 전문인력의 규모도 2016년 기준 약 1,200명에 이르는 것으로 추정된다. 그 밖에 삼성 SDS멀티캠퍼스, SAS코리아 등 민간 교육기관의 빅데이터 전문인력 양성 프로그램을 통해 3,000명 이상의 교육생이 배출되고 있는 것으로 집계되어 매년 새롭게 배출되는 국내 데이터 전문인력은 약 4,500~5,000명으로 예상된다.

이상으로 현재 국내 데이터직무 인력의 수요와 공급 전망을 종합해 보면, 수치상으로는 현재 교육되고 있는 다양한 데이터 전문가 양성과정을 통해 데이터직무 인력의 수요가 대체로 충족될 수 있을 것으로 보인다. 하지만 데이터직무 인력 공급 전망은 다수의 전문가 양성 프로그램에 중복 수강한 교육생이 존재할 수 있어 과대 추정되었을 가능성이 있으며, 대부분의 데이터직무 인력 수요 기업들의 경우 대학이나 전문 교육기관의 프로그램을 이수한 초급 인력이 아닌 실무 경험을 가진 경력자를 원한다는 점에서 데이터직무 인력의 실제 수요와 공급 간에는 불균형이 존재할 것으로 보인다.

제4절 소 결

데이터 경제의 활성화는 우리나라의 기존 주력 산업인 조선, 자동차, 철강 등의 부문에서 생산성과 효율성을 향상시키고, 금융, 물류·유통, 보건·의료, 바이오 등 새로운 산업의 발전을 촉진시킬 성장 동력으로서 역할을 할 수 있을 것으로 전망된다. 이처럼 데이터가 산업의 발전과 새로운 가치 창출의 원동력이 되는 새로운 데이터 경제가 도래함에 따라 세계 주요국들은 데이터 경제를 선도하기 위해 범국가 차원에서 데이터 활용 및 활용 확대 방안, 데이터 분석 인재 양성 방안, 데이터 활용도 정비 등 종합적인 대책을 마련하고 있다(정지원, 2019: 9). 우리나라 역시 국가 차원에서 데이터 경제 활성화를 위한 추진 체계를 마련하려는 노력을 기울이고 있으며, 기업들 역시 데이터 구축 및 활용에 높은 관심을 갖고 향후 관련 사업을 발전시켜 나갈 것으로 보인다.

본 장에서는 국내 데이터 산업 활성화를 위한 현 위치 진단을 목적으로 우리나라 데이터 경제 기반 상황에 대해 전반적인 현황과약을 실시하였다. 국내 데이터 산업 시장의 규모는 2010년 이후 연평균 성장률 7.5%로 꾸준한 성장세를 유지하고 있으며, 이른바 4차 산업혁명의 가속화로 데이터의 가치가 더욱 부각됨에 따라 그 규모는 2022년까지 연평균 성장률 5.3%로 18조 원대에 이를 것으로 전망되고 있다.

이러한 데이터 산업 시장 확대와 기술의 발전으로 인해 기업에서 필요한 데이터를 기획, 개발, 분석, 운영 및 관리하는 데이터직무 인력에 대한 수요 역시 급증하고 있는 것으로 파악되었다. 2017년 국내 데이터 산업에 종사하고 있는 인력 약 29만 명 가운데 데이터직무 인력은 약 25%를 차지하고 있으며 일반 산업을 포함할 경우 데이터직무 인력만 총 11만에 달하는 것으로 조사되었다. 또한, 2017년 전 산업에서 데이터직무 인력 1만 3천여 명이 추가로 필요한 것으로 나타났다.

이처럼 국내 데이터 산업은 시장과 인력의 규모 측면에서 모두 빠른 양적 성장을 이루고 있다. 그러나 전통적인 데이터 산업이 데이터의 축적

과 데이터베이스 구축에 초점을 두었던 것에 반해 최근에는 데이터를 활용 및 분석하는 영역으로 확장·진화되는 추세다. 그러나 이러한 부문의 국내 기술 경쟁력은 여전히 취약한 것으로 평가되고 있다. 뿐만 아니라 데이터직무 인력 또한 데이터의 저장과 자료 축적에서 데이터의 분석 및 활용 중심으로 변화해 가고 있으나 기업이 요구하는 실무 경험을 갖춘 인력에 대한 공급은 충분하지 않은 것으로 진단되고 있어 이에 대한 고민이 필요할 것으로 보인다.

제 3 장

빅데이터 산업

제1절 빅데이터의 개념

1. 빅데이터의 정의

빅데이터(big data)의 사전적 정의는 큰 규모의 데이터를 지칭한다. 그러나 단순히 규모만으로 빅데이터를 정의할 수 없는 것이, 얼마나 큰 데이터를 빅데이터라고 할 수 있는지도 명확하지 않을 뿐만 아니라, 단순한 크기 외에도 데이터의 여러 성격들이 데이터와 빅데이터를 나누는 기준이 될 수 있기 때문이다.

빅데이터라는 용어가 컴퓨터공학계에 처음 등장한 것은 1990년대이며, 당시에는 기존의 데이터 처리 방식대로 분석할 때 분석 시간이 오래 걸리는 데이터를 빅데이터라고 주로 지칭했었다. 최근 주로 사용되는 빅데이터의 정의가 국내에 처음 소개된 것은 대략 2011년경 몇몇 언론사를 통해서였으며, 국내 학술지에도 비슷한 시기에 해당 용어가 등장하기 시작하여, 2012년부터는 일반에도 그 개념이 알려지기 시작하였다. 2000년대 이후의 미국, 그리고 2010년대의 한국에서 빅데이터라는 용어가 널리 사용되기 시작한 이유는, 데이터가 생성되는 속도가 데이터를 저장, 처리, 분석하는 속도보다 빨라진 까닭에, 데이터를 생성, 보관, 처리하는 방법과

기술에 대한 논의가 대두됐기 때문이다.

해당 시기에 데이터 생성량이 많아진 것은 두 가지 측면에서 기인한다. 첫 번째는 데이터의 생성량이 절대적으로 많아졌기 때문이다. 기존에는 소비자들이 특정한 상품을 어디에서 언제 구매했는지에 관한 정보를 기록할 수 없었으며, 설령 기록한다 해도 개별 판매자 선에서 수기로 진행되어 체계적으로 관리되지 않고 추후 활용될 여지도 낮은 편이었다. 판매자 측면에서 보자면, 판매자의 신용도나 평판이 주로 인근 지역 공동체 수준에서 공유되거나 주 거래자들 사이에서 공유될 뿐 여타 집단에서는 접근하기 어려운 상황이었다. 그러나 스마트폰의 보급, 전자상거래의 확산, 사회관계망 서비스(SNS)의 확산 및 보급으로 인해 소비자의 구매 이력, 선호, 시기별 판매 동향 등이 체계적인 정보로 생성되고 있을 뿐만 아니라, 판매자에 대한 각종 정보 및 여러 판매자들의 판매 가격과 조건에 대한 정보들도 쉽게 생성, 공유, 배포되고 있다. 따라서 절대적인 생성 데이터의 양이 증가한 것이 빅데이터의 개념이 등장하게 된 한 가지 이유이다.

다생성되는 데이터의 양이 증가한 또 다른 이유는 기존에 데이터로 간주할 수 없던 것들이 데이터로 간주되기 시작한 데 있다. 대표적인 예로 문자 및 음성 인식을 꼽을 수 있다. 기존에는 서식에 수기(手記)로 작성한 문서는 데이터 분석을 할 수 없었다. 전통적인 분석 기법들은 주로 수치화된 자료에 한정된 것이 다수였으며, 1960년대에 등장한 단어주머니(bag of words) 기법 역시도 오랜 기간 동안 단순히 단어나 특정 문구가 몇 회 등장하였는지에 그쳐, 결국에는 문자 및 음성 데이터를 수치 데이터로 변환함으로써 기존의 데이터 분석 방법에 의존했던 바가 크다. 하지만 이제는 Doc2Vec 등의 기법이 등장함에 따라 음성이나 수기로 된 표현 자체도 데이터 분석이 가능해지고, 이미지나 영상에서 문자를 읽어내거나 추출하는 것은 물론 특정한 이미지나 형태를 데이터베이스화하는 것이 가능해지면서 분석 가능한 데이터의 범위가 크게 확장되었다. 그리고 이러한 형태의 데이터들은 기존의 데이터 처리 방법으로는 분석이 불가능함에 따라 기계학습과 딥러닝(deep learning) 등에 의존하게 되어 데이터 하나하나의 크기가 커지는 결과를 낳았다. 따라서 이들을 분석하는 데 기존의 분석 방법들보다 더 오랜 시간이 소요됨에 따라 빅데이터의 개념

이 등장할 필요가 있었다.

본 장에서는 빅데이터 산업을 이해하는 것이 목적인바, 이를 위해서는 우선 빅데이터의 개념과 특성을 살펴봐야 한다.

가. 빅데이터의 정의

빅데이터는 데이터를 수집, 저장, 분석하던 기존의 기법으로 다루기 힘든 정도의 용량, 속도, 및 다양성을 가진 데이터를 의미하며, 경우에 따라서는 이러한 빅데이터와 관련된 기술이나 산업까지 포함한다. 여기서 빅데이터와 데이터를 구분하는 기준은 기존의 기법으로는 다루기 힘들다는 것이므로, 한 번 빅데이터로 분류된 데이터가 영원히 빅데이터가 되는 것은 아닌 가변성을 가지고 있다.

즉, 만일 컴퓨터의 저장 공간이 보다 커지고 데이터 처리 및 분석 능력이 향상된다면 기존에는 빅데이터로 분류되었던 것도 일반 데이터가 될 수 있다는 것이다. 반면, 기존에는 분석에서 제외되었던 데이터의 어느 일부분이 분석 가능하게 되면서 데이터의 범위에 포함됨에 따라 분석에 오랜 시간이 소요되는 등의 사유로 빅데이터로 재분류될 수도 있다. 따라서 데이터와 빅데이터를 가르는 기준이 명확하다기보다는 가변적이며, 따라서 빅데이터의 개념 역시 유동적이라 할 수 있다.

나. 빅데이터 개념의 정의 주체

빅데이터의 개념이 상대적, 가변적, 유동적이라는 것은 조금 전에 살펴 보았다. 그렇다고 해서 빅데이터의 개념을 개별 주체들이 각각 정의하도록 한다면, 다양한 개념과 기준, 정의가 혼용되어 국제적 의사소통이나 세계적인 단일 표준에 대한 논의가 어려울 수 있고, 학술 연구를 진행하는 데 있어서도 불편함이 발생하게 된다. 따라서 전 세계적으로 빅데이터의 개념을 상황 및 시대적 환경에 따라 정의할 필요가 생겼다.

이러한 요구에 의해서 미국 국립표준기술협회(US National Institute of Standards and Technology; NIST)에서는 빅데이터의 정의와 개념에

대한 논의를 보다 체계적으로 진행하기 위해 NIST Big Data Public Working Group(NBD-PWG)를 만들어 지난 2015년 보고서³⁾를 펴냈으며, 본 연구에서 다루는 빅데이터의 개념 및 특성의 대부분도 해당 보고서에 기반하고 있다.

2. 빅데이터의 특성

빅데이터의 정의가 유동적이기 때문에 빅데이터 간의 공통적인 특성을 규정짓는 것 역시 쉽지 않다. 하지만 학계와 데이터 산업계에서는 앞서 살펴본 빅데이터의 정의에 부합하는 데이터들이 가지는 일반적인 성격들에 대해서 3Vs라는 특성을 꼽고 있으며, 이에 더하여 최근에 두 가지의 Vs가 거론되고 있다. 본 절에서는 전통적인 3Vs와 추가적인 Vs에 대해서 상세히 살펴봄으로써 빅데이터의 특성을 소개할 것이다.

가. 전통적인 3Vs(conventional 3Vs)

빅데이터의 특징으로 가장 널리 꼽히는 것은 Gartner사의 데이터 분석가인 Doug Laney가 꼽은 빅데이터의 세 가지 주요 개념⁴⁾으로, 이들이 각각 용량(volume), 속도(velocity), 다양성(variety)에 관한 것이라 3Vs로 불리곤 한다. 이들 3Vs는 다수의 데이터 전문가들이 동의하는 빅데이터의 특성이므로 보다 상세하게 살펴볼 필요가 있다.

첫 번째 V는 용량에 관한 것으로 Laney는 고용량(high-volume)을 빅데이터의 첫 번째 특성으로 꼽았다. 여기서 용량은 순수한 물리적인 크기의 용량을 지칭하는 것으로, 데이터의 크기가 커야 한다는 의미이다. 일반적으로 데이터의 용량이 커지면 커질수록 더 많은 저장 공간이 필요하게 될 뿐만 아니라 이렇게 저장된 데이터를 처리하고 분석하는 데 더 긴

3) NIST Big Data Interoperability Framework(2015), "NIST Big Data Interoperability Framework", NIST Special Publication 1500.

4) Laney, Doug(2001), "3D data management: Controlling data volume, velocity and variety", META Group Research Note. 6(70).

시간이 소요되게 마련이다. 빅데이터는 이처럼 용량이 큰 데이터이기 때문에 데이터 저장 측면에서 효율을 높이기 위한 여러 기술의 발전을 초래한다. 예를 들어 보다 작은 크기의 디스크에 더 많은 데이터를 저장하는 장치나 서버 등의 개발이 그 예이다.

두 번째 V는 속도에 관한 것으로, Laney에 따르면 고속도(高速度, high-velocity)를 지칭한다. 빅데이터는 데이터의 크기가 크기 때문에 데이터를 수집, 공유, 생성하는 과정에서 과거보다 많은 데이터의 교환 및 전송을 만들어낸다. 특히 최근에는 스마트 기기의 보급 등으로 데이터가 실시간으로 생성되고 저장될 뿐만 아니라 데이터를 주고받는 기기의 절대적인 숫자 및 교환하는 데이터의 크기가 커졌다. 이에 따라 단일 데이터의 크기가 크건 작건 과거의 전통적인 방법으로는 분석할 수 없을 정도의 데이터가 생성되고 처리된다. 따라서 데이터의 이동 속도가 빠르지 않다면 데이터의 유용성이 떨어지게 될 뿐만 아니라, 데이터 분석에 효율적인 기법을 마련하였다 해도 데이터 수집 과정에서 병목현상이 발생하여 데이터 분석에 소요되는 시간이 길어지게 된다. 따라서 빅데이터는 데이터의 전송 속도가 빨라질 것을 요구한다.

마지막 V는 다양성(high-variety)으로, 생성, 수집, 교환, 전송, 분석되는 데이터의 형태와 종류가 다양해지는 것뿐만 아니라, 데이터를 다른 종류로 변환하는 일도 찾아지는 것을 아우르는 개념이다. 과거의 데이터는 주로 수치로 작성된 구조화된 데이터(structured data)였으며, 분석 방법 역시 주로 이러한 구조화된 데이터에 대한 것이었다. 그러나 최근에는 텍스트, 이미지, 음성, 그리고 비디오 등등 비구조화된 데이터(unstructured data) 형태도 많아졌을 뿐만 아니라, 이러한 데이터를 수집하고 분석하기 위한 다양한 기법들이 개발되고 있다. 또한 이렇듯 다양한 형태 간의 변환도 자유롭게 이뤄지고 있는데, 예를 들어 음성 인식 기술로 음성 데이터를 텍스트 데이터로 변환하는 식이다. 또한 문자 인식 기술에 의해 이미지 파일 내의 문자가 텍스트 데이터로 변환되기도 하며, 텍스트 파일을 음성으로 읽어주는 등 다양한 형태의 데이터 간 전환이 이루어지고 있다. 이러한 데이터 형태의 전환은 비단 앞서 언급한 비구조화된 데이터 간에만 일어나는 것이 아니다. Doc2Vec같은 기술의 경우 비구조화 데이터인

텍스트를 수치화시킴으로써 구조화된 데이터로 만들어서 컴퓨터가 분석에 직접 사용할 수 있도록 변환해주는 것이다.

나. 추가적인 Vs(additional Vs)

빅데이터에 대한 연구가 진행되고 빅데이터의 형태와 종류가 다양해짐에 따라 앞서 언급한 3Vs로는 부족한 빅데이터의 특성을 보완한 개념들이 나타난바, 이를 추가적인 Vs로 부르고 있다. 이 중에서 널리 꼽히는 것으로는 신뢰성과 가치를 들 수 있다.

정확성 또는 신뢰성(high-veracity)은 수집된 데이터의 질에 관한 개념이다. 데이터의 질에 대해서는 여러 평가 기준이 있을 수 있으나, 분석을 위해 중요한 것은 데이터가 형식이나 내용에 있어서 얼마나 일관되는가, 수집 과정에서 오류는 없었는가, 그리고 대상이 되는 현상이나 현실의 특성을 데이터가 잘 반영하고 묘사하고 있는가 등이다. 빅데이터는 분석 대상이 방대하고 처리 및 분석에 긴 시간이 소요되기 때문에 데이터의 정확성이 떨어지면 오히려 제외시키는 것이 효율적인 경우가 많다. 흔히 빅데이터 분석에서 주요한 방법으로 간주되는 것이 GIGO(garbage in, garbage out)인데, 빅데이터의 성격상 부정확하거나 잘못된 데이터가 있다면 차라리 해당 부분을 사용하지 않는 것이 나은 때가 있다. 부정확한 데이터를 가지고 하는 분석들은 대개 어떠한 유용한 정보도 창출하지 못하는 경우가 많기 때문이다.

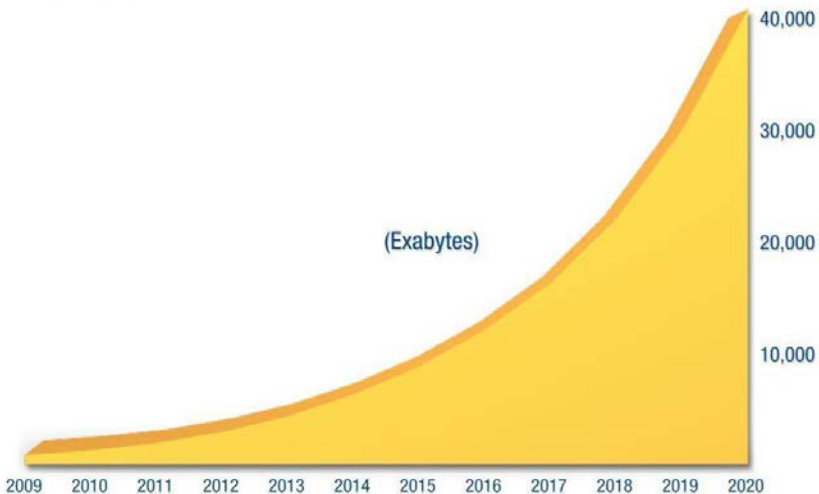
마지막 V는 가치(high-value)인데, 이는 데이터가 내포한 활용가치를 지칭한다. 데이터는 빅데이터건 아니건 그 자체로 가치를 가진다기보다는 이를 수집, 가공, 분석하는 과정을 통해 유용하고 가치있는 정보로 변하게 마련이다. 빅데이터는 분석 대상의 범위가 일반적인 데이터보다 훨씬 넓기 때문에 일반적인 데이터와 비교하여 높은 가치를 가지는 경우가 많으며, 또한 그래야만 긴 시간이 소요되는 빅데이터 분석을 하는 게 적합하다 할 수 있다.

제2절 빅데이터 현황 및 기술

1. 빅데이터 현황

최근 데이터들의 특징은 과거에 비해 단일 데이터의 크기가 커졌을 뿐만 아니라 발생 경로 역시 다양하다는 점이다. 발생 경로가 다양하다는 것은 더 많은 기기에서 보다 많은 데이터 주체들이 데이터를 생성하고 있음을 의미한다. 이로 인해 과거에는 수집된 데이터를 관리의 용이성을 위해 하나의 저장 공간에 저장하도록 해주는 단일 대용량 저장소가 주를 이루었다. 그러나 최근에는 데이터 이동 속도 및 처리 방법이 발전하면서 수집의 용이성을 위해 데이터를 분산된 저장소들에 저장한 후, 이를 네트워크상에서 마치 하나의 저장소로 보관하는 것처럼 접속케 함으로써 전체 데이터에 접근할 수 있게 해주는 분산 파일 저장시스템을 주로 사용하고 있다.

[그림 3-1] 미국의 연도별 데이터 생성량



자료 : EMC Corporation(2012: 4).

[그림 3-1]은 10억 기가바이트에 해당하는 엑사바이트 단위로 표시한 미국 내 데이터 생산량이다. 많은 통계에서 미국 내에는 해마다 전년 대비 약 100~125%가 넘는 양의 데이터가 생성되고 있으며, 125%의 경우 3년 후에 그 양이 두 배가 되는 수준이다. 빠른 네트워크 속도와 사회관계망 서비스의 보급 및 확산, 모바일 및 휴대용 기기의 등장과 많은 판매량, 유비쿼터스의 보편화 등에 따라 데이터 산업의 규모는 빠르게 커지고 있으며, 데이터 생성량 역시 같은 양상을 보이고 있다. 특히 이러한 데이터 생성량의 증가에는, 앞서 언급했듯이, 구조화된 데이터만을 저장하던 데서 벗어나 비구조화된 데이터의 기록을 지원하는 비관계형 모델의 데이터베이스가 활용된 것도 한 원인이다.

비관계형 모델 데이터베이스에서는 데이터가 있는 그대로 저장되어 더 많은 저장 공간이 필요하게 되었다. 따라서 빅데이터의 저장은 수직적으로 구조화된 하나의 데이터 저장소를 설치하여 운용하는 것보다 수평적으로 조직된 여러 분산 데이터 저장소를 사용하는 것이 비용적인 면에서도 필연적으로 유리하게 된 것이다.

관련 정의가 다양하기 때문에 빅데이터에 대한 체계적인 글로벌 통계는 산출하기 어려우나, 디지털화 및 각종 인터넷 기반 서비스가 발달한 한국의 특성상 국내에서도 빅데이터를 포함한 전체 데이터 생성량은 매년 빠르게 증가하고 있는 것으로 보인다.

2. 빅데이터 관련 기술

빅데이터는 단순히 데이터를 생성하고 저장하는 것뿐만 아니라 빠르고 정확하게 분석하고 활용하는 것도 핵심적이다. 하지만 말 그대로 데이터의 크기와 다양성이 증가한 빅데이터는 전통적인 분석 방법이나 인간의 분석 능력으로는 분석과 유용한 정보의 추출이 사실상 불가능하다. 이에 따라 등장한 각종 인공지능과 기계학습 등의 기법은 빅데이터 분석의 한계를 극복하고 기계가 데이터를 대신 분석하도록 하는 시스템이다.

가. 인공지능(Artificial Intelligence)

인공지능은 기계가 주변 환경을 인식하고 분석하여 그에 맞는 최적의 행동을 하도록 만드는 시스템을 의미한다. 최근에는 여기서 더 나아가 주변 환경을 통해 새로운 지식을 학습하고 이렇게 습득한 지식을 향후 각종 분석 및 의사결정에 활용하는 능력으로까지 발전하였다. 인공지능은 이미 빅데이터 분석에서 널리 활용되어, 추론(reasoning), 지식 표현(knowledge representation), 자연어 처리(natural language processing) 등에서 많이 활용되고 있다.

나. 기계학습(Machine Learning)

기계학습은 인공지능의 한 분야로 인간의 개입을 최소화하고 대신 기계나 시스템이 수집 데이터에서 각종 관계와 패턴 등을 파악하여 분석하는 시스템이다. 초창기에 인간이 시스템의 설정 및 학습을 설정하고 지도한 데 반해 현재는 기계나 시스템이 스스로 학습하여 사람이 수행할 경우 많은 시간이 소요되는 데이터의 분석을 빠르게 수행하고 있다.

기계학습은 크게 지도 학습(supervised learning), 반지도 학습(semi-supervised learning), 비지도 학습(unsupervised learning)으로 나눌 수 있는데, 정답을 제시하는 훈련을 위한 데이터의 크기가 어느 정도인가로 구분된다(권배근, 2015: 16). 비지도 학습의 경우 그러한 데이터가 주어진 상황에서 정답이나 의도를 제시하는 것이 아니라 목표함수 설정 후 그 함수값이 최소화 혹은 최대화되도록 기계를 학습시킨다.

다. 딥러닝(Deep Learning)

딥러닝은 기계학습의 하위 개념으로 인간의 신경망을 모방하여 만든 인공 신경망(artificial neural networks)을 이용해서 각종 분석이나 예측을 수행하는 알고리즘이다. 딥러닝의 경우 시스템을 학습시키는 데 많은 연산을 요구하고 학습에 소요되는 시간도 길다는 단점이 있지만, 비교적

높은 예측력과 정확성을 보여주어 컴퓨터 영상 인식, 음성 인식, 자연어 처리, 사회관계망 서비스 필터링, 기계 번역 등에 폭넓게 사용되고 있다. 다음 소절에서 소개될 각종 빅데이터 활용 사례의 다수는 딥러닝에 기반하고 있다 해도 무리가 아니다.

라. 워드 임베딩(Word Embedding)

각종 문자나 서체, 글을 인식하고 처리하는 데 있어서 핵심적인 개념인 워드 임베딩은 딥러닝을 활용한 알고리즘의 일종으로, 최근 자연어 처리 연구에서 가장 널리 사용되고 있다. 기계가 인간의 언어를 인식하고 처리하기 위해서는 자연어를 숫자로 변환하는 과정이 필요하다. 여기에는 여러 가지 방법이 있는데, 초기의 워드 임베딩은 각각의 단어에 고유한 위치를 부여하고 그 위치를 벡터에서 1로 표시하는 방법을 썼다. 예를 들면, 한 문서에 총 네 가지의 단어가 들어가는 경우, 4가지의 단어를 각각 $[1\ 0\ 0\ 0]$, $[0\ 1\ 0\ 0]$, $[0\ 0\ 1\ 0]$, $[0\ 0\ 0\ 1]$ 로 표시하는 것이다. 하지만 이 방식은 메모리를 많이 소요할 뿐만 아니라, 단어의 위치에 따른 의미 변화나 단어 사이의 관계 등을 파악할 수 없다는 단점이 있다.

최근의 워드 임베딩은 초기에 학습한 데이터의 모든 단어에 임의로 벡터를 분배한 후, 한 단어를 둘러싼 주변 어휘들을 가지고 중심에 있는 단어를 예측하도록 학습한다. 이때 주변 어휘나 중심 단어는 모두 벡터화된 값들이다. 학습을 위한 데이터의 크기가 작은 경우에는 이러한 알고리즘을 통한 결과물이 큰 의미를 보이지 않지만, 학습에 사용하는 데이터의 크기가 커지면 각 벡터값이 해당 단어를 표현하는 고유의 값으로 수렴하게 되는데, 이는 학습을 통해 중심 단어와 주변 어휘의 관계를 보다 명확하게 인식하게 된다.

3. 빅데이터 기술의 활용

앞서 설명한 빅데이터 기술은 여러 비구조화된 데이터를 분석하는 것을 가능케 할 뿐만 아니라 매우 큰 데이터의 분석 역시 빠르고 정확하게 할 수 있는 가능성을 열어주었다. 이에 따라 사회 여러 분야에서 빅데이터 기술이 빠르게 적용되고 있는데, 다음은 그 주요 사례들을 꼽아본 것이다.

가. UPS 배송 경로 최적화 시스템

미국 최대 물류회사인 UPS는 46,000대가 넘는 자사의 배송 차량에 원격 추적 장치를 설치하고 이를 통해 수집된 데이터를 분석하여 배송 체계 전반의 효율성을 제고하기 위한 시스템을 개발하였다. 이 시스템은 차량 별로 최적의 경로나 배송 순서를 설정해 주는 것뿐만 아니라, 가장 효율적인 차량 운행 속도 등도 분석한다. 이 시스템을 통해 배송 시간을 단축시켜줄 뿐만 아니라 차량 운행에 소요되는 연료까지 절감하고 배송에 배치되는 인원도 단축할 수 있다. 일례로 이 시스템을 통해 절감하게 된 연료는 연간 약 840만 갤런으로, 이를 비용으로 환산하면 약 240억 원에 해당한다.

나. 넷플릭스의 영상물 추천 시스템

넷플릭스 등 다수의 영상 스트리밍 서비스 업체들은 사용자의 시청 이력을 파악하여 시청자의 기호에 맞는 영상물을 제공하는 시스템을 갖추려 한다. 이러한 시스템을 구축하는 데 있어서 현재 가장 앞서 있는 업체가 넷플릭스이다. 넷플릭스는 기본적으로 사용자가 선택한 영상물에 대한 스트리밍 서비스를 제공하는 것뿐만 아니라, 더 많은 영상물을 시청하도록 사용자 개개인의 기호에 맞는 영상물을 검색하여 추천한다. 이 영상물 추천 시스템은 끊임없이 영상물을 전송하기 위한 스트리밍 시스템과 더불어 넷플릭스의 서비스를 지탱하는 양대 핵심 기술이다. 이 추천 시스

템은 개별 사용자의 검색 이력 및 서비스 사용 기록이 저장된 클라우드의 자료를 회귀분석, 클러스터링, 토픽 모델링 등의 여러 기계학습 기법을 통해 사용자의 기호를 실시간으로 분석하여 그 결과물을 사용자에게 추천한다. 넷플릭스 사는 자사가 보유한 빅데이터를 SQL 및 NoSQL 시스템을 사용해서 아마존 웹 서비스에 저장한다.

다. 운전 기록 데이터

보험사 및 차량 제조사들은 많은 지원자들을 모집하여 차량에 각종 기록 장치를 부착한 후 장기간에 걸쳐 운전 환경이나 차량 운행에 대한 데이터를 수집한다. 이렇게 수집된 빅데이터는 주행 시스템 및 차량 개선에 활용된다.

또한 여러 차량 보험사들이 자사가 제공하는 차량 운행 시스템을 설치할 경우 보험료를 할증해 주는데, 이러한 시스템을 통해 해당 운전자의 운전 경로, 운전 성향, 주행 시간, 주행 환경 등에 대한 다양한 정보를 수집한다. 이러한 빅데이터는 사고 발생과 관련되는 환경 요인, 운전자의 성향, 차량의 특성 등을 분석한다. 이는 추후 개별 보험 가입자들의 사고 확률 및 예상되는 사고 위험을 식별하는 데 쓰여 보험료를 계산하거나 보험 가입 여부 파악 및 상품 추천 등에 활용된다.

라. 전자 의료 기록(Electronic Health Record, EHR) 데이터

건강 관리 산업은 최근 빅데이터를 활용하여 각광받으면서 빠르게 성장하고 있는 산업이다. 개별 환자들의 의료 기록은 개별 의료기관에서 수집하고 관리하지만, 미국에서는 환자의 개인 식별 정보를 삭제하는 경우 연구 목적으로 의료 기록을 공유할 수 있다. 이렇게 축적된 대량의 데이터를 분석하여 사람이 파악하기 힘든 여러 요인들 간의 상관관계를 밝힘으로써 질병을 예방하거나 유전자 정보와 후천적 요인, 각종 유전 질환의 발병 원인 등을 파악하는 데 빅데이터 분석이 활용된다. 이러한 분석 결과를 통해 후천적 유전 질환의 발현을 억제할 수 있고, 응급실에 입원한

환자의 최초 상태를 데이터베이스화하여 48시간 이내 사망률을 예측하는 등 개인의 건강 관리와 증진을 위한 수단으로 활발하게 활용되고 있다.

제3절 빅데이터 산업에서의 고용

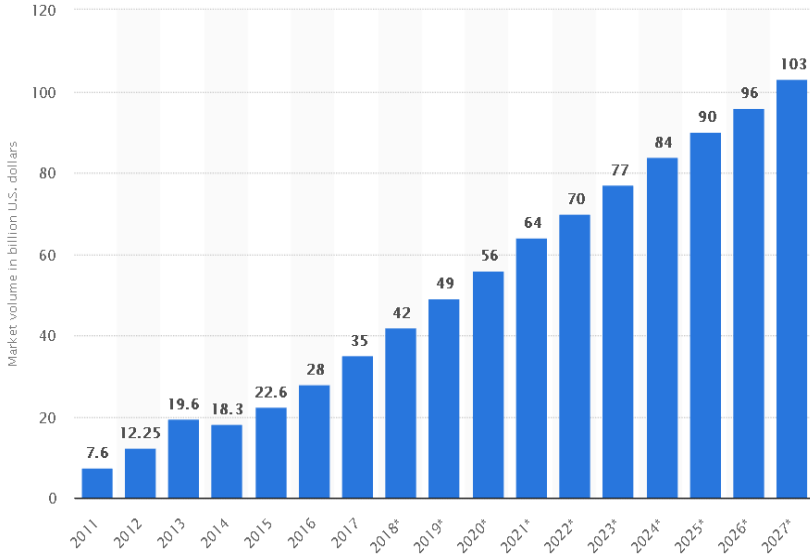
1. 빅데이터 산업에서의 고용 전망

[그림 3-2]에 따르면, 2017년 기준 전 세계 빅데이터 관련 시장 규모는 약 40조 원 정도로 추정되고 있으며, 관련 산업의 규모는 10년 내에 세 배 가량 성장할 것으로 예측되고 있다. 이렇게 시장 규모가 커지게 되면서 빅데이터 산업 종사자 수 및 고용 규모도 크게 증가할 것으로 전망된다. Statista.com에 따르면, 빅데이터 관련 직종 전체의 연평균 고용 성장 규모는 약 4.7%이며, 데이터 분석가가 3.8%, 데이터 관련 연구직군이 6.4%, 데이터 관련 기술직군이 5.8%, 데이터베이스 관리자와 개발자가 각각 3.6%와 4.1%의 비율로 매년 성장할 것으로 예상된다. [그림 3-2]와 [그림 3-3]은 글로벌 빅데이터 시장 규모와 그에 따른 전 세계 빅데이터 관련 직군별 고용 규모를 각각 보여주고 있다.

2. 빅데이터 산업의 주요 직종

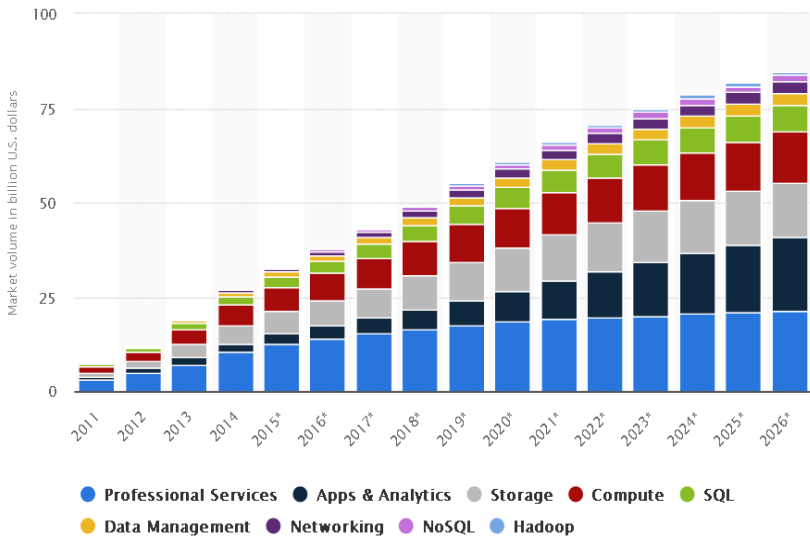
빅데이터 산업이 활성화되면서 빅데이터 관련 직종들이 여러 개 생겨나고 직종별 고용 규모도 커지고 있다. 한국은 빅데이터 및 데이터 산업에서의 직종 분류 체계에 있어 미국과 유사한 분류법을 채택하였으므로, 본 소절에서는 미국 노동부 분류 체계에 따라 빅데이터 산업에서의 주요 직종은 무엇이 있으며, 이들의 업무는 무엇이고 현황은 어떠한지를 살펴보고자 한다.

[그림 3-2] 빅데이터 시장 규모 전망



자료 : Statista(2019), <http://www.statista.com>(접속일자 : 2019. 9. 24).

[그림 3-3] 직군별 빅데이터 고용 규모



자료 : Statista(2019), <http://www.statista.com>(접속일자 : 2019. 9. 24).

미국 노동부는 2016년을 기준으로 2026년까지 전체 직종에서 연평균 일자리가 0.68% 정도 증가할 것으로 전망했다. 이는 10년간 전체 일자리 규모가 약 7% 정도 증가할 것으로 예측됐음을 의미한다. 반면 같은 기간 빅데이터 관련 주요 직종들은 매년 3.6~6.4%까지 성장할 것으로 전망됨에 따라 전체 일자리의 성장률을 크게 웃돌고 있다. 이는 전체 고용에서 빅데이터 관련 고용이 차지하는 비중이 지속적으로 커질 것이라는 전망을 낳는바, 이제부터는 빅데이터 산업의 세부 직종별 주요 업무와 특성을 살펴보고자 한다.

가. 최고 데이터 관리자(chief data officer)

기업의 규모가 커지고 업종이 다양화됨에 따라 CEO(chief executive officer)나 CFO(chief finance officer)와 같은 직책들이 생겨났다. 최근에는 데이터를 수집, 관리, 분석하는 업무와 데이터 관련 보안의 중요성이 강화됨에 따라 일부 기업에서 최고 데이터 관리자가 등장하였다. 최고 데이터 관리자는 사내 데이터 관련 정책을 수립하고 데이터 관련 지침 등을 표준화하며, 데이터와 관련한 문제가 발생하였을 때 당사자 혹은 문제의 주체 간 이해관계를 조정하는 것뿐만 아니라 기업 내외부의 데이터와 관련된 고객을 관리하고 데이터 사업과 관련된 전략을 수립한다. 아울러 지속적으로 데이터를 관리하여 데이터의 질적 가치를 제고하고 활용 방안을 모색함은 물론 데이터 보안 및 관리의 최고 책임도 진다.

최고 데이터 관리자는 몇몇 기업들에서 막 생겨난 직종으로 아직 집계할 만큼 많은 숫자가 있는 것은 아니라서 해당 직종에 종사하는 종사자의 수를 파악할 수는 없으나, 두 자릿수를 넘기지는 않는 것으로 보인다. 이들의 2017년 기준 연간 보수는 \$118,965에서 \$307,673 정도로 한화로는 1억 3천만 원에서 3억 5천만 원 사이이다.

나. 데이터 분석가

데이터 분석가는 수집된 데이터, 고객의 보유 또는 제공 데이터를 분석

하여 어떠한 문제를 해결하거나 의사결정을 돕는 업무를 수행한다. 이들은 주로 단독으로 일하는 경우가 많지만 업무의 규모와 데이터의 범위에 따라서 팀 단위로 일하기도 한다. 하지만 팀 단위로 업무를 한다 하더라도 대부분의 경우 개별 분석가의 업무는 단독성 및 독립성이 보장되는 경우가 많다. 한편 이들은 데이터를 분석하는 것뿐만 아니라 데이터베이스에서 정보 검색 시스템 구축의 업무를 수행하는 경우도 있다. 현재 데이터 분석가들이 주로 활용하는 분석 도구는 마이크로소프트 엑셀, 마이크로소프트 액세스, SPSS이며, 보다 큰 데이터의 경우 SAS나 SQL 등의 프로그래밍 언어가 주로 사용된다.

이들의 2017년 미국 내 추정 일자리 수는 15,228개이며, 연간 예상 일자리 증가율은 약 3.8% 정도로 예상된다. 미국 내 데이터 분석가들의 보수는 연봉 기준으로 \$38,146~83,685로 한화 기준 약 4200만 원 정도에서 1억 원 정도이다.

다. 데이터 연구원(data researcher or data scientist)

데이터 연구원(데이터 사이언티스트)은 여러 원시 데이터를 분석하고 연구하여 새로운 지식이나 모형을 발굴하는 일을 한다. 기업에 소속된 데이터 연구원들은 업무 혁신이나 신제품 개발 및 각종 상품 관련 예측을 하거나 기업 전략에 관한 의사결정에 참여하기도 한다. 이들은 주로 원시 데이터를 어디에 활용할 수 있으며, 어떠한 분석 방법이 가장 적합한지 등을 연구하는 직무를 수행하므로, 앞서의 데이터 분석가들과는 달리 주로 기계학습이나 딥러닝 등을 수행할 수 있는 프로그래밍 언어를 사용하며, 통계 전공자들도 많다.

이들은 연구원으로서 전문적인 연구를 수행하다 보니 석사 이상의 학력을 요구하며, 기계학습과 딥러닝 등을 수행하는 능력이 필수적으로 요구된다. 2017년 기준 미국 내 일자리의 수는 21,098명으로 데이터 분석가보다도 많다. 이는 빅데이터에 대해서 연구가 활발하게 진행되고는 있지만, 아직은 기업들이 이러한 연구를 실제 데이터에 적용시키려는 필요성을 느끼지 못하고 있을 뿐만 아니라, 빅데이터를 활용할 수 있는 범위를

여전히 제한적이고, 무엇보다 빅데이터를 활용하는 데 각종 법적 제도적 규제가 있는 현실에 기인한다. 그러나 빅데이터에 대한 연구가 활발히 진행되는 것이 최근의 추세이기 때문에 미국 노동부는 해당 일자리의 연간 증가율이 약 6.4%에 이를 것으로 추정하고 있다. 이들의 연봉은 \$60,386~145,423이며, 중간값은 \$114,520으로 다수의 연구원들이 높은 수준의 연봉을 받고 있음을 확인할 수 있다.

라. 데이터 엔지니어(data engineer)

데이터 엔지니어들은 회사의 경영진이나 의사결정 집단과 데이터 연구원을 이어주는 가교 역할을 하는 직종이다. 빅데이터는 빅데이터 분석을 위한 각종 기법뿐만 아니라 그 정의 자체도 일반인에게는 생소하고 어려운 개념이기 때문에 데이터 연구원들의 업무나 결과물을 회사 경영진들이 쉽게 이해하기가 어렵다. 따라서 데이터 엔지니어들은 데이터 연구원들에게 회사의 주요 경영 목표나 방침을 설명하여 이들의 연구 방향을 설정하거나 조정하고, 데이터 연구원들의 연구 성과에 대해서 그 결과는 무엇이며 기업이 어떻게 활용할 수 있는지 등에 대해서 설명한다. 한편 이들은 이러한 중개 업무 외에도 원시 데이터를 관리하거나 새로운 데이터 출처에 대해서 평가하는 업무도 수행한다.

데이터 엔지니어 직종은 기본적으로는 경영에 대한 이해와 기술적인 지식을 모두 갖출 것이 요구되며, 석사 이상의 학력을 소지해야 하는 것이 보통이다. 이들이 사용하는 데이터 관련 기술은 하둡(Hadoop), 맵리듀스(MapReduce), 하이브(Hive), 피그(Pig), 데이터 스트리밍, NoSQL 등이며, 이를 위해 DashDB, MySQL, MongoDB, Cassandra와 같은 프로그램들을 주로 사용한다. 2017년 미국 내 이들 직군의 일자리는 36,148개 정도이며, 예상 연간 고용 증가율은 약 5.8%로 추정된다. 연봉은 데이터 연구원과 비슷한데, 2017년 기준 미국 노동부는 \$64,172~138,594 정도로 보고 있다.

마. 데이터베이스 관리자

데이터베이스 관리자는 기업 전체의 데이터베이스를 유지, 보수하는 데이터베이스 환경 관리와 데이터베이스 관련 인력과 예산의 관리 및 조정 업무를 수행한다. 따라서 데이터베이스 관리자는 데이터베이스 기술 전반에 대한 이해가 필요하며, 요구되는 기술 수준 역시 앞서 설명한 데이터 기술자와 유사하다. 하지만 이들은 실제 데이터베이스를 유지하거나 보수하는 업무를 수행하는 것은 아니며, 주된 업무 역시 데이터베이스의 운용 현황을 점검하고 방향을 제시하며 관련 인력과 예산을 관리하는 것이므로 석사가 아닌 학사 학위 소지자들의 비율도 낮지 않다.

2017년 데이터베이스 관리자의 숫자는 16,426개로 추정되며, 예상 연간 일자리 증가율은 3.6%이다. 이들의 연봉 수준은 \$35,286~121,958이며, 중간값은 \$87,020인 것으로 미국 노동부는 보고 있다. 데이터베이스 관리자들의 연봉은 데이터 연구원이나 데이터 엔지니어보다는 낮은 편인데, 이는 데이터 연구원이나 데이터 엔지니어가 보다 전문적인 기술을 요구하기 때문에 대체 가능성이 떨어지며, 최근 빅데이터와 관련한 연구직들의 수요가 급증함에 따라 데이터 연구원과 엔지니어에 대한 수요가 증가한 데서 기인하는 것으로 보인다.

바. 데이터베이스 설계 및 개발자(database architect or database developer)

데이터베이스 설계자는 데이터베이스를 설계하고 표준화를 수행하는 업무를 맡는다. 데이터베이스의 형태가 표준화되어 있지 않으면 개별 데이터를 불러와 결합하는 과정에서 문제가 발생하거나, 데이터 분석에서 표준화를 위한 추가적인 시간이 요구된다. 따라서 데이터베이스 설계자는 데이터베이스를 표준화하고 일원화하여 데이터 연구자나 분석자들이 보다 수월하게 연구 및 분석을 진행할 수 있도록 돕는다.

데이터베이스 개발자는 데이터가 원활하게 이동 및 전송되도록 새로운 데이터베이스에 대한 원형(原型, prototype)을 개발하고 테스트하는 역할

을 수행한다. 여기서 데이터와 데이터베이스의 차이를 간단하게 설명하면, 데이터베이스는 각종 화물을 저장할 수 있는 창고이며, 데이터는 실제 창고를 드나드는 화물이다. 창고 그 자체에는 화물을 적재하거나 적재된 화물을 처리할 수 있는 능력이나 기술이 내포되어 있지 않기 때문에 화물을 직접 적재하고 관리할 데이터 기술자나 연구자들이 필요하다. 그런데 창고 자체가 화물의 규격에 맞지 않거나, 혹은 창고마다 다른 형태를 띠고 있다면 화물을 적재하고 보관하는 데 많은 어려움이 따른다. 따라서 데이터의 형태에 맞게 데이터베이스를 구축하고, 효율적이면서 안정적으로 데이터베이스를 유지해야 데이터의 안전한 수집, 저장, 보관 및 효율적인 처리와 분석이 가능해지는 것이다. 데이터베이스 설계자는 이러한 창고를 설계하는 업무를 수행하며, 개발자는 보다 효율적이고 안전한 창고를 구축하는 업무를 수행한다. 일반적으로 데이터베이스 설계자와 개발자는 분리되어 있지 않으며, 한 사람이 설계자와 개발자의 업무를 동시에 수행한다. 왜냐하면 데이터베이스를 설계한 사람이 직접 만드는 것이 가장 효율적이기 때문이다.

데이터베이스 설계 및 개발자에게 요구되는 기술은 주로 데이터 쿼리와 관련된 SQL이나 NoSQL이며, 이들이 사용하는 분석 도구는 Microsoft SSMS, Oracle SQL Developer, PL/SQL Developer, IDERA Rapid SQL 등이다. 2017년 기준 일자리 규모는 61,110명으로 비교적 많은 편이며, 일자리의 증가율 역시 4.1% 정도로 추정된다. 이들의 연평균 보수는 \$72,012~168,101로 데이터 기술자 및 분석가와 비교해도 높은 수준인데, 이는 빅데이터의 핵심이 데이터를 효율적으로 수집하고 저장하는 데서 출발하는 것이므로 데이터베이스를 잘 구축하는 것이 무엇보다 중요하며, 현재는 빅데이터의 수집과 저장 및 보관과 관리가 중요한 산업 초기 단계여서 데이터베이스 구축을 위한 인력 수요가 많기 때문으로 보인다.

사. 한국과 미국의 빅데이터 노동시장 비교

앞선 두 개의 소절에서 데이터 산업 전반의 고용 현황과 주요 직종을 살펴보았지만, 여기서 한국의 빅데이터 산업 현황에 대해서 간략하게

〈표 3-1〉 미국과 한국의 빅데이터 관련 고용 및 직무 현황 비교표

(단위: 명, %)

구분	미국(2017년 기준)			한국(2018년 기준)		
	인원 (추정)	구성비 (추정)	증가율 (추정)	인원 (실제)	구성비 (실제)	증가율 (실제)
데이터 분석가	34,921	10.2	3.8	5,820	7	21.5
데이터 연구원	48,273	14.1	6.4	1,285	1.6	0.2
데이터 기술자	82,509	24.1	5.8	13,731	16.6	4.6
데이터베이스 관리자	37,317	10.9	3.6	12,615	15.3	6.4
데이터베이스 개발자	139,340	40.7	4.1	35,923	43.4	4.8
합 계	342,360	100	4.7	69,374	100	6.4

주: 인원 단위는 명이며, 구성비와 증가율 단위는 %임.

자료: 미국 - Bureau of Labor Statistics(2018), Occupational Employment Statistics 재구성; 한국 - 한국데이터진흥원(2018), 데이터 산업 현황 조사 재구성.

살펴본 후, 한국과 미국의 빅데이터 관련 노동시장을 비교하고자 한다.

한국의 데이터 산업 전체 시장 규모는 2017년 기준 전년 대비 약 5.6% 정도 성장한 7조 원 정도로 파악되었으며, 이 중 빅데이터 시장은 약 4,547억 원 정도로 동년 미국 빅데이터 시장의 약 11.3% 정도의 규모로 집계되었다. 하지만 한국의 빅데이터 관련 시장은 2014년 이후 연평균 30% 이상의 성장률을 기록하여 지속적으로 빠르게 성장하는 추세를 확인할 수 있었다.

한국의 빅데이터 관련 고용원은 2018년 기준 전체 5,820명으로 추정되며, 이 중 데이터베이스 개발자가 전체의 43%로 가장 많았고, 그다음에 데이터 기술자와 데이터베이스 관리자로 나타났으며, 데이터 분석가와 데이터 연구원은 비율이 매우 낮은 것으로 집계되었다. 이러한 한국 빅데이터 시장의 고용 현황을 미국과 비교한 결과는 <표 3-1>과 같다.

미국과 비교하여 한국의 빅데이터 관련 노동시장은 몇 가지 특성을 보이고 있다. 우선 공통점으로는 양국 모두 데이터베이스 개발자들의 비율이 40%를 넘는다는 점이다. 앞서 언급했던 빅데이터의 특성과 같이 살펴보면, 빅데이터의 특성 중 하나인 데이터의 양이 매우 방대해데다 아직은 빅데이터 산업 초기여서 빅데이터의 수집 및 데이터 구축 자체에 대한 수

요가 많아 현재로서는 데이터베이스 개발에 많은 인력이 투입되고 있음을 알 수 있다.

하지만 미국과 비교하면 한국은 데이터 분석가, 연구원 및 기술자의 비중이 낮은 반면 데이터베이스 관리자와 개발자의 비중은 높아서 전체의 약 60%에 이른다. 특히 데이터 연구원의 경우 한국에서는 빅데이터 전체 종사자에서 약 1.6%만이 데이터 연구원으로 일하고 있으나, 미국의 경우 그 비율이 14%가 넘어 9배 정도의 격차를 보이고 있다. 이는 현재 한국에서 데이터베이스의 구축에 많은 인력을 투입하고 있는 반면, 실제 데이터를 분석하거나 활용을 연구하는 직군의 비율은 낮다는 것을 시사한다.

이러한 상황에 대해서 한편으로는 긍정적으로 해석할 수 있는 반면, 부정적인 해석도 가능하다. 긍정적인 측면으로는 한국의 경우 빠르고 넓은 통신망을 구축한데다 각종 스마트 기기의 보급률도 높아서 데이터의 생성 및 저장 속도가 매우 빠르고, 방대한 양의 데이터가 축적되면서 데이터베이스 구축의 필요성이 상대적으로 크다는 의미일 수 있다. 빅데이터와 데이터 산업에서 궁극적으로 가장 중요한 것은 양질의 데이터를 최대한 많이 체계적으로 구축하고 이를 위한 시스템을 갖추어 놓는 것이다. 따라서 데이터베이스에 많은 인력을 투자함으로써 현재 많이 축적되고 있고 추후 더 많이 축적될 데이터를 위한 좋은 데이터베이스의 구축이야말로 장기적으로 바람직한 현상일 수 있다.

반면, 현재 데이터베이스 산업에 많은 인력이 종사하지만 실제 데이터를 연구, 분석하는 직종의 종사자가 적다는 것은 현재 데이터를 연구하고 분석할 만한 역량을 가진 인력이 한국에 부족하다는 것을 의미할 수 있다. 다른 한편으로는 아직 축적된 데이터를 활용할 만한 필요성을 경영진에서 느끼지 못하거나 빅데이터 분석 기법을 활용할 만한 시장이 성숙되지 않았다는 것을 뜻하기도 한다.

빅데이터 분석과 같이 많은 변화가 짧은 기간 진행되고 기술 진보의 속도가 빠른 분야는 한 번 기술 격차가 발생하면 추격하는 것이 쉽지 않다. 또한, 자칫 많은 양질의 데이터를 구축해 놓은 후 정작 이를 분석하고 활용하는 고부가가치 부문에서는 외국 인력이나 기업에게 의존해야 할 가능성도 있다. 따라서 한국은 데이터 분석을 위한 고도의 역량을 갖춘

인력을 양성하기 위한 시스템과 지원책을 마련해야 할 필요가 있다.

아울러 경영진 등 의사결정 상층부에서 빅데이터 분석의 필요성을 느끼지 못한다면 자칫 장기적으로 글로벌 경쟁력에서 뒤처질 가능성이 있으므로 이들의 의식 개선도 필요하다 하겠다. 또한 내수 시장이 작다 하더라도 해외 시장에서 구축된 데이터까지 활용하여 빅데이터 분석을 통해 생산과 기업 경영 및 조직 전반에 대한 체계적이고 고도화된 분석을 시행해야 한다. 이를 바탕으로 합리적인 의사결정을 하는 시스템을 갖춰야 할 뿐만 아니라, 내수 시장에서 빅데이터를 적극 활용함으로써 앞으로는 내수의 범주를 넘어 빅데이터 산업의 시장 규모를 키울 필요도 있다.

<표 3-1>에서 제시되지는 않았지만 한국의 데이터 산업 노동시장이 미국과 비교하여 여성 참가율이 낮다는 점도 주목할 만하다. 미국의 경우 데이터 분석가를 대상으로 한 조사에서 여성의 비중이 39%였지만 한국의 데이터 산업 여성 종사 인력은 19%에 불과하였다. 물론 미국은 데이터 분석 산업, 한국은 데이터 산업 전반에서 각각 여성 인력비를 조사한 것이어서 비교 대상이 다르긴 하지만, 국내 데이터 산업의 여성 종사 비율이 낮다는 것은 데이터 산업에서의 노동 여건이 여타 고연봉 직종과 비교하여 열악하다는 점을 시사한다.

특히 한국의 경우 소프트웨어 개발 직종 등에서 잦은 야근과 불규칙한 업무 투입 및 일정 준수를 위한 장시간 노동이 잦은 특성이 있다. 최근 이러한 문제를 해결하기 위한 여러 방안이 강구되고 있고, 기업들 스스로도 조직 문화를 개선하고 근로시간을 조정하려는 등의 노력을 기울이고 있어서 이러한 문제들이 장기적으로는 해소될 수 있을 것으로 전망되나, 그 속도를 높일 필요가 있다 하겠다.

제4절 데이터 현지화 조치의 함의와 노동시장 영향 분석

1. 데이터 현지화 조치의 개념 및 현황

데이터 그 자체는 국경의 개념이 없는 서비스 상품이다. 그러나 데이터는 해당 데이터를 생성한 생산자가 존재하고, 데이터를 분석할 사용자가 존재하며, 궁극적으로 데이터를 수집, 관리 및 저장할 서버도 필요하다. 데이터 현지화(data localization) 조치는 특정 목적으로 생성하였거나 특정 국가가 보관하는 데이터를 물리적으로 관리하기 위한 서버를 어디에 구축하고 데이터를 어디까지 어떻게 공개 또는 통제할지를 결정하는 문제에 대한 각국 정부의 해답이다.

데이터 현지화 조치가 특히 최근에 주목을 받는 것은 빅데이터의 핵심 중 하나가 데이터를 수집하고 저장하는 작업을 과거처럼 단일 대용량 저장소에서 처리하지 않고 분산된 저장소에서 네트워크를 통해 연결시키는 분산 파일 저장 시스템을 주로 활용하기 때문이다. 이러한 분산 파일 저장 시스템의 효율이 높아지기 위해서는 개별 저장소에 저장된 데이터에 대해 외부에서 자유롭게 접근 가능해야 한다. 하지만 데이터 현지화 조치는 저장소에 보관되어 있는 데이터에 국경 밖에서 자유롭게 접근하는 것을 제한하기 때문에 분산 파일 저장 시스템의 효율이 떨어지게 된다. 이로 인해 데이터의 접근성과 전송 속도가 떨어지게 되면 궁극적으로 빅데이터의 활용도가 낮아지거나 데이터의 질이 떨어지는 문제가 발생할 수 밖에 없다.

데이터 산업이 일정 수준 발전한 국가에서는 개인 혹은 기업 등 사회 주체의 개인 정보와 이들이 생성한 데이터를 보호하기 위한 조치를 어느 정도 취하고 있다. 데이터를 무작정 개방하는 경우, 개인의 사생활이 보호받지 못할 뿐만 아니라 악용될 소지마저 다분하다. 또한 데이터를 생성한 주체가 데이터를 관리할 권한을 갖지 못하는 경우 발생할 수 있는 소유권의 문제나 데이터 생성 과정에서 발생한 비용에 대한 보상 및 사용료

지불 문제 등이 발생하므로 일정 정도 데이터의 사용을 제약하거나 특정한 조건을 갖출 것을 요구하는 것은 당연하다 하겠다. 또한 이러한 규제는 데이터라는 상품이 거래되는 데 있어 국제적 표준을 설립하는 역할을 하여 오히려 데이터의 안전한 이동과 효율적인 거래를 위한 국제적 시장을 형성하는 데 긍정적인 역할을 할 수도 있다.

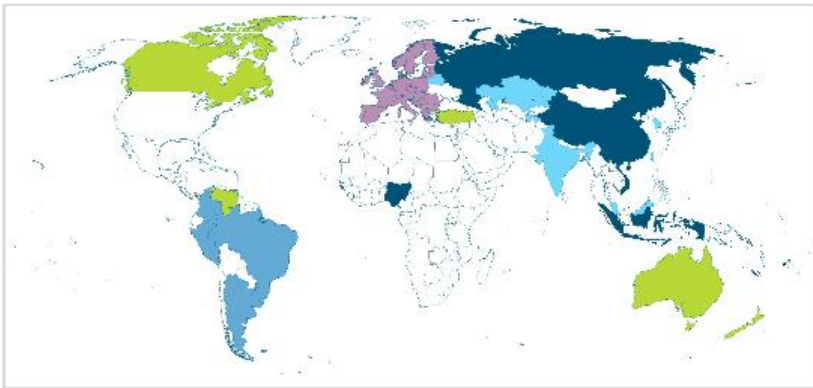
가. 데이터 현지화 조치의 현황

데이터 현지화 조치에 대한 국가별 현황은 [그림 3-4]에서 확인할 수 있다. 데이터 현지화와 관련하여 아무런 조치를 취하지 않는 나라는 다수의 아프리카 및 중동 국가와 미국, 일본이 있다. 이 중 앞선 그룹의 국가들은 데이터 산업의 발전이 미비하여 아직 관계 법령 등을 갖추지 않은 것으로 보이며, 미국과 일본에서는 데이터 현지화와 관련하여 어떠한 제약도 두고 있지 않다. 유럽연합은 개인정보보호법(GDPR)을 발효하면서 개인 정보 자체의 국외 이전을 제한하고 있으나, 데이터를 보관하는 저장소에 대해서는 지역의 규제를 따로 두고 있지는 않다. 반면, 중국과 러시아, 베트남, 인도네시아 등은 강한 데이터 현지화 조치를 실시하고 있으며, 한국의 경우 인도 등과 비슷하게 국외로 이전할 수 없는 정보의 종류를 지정하고 있다.

한편 현지화 조치의 경우 그 대상이 비단 데이터 저장 시설 자체나 저장 설비만을 대상으로 하지는 않는다. 예를 들어, 소비자나 데이터 생산자의 동의가 있다면 국외 이전을 허락하는지, 그리고 이러한 현지화 조치가 실제로 강력하게 실행되고 있는지 등에 따라 규제의 모습은 다양하다. <표 3-2>는 주요 국가와 한국에서 취하고 있는 데이터 현지화 조치의 세부적인 규제 수준을 보여주고 있다. 이규엽 외(2018)에 의하면 미국은 데이터 현지화와 관련하여 거의 규제가 없는 것으로 나타났다. 이는 미국이 데이터 산업과 관련된 각종 기술이 빠르게 나타나고 있으나, 데이터 산업 관련 주요 대기업들이 미국 회사인 관계로 자국 기업들의 해외 진출과 글로벌 영업을 돕는 측면에서도 특별한 규제를 가하지는 않는 것으로 보인다. 중국은 모든 조치 대상에서 강한 규제를 펼치는 것으로 나타나

표에 제시된 국가들 중 가장 강력한 데이터 현지화 정책을 펴고 있음이 확인된다. 유럽연합은 데이터를 직접적으로 국외로 이전하는 것에 대해서는 한국과 비슷한 수준의 규제를 가하고 있으나 데이터 저장 설비에 대해서는 미국과 유사한 낮은 수준의 규제를 가하고 있다. 한국은 전면적인 현지화 조치를 취하기보다는 부분적인 조치를 취하고 있으며, 이에 대한 상세한 내용은 다음 소절에서 살펴보고자 한다.

(그림 3-4) 데이터 현지화 조치의 국가별 현황



지도상의 색	데이터 해외 이전에 대한 제한의 정도	해당 국가
■	[매우 엄격] 개인정보는 국내 설치된 서버에 보관되어야 함	러시아, 중국, 인도네시아, 베트남, 브루나이, 나이지리아
■	[사실상 엄격] 개인정보의 국외 이전을 법령으로 제한함으로써 사실상 데이터 현지화가 이루어지고 있음	유럽연합(EU)
■	[부분적 제한] 국외 이전할 수 없는 정보의 종류를 지정하거나 정보의 국외 이전에 정보주체의 동의를 요함	한국, 인도, 말레이시아, 카자흐스탄, 벨라루스
■	[완 화] 일정한 조건하에서만 정보의 국외 이전을 제한함	아르헨티나, 브라질, 콜롬비아, 페루, 우루과이
■	[특정분야 제한] 헬스케어, 이동통신, 금융, 국가안보 등의 분야에 한하여 정보의 국외 이전을 제한함	타이완, 캐나다, 오스트레일리아, 뉴질랜드, 터키, 베네수엘라
□	[제한 없음] 데이터 현지화 의무가 없음	미국, 일본 등 기타 국가

자료 : 박환일(2017), 「개인정보의 현지화에 관한 연구」, 『경희법학』 Vol.52 No.4. pp. 129~164.

〈표 3-2〉 주요 국가의 항목별 데이터 현지화 조치

구분	약함	중간	강함
데이터 국외 이전에 관한 직접적인 제한 조치	미국	한국 유럽연합	중국
데이터 저장 설비의 지역화	미국 유럽연합	한국	중국
소비자 동의에 따른 국외 이전	미국		한국 중국 유럽연합
규제의 실효성	한국	유럽연합	중국

자료: 이규엽 외(2018), 「국경 간 데이터 이동에 관한 국제적 논의 동향과 대응 방향」, 대외경제정책연구원.

나. 개인 정보 보호에 관한 국제 비교

이러한 데이터 현지화 정책의 궁극적인 목적은 크게 두 가지로, 첫 번째는 인터넷 공간에서의 개인 정보 보호이며, 두 번째는 각국의 데이터 및 개인 정보에 대한 통제권 확보이다. 앞에서 데이터 현지화 정책에 대한 여러 국가의 정책 방향에 대해서 살펴보았으나, 논의를 더 진행하기에 앞서 개인 정보 보호 및 데이터 이동에 관한 각국의 입장을 상세하게 살펴보기로 하자.

유럽연합의 GDPR(General Data Protection Regulation)은 국적에 상관없이 유럽연합에서 활동하는 모든 데이터 관련 기업에 적용되는 규칙으로, 유럽연합 외에서 설립된 기업의 경우 반드시 유럽연합 내부의 사업 대리인을 정하여 GDPR의 준수를 실행하고 유럽개인정보보호이사회(EDPB: European Data Protection Board)의 감독을 받아야 한다. GDPR이 보호하는 대상은 온라인에서 개인을 식별할 수 있는 모든 정보를 포괄한다. 대표적으로 IP 주소나 쿠키 정보, RFID가 포함되며, 이러한 개인 정보를 최소한으로 수집하고 적법하며 투명하게 처리해야 함은 물론 명시된 목적 외의 용도로 활용할 수 없게 하고 있다. 또한 개인 식별 정보를 수집, 활용하기 위해서는 해당 정보를 수집하고 접근하는 데 개별 주체의 동의를 반드시 받도록 규정하고 있다. 만일 유럽연합 내의 사업자에 의해

개인 정보가 유출된 경우, 해당 사업자는 유출 사실을 인지한 시점으로부터 72시간 내에 유럽개인정보보호이사회 혹은 여타 감독 당국에 유출 사실을 신고해야 하며, 유출된 정보가 해당 개인에게 큰 위험을 초래할 수 있는 경우 해당 주체에게도 바로 통보해야 한다. 이처럼 유럽연합의 GDPR은 정보의 제공 및 활용에 대한 동의에 있어서 개인 정보의 주체가 해당 정보의 통제권을 가지도록 하는 것에 주안점을 두고 있다.

반면 중국의 경우, 개인 정보에 대한 국가의 주권과 국경 내 저장 및 처리에 방점을 두고 있는 모습이다. 중국 법률에서 규정하는 개인 정보는 전자적으로 혹은 다른 방법으로 기록되어 단독으로든 다른 정보와 결합해서든 개인의 신분을 식별할 수 있는 각종 정보이다(이규엽 외, 2018: 63). 개인 정보의 예로는 성명과 출생일자, 신분증상의 일련 번호나 개인의 계좌 정보, 주소나 전화번호, 지문 및 홍채 등의 신체 정보를 망라한다(사이버보안법 제76조). 이 중에서 신분증 번호와 은행의 카드 번호, 건강 관련 기록과 생물학적으로 식별 가능한 정보, 즉 홍채나 지문 데이터 등이 유출되거나 공개되는 경우, 혹은 당국 및 자연인 개인의 허가를 받았다 하더라도 남용 시 개인의 인신이나 재산에 해가 되거나 명예 등에 손해를 끼치고 차별의 소지로 악용될 수 있을 경우는 해당 정보를 특히 개인 민감 정보로 분류하고 있다. 아울러, 관련 조직이나 기구 및 개인이 국내에서 수집하거나 생성하였고, 국가 기밀은 아닐지라도 국가의 안전과 경제 발전이나 공공 이익과 밀접하게 연관된 데이터들은 원자료이건 파생 자료이건 구분하지 않고 중요 데이터로 분류하고 있다. 중국 정부는 이러한 개인 정보와 중요 데이터는 모두 국경 내에 저장해야 하며, 일체의 예외 조항도 허용하지 않는 정책을 취하고 있다. 또한 해당 정보의 국외 송출은 정부의 주관 부서나 관리 감독 부서의 보안평가를 신청한 경우에만 가능하도록 하고 있다.

중국에서의 개인 정보와 관련된 무단 유출 및 송출 관련 사례로 심천 화파 유전자과학기술서비스와 화산병원의 사례를 들 수 있다. 화파 유전자과학기술서비스는 영국의 옥스퍼드 대학교와 함께 중국 인류 유전자원 국제 협력 연구를 수행하던 중이었다. 해당 판결에서 화파 유전자과학기술서비스는 당국의 허가를 받지 아니하고 온라인상으로 중국인의 유전자

자원 일부, 즉 유전자 정보를 옥스퍼드 대학교에 제공하였는데, 이에 대해서 중국 정부는 당국의 허가를 받지 아니하고 개인 정보를 송출한 것으로 판단하였다. 그에 따라 화따 유전자과학기술서비스와 화산병원은 네트워크 보안법과 유사한 인류유전자자원관리 입사방법 규정 위반으로 처벌되었다.

미국의 경우, 연방 정부에서 개인 정보 보호에 관련된 법 제도를 갖추고 있으나, 각 주별로 세부적인 사항이나 집행에 있어 차이를 보이기 때문에 일괄적으로 특정하기는 힘들다. 연방 정부의 입장은 기본적으로 일부 분야를 제외한다면 개인 정보 규제를 최소화하고 있다. 하지만 일부 주에서는 주 정부 차원에서 정보 주체의 개인 정보를 보호하기 위한 규정을 두고 있는데, 대표적인 경우가 캘리포니아 소비자 프라이버시법이다.

캘리포니아 소비자 프라이버시법은 캘리포니아 주민에게 자신의 정보에 접근하고 삭제를 요청할 수 있는 권리를 부여하고, 자신의 정보를 제3자에게 판매하거나 공유하는 것을 중지할 수 있는 권리를 보장하고 있다. 하지만 미국은 전형적인 영미법 체계를 따르고 있어 개별 소송을 통해 보호받는 개인 정보의 범위, 공유와 판매 및 가공이 가능한 영역이 어디까지인지를 판별해가고 있는 중이다.

베트남 또한 개인 정보에 대한 국가의 소유를 중국 못지않게 엄격히 규정해놓은 국가이다. 2019년 1월 1일부터 발효되는 새로운 사이버 보안법은 글로벌 기업들이 데이터 관련 사업을 하거나 데이터를 생성, 저장, 관리하는 경우 베트남 현지 사무소의 설립을 의무화하도록 요구하고 있는데, 이는 궁극적으로 모든 기업의 네트워크를 베트남 공공보안부가 조사 및 관찰할 수 있도록 하는 데 있다. 여기서 사이버 보안법상의 보호 대상인 개인 정보는 개인과 기업의 금융 기록과 생체 인식 데이터, 민족 정보와 정치적 견해에 관련된 정보까지 개인과 기업에 의해 생성된 대부분의 정보에 적용된다. 베트남 사이버 보안법은 개인의 정보에 대한 개인의 권리를 보호하는 데 초점을 맞추기보다는 정부가 정보의 흐름을 통제하고 핵심 인프라를 보호할 수 있도록 하는 데 초점이 있다.

새로운 사이버 보안법에 따르면 페이스북 등 글로벌 사회적 관계망 사업자들에게 대해서도 베트남 정부가 명예훼손 등을 이유로 삭제를 요청하는

콘텐츠에 대해서는 해당 사업자가 24시간 이내에 어떠한 조치를 취해야 하며, 아무런 조치도 취하지 않을 경우 당국에 해당 사유를 제출해야만 한다. 그렇지 않은 경우, 베트남 정부는 해당 사업자의 베트남 내에서의 사업 정지나 각종 규제를 가하게 된다. 해당 규제에 적용을 받은 첫 사례가 바로 페이스북의 경우였다. 베트남 당국에 의해서 특정 콘텐츠의 삭제를 요구받았으나 페이스북 측이 본사의 정책을 이유로 이를 거부하자, 이에 베트남 당국은 해당 건과 관련한 경제적, 기술적 조치를 검토하겠다고 답한 바 있다.

하지만 이러한 과도한 규제에 대해서 대내외적인 반발이 있자 베트남 정부는 모든 국외 기업에게 현지 사무소의 설립을 요구하던 조항을 완화하여 해당 의무가 적용되는 기업의 범위를 제한하게 된다. 그렇기는 해도 앞서 살펴보았듯이 페이스북 등 다수의 데이터 관련 기업들은 여전히 적용 대상으로 남아있다.

다. 대한민국에서의 데이터 현지화 조치

한국은 현재 국외로 이전할 수 없는 정보를 제한하는 소극적 현지화 조치를 취하고 있다. 따라서 원론적으로는 데이터의 국외 이전 및 사용이 허가되고 있으나, 국토 정보 및 실시간 스트리밍 서비스(OTT)하에서 스트리밍 대상이 되는 데이터는 국내의 서버에 보관해야 한다고 규정하고 있다. 그리고 이러한 예외적인 이전 제한 조치에 대해서 필요한 경우에는 허용 단서를 두는 형태를 취하고 있다.

한국 정부가 이전을 제한하는 영역은 국토 정보와 관련된 분야로, 해당 분야는 원칙적으로는 반출이 금지되어 있으나, 국토교통부 장관 및 장관 산하 협의체의 심의를 거쳐 허용하는 예외 조항이 있어 조건부로 반출이 가능하다. 즉, 모든 데이터를 현지화할 것을 요구하지는 않으나 지리 정보와 관련된 일부에 대해서는 정부의 허락을 득하는 경우에만 이전 및 반출이 가능하다.

2013년 구글이 지도 및 그와 연계되는 서비스를 제공하기 위하여 국내 지리 정보의 반출을 요청하였다. 하지만 국토교통부는 남북 대립의 측면

등을 고려하여 특수 시설 등에 대한 블러(blur) 처리가 동반된 조건부 허용만 가능하다는 입장을 밝혔으며, 구글은 이에 대해서 불가하다는 답변을 하였다.

특히 최근 많이 사용되는 OTT와 사회관계망 서비스 등이 생활에 보다 밀접하게 연결되면서 해당 정보들을 저장하는 서버를 국내에 설치할 것을 의무화하는 법인이 현재 국회에 계류 중이어서 한국의 경우에도 데이터 현지화 정책은 보다 강화되는 추세라고 말할 수 있다.

한국 데이터 현지화의 특징은 규제의 정도가 열려 있다는 점이다. 한국은 국가 안보 등의 분야와 관련하여 데이터 해외 이전에 대한 제한의 정도를 현행 법령을 기준으로 판단하는 것이 어려운 상태로, 행정 주체의 해석에 따라서 매우 엄격하게 데이터 현지화 조치를 실시할 수도 있는 반면, 한편으로는 행정적인 판단에 따라 사실상 실행하지 않는 데이터 현지화 조치를 취하는 것도 가능하기 때문이다.

반면 가장 강력하게 데이터 현지화 조치를 취하는 중국의 경우에는 중국 내에서 생성된 정보를 저장하기 위한 서버를 중국 내에 두도록 할 뿐만 아니라 해당 정보를 관리하거나 삭제할 권한까지도 행정 당국에게 부여하고 있다. 또한 중국에서 데이터 현지화 정책과 관련된 법률은 국가안전법, 네트워크 안전법(中國人民共和國網絡安全法), 개인정보와 중요데이터 국외송출 보안평가 방법(個人信息和重要數據出境安全評估辦法(征求意见稿)), 정보안전기술 데이터 국외송출 보안평가 지침(1차, 2차 의견수렴안)((信息安全技術 數據出境安全評估指南(征求意见稿, 第二稿) 등으로 다양하다. 데이터 산업에서 발생하는 여러 기업 행위들이 상술한 법이나 지침 중 적어도 하나에 해당할 가능성이 매우 높아, 기업들은 높은 수준의 데이터 현지화 조치를 취할 것을 요구받게 된다.

2. 데이터 현지화 조치의 고용 영향

가. 데이터 현지화 조치의 직·간접 고용 효과

데이터 현지화 조치가 노동시장에 미치는 영향은 단선적이지 않다. 데

이터 현지화 조치는 일종의 비관세 무역장벽이라 할 수 있는데, 비관세 무역 장벽이 고용을 늘리는지 줄이는지에 대해서는 해답이 분명치 않다.

우선 데이터 현지화 조치는 비관세 무역 장벽으로서 자국 내 산업의 육성을 도모할 수 있기에 고용 증진 효과가 있을 수 있다. 데이터 산업에서 후발 국가들은 선도 국가에서 활동 중인 기업들에 대해서 데이터 현지화 조치를 요구함으로써 추가적인 비용을 발생시켜 자국 기업에게 유리한 편의를 제공한다. 이러한 편의를 통해 선도 국가 기업들과 비교하여 미성숙한 단계에 있는 자국 기업들의 상품과 서비스의 질이 떨어지고 생산 비용이 높다 하더라도 이러한 비관세 장벽을 통해 유치 산업 보호를 함으로써 대외 경쟁력을 갖출 수 있는 시기까지 내수 시장을 기반으로 자국 기업들의 활동을 유지시킬 수 있는 것이다. 비관세 무역 장벽은 자국 내 데이터 산업에서 국내 기업의 활동을 촉진시켜 자국 기업들에 의한 직접 고용을 증가시킨다. 따라서 비관세 무역 장벽으로서의 데이터 현지화 조치는 유치 산업 단계에서 자국 내 데이터 산업의 직접 고용을 증가시킬 것이다.

반면, 비관세 무역 장벽이 존재하기 때문에 외국의 값싸고 좋은 데이터 관련 서비스의 비용이 올라감에 따라 국내 여타 유관 산업들에서 관련 서비스를 제공하지 못하게 될 수가 있다. 설령 제공한다 하더라도 높은 비용으로 인해 그 범위가 줄거나 질적 저하를 불러올 것이므로 결국 국내 여타 산업에서 간접 고용을 줄이는 효과를 낳는다. 즉, 값싼 양질의 서비스를 통해서 데이터 산업 외의 산업에서 신규 상품이나 서비스가 출현하고 이로 인해 고용이 늘어날 수 있는 기회가 줄어들거나 없어지게 되는 셈이다.

나. 산업별 취업유발계수

비관세 무역 장벽이 가지는 일반적인 두 가지의 고용 효과, 양의 직접 고용 효과와 음의 간접 고용 효과 중 무엇이 더 큰지에 대해서 살펴보는 것이 본 장의 목적이다. 이를 위해 우선 <표 3-3>의 취업유발계수를 활용할 것이다.

〈표 3-3〉 한국의 산업별 취업유발계수

(단위: 명/10억 원)

구분	2000	2005	2010	2015
농림수산물	67.4	50.1	36.1	26.9
광산품	14.1	11.6	9.6	8.6
공산품	20.0	14.8	8.6	8.0
소비재 제품	34.6	25.5	13.7	11.6
음식료품	42.9	27.6	18.7	14.4
섬유 및 가죽 제품	24.3	21.0	12.8	10.1
목재 및 종이, 인쇄	21.2	18.0	11.2	10.6
기타 제조업 제품	49.9	35.4	12.3	11.4
기초 소재 제품	12.2	9.0	5.5	5.7
석탄 및 석유화학	1.6	1.3	1.3	1.4
화학 제품	12.2	9.4	6.1	5.5
비금속 광물 제품	16.4	12.6	8.0	8.7
1차 금속 제품	9.4	6.6	4.2	5.1
금속 가공 제품	21.4	15.2	8.0	7.8
조립 가공 제품	15.3	11.4	7.4	7.3
컴퓨터, 전자 및 광학 기기	9.3	7.7	5.2	4.7
전기 장비	17.7	13.1	8.2	7.7
기계 및 장비	18.6	14.0	8.8	8.6
운송 장비	15.6	10.9	7.5	8.2
전력, 가스, 수도 및 폐기물	9.5	6.6	5.0	3.8
건설	23.1	17.6	13.4	12.5
서비스	29.9	24.5	18.2	15.2
도소매 및 운송	35.6	26.1	17.5	16.8
도소매 및 상품 증개 서비스	47.0	33.8	21.3	18.9
운송 서비스	24.3	18.4	13.7	14.7
생산자 서비스	18.6	17.2	14.1	11.3
정보통신 및 방송 서비스	16.3	14.9	13.0	10.0
금융 및 보험 서비스	16.7	13.0	11.0	8.7

〈표 3-3〉의 계속

구분	2000	2005	2010	2015
부동산 서비스	10.5	9.3	6.0	5.3
전문, 과학기술 서비스	25.0	17.6	16.7	12.8
사업 지원 서비스	24.5	31.4	23.8	19.5
사회 서비스	27.8	21.2	17.1	15.3
공공행정, 국방 및 사회보장	22.7	17.3	12.7	10.5
교육 서비스	35.4	26.0	20.3	17.3
보건 및 사회복지 서비스	25.4	20.4	18.2	18.1
소비자 서비스	46.8	38.6	26.6	20.5
음식점 및 숙박 서비스	52.9	41.9	26.5	22.4
예술, 스포츠 및 여가 관련 서비스	37.1	26.0	17.1	14.7
기타 서비스	50.3	47.8	36.2	24.5
기타	34.9	31.2	-	11.4
평균	25.7	20.3	13.8	11.8

자료: 한국은행(2019), 『2015년 산업연관표』.

취업유발계수란 특정 상품 및 서비스에 대한 최종 수요가 1단위 발생할 경우 해당 상품을 포함한 모든 상품에서 직·간접적으로 유발되는 취업자의 수를 의미한다. 취업유발계수는 취업계수에 최종 수요 1단위당 직·간접적 생산 유발 효과를 나타내는 생산유발계수를 곱하여 계산한다.⁵⁾

한국은행에서 2019년 발표한 『2015년 산업연관표』에 따르면, 전체 산업에서의 취업유발계수(10억 원당, 명목 금액 기준)는 2015년 기준 약 11.8명이며, 세부 업종별로는 농림수산업이 가장 높고, 그다음이 서비스업종, 제조업, 광업 순서였다. 여기서 우리가 관심을 가져야 하는 산업은 데이터 산업과 관계된 정보통신 및 방송 서비스와 전체 평균이다.

정보통신 및 방송 서비스의 경우 취업유발계수가 2000년부터 16.3명, 14.9명, 13.0명, 10.0명으로 2015년 기준 10억 원 정도 시장 규모가 커질 때 10명 정도의 고용 증가가 있음이 확인 가능하다. 한편 전체 평균으로는 10억 원 정도의 시장 규모 변화에 약 11.8명 정도의 고용 변화가 발생

5) 한국은행(2019: 230).

하는 것으로 나타났다.

비관세 무역 장벽도 엄연한 무역 장벽인만큼, 일정 정도 국내총생산을 줄이는 역할을 한다. 따라서 여기서는 편의상 데이터 산업에서의 고용원 수와 데이터 현지화 조치에 따른 국내총생산 감소에 의한 취업 유발 효과를 비교하여 데이터 현지화 조치에 따른 고용의 직간접 효과를 비교해보고자 한다.

다. 데이터 현지화 정책에 따른 국내총생산 및 고용 변화 추계

데이터 현지화 조치에 따른 거시경제 효과는 본 연구의 핵심적인 요소는 아니므로 여기서는 데이터 현지화 조치가 국내총생산에 미치는 영향에 대한 선행 연구의 결과를 빌리고자 한다. 이 분야에서 가장 많이 거론되는 연구 결과는 Bauer et al.(2014)로서, 연산 가능 일반균형모형(CGЕ: computational general equilibrium)을 사용하여 데이터 현지화 정책이 거시경제에 미치는 효과를 분석하였다. 여기에 따르면 한국의 현재 데이터 현지화 조치는 전체 국내총생산의 0.4%를 감소시키고 있으며, 전면 현지화 정책을 시행할 경우 그 값은 1.1%로 늘어날 것이라고 추정된다. 이로 인해 경제성장률은 2.8%에서 1.7%로 하락하고, 투자는 두 경우 각각 0.5%와 3.6% 감소하는 것으로 나타났다.

해당 연구가 진행되었던 2014년 기준으로 한국의 국내총생산은 국제통화기금 집계로 약 1조 4,044억 달러, 세계은행 집계로는 1조 3,779 달러이다. 편의상 해당 값을 1조 4,000억 달러라 상정하면 0.4%는 56억 달러, 1.1%는 154억 달러이다. 이를 원화로 환산하면 달러당 1,200원 기준으로 각각 6조 7,200억 원과 18조 4,800억 원에 해당한다.

이를 이용하여 두 경우의 전체 고용 변화를 살펴보면, $6,720 \times 11.8 = 79,296$ 명이며, $18,480 \times 11.8 = 218,064$ 명이다. 다시 말해서 현재 수준의 데이터 현지화 조치는 전체 거시경제에서 약 8만여 명 정도의 고용 감소 효과를 유발하며, 전면적인 데이터 현지화 조치는 약 22만 명 정도의 고용 감소를 초래한다.

물론 이러한 계산은 두 가지 가정을 포함한다. 첫 번째로 2014년의 국

내총생산 변화와 2015년의 취업유발계수를 연결짓는 데 따른 일관성이다. 국내총생산의 영향은 2014년의 기준인 반면 취업유발계수는 2015년이며, 취업유발계수는 지속적으로 감소하는 추세를 보여주고 있으므로 보아 이 두 가지를 통한 계산이 정확한 값을 내지 못할 수 있다. 다른 하나는 연산 가능 일반균형의 결과에서 추정된 값에 오쿤의 법칙(Okun's law)을 적용하는 것이 타당한가의 문제이다. 특히나 연산 가능 일반균형모형을 통해 값을 얻었을 경우에는 총고용의 변화 역시 모형 내에서 산출 가능할 수 있는데 국내총생산만을 추계한 후 여기에 취업유발계수를 적용하는 것의 적합성에 관한 가정이다.

여기서 첫 번째 가정은 2014년과 2015년이 긴 시차가 있는 것이 아니며, 2010년에서 2015년 사이 취업유발계수의 변화가 동일한 시차의 여타 기간에 비해서 급격하거나 특이한 양상을 보인 것은 아니므로 큰 무리가 없다 할 수 있다. 또한 2015년 취업유발계수가 2014년에 비해 작을 수 있으므로, 본 연구에서 얻은 고용 감소치는 최소값에 가까울 것이며, 실제로는 이보다 더 클 수 있다는 사실을 열어놓는다면 큰 문제가 되지 않을 것이다.

두 번째 가정은 첫 번째보다 더 강한 조건을 내포한다. 연산 가능 일반균형에서 직접 고용 변화를 계산하지 않고 국내총생산에 대한 변화값을 가지고 오쿤의 법칙을 다시 사용하는 것은 실제 연산 가능 일반균형에서 얻은 값과 다르기 때문이다. 우선 Bauer et al.(2014)에서는 고용 변화에 대해서 연산을 하지 않았거나, 결과치를 보여주지 않았다. 해당 논문은 GTAP8 모형을 사용하여 국내총생산의 변화, 투자, 수출, 후생 감소 등을 모두 보여주었지만 고용 변화는 보여주지 않았다. 이는 GTAP8이 총인구와 노동력을 외생변수로 처리하는 모형이기 때문일 가능성이 높다. 따라서 이들 모형에서는 국내총생산의 변화나 데이터 현지화 정책에 따른 고용 변화를 잡아내기 어려웠을 것이다. 따라서 부득이하게 추계된 국내총생산 변화 추정치를 이용하여 고용 변화를 추정할 수밖에 없었다. 이렇게 추정된 국내총생산 변화에 오쿤의 법칙을 적용하는 것이 논리적으로 매끄럽게 연결되지는 않지만 차선택으로 이러한 방법을 선택하였음을 밝힌다.

라. 데이터 현지화 정책의 고용 효과

현재의 데이터 현지화 정책이 약 8만여 명 정도, 전면적인 데이터 현지화 정책이 약 22만 명 정도의 고용 감소 효과를 일으킬 수 있으며, 취업유발계수의 특성상 이 값은 실제 값보다 적게 추계되었을 가능성을 배제할 수 없다. 한편 앞선 1장에서 살펴본 데이터 산업의 고용량은 2015년 280,323명, 2016년 288,621명, 2017년 294,753명이며, 이 중 데이터 산업에서의 직접적인 고용은 각각 70,338명, 73,256명, 77,105명이다.

데이터 산업에서의 총고용 수치는 2015년 기준 7만여 명인데, 이는 앞서 추계한 현재의 데이터 현지화 정책하에서의 간접적인 고용 감소량인 79,296명보다 작다. 따라서 현재는 추계적 고용 감소 효과보다 산업 내 직접 고용량이 적으므로, 비관세 무역 장벽으로서의 데이터 현지화 정책의 순고용 효과는 음(-)의 값일 가능성이 높다고 할 수 있다.

국내 모든 데이터 산업이 유치 산업으로 수혜를 입어서 직접 고용량의 100%가 비관세 무역 장벽의 효과라 하더라도 그 값이 간접 고용에서의 음의 값을 초과하지 못하는 것이다. 또한 비관세 무역 장벽에서 나오는 음의 간접 고용 효과가 데이터 산업에서는 나타나지 않으므로 데이터 산업을 제외한 여타 산업들만을 가지고 취업유발계수를 적용할 경우 고용 감소량은 더 크게 증가한다. 이는 앞서 살펴본 <표 3-3>에서 정보통신 및 방송 서비스의 취업유발계수가 전 산업 평균보다 낮은 터여서, 해당 산업의 취업유발계수를 제외한 나머지 산업을 이용한 전 산업 평균값은 해당 표에서 얻은 값보다 더 커질 것이며, 이는 국내총생산 감소에 대한 고용 감소 효과를 더 크게 할 것이라는 의미이다.

이러한 결과의 한 이유는 본 장의 초반에서 언급한 빅데이터의 특성에서 기인하는 것일 수 있다. 데이터의 성격상 양질의 데이터가 더 많아지면 많아질수록 그 가치도 올라가고 정확성 혹은 신뢰성도 상승한다. 하지만 데이터 현지화 조치는 이를 위한 자유로운 데이터의 이동을 방해하므로 국내총생산 감소 효과가 크게 나타나는 것으로 보인다.

하지만 본 연구에서 살펴본 기간 동안에는 미국과 비교하여 후발 주자인 국내 데이터 산업 내 기업들이 유치 기간을 충분히 가지지 못함으로

인해 현재도 성장 단계일 뿐만 아니라 보다 규모가 커지면서 양의 직접 고용 효과가 향후 얼마든지 더 크게 나올 수 있기 때문에 본 연구의 결과가 선불리 데이터 현지화 정책의 즉각적인 폐지를 위한 주장으로 이어지거나 데이터 현지화 정책이 단기적으로건 장기적으로건 고용을 감소시키는 효과만을 낳을 뿐이라는 결론에 이르는 것은 성급하다 하겠다.

한편 본 소절의 연구 결과는 일반적인 비관세 무역 장벽의 고용 효과에 대한 논의와도 연결지을 수 있다. 데이터 현지화 정책 역시 전면 현지화와 전면 자유화 사이의 여러 정책이 있을 수 있으므로 명확하게 구분짓고 수치화할 수 있는 비관세 무역 장벽은 아니지만, 여타의 다른 비관세 무역장벽과 비교하여 상대적으로 명확하게 정도를 구분할 수 있는 편이다. 비관세 무역 장벽의 가장 큰 논리적 지지대는 유치 산업 보호와 자국 내 고용 증가이다. 이 중 유치 산업 보호는 본 연구의 초점이 아니므로 차치한다 해도, 본 연구에 따르면, 비관세 무역 장벽으로 자국 내 고용 증가 효과를 볼 수는 없는 것으로 나타났다. 오히려 비관세 무역 장벽을 제거하는 것이 해외 사업자들과 국내 사업자들 간의 경쟁을 촉발시키고 연관 산업에서의 비용 감소 및 중간재 투입물의 품질 증가를 가져와 거시경제의 성장에 도움이 될 여지가 있다.

제5절 소 결

본 장에서는 최근 급성장하는 데이터 산업에서 가장 주목받고 있는 빅데이터 산업의 노동시장을 살펴보았다. 빅데이터의 핵심적 특성으로는 고용량, 고속의 이동, 다양성, 신뢰성, 고부가가치를 꼽을 수 있는데, 이를 위해 다양한 정보를 취합하고 분석하는 데 소요되는 시간을 단축시키고 분석의 정확도를 올리기 위한 노력을 기울일 필요가 있다. 한편 이러한 고용량의 특성상 기존의 데이터 분석 방법이 빅데이터에서는 비효율적이거나 적용 불가능한 경우가 발생하므로 새로운 데이터 분석 및 처리 기법에 익숙한 인력들에 대한 수요 역시 급증할 것이 예측된다.

빅데이터 산업은 데이터 산업 내에서도 성장 속도가 빠를 뿐만 아니라 비교적 새로운 직무와 직종이 등장하는 등 급변하는 노동시장 중 하나이다. 특히 최고 데이터 관리자와 같은 직급의 등장은 기존의 기업 조직도 내에서도 빅데이터의 수집과 활용이 중요해지고 있음을 시사하며, 아울러 데이터 산업으로 인해 기존의 기업 조직도 변하고 있음을 보여준다.

빅데이터 관련 직종에서의 연봉 및 학력 수준을 살펴보면, 대체로 고학력 고임금의 직종이 많음을 알 수 있다. 또한 산업의 크기만큼이나 고용원 숫자 혹은 노동력 수요도 추후 빠르게 늘어날 것임을 알 수 있다. 그러나 미국과 비교하면 데이터 산업 분야의 고도화가 아직 이루어지지 않아 실제 데이터를 분석하고 활용하는 직군의 비율이 낮으므로 이러한 능력을 갖춘 인력을 양성하기 위한 방안을 찾아야 할 것이다.

한편 빅데이터의 핵심인 고용량과 빠른 이동 속도로 인해 많은 경우 분산 파일 저장 시스템이 효율적이나, 데이터 현지화 조치 등은 이 시스템의 핵심인 빠른 이동 속도에 제동을 건다. 물론 이러한 조치는 일정 부분 데이터 소유권을 명확히 하고 개인 정보를 보호하기 위한 취지를 담고 있어 무조건적인 데이터 이동 자유화가 바람직한 것만은 아니다. 하지만 데이터 현지화 조치의 목적이 유치 산업을 보호하고 해당 산업에서 직접 고용을 늘리는 데 있는 것이라면 적어도 본 장의 연구 결론상 데이터 현지화 조치에 대해 긍정적인 평가를 내리긴 어렵다. 간접 고용의 감소 효과가 직접 고용의 증가 효과를 상쇄하고도 남음에 따라 원하는 고용 증가 효과를 얻지 못할 가능성이 높기 때문이다. 따라서 데이터 현지화 정책을 완전히 폐지하는 것이 능사는 아니지만 그럼에도 데이터 현지화 정책의 확대는 신중하게 선택해야 하며, 기존의 현지화 정책에 대해서도 행정 주체의 자의적인 해석에 의하지 않고 법에 의해서 명확하게 현지화 범위를 설정하는 등의 노력을 기울여야 할 필요가 있을 것이다.

제 4 장

데이터 산업의 고용관계 이해

제1절 들어가는 말

데이터가 4차 산업혁명을 견인하는 핵심으로 떠오르면서 데이터 산업이 주목받고 있다. 데이터가 자본과 노동과 같은 기존의 생산요소를 능가하는 경쟁력으로 주목받으면서, 대규모 데이터를 보유하거나 이를 이용할 수 있는 기술을 가지고 있는 기업이 시장 혁신을 주도할 것이라고 보고 있다. 이는 우리나라뿐 아니라 세계 각국의 전반적인 흐름으로 최근 한국은 데이터 산업을 육성하기 위한 다양한 국가 정책을 수립하고 있다. 데이터 이용 제도의 패러다임을 전환하고, 양질의 데이터를 구축하기 위한 방안을 모색하며, 빅데이터와 같은 특정 분야의 육성을 위하여 산업 생태계 조성에 힘쓰는 등 다양한 국가 정책이 제시되고 있다(4차산업혁명위원회, 2018).

이와 같이 데이터 산업 발전을 유인하기 위해서 다양한 정책 지원들이 이루어져 있고 이에 따라 산업이 확장되고 있으며, 이에 기여하는 사업체와 종사자가 지속적으로 증가하고 있다. <표 4-1>을 보면, 데이터 산업의 사업체 수는 2017년 기준 총 약 23,000여 개에 다다르며, 관련 종사자 수 또한 30만 명이 넘는 것으로 나타난다. 그 안에 종사자 수도 지속적으로 증가하고 있다는 것을 알 수 있다. 그런데 현재 다수의 조사와 연구는

〈표 4-1〉 데이터 산업 사업체 수 및 종사자 수

(단위: 명)

구분	2016		2017	
	사업체 수	종사자 수	사업체 수	종사자 수
전 체	22,336	303,118	22,793	309,560
1~4명	12,051	24,569	12,887	25,560
5~9명	4,676	30,576	4,400	28,787
10~19명	2,866	37,953	2,727	36,215
20~49명	1,764	52,655	1,789	53,211
50~99명	542	36,796	556	37,331
100~299명	351	55,568	332	52,938
300~499명	54	21,045	65	24,832
500~999명	21	13,951	24	15,420
1000명 이상	11	24,833	13	29,938

자료: 국가통계포털(통계청)을 이용하여 저자 가공.

데이터 산업에 얼마의 인력이 필요하고, 인력 수급이 어느 정도 이루어지는지와 같은 규모 추정에 중점을 두고 있다. 그러나 무엇보다 데이터 산업에서 기업의 인력 수요 대비 공급이 원활하게 이루어지고 있지 않으며, 또한 인력이 공급된다 하여도 이직률이 높다는 점에서(2017년 기준, 중소기업 대상, 컴퓨터 프로그래밍 시스템 통합관리 및 관리업 12.1%, 정보서비스업 15.3% - 중소기업중앙회, 중소기업 실태조사, 2018: 278과 2019: 242), 인력 미스매치 문제가 대두되고 있는 상황이며, 이에 현 데이터 산업의 고용관계를 파악하고 이해하려는 시도가 필요하다.

데이터 산업을 대상으로 고용관계가 어떻게 전개되고 있는지 파악하기 위하여 데이터 산업에 속한 사업체를 대상으로 설문조사를 실시하였다. 설문조사를 실시한 이유는 통계청 조사 또는 각 기관들이 발간하는 「사업체패널조사」나 「기업인적자본패널조사」와 같은 공개 패널 데이터에서 데이터 산업을 분류할 수 없고 또한 데이터 산업에 국한하여 고용관계를 조사하고 있지 않기 때문이다. 설문조사를 통하여 데이터 산업에 속한 기업의 특성을 파악하고, 고용 규모를 비롯하여 근로자 구성, 근로 형태 등 고

용 구성에 대해 살펴보고자 하며, 사업체에 속한 근로자들을 어떠한 방식으로 채용하고 육성하고 동기부여하고 있는지 분석하고자 한다. 단, 현재까지의 산업 분류 방식에서 일반 데이터 산업과 빅데이터 산업을 분류하는 명확한 기준이나 분류 체계가 없으므로 본 장에서 구성되는 설문조사는 앞선 3장에서 살펴본 빅데이터 산업과 일반 데이터 산업을 따로 구분하지 않고 전체 데이터 산업에 대해서 진행하였음을 밝혀둔다.

제2절 설문조사 개요와 표본의 특성

1. 조사 방법 및 대상

조사는 데이터 산업에 속한 사업체를 대상으로 하였다. 통계청의 「한국 표준산업분류(10차)」에 따르면, 데이터 산업은 다음과 같다. 세세분류 기준을 적용하여, 시스템·응용 소프트웨어 개발 및 공급업, 컴퓨터 프로그래밍 서비스업, 컴퓨터 시스템 통합 자문, 구축, 및 관리업, 기타 정보기술 및 컴퓨터 운영 관련 서비스업, 자료 처리, 호스팅 및 관련 서비스업, 포털 및 기타 인터넷 정보 매개 서비스업, 뉴스 제공업, 그 외 기타 정보 서비스업 등을 포함한다.

- (5822) 시스템·응용 소프트웨어 개발 및 공급업
- (6201) 컴퓨터 프로그래밍 서비스업
- (6202) 컴퓨터 시스템 통합 자문, 구축, 및 관리업
- (6209) 기타 정보기술 및 컴퓨터 운영 관련 서비스업
- (6311) 자료 처리, 호스팅 및 관련 서비스업
- (6312) 포털 및 기타 인터넷 정보 매개 서비스업
- (6391) 뉴스 제공업
- (6399) 그 외 기타 정보 서비스업

조사는 고용보험데이터를 활용하여 데이터 산업 사업체의 모집단을 추출하였고 이를 기준으로 표본을 구성하였다. 조사는 2019년 9월 1일부터 10월 31일까지 약 2달간 이루어졌으며, 응답은 각 사업체의 인사담당자를 통하여 이루어졌다.

2. 표본의 특성

조사된 표본의 특성을 보면 다음과 같다. 먼저, 업종별 표본 수를 보면, 시스템 소프트웨어 개발 및 공급업이 39.5%로 가장 많은 것으로 나타나며, 다음으로 응용 및 소프트웨어 개발 및 공급업이 35.5%로 많은 것으로 나타났다(표 4-2 참조).

지역별 분포를 보면, 서울에 있는 사업체(69.8%)가 가장 많은 것으로 나타났으며, 다음으로 경기 지역(14.4%)에 사업체가 많은 것으로 나타났다(표 4-3 참조).

〈표 4-2〉 업종별 표본 수

(단위: 개, %)

업종	표본 수	비율
시스템 소프트웨어 개발 및 공급업	178	39.5
응용 소프트웨어 개발 및 공급업	160	35.5
컴퓨터 프로그래밍 서비스업	13	2.9
컴퓨터 시스템 통합 자문 및 구축 서비스업	22	4.9
기타 정보기술 및 컴퓨터 운영 관련 서비스업	27	6.0
자료 처리업	2	0.4
호스팅 및 관련 서비스업	2	0.4
포털 및 기타 인터넷 정보 매개 서비스업	10	2.2
뉴스 제공업	4	0.9
데이터베이스 및 온라인 정보 제공업	4	0.9
그 외 기타 정보 서비스업	29	6.4
전 체	451	100

자료: 저자 작성.

80 데이터 산업의 노동시장 분석

그리고 기업 규모 분포를 보면(표 4-4 참조), 종업원 수를 기준으로 볼 때 10~99인 사업체(46.8%)가 가장 많은 것으로 나타났으며, 다음으로 1~9인 사업체가 32.2%를 차지하는 것으로 나타났다. 그리고 300인 이상 사업체는 전체 표본 중 1.3%에 머무는 것을 알 수 있다.

〈표 4-3〉 지역별 표본 수

(단위: 개, %)

지역	표본 수	비율
서울	315	69.8
부산	12	2.7
대구	14	3.1
인천	4	0.9
광주	6	1.3
대전	16	3.5
경기	65	14.4
강원	2	0.4
충북	1	0.2
충남	2	0.4
전북	2	0.4
전남	2	0.4
경북	2	0.4
경남	7	1.6
제주	(1)	0.2
전체	451	100

자료: 저자 작성.

〈표 4-4〉 기업 규모별 표본 수

(단위: 개, %)

기업 규모	표본 수	비율
1~9인	145	32.2
10~99인	211	46.8
100~299인	89	18.7
300인 이상	6	1.3
전체	451	100.0

자료: 저자 작성.

마지막으로 <표 4-5>에서 표본 사업체의 업력 구성을 보면, 전체 사업체의 업력은 11.9, 약 12년 정도 되는 것으로 나타났다. 기업 규모별로 보면, 1~9인 사업체가 약 8년, 10~99인 사업체가 약 13년, 100~299인 사업체가 약 15년, 그리고 300인 이상 사업체가 약 17년으로 기업 규모의 증가와 업력의 증가가 함께 이루어짐을 알 수 있다.

<표 4-5> 표본 사업체 업력

(단위:년)

구분	평균	표준편차	사업체 수
전 체	11.9	6.90	451
기업 규모	1~9인	8.3	145
	10~99인	12.9	211
	100~299인	14.8	89
	300인 이상	17.1	6

자료: 저자 작성.

제3절 데이터 산업 사업체 특성

본격적으로 데이터 산업의 고용관계를 파악하기 전에 고용관계에 대한 이해를 높이기 위하여 이에 속한 사업체들이 어떠한 특성을 가지고 있는지 살펴보았다. 사업체의 속성은 고용관계에 영향을 미칠 수 있으며 또한 고용관계를 이해하는 데 맥락적 요인으로 작용한다는 점에서 살펴볼 필요가 있다.

1. 경영 체계

먼저, 사업체의 거버넌스를 파악하기 위해 의사결정 권한이 누구에게 있는지 살펴보았다(표 4-6 참조). 전체 사업체 중 소유 경영 체제를 가지고 있는 사업체가 94.2%인 것으로 나타났으며, 전문경영인에게 회사 경영

〈표 4-6〉 사업체 경영 체제

(단위: 개, %)

항목	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
소유 경영 체제	425	94.2	145	100	206	97.5	73	82.0	1	16.7
전문경영인이 말지만, 권한 이양이 미약한 소유주 중심 경영 체제	2	0.4	0	0	1	0.5	1	1.1	0	0
전문경영인에서 대폭 권한을 이양 하였지만, 여전히 경영 문제에 대한 결정권을 소유주가 지닌 체제	13	2.9	0	0	3	1.4	9	10.1	1	16.8
소유와 경영이 독립된 전문 경영 체제	11	2.4	0	0	1	0.5	6	6.7	4	66.7
전 체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

권한을 대폭 이양하고 있지만, 임원 인사, 신규 투자 등 주요 경영 문제에 대한 결정권은 여전히 소유주가 지니고 있는 체제가 2.9%로 다음으로 많은 것으로 나타났다. 이러한 결과는 기업 규모별로 구분하여 보았을 때도 동일한 경향을 보인다고 볼 수 있는데, 1~9인 사업체는 100%가 소유 경영 체제라고 응답하였으며, 10~99인 이상 사업체는 97.5%, 100~299인 사업체는 82.0%가 소유 경영 체제라고 응답하였다. 그러나 300인 이상 사업체의 경우는 6개 기업 중 66.7%가 소유와 경영이 독립된 전문 경영 체제를 가진다고 응답하여 299인 이하 규모의 사업체와 의사결정 거버넌스가 다르게 운영되고 있음을 알 수 있다.

다음으로 해당 사업체의 외국인 자본이 어느 정도 투자되고 있는지 조사한 결과(표 4-7 참조), 전체적으로 외국인 지분 비율은 0.54%인 것으로 나타나 외국인 자본이 투자되고 있는 비중이 매우 작은 것을 알 수 있다.

〈표 4-7〉 외국인 지분율

(단위: %)

구분		평균	표준편차	최소	최대
전 체		0.54	5.72	0	80
기업 규모	1~9인	0.55	6.64	0	80
	10~99인	0.40	5.52	0	80
	100~299인	0.87	4.72	0	40
	300인 이상	0	0	0	0

자료: 저자 작성.

이는 기업 규모로 구분하여 보아도 크게 다른 경향을 보이지 않는 것을 알 수 있으며, 300인 이상 사업체의 경우 외국인 지분 비율이 0%인 것으로 나타났다.

2. 하도급 거래 여부 및 기업 간 네트워크

사업체가 기업 간 네트워크에서 어떠한 위치에 있는지 알아보기 위하여 거래관계에 대하여 조사하였다. 먼저, 본 기업이 거래관계에 있어 원청 기업에 해당하는지 아니면 하청 기업일 경우 어느 위치에 존재하는지 물은 결과(표 4-8 참조), 전체 사업체 중 51.29%는 원청 기업에 해당하는 것으로 나타났다. 그리고 그 외 기업은 26.8%가 1차 벤더이고, 12.4%가 2차 벤더인 것으로 나타났다. 기업 규모별로 살펴보면, 1~9인 사업체 중 32.4%는 원청 기업이라고 응답하였으며, 29%는 3차 이상 하청 기업이고, 26.9%가 2차 벤더에 해당하는 것으로 나타났다. 10~99인 이상 사업체를 보면 50.2%가 원청 기업에 해당하며, 다음으로 1차 벤더(41.2%)가 많은 것으로 나타났다. 100~299인 이상 사업체도 유사한데 전체 중 80.9%가 원청 기업에 해당하며, 나머지 19.1%는 1차 하청 기업이라고 응답하였다. 300인 이상 사업체는 모두 원청 기업인 것으로 나타났다. 1~9인 사업체도 원청 기업의 비율이 높기는 하지만 다른 규모 기업군과 비교하면 2, 3차 벤더인 경향이 더욱 높은 것을 알 수 있다.

〈표 4-8〉 하도급 거래

(단위: 개, %)

규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
원청	231	51.2	47	32.4	106	50.2	72	80.9	6	100
1차 벤더	121	26.8	17	11.7	87	41.2	17	19.1	0	0
2차 벤더	56	12.4	39	26.9	17	8.1	0	0	0	0
3차 이상	43	9.5	42	29.0	1	0.5	0	0	0	0
전체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

〈표 4-9〉 원청 기업에 대한 매출 비중

(단위: %)

규모	전체		1~9인		10~99인		100~299인	
	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
0~20%	17	7.7	11	11.2	4	3.8	2	11.8
20~40%	20	9.1	12	12.2	7	6.7	1	5.9
40~60%	46	20.9	26	26.5	8	7.6	12	70.6
60~80%	36	16.4	18	18.4	18	17.1	0	0
80%	101	45.9	31	31.6	68	64.8	2	11.8

자료: 저자 작성.

이어서, 1, 2, 3차 하청기업이라고 응답한 사업체를 대상으로 주요 원청 기업에 대한 매출이 전체 매출액 중 차지하는 비중을 조사하였다. 조사 결과(표 4-9 참조), 전체를 보았을 때 매출액의 80% 이상을 차지한다고 응답한 기업이 45.9%로 가장 많은 것으로 나타났으며, 다음으로 40~60% 정도 매출액을 차지한다고 응답한 기업이 20.9%로 많은 것으로 나타났다. 기업 규모별로 구분하여 살펴보면, 1~9인의 경우, 매출액의 80% 이상을 차지한다고 응답한 비율이 31.6%이며, 10~99인 기업도 마찬가지로 매출액의 80% 이상 차지한다고 응답한 사업체 비중이 64.8%인 것으로 높게 나타났다. 이는 전체와 유사한 경향을 보인다고 볼 수 있다. 이와

달리 100~299인은 매출액 비중이 40~60%라고 응답한 사업체가 70.6%로 가장 많은 것으로 나타났다.

3. 노사관계 및 고용안정 수준

사업체의 노사관계를 알아보기 위하여(표 4-10 참조), 먼저 노동조합이 있는지 조사하였다. 조사 결과, 노동조합이 있다고 응답한 사업체는 4.4%이고, 노동조합이 없다고 응답한 사업체가 95.6%인 것으로 나타나 다수의 사업체에 노동조합이 없다는 것을 알 수 있다. 기업 규모별로 보면, 1~9인인 경우 노동조합이 있다고 응답한 사업체가 없으며, 10~99인인 경우 2.8% 사업체가 노동조합이 있다고 응답하였다. 100~299인의 경우는 11.2%만 노동조합이 있다고 응답하였다. 이는 전체의 경향과 다르지 않은 것으로 볼 수 있다. 그러나 300인 이상 경우는 전체 중 66.7%가 노동조합이 있다고 응답하였으며, 33.3% 사업체가 노동조합이 없다고 응답하여 전체 경향과 다르다고 볼 수 있다.

그리고 사업체의 노사관계가 어떠한지 인식을 묻은 결과, 전체를 보면, 매우 나쁨이나 나쁜 편이라고 응답한 사업체는 없는 것으로 나타났다. 전체 중 46.6%가 노사관계는 보통이라고 응답하였으며, 42.4%가 노사관계는 좋은 편이라고 응답하였다. 기업 규모별로 보면, 1~9인인 경우 44.1%가 노사관계가 좋은 편이라고 응답하였으며, 10~99인 경우 53.1%가 좋은 편이라고 응답하였다. 100~299인은 전체 중 87.6%가 노사관계는 보통이라고 응답하였으며, 300인 이상 사업체의 경우 66.7%가 노사관계가 좋은 편이라고 응답하였다. 전체적으로 노사관계는 긍정적으로 평가하고 있는 것을 알 수 있다.

다음으로 최근 3년간 정규직 근로자에 대한 인력감축을 실시한 적 있는지 묻은 결과(표 4-11 참조), 전체를 보면, 99.6%가 정리해고를 실시한 적이 없다고 응답하였다. 기업 규모별로 보면, 1~9인인 경우 99.3%가 정리해고를 실시한 적이 없다고 응답하였으며, 10~99인은 99.5%, 100~299인 100%, 300인 이상 사업체도 100%라고 동일하게 응답하였다. 전체적으로 정리해고는 이루어지지 않는다고 볼 수 있다.

〈표 4-10〉 노동조합 및 노사관계 평가

(단위: 개, %)

규모		전체		1~9인		10~99인		100~299인		300인 이상	
항목		사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
노동조합	있음	20	4.4	0	0	6	2.8	10	11.2	4	66.7
	없음	431	95.6	145	100	205	97.2	79	88.8	2	33.3
노사관계	매우 나쁨	0	0	0	0	0	0	0	0	0	0
	나쁜 편임	0	0	0	0	0	0	0	0	0	0
	보통임	210	46.6	54	29.7	87	41.2	78	87.6	2	33.3
	좋은 편임	191	42.4	64	44.1	112	53.1	11	12.4	4	66.7
	매우 좋음	50	11.1	38	26.2	12	5.7	0	0	0	0

자료: 저자 작성.

〈표 4-11〉 고용안정

(단위: 개, %)

규모		전체		1~9인		10~99인		100~299인		300인 이상	
항목		사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
정리해고	실시	2	0.4	1	0.7	1	0.5	0	0	0	0
	미실시	449	99.6	144	99.3	210	99.5	89	100	6	100
합의사항	있음	79	17.5	15	10.3	61	28.9	4	3.4	0	0
	없음	372	82.5	130	89.7	150	71.1	86	96.6	6	100

자료: 저자 작성.

추가적으로 해당 사업체에 정규직 근로자를 경영상의 이유로 해고하지 않는다는 경영 정책이나 합의 사항을 별도로 두고 있는지 묻은 결과, 전체를 보면, 82.5%가 관련 정책이나 합의 사항이 없다고 응답하였고, 17.5%가 관련 사항을 두고 있다고 응답하였다. 기업 규모별로 보면, 1~9인은 89.7%, 100~299인은 96.6%, 300인 이상은 100%가 관련 정책이나 합의 사항이 없다고 응답하였다. 이를 종합하여 보면, 고용안정에 관하여 정책적으로 보호하고 있지는 않으나 사업체에서 고용안정이 이루어지고 있다고 볼 수 있다.

제4절 데이터 산업 사업체 고용 현황

다음으로 데이터 산업의 사업체 내 인력 구성을 살펴보기 위하여 정규직, 비정규직 등 고용 형태와 데이터 관련⁶⁾, 데이터 비관련 직무⁷⁾로 구분하여 인력 현황을 파악하였다.

1. 고용 구성

가. 상용직 근로자 구성

먼저, 정규직 근로자의 고용 구성을 살펴보았다(표 4-12 참조). 전체를 보면, 고용 규모는 평균으로 보았을 때 2016년 58.4명, 2017년 59.9명, 2018년 60.8명인 것으로 나타나 큰 변화가 없는 것을 알 수 있다. 성별로 보면, 남성 근로자가 41.7명, 여성 근로자가 16.8명으로 남성 근로자 비중이 높은 것을 알 수 있으며, 연령을 구분해 보면 20~39세가 약 40명 정도 되고, 40~59세가 약 18명, 그리고 60세 이상이 1명도 안 되는 것으로 나타났다. 외국인 근로자 또한 1명이 안 되는 것으로 나타났다.

기업 규모를 나누어 살펴보면, 1~9인 사업체의 경우, 남성이 약 3.7명, 여성이 1명이 안 되는 것으로 나타났다. 그리고 연령을 보면 20~39세가 약 3명, 40~59세가 1.5명, 60세 이상이 1명도 되지 않는 것으로 나타났으며, 외국인 근로자는 없는 것으로 나타났다. 10~99인 사업체를 보면, 남성 근로자가 약 37명, 여성 근로자가 약 15명 정도인 것으로 나타났다. 연령을 보면, 20~39세가 약 38명, 40~59세가 약 14명, 60세 이상이 1명도 되지 않는 것으로 나타났고, 외국인 근로자는 없는 것으로 나타났다. 100~299인을 보면, 조사 연간에 걸쳐 남성 근로자가 86~98명 정도 되고,

6) 데이터 아키텍트, 데이터 개발자, 데이터 엔지니어, 데이터 분석가, 데이터베이스 관리자, 데이터 과학자, 데이터 컨설턴트, 데이터 기획자를 가리킨다.

7) 데이터 관련 업무 이외 모든 업무로 예를 들어, 인사, 회계/재무, 구매, 영업 등을 의미한다.

여성 근로자가 34~39명 정도 되는 것으로 나타났다. 연령을 보면, 20~39세가 85~94명, 40~59세가 35~41명, 60세 이상이 1명이 안 되는 것으로 나타났으며, 외국인 근로자는 없다. 마지막으로, 300인 이상 사업체의 경우, 조사 연간에 걸쳐 남성 근로자가 342~346명, 여성 근로자가 176~179명 정도인 것으로 나타났다. 연령을 구분하여 보면, 20~39세가 313~317명, 40~59세가 200~203명, 60세 이상이 5~6명 정도 되는 것으로 나타났으며, 외국인 근로자는 없는 것으로 나타났다.

종합하여 보면, 사업체는 남성 근로자가 2/3 정도 구성하는 것을 알 수 있고, 연령대는 20~39세가 많아 조직이 젊으며, 외국인 근로자는 없는 것을 알 수 있다.

나. 비정규직 고용

다음으로 비정규직 고용을 살펴보았다(표 4-13 참조). 비정규직 고용 규모는 2016년 0.8명, 2017년 1.11명, 2018년 1.23명으로 사업체당 약 한 명 정도 비정규직을 고용하고 있는 것으로 나타났다. 그리고 상용직 대비 비정규직 근로자비율은 2017년 2.70%, 2017년 2.74%, 2018년 2.68%인 것으로 나타났다. 전반적으로 비정규직 규모가 매우 작다는 점에서 성별이나 연령별 구성을 살펴보지 않고, 비정규직 규모와 상용직 대비 비정규직 근로자 비율을 살펴보도록 한다.

기업 규모별로 보면, 1~9인 사업체의 비정규직 수는 2016년 약 0.14명 정도이고 비정규직 비율은 2016년 6.73%인데, 2017년 5.90%, 2018년 5.66%로 감소하는 것으로 나타났다. 10~99인 사업체를 보면, 비정규직 근로자 수는 약 0.50정도로 1명이 안 되는 것으로 나타났으며, 상용직 근로자 대비 비정규직 근로자 비율은 2016년 0.73%, 2017년 0.80%, 2018년 0.84%로 비중이 매우 작고 변화도 없는 것으로 나타났다. 100~299인 사업체의 비정규직 근로자 수는 2016년 2.25명, 2017년 3.79명, 2018년 4.21명으로 증가하는 것으로 나타났으며, 비율도 2017년 1.61%, 2017년 2.52%, 2018년 2.34%로 증가하는 것으로 나타났다. 마지막으로 300인 이상 사업체를 보면, 비정규직 근로자 수는 약 5명 정도 되는 것으로 나타

〈표 4-12〉 상용직 근로자 구성

(단위: 명)

구분	연도		2106		2017		2018	
	전체		평균	표준편차	평균	표준편차	평균	표준편차
전체	성별	남성	41.7	50.7	42.0	51.4	43.1	52.6
		여성	16.8	25.9	17.1	26.2	17.7	26.6
	연령	20~39세	40.9	48.6	41.0	48.6	42.3	49.9
		40~59세	17.3	27.8	17.8	28.6	18.3	28.7
		60세 이상	0.2	1.3	0.3	1.5	0.2	1.3
외국인		0.0	0.1	0.2	3.9	0.1	0.8	
1~9인	성별	남성	3.7	1.9	3.7	2.0	3.8	2.1
		여성	0.8	1.1	0.9	1.1	0.9	1.2
	연령	20~39세	3.0	1.9	3.2	2.0	3.3	2.2
		40~59세	1.5	1.2	1.4	1.1	1.5	1.1
		60세 이상	0.0	0.1	0.0	0.1	0.0	0.1
외국인		0.0	0.0	0.0	0.0	0.0	0.0	
10~99인	성별	남성	36.9	19.7	37.4	19.8	38.5	20.6
		여성	14.3	10.6	14.9	10.8	15.5	11.2
	연령	20~39세	37.3	18.5	37.9	18.8	39.2	19.4
		40~59세	13.7	8.1	14.0	8.1	14.6	8.1
		60세 이상	0.2	0.7	0.3	1.4	0.2	0.7
외국인		0.0	0.1	0.5	5.6	0.1	1.1	
100~299인	성별	남성	86.9	24.2	91.9	25.5	97.7	27.1
		여성	34.7	13.6	36.4	15.4	39.4	16.7
	연령	20~39세	85.6	22.2	88.8	22.1	94.9	23.7
		40~59세	35.7	14.3	39.3	17.7	41.9	18.6
		60세 이상	0.2	1.0	0.2	1.2	0.2	1.2
외국인		0.0	0.1	0.0	0.1	0.0	0.2	
300인 이상	성별	남성	342.5	91.9	344.7	93.1	346.8	93.4
		여성	176.0	75.0	178.5	75.3	179.0	75.6
	연령	20~39세	313.5	120.2	314.2	119.4	317.0	118.0
		40~59세	200.0	74.2	203.7	75.4	203.2	73.2
		60세 이상	5.0	8.4	5.3	7.3	5.7	8.1
외국인		0.0	0.0	0.0	0.0	0.0	0.0	

주: 2016 N=418, 2017 N=438, 2018 N=451

자료: 저자 작성.

〈표 4-13〉 비정규직 근로자 구성

(단위: 명)

연도		2106		2017		2018	
구분		평균	표준편차	평균	표준편차	평균	표준편차
비 정규직 수	전 체	0.8	5.35	1.11	9.88	1.23	10.37
	1~9인	0.14	0.50	0.13	0.50	0.13	0.53
	10~99인	0.46	3.36	0.51	3.46	0.62	3.62
	100~299인	2.25	10.0	3.79	21.0	4.21	22.2
	300인 이상	5.00	12.2	4.83	11.8	5.00	12.2
비 정규직 비율	전 체	2.70	15.1	2.74	14.0	2.68	13.4
	1~9인	6.73	26.1	5.90	21.7	5.66	21.3
	10~99인	0.73	4.63	0.80	4.66	0.84	4.29
	100~299인	1.61	6.82	2.52	13.2	2.34	10.6
	300인 이상	0.78	1.93	0.76	1.86	0.78	1.91

주: 2016 N=418, 2017 N=438, 2018 N=451

자료: 저자 작성.

났으며, 상용직 대비 비중을 보면, 대략 0.78%로 낮고 이 또한 연도에 따른 변화가 없는 것으로 나타났다.

종합적으로 보면, 100~299인 사업체를 제외하고는 사업체의 상용직 근로자 대비 비정규직 활용 비중이 매우 낮고 규모의 변화가 크게 없거나 감소하는 것을 알 수 있다.

다. 데이터 관련 직무 근로자 구성

〈표 4-14〉는 데이터 관련 직무를 수행하는 근로자의 구성에 관한 조사 결과이다. 먼저 전체를 보면, 사업체별로 평균 2016년 34명, 2017년 35명, 2018년 36명 정도 되는 것으로 나타났으며, 전체 상용직 근로자 중 비중을 보면 2016년 59.4%, 2017년 59.2%, 2018년 59.1%로 전체 근로자 중 데이터 관련 직무를 수행하는 근로자 비중의 변화는 절반이 넘고, 변화는 거의 없는 것으로 나타났다. 성별로 보면, 조사 연간에 걸쳐 남자가 약 28~

29명, 여자가 약 6~7명으로 남성 근로자가 80%가량 차지하는 것으로 나타났다. 연령으로 보면, 20~39세가 26~27명, 40~59세가 8명, 60세 이상이 0명으로 나타나 근로자 중 다수가 20~39세인 것으로 나타났다.

기업 규모별로 살펴보면, 1~9인 사업체의 데이터 관련 직무 수행자는 약 3명인 것으로 나타났으며, 비중으로 보면, 2016년 75.9%, 2017년 73.9%, 2018년 74.0%인 것으로 나타났다. 성별을 구분하여 보면, 남성이 약 3명, 여성이 1명도 안 되는 것으로 나타났으며, 연령으로 보면 20~39세가 약 2명, 40~59세가 약 1명, 60세 이상이 0명인 것으로 나타났다. 또한 외국인도 없는 것으로 나타났다. 10~99인 사업체의 데이터 관련 직무 수행자는 약 22명인 것으로 나타났으며, 비중으로 보면, 2016년 21.8%, 2017년 22.7%, 2018년 22.8%인 것으로 나타났다. 성별을 구분하여 보면, 조사 연간에 걸쳐 남성이 약 18~19명, 여성이 3~4명인 것으로 나타났으며, 연령으로 보면 20~39세가 약 17~18명, 40~59세가 약 4~5명, 60세 이상이 0명인 것으로 나타났고, 외국인도 없는 것으로 나타났다. 100~299인 사업체의 데이터 관련 직무 수행자는 조사 연간에 걸쳐 약 89~98명인 것으로 나타났으며, 비중으로 보면, 2016년 72.5%, 2017년 72.1%, 2018년 71.6%인 것으로 나타났다. 성별을 구분하여 보면, 남성이 약 71~74명, 여성이 17~19명인 것으로 나타났으며, 연령으로 보면 20~39세가 약 67~73명, 40~59세가 약 21~24명, 60세 이상이 없거나 1명있는 것으로 나타났다. 외국인은 없는 것으로 나타났다. 마지막으로, 300인 이상 사업체의 데이터 관련 직무 수행자는 약 330명인 것으로 나타났으며, 비중으로 보면, 2016년 61.8%, 2017년 61.8%, 2018년 61.5%인 것으로 나타났다. 성별을 구분하여 보면, 조사 연간에 걸쳐 남성이 약 257~259명, 여성이 72~74명인 것으로 나타났으며, 연령으로 보면 20~39세가 약 243~245명, 40~59세가 약 84~91명, 60세 이상이 없거나 1명 있는 것으로 나타났다. 외국인은 없는 것으로 나타났다.

종합적으로 보면, 전체 상용직 근로자 대비 데이터 관련 직무 수행 근로자 비율은 10~99인 사업체를 제외하고 절반이 훨씬 넘는 것을 알 수 있으며, 데이터 직무 수행 근로자의 2/3가 남성이고, 20~30세라고 볼 수 있다.

〈표 4-14〉 데이터 관련 직무 근로자 구성

(단위: 명)

연도		2016		2017		2018		
구분		평균	표준편차	평균	표준편차	평균	표준편차	
전체	성별	남성	28.1	44.0	28.5	44.3	29.3	45.2
		여성	6.5	12.5	6.5	12.6	6.7	13.0
	연령	20~39세	26.4	41.6	26.5	41.3	27.3	42.0
		40~59세	8.1	14.6	8.5	15.4	8.7	15.8
		60세 이상	0.0	0.2	0.0	0.2	0.0	1
	외국인		0.0	0.0	0.0	0.0	0.0	0
	전체		34.5	54.5	35.1	54.9	36.1	56.1
	데이터 관련 직무 비율		59.4	31.9	59.2	31.6	59.1	31.7
1~9인	성별	남성	2.9	1.7	2.9	1.8	3.0	1.9
		여성	0.3	0.6	0.3	0.6	0.3	0.6
	연령	20~39세	2.3	1.6	2.3	1.7	2.4	1.9
		40~59세	0.9	0.9	0.9	0.8	0.9	0.8
		60세 이상	0.0	0.1	0.0	0.1	0.0	0.1
	외국인		0.0	0.0	0.0	0.0	0.0	0.0
	전체		3.2	1.9	3.2	1.9	3.3	2.0
	데이터 관련 직무 비율		75.9	26.7	73.9	26.5	74.0	26.0
10~99인	성별	남성	18.4	19.0	18.5	11.5	19.8	12.2
		여성	3.46	5.42	3.64	5.50	3.75	5.70
	연령	20~39세	17.1	18.0	17.8	18.4	18.7	19.3
		40~59세	4.62	5.58	4.86	5.86	4.95	5.90
		60세 이상	0.0	0.2	0.0	0.0	0.0	0.0
	외국인		0.0	0.0	0.0	0.0	0.0	0.0
	전체		21.8	22.6	22.7	23.2	22.8	22.2
	데이터 관련 직무 비율		43.4	32.5	43.8	32.9	43.5	33.2

〈표 4-14〉의 계속

연도		2016		2017		2018		
구분		평균	표준편차	평균	표준편차	평균	표준편차	
100~ 299인	성별	남성	71.2	24.6	74.2	29.3	74.8	31.5
		여성	17.8	11.3	18.5	11.5	19.8	12.2
	연령	20~39세	67.8	22.3	69.9	22.5	73.6	24.6
		40~59세	21.3	13.2	23.5	14.8	24.9	15.2
		60세 이상	0.0	0.0	0.7	1.0	0.7	1.0
	외국인	0.0	0.0	0.0	0.0	0.0	0.0	
	전체	89.1	29.0	93.2	29.2	98.6	31.7	
데이터 관련 직무 비율	72.5	15.6	72.1	15.7	71.6	15.9		
300인 이상	성별	남성	257.3	148.1	259.0	149.3	259.3	150.9
		여성	72.8	28.6	74.2	29.3	74.8	31.5
	연령	20~39세	245.0	132.9	244.2	132.4	242.2	132.5
		40~59세	84.8	41.9	88.3	40.4	91.3	39.5
		60세 이상	0.3	0.5	0.7	1.0	0.7	1.0
	외국인	0.0	0.0	0.0	0.0	0.0	0.0	
	전체	330.2	160.1	333.2	161.3	334.2	165.3	
데이터 관련 직무 비율	61.8	25.5	61.8	25.5	61.5	26.3		

주: 2016 N=418, 2017 N=438, 2018 N=451

자료: 저자 작성.

2. 근로자 학력 수준, 근속기간 및 이직률

가. 정규직 근로자 학력 수준

근로자 학력 구성은 데이터 관련 직무를 수행하는 근로자와 데이터 비관련 직무를 수행하는 근로자의 학력 수준을 구분하여 살펴보았다. 먼저, 데이터 관련 직무 수행 근로자의 학력 구성을 보면(표 4-15 참조), 대졸 학력(24.0명)로 다수를 차지하는 것으로 나타났으며, 다음으로 전문대졸 학력(7.3명)이 많은 것으로 나타났다. 기업 규모별로 보면, 1~9인 경우 대졸

〈표 4-15〉 데이터 관련 직무 수행 근로자 학력 수준

(단위: 명)

구분	고졸 이하		대졸		전문대졸		국내 석사		국내 박사		해외 석박사		
	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	
전 체	1.0	2.5	24.0	37.0	7.3	13.5	2.6	9.7	1.2	10.2	0.1	0.7	
기업 규모	1~9인	0.2	0.6	1.6	1.8	1.3	1.3	0.1	0.4	0.0	0.1	0.0	0.4
	10~99인	0.5	2.1	17.6	18.8	4.3	6.2	1.2	1.9	0.1	0.4	0.0	0.1
	100~299인	3.2	2.9	64.6	26.1	22.3	16.5	6.2	6.7	2.2	4.1	0.1	0.6
	300인 이상	6.0	7.9	184.7	141.5	34.7	61.1	55.5	60.9	50.2	77.7	3.2	4.7

자료: 저자 작성.

학력이 1.6명으로 가장 많으며, 전문대졸이 1.3명으로 그다음인 것으로 나타났다. 10~99인 경우 대졸 학력이 17.6명으로 가장 많으며, 다음으로 전문대졸이 4.3명으로 많은 것으로 나타났다. 100~299인을 보면 대졸 학력이 64.6명으로 가장 많으며, 전문대졸이 22.3명으로 다음을 차지하였다. 300인 이상 사업체는 대졸 학력이 184.7명으로 가장 많으며, 다음으로 국내 석사가 55.5명, 국내 박사가 50.2명으로 많은 것으로 나타났다. 300인 이상 사업체를 제외하고 정규직 근로자들의 학력 수준은 유사한 것을 알 수 있다.

다음으로, 데이터 비관련 직무 수행 근로자의 학력 구성을 보면(표 4-16 참조), 대졸 학력(18.9명)이 다수를 차지하는 것으로 나타났으며, 다음으로 전문대졸 학력(6.6명)이 많은 것으로 나타났다. 기업 규모별로 보면, 1~9인 경우 전문대졸 학력이 1.2명으로 가장 많으며, 10~99인 경우 대졸 학력이 20.8명으로 가장 많으며, 다음으로 전문대졸이 7.2명으로 많은 것으로 나타났다. 100~299인을 보면 대졸 학력이 28.1명으로 가장 많으며, 전문대졸이 10.8명으로 다음을 차지하였다. 300인 이상 사업체는 대졸 학력이 125.8명으로 가장 많으며, 다음으로 국내 석사가 33.7명, 국내 박사가 10.8명으로 많은 것으로 나타났다. 데이터 관련 직무 수행 근로자와 달리 전반적으로 학력 수준은 유사하다고 볼 수 있다.

〈표 4-16〉 데이터 비관련 직무 수행 근로자 학력 수준

(단위: 명)

구분	고졸 이하		대졸		전문대졸		국내 석사		국내 박사		해외 석박사	
	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
전 체	1.6	2.8	18.9	26.8	6.6	6.5	1.4	6.5	0.2	2.0	0.0	0.5
기업 규모	1~9인	0.2	0.4	0.8	1.1	1.2	1.3	0.0	0.0	0.0	0.0	0.0
	10~99인	1.6	2.8	20.8	18.3	7.2	5.6	1.4	2.0	0.1	0.4	0.0
	100~299인	2.8	2.2	28.1	30.9	10.8	7.0	0.9	1.6	0.1	0.4	0.0
	300인 이상	8.7	9.0	125.8	75.2	15.5	14.1	33.7	44.4	10.8	13.1	2.2

자료: 저자 작성.

나. 정규직 근로자 근속기간 및 이직률

근로자들의 조직 유지 경향을 알아보기 위하여 근속기간과 이직률을 조사하였다(표 4-17 참조). 먼저, 데이터 관련 직무 근로자의 근속기간은 2016년 4.14년 2017년 4.45년 2018년 3.83년으로 2018년에 줄어든 것으로 나타났다. 기업 규모별로 보면, 1~9인은 2016년 3.47년에서 2018년 4.39년으로 증가한 것으로 나타났으며, 10~99인은 2016년 4.53년, 2017년 3.95년, 2018년 5.39년으로 2017년에는 감소하였지만 2018년에 증가한 것으로 나타났다. 100~299인은 2016년 4.99년에서 2018년 4.12년으로 큰 변화가 없었고, 300인 이상은 근속연수가 7.5년으로 변화가 없는 것으로 나타났다.

데이터 비관련 직무 근로자의 근속기간을 보면, 2016년 4.27년 2017년 4.83년 2018년 4.70년으로 큰 변화가 없는 것으로 나타났다. 기업 규모별로 보면, 1~9인은 2016년 3.32년에서 2018년 3.80년으로 변화가 없는 것으로 나타났으며, 10~99인은 2016년 4.21년, 2017년 4.95년, 2018년 4.88년으로 감소 폭은 있지만 변화는 없는 것을 알 수 있다. 100~299인은 2016년 5.04년에서 2018년 5.09년으로 변화가 없었고 300인 이상은 근속연수가 8.0~8.1년으로 변화가 없는 것으로 나타났다.

전체적으로 데이터 관련 직무 수행 근로자의 근속연수와 데이터 비관련 직무 수행 근로자의 근속연수는 크게 차이가 없지만 2018년을 기준으로 보면 데이터 비관련 직무 수행 근로자의 근속연수가 조금 높은 것을 알 수 있다. 그리고 앞서 살펴본 사업체의 업력과 근속연수를 함께 보면, 사업체의 업력이 전체적으로 약 12년이고 기업 규모가 커지면서 업력이 증가하고 있는데, 이를 감안해서 보면 근속연수가 높지 않다고 볼 수 있다.

한편, <표 4-18>에서 데이터 관련 직무 근로자의 이직률을 보면, 2016년 7.85%, 2017년 7.65%, 2018년 6.65%로 점차 줄어드는 것으로 나타났다. 기업 규모별로 보면, 1~9인은 2016년 6.19%에서 2018년 3.44%로 줄어든 것으로 나타났으며, 10~99인은 2016년 8.50%, 2018년 7.79%로 줄어든 것으로 나타났다. 100~299인은 2016년 9.05%에서 2018년 9.40%로 약간 증가한 것으로 나타났으며, 300인 이상은 2016년 3.16%에서 2018년 3.50%로 약간 증가한 것으로 나타났다.

데이터 비관련 직무 근로자의 이직률은 2016년 3.86%, 2017년 6.63%, 2018년 6.06%로 증가하는 것으로 나타났다. 기업 규모별로 보면, 1~9인은 2016년 4.90%에서 2018년 2.89%로 줄어든 것으로 나타났으며, 10~99인은 2016년 8.53%에서 2018년 8.16%로 증가한 것으로 나타났다. 100~299인은 2016년 4.97%에서 변화가 없는 것으로 나타났으며, 300인 이상 경우도 2016년 2.33%에서 변화가 없는 것으로 나타났다.

종합적으로 보면, 기업 규모별로 차이를 보이기는 하지만 데이터 관련 직무 수행 근로자의 이직률이 약간 감소한 것을 알 수 있으며, 이와 달리 데이터 비관련 직무 수행 근로자의 이직률은 약 1.5배 정도 증가한 것을 알 수 있다.

〈표 4-17〉 근로자 근속기간

(단위:년)

직무	데이터 관련 직무						데이터 비관련 직무						
	2016		2017		2018		2016		2017		2018		
구분	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	
전 체	4.14	2.24	4.45	2.40	3.83	2.52	4.27	2.13	4.83	2.52	4.70	2.25	
기업 규모	1~9인	3.47	2.58	3.81	2.78	4.39	2.93	3.32	2.72	3.56	2.89	3.80	3.08
	10~99인	4.53	2.02	3.95	2/18	5.39	2.34	4.21	1.74	4.95	2.18	4.88	1.69
	100~299인	4.99	1.62	4.06	1.60	4.12	1.56	5.04	1.83	5.05	1.82	5.09	1.74
	300인 이상	7.50	3.98	7.50	3.98	7.50	3.98	8.0	3.68	8.0	3.68	8.1	3.60

자료: 저자 작성.

〈표 4-18〉 근로자 이직률

(단위:%)

직무	데이터 관련 직무						데이터 비관련 직무						
	2016		2017		2018		2016		2017		2018		
구분	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	
전 체	7.85	8.90	7.65	9.20	6.65	8.07	3.86	6.23	6.63	5.91	6.06	5.56	
기업 규모	1~9인	6.19	10.0	5.81	10.9	3.44	8.06	4.90	8.21	3.63	5.65	2.89	4.99
	10~99인	8.50	9.31	8.49	0.09	7.79	8.13	8.53	5.95	8.84	6.34	8.16	5.84
	100~299인	9.05	5.28	9.39	5.61	9.40	6.35	4.97	2.79	4.97	2.63	4.97	2.90
	300인 이상	3.16	2.48	3.67	2.04	3.50	3.27	2.33	1.36	2.33	1.36	2.33	1.36

자료: 저자 작성.

제5절 데이터 산업 사업체 고용관리

데이터 산업 사업체의 고용관리 현황을 파악하기 위하여, 인사관리 체계에서부터, 채용, 교육훈련, 보상까지 제도들을 통하여 고용관리가 어떻게 이루어지고 있는지 살펴보았다.

1. 인사관리 체계 및 방향

가. 인력계획 수립

사업체에서 정기적으로 소요인원 상정 및 충원계획 등과 관련한 인력 계획을 수립하는지, 그리고 얼마 단위로 계획을 수립하는지 조사하였다. 전체를 보면, 계획을 수립하지 않는 사업체는 128개로 전체 중 28.4%를 차지하는 것으로 나타났으며, 그 외 사업체는 인력계획을 수립하고 있고 그 중 1년 단위로 계획을 수립하는 사업체(69.2%)가 가장 많은 것으로 나타났다. 기업 규모별로 보면, 1~9인 사업체에서는 계획을 수립하지 않는 사업체가 28.3%이며, 그 외 사업체는 인력계획을 수립하고 있고 모두 1년 단위로 계획을 수립(71.7%)하고 있는 것으로 나타났다. 10~99인을 보면, 계획을 수립하지 않는 사업체가 38.4%이며, 그 외 사업체는 인력계획을 수립하고 있고 그중 1년 단위로 계획을 수립하는 사업체가 56.4%로 가장 많은 것으로 나타났다. 100~299인 사업체는 계획을 수립하지 않는 사업체가 5.6%이며, 나머지 사업체는 모두 1년 단위로 인력계획을 수립(94.4%)하는 것으로 나타났다. 마지막으로, 300인 이상 사업체를 보면, 인력 계획을 수립하지 않는 사업체는 16.7%이고, 나머지 사업체는 1년 단위로 인력계획을 수립(83.3%)하고 있다고 응답하였다.

종합적으로 보면, 사업체들은 인력계획을 수립하고 있으며, 그 단위는 1년 단위가 가장 많은 것을 알 수 있고, 이는 사업체들이 단기적 관점을 가지고 인력운영을 하고 있다고 볼 수 있게 한다.

〈표 4-19〉 인사계획 수립 단위

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
1년 단위	312	69.2	104	71.7	119	56.4	84	94.4	5	83.3
2년 단위	10	2.2	0	0	10	4.7	0	0	0	0
4~5년 단위	1	0.2	0	0	1	0.5	0	0	0	0
계획 수립하지 않음	128	28.4	41	28.3	81	38.4	5	5.6	1	16.7
전 체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

나. 인사정책

다음으로, 데이터 산업 사업체의 전반적인 인사관리 특성을 파악하기 위하여 인사관리 목표, 고용 및 근로자 육성 방향, 근로자 활용 방향, 그리고 주요 고용 형태에 관하여 5점을 기준으로 조사하였다.

먼저, 인사관리 목표를 보면, 인사관리의 일차적 목표가 근로자의 기업에 대한 충성심과 애착심을 높이는 데 있으면 5점에 가깝게, 그렇지 않고 가능한 고정적인 인건비 절감에 있다면 1점에 가깝게 응답하라고 한 결과, 3.37점으로 사업체들은 근로자에 기업에 대한 충성심과 애착심을 높이는 방향으로 인사관리 목표를 수립하고 있지만 그 정도는 높지 않은 것으로 나타났다. 두 번째로, 고용 방향을 보면, 필요한 자격을 갖춘 사람을 장기고용을 통해 내부에서 육성하면 5점에 가깝게, 그렇지 않고 외부에서 충원하고 불필요할 경우 해고하면 1점에 가깝게 응답하라고 한 결과, 3.39점으로 사업체들은 장기고용을 통해 내부육성을 하는 방향으로 고용을 하고 있으나 그 정도는 높지 않은 것으로 나타났다. 세 번째로 근로자 활용 방향을 보면, 가능한 정규직 근로자를 활용한다면 5점에 가깝게, 그렇지 않고 가능한 비정규직 근로자를 활용하면 1점에 가깝게 응답하라고 한 결과, 4.34점으로 사업체들이 가능한 정규직 근로자를 주로 활용하고 있는 것으로 나타났다. 마지막으로, 근로자 육성을 보면, 인사관리가 근로자의

〈표 4-20〉 인사관리 특성

(단위: %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
인사관리 목표	3.37	0.61	3.59	0.58	3.36	0.67	3.06	0.31	3.00	0
고용 방향	3.39	0.62	4.09	0.62	4.00	0.65	3.79	0.53	3.83	0.40
근로자 활용 방향	4.34	0.59	4.25	0.67	4.53	0.56	4.03	0.35	4.33	0.51
근로자 육성	3.39	0.53	3.55	0.49	3.40	0.58	3.12	0.36	3.33	0.51

자료: 저자 작성.

장기적인 육성과 개발을 위한 방향으로 운영되면 5점에 가깝게, 그렇지 않고 근로자의 단기적 성과와 업적을 최대한 높이는 방향으로 운영되면 1점에 가깝게 응답하라고 한 결과, 3.39점으로 사업체들이 근로자의 장기적인 육성과 개발을 위한 방향으로 운영하고는 있지만 그 수준은 높지 않은 것으로 나타났다.

기업 규모별로 보면, 1~9인의 경우 전반적으로 전체 경향과 일치하고 수준도 약간 높은 것으로 나타났다. 10~99인을 보면, 전체 경향과 일치하고 그 수준 또한 유사한 것으로 나타났다. 100~299인은 전체 경향과 일치하되 고용 방향을 제외하면 수준이 낮은 것으로 나타났다. 300인 이상 또한 전체 경향과 일치하지만 수준이 낮은 것으로 나타났다.

종합적으로 보면, 사업체들은 고용안정 수준을 높이고 장기적으로 근로자를 육성하며 조직에 몰입케 하는 방향으로 인력 운영을 하고 있지만, 그 수준은 크게 높지 않은 것을 알 수 있다.

2. 인력 총원

가. 데이터 관련 직무 채용 시도

데이터 관련 직무를 지난 1년간 어떻게 채용하였는지 조사하였다. 먼저, 지난 1년간 채용 시도가 없다고 응답한 사업체가 전체 중 67.4%인 것으로 나타났다. 그 외 사업체에서 어떻게 채용을 시도하였는지 물은 결과

〈표 4-21〉 채용 시도

(단위: 개, %)

구분	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
정기공채	5	1.1	0	0	0	0	4	4.5	1	16.7
수시채용	81	18.0	22	15.2	55	26.1	4	4.5	0	0
정기공채 및 수시채용 동시	61	13.5	0	0	30	14.2	27	30.3	4	66.7
시도 없음	304	67.4	123	84.8	126	59.7	54	60.7	1	16.7
전체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

수시채용이 18.0%, 정기공채와 수시채용을 동시에 진행한 사업체가 13.5%, 그리고 정기공채만 진행한 사업체가 1.1%인 것으로 나타났다. 기업 규모별로 보면, 1~9인은 지난 1년간 채용 시도가 없다고 응답한 사업체가 전체 중 84.8%인 것으로 나타났다. 그 외 사업체 15.2%는 수시채용만 진행한 것으로 나타났다. 10~99인을 보면, 지난 1년간 채용 시도가 없다고 응답한 사업체가 전체 중 59.7%인 것으로 나타났다. 그 외 사업체에서 어떻게 채용을 시도하였는지 묻은 결과 수시채용이 26.1%, 정기공채와 수시채용을 동시에 진행한 사업체가 14.2%인 것으로 나타났다. 100~299인은 지난 1년간 채용 시도가 없다고 응답한 사업체가 전체 중 60.7%인 것으로 나타났다. 그 외 사업체에서 어떻게 채용을 시도하였는지 묻은 결과 정기공채와 수시채용을 동시에 진행한 사업체가 30.3%, 정기공채 4.5%, 수시채용 4.5%인 것으로 나타났다. 마지막으로, 300인 이상 사업체는 지난 1년간 채용 시도가 없다고 응답한 사업체가 전체 중 16.7%로 가장 적은 것을 알 수 있으며, 그 외 사업체 중 66.7%가 정기공채와 수시채용을 동시에 진행하였다고 응답하였으며, 16.7%는 정기공채를 진행하였다고 응답하였다.

종합적으로 보면, 전체를 보았을 때 채용 시도를 하지 않은 사업체가 많다고 볼 수 있지만 기업 규모별로 구분하여 보면, 사업체 규모가 작은 경우 채용 시도가 적다는 것을 알 수 있다. 또, 채용 방식은 규모가 클수

록 정기공채와 수시채용을 동시에 사용하고, 규모가 작은 사업체의 경우 수시채용을 주로 하는 것을 알 수 있다.

나. 신입직, 경력직 채용 경로

데이터 관련 직무 신입사원과 경력사원을 채용하기 위해 주요하게 활용되는 구인 수단을 조사하였다.

먼저, 신입직원 채용할 때 활용하는 구인 수단을 보면, 인터넷 게시(56.1%)가 가장 많았으며, 다음으로는 노동부 워크넷(25.8%)을 활용하는 것으로 나타났다. 기업 규모별로 보면, 1~9인 사업에서는 인터넷 게시(53.6%)가 가장 많았고, 다음으로 직원 추천(23.9%)을 활용하는 것으로 나타났다. 10~99인 사업체는 인터넷 게시(54.0%)가 가장 많았고 노동부 워크넷(31.4%)을 다음으로 많이 활용하는 것으로 나타났다. 100~299인 사업체는 인터넷 게시(69.6%)가 가장 많았고, 노동부 워크넷(20.5%)을 그 다음으로 활용하는 것으로 나타났다. 마지막으로, 300인 이상 사업체도 인터넷 게시(66.7%)를 가장 많이 사용하고 다음으로는 신문, 라디오 등 매체 활용(22.2%)을 많이 하는 것으로 나타났다.

다음으로 경력직을 보면, 인터넷 게시(49.6%)가 가장 많았으며, 다음으로는 노동부 워크넷(29.0%)을 활용하는 것으로 나타났다. 기업 규모별로 보면, 1~9인 사업체에서는 인터넷 게시(49.6%)가 가장 많았고, 다음으로 직원 추천(31.9%)을 활용하는 것으로 나타났다. 10~99인 사업체는 인터넷 게시(48.9%)가 가장 많았고 노동부 워크넷(31.0%)을 다음으로 많이 활용하는 것으로 나타났다. 100~299인 사업체는 인터넷 게시(50.7%)로 가장 많았고, 노동부 워크넷(45.3%)을 활용하는 것으로 나타났다. 마지막으로, 300인 이상 사업체도 인터넷 게시(71.4%)를 가장 많이 사용하고 다음으로 노동부 워크넷(14.3%)을 활용하고 있는 것으로 나타났다.

종합적으로 보면, 신입, 경력직에 관계없이 인터넷에 게시하여 필요한력을 확보하고 있는 것을 알 수 있으며, 다음으로 노동부 워크넷을 많이 활용하며, 기업 규모가 작은 기업의 경우는 직원 추천도 활용하고 있는 것을 알 수 있다.

〈표 4-22〉 구인 수단

(단위: 개, %)

기업 규모		전체		1~9인		10~99인		100~299인		300인 이상	
		사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
신입	신문, 라디오 등 매체	16	2.0	1	0.4	2	0.5	11	9.8	2	22.2
	인터넷 게시	444	56.1	150	53.6	210	54.0	78	69.6	6	66.7
	노동부 워크넷	204	25.8	57	20.4	122	31.4	23	20.5	1	11.1
	직원 추천	91	11.5	67	23.9	24	6.2	0	0.0	0	0.0
	관련 학원 및 업체 이용	36	4.6	5	1.8	31	8.0	0	0.0	0	0.0
	공공 훈련기관	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	합 계	791	100	280	100	389	100	112	100	9	100
경력	신문, 라디오 등 매체	9	1.1	1	0.4	3	0.8	4	2.7	1	14.3
	인터넷 게시	412	49.6	137	49.6	194	48.9	76	50.7	5	71.4
	노동부 워크넷	241	29.0	49	17.8	123	31.0	68	45.3	1	14.3
	직원 추천	153	18.4	88	31.9	63	15.9	2	1.3	0	0.0
	관련 학원 및 업체 이용	14	1.7	0	0.0	14	3.5	0	0.0	0	0.0
	공공 훈련기관	1	0.1	1	0.4	0	0.0	0	0.0	0	0.0
	합 계	830	100	276	100	397	100	150	100	7	100

주: 1) 중복 응답

2) 신입 구인 수단 기타로는 학교추천제가 있음

자료: 저자 작성.

다. 채용 애로사항

데이터 관련 직무 신입사원과 경력사원을 채용하는 데 사업체가 경험하는 어려움이 무엇인지 조사하였다. 전체적으로, 가장 큰 어려움은 지원자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다(46.3%)는 것이었으며, 다음으로는 지원자가 없다(10.4%)는 것이었다. 그리고 별다른 어려움이 없다고 응답한 사업체도 29.5%나 되는 것으로 나타났다. 기업 규모 별로 보면, 1~9인 사업체에서 경험하는 채용 시 가장 큰 어려움은 지원

자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다(35.2%)였으며, 다음으로 지원자가 없다(26.9%)는 것이었다. 그러나 전체 중 30.3%가 채용에 어려움이 없다고 보기도 하였다. 10~99인 사업체도 지원자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다(63.5%)는 것을 가장 큰 어려움으로 응답하였으며, 다음으로는 조직적응에 적합한 태도를 가진 지원자가 없다(13.3%)를 어려움으로 꼽았다. 전체 중 15.2%는 채용에 어려움이 없다고 응답하였다. 100~299인 사업체는 채용에 어려움이 없다는 응답(59.6%)가 가장 많았으며, 채용의 어려움으로는 지원자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다(24.7%)는 것으로 응답하였다. 300인 이상 기업도 마찬가지로 채용에 어려움이 없다는 응답(66.7%)을 가장 많이 하였으며, 채용의 어려움으로는 지원자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다(33.3%)는 응답을 주었다.

종합적으로 보면, 기업의 규모가 커질수록 채용의 어려움은 없다고 볼 수 있다. 하지만 채용상 겪는 어려움을 꼽으라면 공통적이라고 볼 수

〈표 4-23〉 채용 시 애로사항

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
지원자 없음	47	10.4	39	26.9	8	3.8	0	0	0	0
지원자 있으나 적합자 없음	209	46.3	51	35.2	134	63.5	22	24.7	2	33.3
조직 적응에 적합한 태도 가진 지원자 없음	36	8.0	7	4.8	28	13.3	1	1.1	0	0
신입직원 근속기간 짧음	24	5.3	3	2.1	8	3.8	13	14.6	0	0
새로운 학문 분야 등으로 관련 교육 기관 국내에 드물	2	0.4	1	0.7	1	0.5	0	0	0	0
없음	144	29.5	44	30.3	21	15.2	53	59.6	4	66.7
전체	451	100	145	211	100	100	89	100	6	100

자료: 저자 작성.

있는데, 이는 지원자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다는 것이다.

라. 인턴십 제도 운영 및 기간

사업체에서 데이터 관련 직무 채용 시 채용연계형 인턴십 제도를 실시하고 있는지, 그리고 실시한다면 그 기간은 어느 정도 되는지를 조사하였다.

전체적으로, 인턴십 제도를 실시하지 않는 사업체가 80.9%, 실시하는 사업체가 19.1%인 것으로 나타났다. 그리고 실시하는 사업체를 대상으로 기간을 물어본 결과, 3개월 단위 운영 사업체가 7.3%로 가장 많았으며, 다음으로 1년 단위(6.4%)가 많았다. 기업 규모별로 보면, 1~9인은 인턴십 제도를 실시하지 않는 사업체가 97.9%, 실시하는 사업체가 2.1%인 것으로 나타났다. 그리고 실시하는 사업체를 대상으로 기간을 물어본 결과, 1년 단위로 운영하고 있는 것으로 나타났다. 10~99인은 인턴십 제도를 실시하지 않는 사업체가 78.7%, 실시하는 사업체가 21.3%인 것으로 나타났다. 그리고 실시하는 사업체를 대상으로 기간을 물어본 결과, 1년 단위 운영 사업체가 42.2%로 가장 많았으며, 다음으로 3개월 단위 운영(37.8%)이 많은 것으로 나타났다. 100~299인은 인턴십 제도를 실시하지 않는 사업체가 62.9%, 실시하는 사업체가 37.1%인 것으로 나타났다. 그리고 실시하는 사업체를 대상으로 기간을 물어본 결과, 3개월 단위 운영 사업체가 45.4%로 가장 많았으며, 다음으로 6개월 단위(33.3%)가 많은 것으로 나타났다. 마지막으로 300인 이상을 보면, 인턴십 제도를 실시하지 않는 사업체가 16.7%, 실시하는 사업체가 83.3%인 것으로 나타났다. 그리고 실시하는 사업체를 대상으로 기간을 물어본 결과, 6개월 단위 운영 사업체가 60.0%로 가장 많았으며, 3개월과 1년 단위 운영이 각각 20.0%인 것으로 나타났다.

종합적으로 보면, 300인 이상 사업체를 제외하고 인턴십 제도의 실시는 활발하지 않다고 볼 수 있으며, 시행하는 기간 또한 사업체별로 차이를 보이는 것을 알 수 있다.

〈표 4-24〉 인턴십 제도

(단위: 개, %)

기업 규모		전체		1~9인		10~99인		100~299인		300인 이상	
		사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
제도	있음	86	19.1	3	2.1	45	21.3	33	37.1	5	83.3
	없음	365	80.9	142	97.9	166	78.7	56	62.9	1	16.7
기간	3개월	33	7.3	0	0	17	37.8	15	45.4	1	20.0
	6개월	21	4.7	0	0	7	15.6	11	33.3	3	60.0
	1년	29	6.4	3	100	19	42.2	6	18.2	1	20.0
	2년 이상	3	0.7	0	0	237.8	4.4	1	3.0	0	0

자료: 저자 작성.

3. 근로시간, 휴가제도 및 근무제도 운영

데이터 산업 사업체에서 근로시간 관련 제도가 어떻게 운영되는지 살펴보았다.

가. 초과근로시간

먼저, 정규직 근로자들의 주당 초과근로시간⁸⁾을 조사하였다. 데이터 관련 직무와 비관련 직무로 구분하여 살펴보았다. 먼저 데이터 관련 직무를 보면, 평균 주당 초과근로시간은 1.29시간으로 나타났으며, 데이터 비관련 직무 또한 1.21시간으로 차이가 나지 않는 것으로 나타났다. 기업 규모별로 보면, 1~9인의 경우 데이터 관련 직무는 평균 주당 초과근로시간이 0.77시간이고 데이터 비관련 직무는 0.73시간인 것으로 나타났다. 10~99인은 데이터 관련 직무가 평균 주당 초과근로시간이 0.84시간이고 데이터 비관련 직무는 0.74시간인 것으로 나타났다. 100~299인은 데이터 관련 직무는 평균 주당 초과근로시간이 3.12시간이고 데이터 비관련 직무는 2.79시간인 것으로 나타났다. 300인 이상은 데이터 관련 직무의 경우 평균 주당 초과근로시간이 2.50시간이고 데이터 비관련 직무 역시 2.50시간인 것으로 나타났다.

8) 주중 연장근로시간과 휴일근로시간을 모두 합친 시간이다.

〈표 4-25〉 초과근로시간

(단위: 시간)

구분		데이터 관련 직무		데이터 비관련 직무	
		평균	표준편차	평균	표준편차
전 체		1.29	2.20	1.21	2.13
기업 규모	1~9인	0.77	1.84	0.73	1.75
	10~99인	0.84	1.82	0.74	1.70
	100~299인	3.12	2.41	2.79	2.47
	300인 이상	2.50	4.18	2.50	4.18

주: N=408

자료: 저자 작성.

종합적으로 보면, 기업 규모가 커질수록 초과근로시간이 증가하는 것을 알 수 있으며, 100~299인 사업체를 제외하고는 데이터 관련 직무 수행 근로자와 데이터 비관련 직무 수행 근로자의 초과근로시간에 큰 차이가 없는 것을 알 수 있다.

나. 연차휴가 소진율

직무를 구분하여 연차휴가 소진율을 조사하였다. 데이터 관련 직무 수행 근로자의 연차소진율은 87.7%이고, 데이터 비관련 직무 수행 근로자의 연차소진율은 79.5%로 차이가 있는 것으로 나타났다. 기업 규모별로 보면, 1-9인은 데이터 관련 직무 수행 근로자의 연차소진율은 81.6%이고, 데이터 비관련 직무 수행 근로자의 연차소진율은 85.7%인 것으로 나타났다. 10-99인은 데이터 관련 직무 수행 근로자의 연차소진율은 87.2%이고, 데이터 비관련 직무 수행 근로자의 연차소진율은 87.1%인 것으로 나타났다. 100-299인은 데이터 관련 직무 수행 근로자의 연차소진율은 98.2%이고, 데이터 비관련 직무 수행 근로자의 연차소진율은 98.6%인 것으로 나타났다. 300인 이상은 데이터 관련 직무 수행 근로자의 연차소진율은 100%이고, 데이터 비관련 직무 수행 근로자의 연차소진율 또한 100%인 것으로 나타났다.

〈표 4-26〉 연차휴가 소진율

(단위: %)

구분		데이터 관련 직무		데이터 비관련 직무	
		평균	표준편차	평균	표준편차
전 체		87.7	15.9	79.5	12.9
기업 규모	1~9인	81.6	19.3	85.7	16.4
	10~99인	87.2	13.5	87.1	11.2
	100~299인	98.2	7.77	98.6	6.43
	300인 이상	100	0	100	0

자료: 저자 작성.

종합적으로 보면, 기업 규모가 커질수록 연차소진율이 상승하는 것을 알 수 있으며, 전체에서 보면 데이터 관련 직무와 비관련 직무 수행 근로자의 연차소진율이 큰 차이가 나는 것으로 보이지만, 규모로 보았을 때 1~9인을 제외하고는 차이가 없는 것을 알 수 있다.

다. 유연근무제

사업체에서 어떠한 종류의 유연근무제를 시행하고 있고 근로자들의 활용 여부를 4점 기준으로 조사하였다. 먼저 선택적 근무시간제를 보면, 실시하는 사업체가 10.6%, 미실시하는 사업체가 89.4%이고, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 물은 결과 2.33으로 응답하여 활용이 조금 이루어지는 것으로 나타났다. 재택 및 원격 근무제를 보면, 실시하는 사업체가 2.2%, 미실시하는 사업체가 97.8%이고, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 물은 결과 2.60으로 응답하여 활용이 조금 이루어지는 것으로 나타났다. 집중근무시간제를 보면, 실시하는 사업체가 4.4%, 미실시하는 사업체가 95.6%이고, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 물은 결과 2.20으로 응답하여 활용이 조금 이루어지는 것으로 나타났다. 재량근무제를 보면, 실시하는 사업체가 1.1%, 미실시하는 사업체가 98.9%이고, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 물은 결과 2.0으로 응답하여 활용이 조금 이루어지

는 것으로 나타났다. 마지막으로 교육휴가제를 보면, 실시하는 사업체가 0.4%, 미실시하는 사업체가 99.0%이고, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 2.50으로 응답하여 활용이 조금 이루어지는 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체 경우, 선택적 근무시간제는 0.7% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 3.0으로 응답하여 활용이 어느 정도 이루어지는 것으로 나타났다. 재택 및 원격근무제는 2.1% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 3.33으로 응답하여 활용이 어느 정도 이루어지는 것으로 나타났다. 집중근무시간제, 재량근무제, 교육휴가제는 실시하는 사업체가 없는 것으로 나타났다. 10~99인 사업체를 보면, 선택적 근무시간제는 13.7% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 1.76으로 응답하여 조금 활용하고 있는 것으로 나타났다. 재택 및 원격근무제는 1.4% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 1.33으로 응답하여 거의 활용이 제대로 이루어지지 않는 것으로 나타났다. 집중근무시간제는 4.7% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 1.40으로 응답하여 활용이 거의 제대로 이루어지지 않는 것으로 나타났다. 재량근무제는 1.9% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 1.75로 응답하여 조금 활용하는 것으로 나타났다. 교육휴가제는 0.5% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 1.0으로 응답하여 활용이 거의 제대로 이루어지지 않는 것으로 나타났다. 100~299인 사업체의 경우, 선택적 근무시간제는 18.0% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 3.13으로 응답하여 어느 정도 활용하고 있는 것으로 나타났다. 재택 및 원격근무제는 3.4% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 2.67로 응

답하여 조금 활용하고 있는 것으로 나타났다. 집중근무시간제는 7.9% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 2.71으로 응답하여 조금 활용하고 있는 것으로 나타났다. 재량근무제는 1.1% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 4.0으로 응답하여 많이 활용하고 있는 것으로 나타났다. 교육휴가제는 실시하는 사업체가 없는 것으로 나타났다. 마지막으로, 300인 이상 사업체를 보면, 선택적 근무시간제는 33.3% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 묻은 결과 4.0으로 응답하여 많이 활용하고 있는 것으로 나타났다. 재택 및 원격근무제는

〈표 4-27〉 유연근무제 시행 및 활용

(단위: 개, %, 점)

기업 규모		전체		1~9인		10~99인		100~299인		300인 이상	
		사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
선택적 근무제	실시	48	10.6	1	0.7	29	13.7	16	18.0	2	33.3
	미실시	403	89.4	144	99.3	182	86.3	73	82.0	4	66.7
	활용도	2.33		3.0		1.76		3.13		4.0	
재택 및 원격 근무제	실시	10	2.2	3	2.1	3	1.4	3	3.4	1	16.7
	미실시	441	97.8	142	97.9	208	98.6	86	96.6	5	83.3
	활용도	2.60		3.33		1.33		2.67		4.0	
집중근무 시간제	실시	20	4.4	0	0	10	4.7	7	7.9	3	50.0
	미실시	431	95.6	145	100	201	95.3	82	92.1	3	50.0
	활용도	2.20		0		1.40		2.71		3.67	
재량 근무제	실시	5	1.1	0	0	4	1.9	1	1.1	0	0
	미실시	446	98.9	145	100	207	98.1	88	98.9	6	100
	활용도	2.0		0		1.75		4.0		0	
교육 휴가제	실시	2	0.4	0	0	1	0.5	0	0	1	16.7
	미실시	449	99.0	145	100	210	99.5	89	100	5	83.3
	활용도	2.50		0		1.0		0		4.0	

자료: 저자 작성.

16.7% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 물은 결과 4.0으로 응답하여 응답하여 많이 활용하고 있는 것으로 나타났다. 집중근무시간제는 50.0% 사업체가 실시하는 것으로 나타났으며, 실시하는 사업체를 대상으로 활용도를 4점 기준으로 물은 결과 3.67로 응답하여 활발하게 활용하고 있는 것으로 나타났다. 재량근무제는 시행하고 있지 않았으며, 교육휴가제는 16.7% 사업체가 실시한다고 응답하였으며, 활용도도 4.0으로 응답하여 많이 활용하고 있는 것으로 나타났다.

종합적으로 보면, 유연근무제를 활용하는 사업체 비중이 크지 않으며, 실시하여도 활용도가 높지 않다는 것을 알 수 있다. 그러나 300인 이상 사업체의 경우 제도 시행률도 높아지고 제도 활용도 활발하게 이루어지는 것을 알 수 있다.

4. 교육훈련 및 경력개발

데이터 산업 사업체에서 근로자 역량 향상을 위한 교육훈련과 경력개발을 어떻게 전개하고 있는지 살펴보았다.

가. 교육훈련 프로그램

먼저, 데이터 관련 직무 수행자의 직업능력 개발을 위하여 사업체에서 활용하고 있는 교육훈련 프로그램이 무엇인지 조사하였다. 전체적으로 보면, 별다른 교육훈련 프로그램을 실시하고 있지 않다는 응답이 43.2% 절반이 조금 못 되는 것으로 나타났다. 그 외에 실시하는 프로그램을 보면, 인터넷 학습이 23.9%로 가장 많았으며, 다음으로 사내 직업훈련이 17.7%로 많이 진행되는 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서 69.7%가 별도의 교육훈련을 실시하지 않는 것으로 응답하였으며, 실시하는 프로그램 중에서는 사내 직업훈련(19.3%)을 활용하는 것으로 나타났다. 10~99인 사업체를 보면, 전체 중 39.8%가 교육훈련 프로그램을 시행하지 않는다고 응답하였으며, 그 외

〈표 4-28〉 교육훈련 프로그램

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
사내 직업훈련	80	17.7	28	19.3	43	20.4	7	7.9	2	33.3
외부 직업훈련	40	8.9	4	2.8	25	11.8	11	12.4	0	0
인터넷 학습	108	23.9	8	5.5	37	17.5	60	67.4	3	50.0
현장훈련	28	6.2	4	2.8	22	10.4	2	2.2	0	0
실시하지 않음	195	43.2	101	69.7	84	39.8	9	10.1	1	16.7

자료: 저자 작성.

기업에서는 사내 직업훈련(20.4%)이 가장 많이 실시되고 있으며, 다음으로는 인터넷 학습(17.5%)이 시행되는 것으로 나타났다. 100~299인 사업체에서 교육훈련을 실시하지 않는 사업체는 10.1%로 매우 적은 것으로 나타났으며, 주로 인터넷 학습(67.4%)을 많이 시행하는 것으로 나타났다. 마지막으로 300인 이상 사업체를 보면, 교육훈련을 시행하지 않는 사업체는 16.7%로 많지 않으며, 주로 인터넷 학습(50.0%)을 시행하는 것으로 나타났다.

종합적으로 보면, 사업체 규모가 커질수록 사업체 단위에서 교육훈련을 실시하고 있는 것으로 보이며, 많은 사업체에서 인터넷 학습을 통해 교육훈련하고 있는 것을 알 수 있다.

나. 교육훈련 시간

2018년을 기준으로 일인당 연평균 공식적으로 교육훈련을 받은 시간을 조사하였다. 먼저, 데이터 관련 직무 수행 근로자의 경우, 평균 32.7시간 교육훈련 받은 것으로 나타났으며, 데이터 비관련 직무 수행 근로자는 29.6시간인 것으로 나타나 데이터 관련 직무 수행 근로자에 비하여 교육훈련 시간이 낮은 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서 데이터 관련 직무 수행 근로자의

〈표 4-29〉 연평균 인당 공식 교육훈련 시간

(단위: 1인당 연평균 시간)

직무 분야		데이터 관련 직무		데이터 비관련 직무	
		평균	표준편차	평균	표준편차
전 체		32.7	16.5	29.6	13.8
기업 규모	1~9인	33.5	14.4	29.4	13.1
	10~99인	37.0	17.5	21.6	13.6
	100~299인	24.9	12.7	24.3	12.7
	300인 이상	42.0	20.4	40.0	18.7

자료: 저자 작성.

평균 교육훈련 시간은 33.5시간이고 데이터 비관련 직무 수행 근로자의 평균 교육훈련 시간은 29.4시간으로 데이터 비관련 직무 수행 근로자의 교육훈련 시간이 더 낮은 것으로 나타났다. 10~99인 사업체에서 데이터 관련 직무 수행 근로자의 평균 교육훈련 시간은 37.0시간이고 데이터 비관련 직무 수행 근로자의 평균 교육훈련 시간은 21.6시간으로 데이터 비관련 직무 수행 근로자의 교육훈련 시간이 더 낮은 것으로 나타났다. 100~299인 사업체에서 데이터 관련 직무 수행 근로자의 평균 교육훈련 시간은 24.9시간이고 데이터 비관련 직무 수행 근로자의 평균 교육훈련 시간은 24.3시간으로 다른 규모와 달리 둘 사이 차이가 없는 것으로 나타났다. 300인 이상 사업체에서 데이터 관련 직무 수행 근로자의 평균 교육훈련 시간은 42.0시간이고 데이터 비관련 직무 수행 근로자의 평균 교육훈련 시간은 40.0시간으로 이 또한 둘 사이 교육훈련 시간의 차이는 없는 것으로 나타났다.

종합적으로 보면, 통계청에서 매년 실시하는 「기업 직업훈련 실태조사」 결과와 비교해 볼 필요가 있는데 전체 기업들이 2016년 집체교육훈련 시간만 연평균 18.5시간, 2017년 20.2시간 실시한 것과 비교하면 데이터 산업 사업체에서 교육훈련은 활발하게 이루어진다고 볼 수 있다. 그리고 100~299인 규모를 제외하고는 기업 규모가 커질수록 인당 평균 교육훈련 시간이 증가하는 것을 알 수 있으며, 기업 규모가 커질수록 두 직무 간의 교육훈련 시간에는 차이가 없는 것을 알 수 있다.

다. 교육훈련 비용

2018년 기준 사업체가 정규직 근로자들에게 지출한 교육훈련 비용을 조사하였다. 먼저 총액을 보면, 평균 146만 원, 인건비 대비 교육훈련비 지출 비율은 0.14%로 비용 규모가 매우 작은 것으로 나타났다. 기업 규모를 구분하여 보면, 1~9인 사업체의 평균 0.24만 원, 인건비 대비 교육훈련비 지출 비율은 0%이며, 10~99인 사업체는 평균 4.30만 원, 인건비 대비 교육훈련비 지출 비율은 0.06%인 것으로 나타났다. 100~299인 사업체는 평균 30.3만 원, 인건비 대비 교육훈련비 지출 비율은 0.27%이며, 300인 이상 사업체는 평균 144.1만 원, 인건비 대비 교육훈련비 지출 비율은 0.30%인 것으로 나타났다.

종합적으로 보면, 통계청에서 매년 실시하는 「기업 직업훈련 실태조사」 결과와 비교해 볼 필요가 있는데 전체 기업들이 지출한 인당 연평균 집체 훈련 비용이 2015년 22만 원, 2016년 12.6만 원인 것과 비교하면, 데이터 산업 사업체에서 지출하는 교육훈련 비용 규모가 꽤 큰 것을 알 수 있다. 그리고 사업체 규모가 커질수록 비용은 증가하고 인건비 대비 비용 비율도 증가하는 것을 알 수 있다.

〈표 4-30〉 교육훈련 비용

(단위: 만 원)

구분	총액		인건비 대비 교육훈련비 비율	
	평균	표준편차	평균	표준편차
전 체	146	77.9	0.14	0.71
기업 규모	1~9인	0.24	0.60	0
	10~99인	4.30	7.51	0.06
	100~299인	30.3	120.2	0.27
	300인 이상	144.1	251.6	0.30

자료: 저자 작성.

라. 경력상담제도

데이터 관련 직무 수행 근로자들에게 개인 경력을 개발할 수 있는 상담제도를 시행하는지 조사하였다. 전체적으로 보면, 경력상담제도를 실시하고 잘 활용한다는 사업체는 14.2%이며, 실시하지만 활용되지 않는다는 사업체도 14.4% 되는 것으로 나타났다. 그리고 실시하지 않는다는 사업체가 71.4%인 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서 경력상담제도를 실시하지 않는 사업체가 75.2%이고, 실시하고 잘 활용된다는 응답이 15.9%인 것으로 나타났다. 10~99인 사업체에서 경력상담제도를 실시하지 않는 사업체가 58.8%이며, 실시하지만 잘 활용되지 않는다는 응답이 22.7%로 다음으로 나타났다. 100~299인 사업체에서 경력상담제도를 실시하지 않는다는 사업체가 94.4%로 다수이며, 실시하지만 잘 활용되지 않는다는 응답이 3.4%로 나타났다. 마지막으로, 300인 이상 사업체를 보면, 실시하지 않는 사업체가 83.3%로 가장 많았고, 실시하지만 잘 활용되지 않는다는 사업체가 16.7%인 것으로 나타났다.

종합적으로 보면, 경력개발과 관련한 경력상담제도는 전체적으로 잘 시행되지 않는다고 볼 수 있다.

〈표 4-31〉 경력상담제도 시행

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
실시하지만 잘 활용되지 않음	65	14.4	13	9.0	48	22.7	3	3.4	1	16.7
실시하고 잘 활용됨	64	14.2	23	15.9	39	18.5	2	2.2	0	0
실시하지 않음	322	71.4	109	75.2	124	58.8	84	94.4	5	83.3
전체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

5. 보상관리

사업체에서 인사평가와 관련 보상제공을 어떻게 하고 있는지 살펴보았다.

가. 근로자 평균 초임

데이터 관련 직무 수행 근로자와 데이터 비관련 직무 수행 근로자를 구분하여 평균 초임을 조사하였다. 데이터 관련 직무 수행 근로자의 2016년 초임은 296만 원, 2017년 305만 원, 2018년 311만 원으로 2017년 크게 증가하다 2018년 떨어진 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 293만 원, 2017년 301만 원, 2018년 308만 원으로 증가하는 경향인 것으로 나타났으며, 데이터 관련 직무 수행 근로자와 임금수준을 비교하여 보면, 낮은 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서 데이터 관련 직무 수행 근로자의 2016년 초임은 292만 원, 2017년 299만 원, 2018년 302만 원으로 크지는 않지만 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 291만 원, 2017년 297만 원, 2018년 302만 원으로 이 또한 크지는 않지만 증가하는 것으로 나타났으며, 데이터 관련 직무 수행 근로자와 임금 수준이 큰 차이 없는 것으로 나타났다. 10~99인 사업체에서 데이터 관련 직무 수행 근로자의 2016년 초임은 295만 원, 2017년 305만 원, 2018년 311만 원으로 크지는 않지만 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 289만 원, 2017년 298만 원, 2018년 304만 원으로 이 또한 크지는 않지만 증가하는 것으로 나타났으며, 데이터 관련 직무 수행 근로자의 임금수준과 큰 차이는 없지만 데이터 비관련 직무 수행 근로자 초임이 조금 낮은 것으로 나타났다. 100~299인 사업체에서 데이터 관련 직무 수행 근로자의 2016년 초임은 302만 원, 2017년 312만 원, 2018년 323만 원으로 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 301만 원, 2017년 312만 원, 2018년 322만 원으로 이 또한 증가하는 것으로 나타났으며, 데이터 관련 직무 수행 근로자의 임금 수준과 큰 차이 없는 것으로 나타났다. 마지막으로,

〈표 4-32〉 초임 수준

(단위: 천 원)

직무 구분	데이터 관련 직무						데이터 비관련 직무						
	2016		2017		2018		2016		2017		2018		
구분	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	
전 체	2,965.9	370.5	3,052.2	383.4	3,115.5	390.8	2,931.2	344.3	3,018.0	359.6	3,083.1	369.5	
기 업	1~9인	2,928.9	467.2	2,991.0	470.1	3,027.9	483.8	2,919.7	393.2	2,970.4	407.9	3,023.4	435.1
	10~99인	2,956.9	308.2	3,050.6	316.6	3,115.2	312.5	2,890.7	317.1	2,984.4	312.8	3,040.9	590.1
규 모	100~299인	3,021.5	340.9	3,129.7	360.2	3,232.5	356.6	3,014.6	340.3	3,124.3	360.8	3,222.9	344.2
	300인 이상	3,237.1	292.6	3,370.3	344.1	3,512.8	369.0	3,230.0	286.3	3,368.3	341.8	3,510.0	365.5

주: 1) 전체 N=451, 1~99인 N=145, 10~99인 N=211, 100~299인 N=89, 300인 이상 N=6

2) 세진 총액 기준

자료: 저자 작성.

300인 이상 사업체에서 데이터 관련 직무 수행 근로자의 2016년 초임은 323만 원, 2017년 337만 원, 2018년 351만 원으로 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 323만 원, 2017년 336만 원, 2018년 351만 원으로 증가하는 것으로 나타났으며, 데이터 관련 직무 수행 근로자와 임금 수준이 큰 차이 없다고 볼 수 있다.

종합적으로 보면, 데이터 비관련 직무 수행 근로자의 초임 수준이 데이터 관련 직무 수행 근로자보다 낮은 듯 보이지만, 기업 규모별로 세분화해 보았을 때, 10~99인을 제외하고는 큰 차이가 없는 것으로 나타나 결론적으로 직무에 따른 초임 수준의 차이는 없다고 볼 수 있다.

나. 근로자 평균 임금

다음으로, 데이터 관련 직무 수행 근로자와 데이터 비관련 직무 수행 근로자를 구분하여 평균 임금을 조사하였다. 데이터 관련 직무 수행 근로자의 2016년 평균 임금은 363만 원 2017년 372만 원, 2018년 380만 원으

로 계속 증가하고 있는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 350만 원, 2017년 359만 원, 2018년 367만 원으로 증가하는 것으로 나타났으며, 임금 수준을 비교하여 보면 데이터 관련 직무 수행 근로자보다 낮은 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서 데이터 관련 직무 수행 근로자의 2016년 평균임금은 318만 원, 2017년 324만 원, 2018년 328만 원으로 크지는 않지만 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 324만 원, 2017년 312만 원, 2018년 317만 원으로 조금 감소한 것으로 나타났으며, 2018년 기준으로 데이터 관련 직무 수행 근로자와 임금수준을 비교하여 보면 임금이 더 낮은 것으로 나타났다. 10~99인 사업체에서 데이터 관련 직무 수행 근로자의 2016년 평균임금은 379만 원, 2017년 388만 원, 2018년 397만 원으로 크지는 않지만 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 349만 원, 2017년 359만 원, 2018년 367만 원으로 증가하는 것으로 나타났으며, 데이터 관련 직무 수행 근로자 임금과 비교하였을 때 수준이 다소 낮은 것으로 나타났다. 100~299인 사업체에서 데이터 관련 직무 수행 근로자의 2016년 평균임금은 388만 원, 2017년 405만 원, 2018년 419만 원으로 크지는 않지만 증가하는 것으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 386만 원, 2017년 402만 원, 2018년 417만 원으로 조금 감소한 것으로 나타났으며, 2018년 기준으로 데이터 관련 직무 수행 근로자와 임금 수준을 비교하여 보면 임금이 더 낮은 것으로 나타났다. 마지막으로, 300인 이상 사업체에서 데이터 관련 직무 수행 근로자의 2016년 평균 임금은 452만 원, 2017년 474만 원, 2018년 503만 원으로 나타났다. 데이터 비관련 직무 수행 근로자를 보면, 2016년 449만 원, 2017년 471만 원, 2018년 497만 원으로 증가하는 것으로 나타났으며, 데이터 관련 직무 수행 근로자의 평균 임금이 조금 높지만 큰 차이는 없는 것으로 나타났다.

종합적으로 보면, 임금의 차이는 직무의 차이보다 기업 규모에서 오는 차이가 더 큰 것으로 나타난다고 볼 수 있으며, 데이터 비관련 직무 수행 근로자의 평균 임금 수준이 다소 낮기는 하지만 큰 차이는 보이지 않는다고 볼 수 있다.

〈표 4-33〉 평균 임금

(단위: 천 원)

직무 구분		데이터 관련 직무						데이터 비관련 직무					
		2016		2017		2018		2016		2017		2018	
구분		평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
전 체		3,639.6	668.7	3,728.7	689.7	3,808.7	728.1	3,504.6	592.5	3,594.1	632.1	3,679.1	669.3
기업 규모	1~9인	3,181.7	560.2	3,243.4	553.8	3,286.5	579.9	3,243.4	411.4	3,127.2	433.4	3,179.0	475.6
	10~99인	3,790.9	590.1	3,881.2	583.7	3,970.4	607.4	3,490.6	489.0	3,591.1	499.6	3,676.2	515.3
	100~299인	3,889.4	639.0	4,056.1	682.1	4,194.5	698.8	3,864.8	647.7	4,028.8	684.8	4,171.5	705.6
	300인 이상	4,526.6	679.9	4,749.3	734.6	5,031.6	706.8	4,491.6	674.8	4,716.6	760.7	4,978.3	732.1

주: 1) 전체 N=451, 1~99인 N=145, 10~99인 N=211, 100~299인 N=89,
300인 이상 N=6

2) 세진 총액 기준

자료: 저자 작성.

다. 인사사고과 방식 차이

먼저, 데이터 관련 직무와 데이터 비관련 직무 간에 인사사고과 방식의 차이가 있는지 조사하였다. 차이를 두고 있다는 사업체는 10.6%로 적었으며, 89.4% 사업체가 직무에 관계없이 동일한 인사사고과 방식을 적용하고 있는 것으로 나타났다.

기업 규모별로 구분하여 보면, 1~9인 사업체는 직무에 관계없이 동일한 인사사고과 방식을 사용하는 사업체가 89.0%이며, 10~99인 사업체에서 직무에 관계없이 동일한 인사사고과 방식을 사용하는 사업체가 86.7%인 것으로 나타났다. 100~299인 이상 사업체에서는 직무에 관계없이 동일한 인사사고과 방식을 사용하는 사업체가 95.5%이고, 300인 이상 사업체에서는 100%인 것으로 나타났다.

종합적으로 보면, 사업체들에서는 직무에 따른 인사사고과 방식의 차이를 두고 있지 않는 것을 알 수 있다.

〈표 4-34〉 인사고과 방식

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
동일함	403	89.4	129	89.0	183	86.7	85	95.5	6	100
차이 있음	48	10.6	16	11.0	28	13.3	4	4.5	0	0
전체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

라. 인사고과 구성

방식은 구분하고 있지 않지만, 세부적으로 평가하는 항목에는 차이가 있을 수 있다는 점에서 인사고과를 구성하는 항목당 비율을 조사하였다. 데이터 관련 직무의 경우 성과항목이 25.6%, 능력항목이 41.7%, 태도항목이 32.6%, 기타가 0.1%인 것으로, 능력이 가장 중요한 인사고과 기준인 것으로 나타났다. 데이터 비관련 직무의 경우 성과항목이 25.0%, 능력항목이 38.3%, 태도항목이 36.6%, 기타가 0.1%인 것으로, 이 또한 능력 항목이 가장 중요한 인사고과 기준인 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서 데이터 관련 직무의 항목은 성과 22.4%, 능력 41.6%, 태도 36.0%로 능력항목이 가장 큰 비중을 차지하며, 데이터 비관련 직무의 항목은 성과 22.0%, 능력 35.2%, 태도 42.8%로 태도항목이 가장 큰 비중을 차지하는 것으로 나타났다. 10~99인 사업체에서 데이터 관련 직무의 항목은 성과 30.8%, 능력 40.0%, 태도 29.1%로 능력항목이 가장 큰 비중을 차지하며, 데이터 비관련 직무의 항목은 성과 29.2%, 능력 36.5%, 태도 34.0%, 기타 0.2%로 능력항목이 가장 큰 비중을 차지하는 것으로 나타났다. 100~299인 사업체에서 데이터 관련 직무의 항목은 성과 18.8%, 능력 46.1%, 태도 35.1%로 능력항목이 가장 큰 비중을 차지하며, 데이터 비관련 직무의 항목은 성과 18.7%, 능력 45.6%, 태도 35.7%로 능력항목이 가장 큰 비중을 차지하는 것으로 나타났다. 마지막으로, 300인 이상 사업체에서 데이터 관련 직무의 항목은 성과 22.5%, 능력

〈표 4-35〉 인사고과 구성 항목

(단위: %)

기업 규모		전체		1~9인		10~99인		100~299인		300인 이상	
직무 구분	항목	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
데이터 관련 직무	성과(업적)	25.6	18.2	22.4	9.4	30.8	24.1	18.8	6.3	22.5	5.2
	능력(역량)	41.7	14.9	41.6	14.8	40.0	17.4	46.1	5.2	43.3	7.5
	태도(자질)	32.6	12.8	36.0	12.4	29.1	14.4	35.1	5.8	34.2	6.6
	기 타	0.1	0.9	0.0	0.0	0.2	1.4	0.0	0.0	0.0	0.0
	전 체	100		100		100		100		100	
데이터 비관련 직무	성과(업적)	25.0	19.4	22.0	9.2	29.2	25.2	18.7	6.2	22.5	5.2
	능력(역량)	38.3	14.0	35.2	11.6	36.5	16.4	45.6	5.7	43.3	7.5
	태도(자질)	36.6	14.3	42.8	12.8	34.0	16.7	35.7	5.8	34.2	6.6
	기 타	0.1	1.3	0.0	0.0	0.2	1.8	0.0	0.0	0.0	0.0
	전 체	100		100		100		100		100	

자료: 저자 작성.

43.3%, 태도 34.2%로 능력항목이 가장 큰 비중을 차지하며, 데이터 비관련 직무의 항목은 성과 22.5%, 능력 43.3%, 태도 34.2%로 능력항목이 가장 큰 비중을 차지하는 것으로 나타났다.

종합적으로 보면, 직무에 따라 평가항목 간 큰 차이는 없는 것을 알 수 있다. 그리고 1~9인 사업체의 데이터 비관련 직무를 제외하고는 모두 고과항목 중 능력 즉, 역량을 가장 중요하게 보는 것을 알 수 있다.

마. 인사고과 보상 연계

인사점수에 따라 임금인상에 차등을 두고 있는지 평가와 보상 간의 연계를 조사하였다. 인사점수에 따라 임금인상에 차등을 둔다는 사업체가 47.0%인 것으로 나타났으며, 실시하고 있지 않다는 사업체도 37.9%인 것으로 나타났다.

기업 규모별로 보면, 1~9인 사업체에서는 임금인상 시 인사점수에 따른 차등 적용을 실시하지 않는 사업체가 71.1%로 가장 많았고, 실시하고 있다는 사업체가 24.1%로 다음으로 나타났다. 10~99인 사업체에서는 임금

〈표 4-36〉 인사점수에 따른 임금 인상 차등

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
실시함	212	47.0	35	24.1	98	46.4	74	83.1	5	83.3
실시하고 잘 활용됨	68	15.1	6	4.1	48	22.7	13	14.6	1	16.7
실시하지 않음	171	37.9	104	71.1	65	30.8	2	2.2	0	0
전체	451	100	145	100	211	100	89	100	6	100

자료: 저자 작성.

인상에 반영하는 사업체가 46.4%로 가장 많았고, 실시하지 않는다는 사업체가 30.8%로 다음으로 나타났다. 100~299인 사업체에서는 임금인상에 반영하는 사업체가 83.1%로 가장 많았고, 실시하고 잘 활용된다는 사업체가 14.6%로 다음으로 나타났다. 마지막으로, 300인 이상 사업체에서는 임금인상에 반영하는 사업체가 83.3%로 가장 많았고, 실시하고 잘 활용된다는 사업체가 16.7%로 다음으로 나타났다.

종합적으로 보면, 1~9인 사업체를 제외하고는 사업체들에서 인사평가와 보상을 연계한 보상 제도를 운영하고 있는 것을 알 수 있다.

바. 기본급 결정 기준

임금체계를 파악하기 위하여 먼저, 기본급을 어떠한 기준으로 결정하는지 조사하였다.

전체적으로 보면, 근속연수를 기본급 지급의 기준으로 보는 사업체가 46.2%로 가장 많았으며, 다음으로 직무를 기준으로 기본급을 지급(36.7%)하는 경우가 많은 것으로 나타났다.

기업규모별로 보면, 1~9인 사업체의 경우 근속연수를 기준으로 기본급을 지급하는 사업체가 61.1%이고 다음으로는 직능을 기준으로 기본급을 지급(16.1%)하는 것으로 나타났다. 10~99인 사업체는 근속연수를 기준으로 기본급을 지급하는 사업체가 42.0%로 가장 많은 것으로 나타났으며

〈표 4-37〉 기본급 지급 기준

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
근속연수	295	46.2	110	61.1	178	42.0	7	7.1	0	0
직무	234	36.7	27	15.0	176	41.5	88	89.8	6	100
직능	75	11.8	29	16.1	45	10.6	2	2.0	0	0
능력+성과	34	5.3	14	7.8	25	5.9	1	1.0	0	0
전 체	738		180		424		98		6	100

주: 중복 응답

자료: 저자 작성.

직무를 기준으로 기본급을 지급한다는 사업체도 41.5%로 수준이 유사한 것으로 나타났다. 100~299인 사업체를 보면, 직무를 기준으로 기본급을 지급한다는 사업체가 89.9%로 가장 많은 것으로 나타났으며, 다음으로 근속연수를 기준으로 기본급을 지급하는 사업체(7.1%)가 많은 것으로 나타났다. 마지막으로 300인 이상을 보면, 사업체 모두 기본급을 지급하는 기준에 대해 직무라고 응답하였다.

종합적으로 보면, 전체를 보았을 때 근속연수로 기본급을 지급하는 것으로 나타나지만, 기업 규모를 나누어 보면 기업의 규모가 커질수록 근속연수가 아닌 직무를 기준으로 임금을 산정하고 있는 것을 알 수 있다.

사. 기본급 이외 임금 항목

다음으로, 기본급 이외에 상여금 또는 수당 형태로 지급되는 항목들을 조사하였다. 소외 수당으로 명명되는 것이 무엇인지 조사한 것이다.

가장 많이 지급되는 수당 항목은 복지 관련 수당(31.2%)인 것으로 나타났다. 다음으로는 직무수당(30.2%)과 직책수당(21.2%)가 많은 것으로 나타났다.

기업 규모별로 보면, 1~9인은 직무수당(37.1%)로 가장 많은 것으로 나타났다. 다음으로 복지 관련 수당(36.0%)인 것으로 나타났다. 10~99인은 복지 관련 수당(32.1%)이 가장 많았고 다음으로 직무수당(28.8%)인 것으로

〈표 4-38〉 상여금 또는 수당 항목

(단위: 개, %)

기업 규모	전체		1~9인		10~99인		100~299인		300인 이상	
	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
직무수당	416	30.2	138	37.1	186	28.8	86	24.0	6	25.0
직책수당	292	21.2	58	15.6	144	22.3	85	23.7	5	20.8
자격수당	18	1.30	2	0.5	9	1.4	7	1.9	1	4.2
위험수당	8	0.58	1	0.3	2	0.3	5	1.4	0	0.0
기술수당	40	2.90	11	3.0	15	2.3	11	3.1	3	12.5
정기상여금	172	12.4	28	7.5	82	12.7	79	22.0	5	20.8
복지 관련 수당 ⁹⁾	431	31.2	134	36.0	207	32.1	86	24.0	4	16.7
전체	137	100	372	100	645	100	359	100	24	100

주: 기타 응답으로는 스톡옵션이 있음.

자료: 저자 작성.

나타났다. 100~299인 사업체를 보면 복지수당(24.0%)과 직무수당(24.0%)이 가장 많은 것으로 나타났으며 다음으로는 직책수당(23.7%)과 정기상여금(22.0%)이 많은 것으로 나타났다. 마지막으로, 300인 이상 사업체를 보면, 직무수당이 25.0%로 가장 많은 것으로 나타났으며, 다음으로 정기상여금(20.7%)과 직책수당(20.7%)이 많은 것으로 나타났다.

종합하여 보면, 사업체에서 직무수당을 강조하는 것을 알 수 있으며, 이러한 경향은 기업 규모에 관계없이 나타난다고 볼 수 있다. 이는 앞서 기본급에서도 확인하였듯이 데이터 산업 사업체의 임금체계에서 직무가 매우 중요한 보상 기준으로 작용하고 있다는 것을 의미한다고 볼 수 있다.

아. 성과급 운영

마지막으로, 데이터 관련 직무 근로자를 대상으로 변동성과급제도를 어떻게 운영하고 있는지 알아보기 위하여 성과급이 어떠한 기준으로 운

9) 가족수당, 주택수당, 결혼수당, 통근수당 등을 말한다.

영되고 있으며, 또한 전체 임금에서 어느 정도 비중을 차지하는지에 대해 조사하였다.

개인 성과급을 실시하는 사업체는 25.7%, 팀 성과급을 실시하는 사업체는 19.1%, 사업부 성과급을 실시하는 사업체는 1.6%, 그리고 전사 성과급을 실시하는 사업체는 6.4%인 것으로 나타났다. 그리고 기본급 대비 성과급의 수준은 17% 정도 되는 것으로 나타났다.

기업 규모별로, 먼저 1~9인 사업체를 보면, 개인 성과급을 실시하는 사업체는 3.4%, 팀 성과급을 실시하는 사업체는 1.4%, 사업부 성과급을 실시하는 사업체는 0%, 그리고 전사 성과급을 실시하는 사업체는 2.1%인 것으로 나타났다. 그리고 기본급 대비 성과급의 수준은 14% 정도 되는 것으로 나타났다. 10~99인 사업체를 보면, 개인 성과급을 실시하는 사업체는 20.9%, 팀 성과급을 실시하는 사업체는 17.5%, 사업부 성과급을 실시하는 사업체는 3.3%, 전사 성과급을 실시하는 사업체는 9.0%인 것으로 나타났다. 그리고 기본급 대비 성과급의 수준은 13.5% 되는 것으로 나타났다. 100~299인 사업체를 보면, 개인 성과급을 실시하는 사업체는 68.5%, 팀 성과급을 실시하는 사업체는 47.2%, 사업부 성과급을 실시하는 사업체는 0%, 그리고 전사 성과급을 실시하는 사업체는 7.9%인 것으로 나타났다. 그리고 기본급 대비 성과급의 수준은 20.7% 정도 되는 것으로 나타났다. 마지막으로 300인 이상 사업체를 보면, 개인 성과급을 실시하는 사업체는 100%, 팀 성과급을 실시하는 사업체는 83.3%, 사업부 성과급을 실시하는 사업체는 0%, 그리고 전사 성과급을 실시하는 사업체는 0%인 것으로 나타났다. 그리고 기본급 대비 성과급의 수준은 20% 정도 되는 것으로 나타났다.

종합적으로 보면, 기본급 대비 변동 성과급을 지급하는 비율은 크지 않다고 볼 수 있으며, 또한 성과급을 지급하는 데에서 개인 성과를 기준으로 하는 경향을 보이며 이는 사업체 규모가 커질수록 더욱 강하게 나타나는 것을 알 수 있다.

〈표 4-39〉 성과급 실시 현황

(단위: 개, %)

기업 규모		전체		1~9인		10~99인		100~299인		300인 이상	
실시 현황		사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율	사업체 수	비율
개인 성과급	실시	116	25.7	5	3.4	44	20.9	61	68.5	6	100
	미실시	335	74.3	140	96.6	167	79.1	28	31.5	0	0
팀 성과급	실시	86	19.1	2	1.4	37	17.5	42	47.2	5	83.3
	미실시	364	80.9	143	98.6	174	82.5	47	52.8	1	16.7
사업부 성과급	실시	7	1.6	0	0	7	3.3	0	0	0	0
	미실시	444	98.4	145	100	204	96.7	89	100	6	100
전사 성과급	실시	29	6.4	3	2.1	19	9.0	7	7.9	0	0
	미실시	422	93.6	143	97.9	192	91.0	82	92.1	6	100
기본급 대비 성과급 비율	평균	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
		17.08	9.014	14.5	8.57	13.5	9.18	20.7	6.36	20.0	8.31

자료: 저자 작성.

제6절 소 결

데이터 산업을 대상으로 고용관계가 어떻게 전개되고 있는지 파악하기 위하여 데이터 산업에 속한 사업체를 대상으로 설문조사를 실시하여 데이터 산업에 속한 기업의 특성을 파악하고, 고용 규모를 비롯하여 근로자 구성, 근로 형태 등 고용 구성에 대해 살펴보았으며, 사업체에 속한 근로자들을 어떠한 방식으로 채용하고 육성하고 동기부여하고 있는지도 살펴보았다.

주요 결과를 보면 데이터 산업 사업체들은 전반적으로 소유 경영 체제 기업이 많고, 외국인 지분률이 아주 낮으며, 원청 기업이 다수이지만 하청 기업의 경우 종속적 거래 관계를 가지는 것을 알 수 있다. 그리고 노동

조합은 없지만 노사관계를 매우 협력적으로 평가하는 경향을 보인다는 것을 알 수 있다.

사업체의 인적 구성을 보면, 남성 근로자가 2/3를 차지하고 있으며, 연령별로 20~30대가 많은 것으로 나타나 젊은 조직 분위기를 가지고 있다고 볼 수 있다. 이러한 경향은 데이터 관련 직무 수행자에게도 동일하게 나타나 데이터 산업을 나타내는 매우 큰 특징이라고 볼 수 있다. 그리고 비정규직 활용률이 매우 낮은 것으로 나타났으며 고용은 안정적이라고 평가할 수 있다. 그러나 사업체 업력과 근속연수를 함께 보면, 업력에 비하여 근로자 근속연수가 높지 않다고 볼 수 있지만, 전체적으로 이직률이 연간 4%이며 계속 감소하는 것으로 나타난다. 이는 앞선 결과에 이어 고용이 안정적으로 이루어지고 있다고 볼 수 있게 하는 근거이다.

그러나 인력 운영은 다소 단기적인 관점에서 이루어진다고 볼 수 있고 이러한 경향은 인사 정책 방향에도 나타나는데 사업체들은 고용안정 수준을 높이고 장기적으로 근로자를 육성하고 조직에 몰입하게 하는 방향으로 인력 운영을 하고 있지만, 그 수준이 크게 높지 않아 사업체의 방향과 현재 운영 면에 있어 괴리가 발생하고 있다고 볼 수 있다.

세부적으로 보면, 채용의 경우, 전체적으로 보았을 때 채용시도를 하지 않은 사업체가 많다고 볼 수 있지만 기업 규모별로 구분하여 보면 이는 사업체 규모가 작은 기업에서 기인한 현상으로 보인다. 채용 방식은 규모가 클수록 정기공채와 수시채용을 동시에 사용하지만 규모가 작은 사업체의 경우 수시채용을 주로 하는 것으로 나타나, 데이터 산업에 규모가 작은 기업이 많다는 점을 고려하면 채용이 매년 이루어지지 않을 수 있으며 인력 필요 시 수시로 진행된다고 볼 수 있다. 그리고 주목할 점은 채용을 인터넷 공고와 노동부 워크넷을 주로 이용한다는 것인데, 인터넷 공고가 다수의 산업과 기업에서 사용되는 방식이며, 데이터 산업에서 노동부 워크넷이 많이 이용된다는 것에 주목할 필요가 있다. 그리고 회사가 채용을 하고 싶어도 채용하지 못하는 이유로는 지원자가 있지만 회사에서 원하는 자질을 갖춘 사람이 없다는 점을 들고 있다. 데이터 산업 인력 양성에 대한 국가적 고민이 필요하다고 볼 수 있다.

교육훈련을 보면, 산업 전체와 비교하였을 때 상대적으로 인당 교육훈

런 시간과 교육훈련 비용이 높게 나타나 사업체에서 교육훈련이 활발하게 이루어진다고 볼 수 있다. 마지막으로 보상의 경우, 먼저 임금 수준을 보면, 초임은 직무에 관계없이 수준이 유사하며 약 300만 원 초반대인 것을 알 수 있다. 평균 임금도 직무의 차이보다는 기업 규모의 영향이 더 큰 것으로 나타나며 300만 원 후반대로 형성되어 있는 것을 알 수 있다. 또한 임금 체계를 보면, 직무를 기준으로 기본급을 산정하고 있는 경향을 발견할 수 있으며, 수당 또한 직무에 관한 수당의 비중이 높은 것으로 나타나 보상 지급에 직무를 매우 중요한 기준으로 두고 있는 것을 알 수 있다. 그리고 전체 임금에서 변동 상여금이 차지하는 비중은 크지 않다고 볼 수 있다. 마지막으로, 근로자 평가 시에는 근로자가 가진 능력 즉, 역량을 가장 중요하게 보는 것으로 나타났으며 평가제도를 보상과 연동하여 운영하는 것을 알 수 있다.

이상으로 데이터 산업에 속한 사업체에서 고용관리가 어떻게 이루어지는지 살펴보았다. 앞서의 결과에서도 나타나듯 데이터 산업의 고용관계는 인적 구성이나 임금 체계 등에서 데이터 산업이라는 산업 특성상 기인하는 고용관계의 특성도 보이지만, 데이터 산업에 규모가 작은 사업체가 많다는 점에서는 기업 규모에서 오는 고용관계의 특성도 함께 가진다고 볼 수 있다. 이러한 점에서 데이터 산업 고용관계에 대한 정책 지원 시 기업 규모가 작음에서 오는 고용관리의 영세성을 함께 고려하여 진행할 필요가 있다고 볼 수 있다. 그리고 데이터 산업의 사업체와 종사자 수가 지속적으로 증가하고 있고 국가적으로 중요한 산업이 되고 있다는 점에서, 데이터 산업에 종사하는 인력을 사업체에 유지할 수 있는 고용 지원 방안 에 대한 고민도 함께 필요하다고 볼 수 있다.

제 5 장

데이터 관련 기업 사례 조사

제1절 데이터 서비스 수요 기업 사례 조사

데이터 서비스는 소비자를 대상으로 하는 경우도 있으나, 최종 생산품 생산자들에게 제공되거나 구매되어 최종 생산품의 생산에 활용되기도 한다. 앞선 장에서 살펴본 빅데이터 산업에서의 인력 수급은 대부분이 데이터 산업 내에서의 이야기에 한정된 것이라면 이번 장에서는 데이터 서비스를 수요하는 기업과 공급하는 기업들에서의 인력 현황이나 노동 수요 및 공급에 대해서 사례 조사를 통해 살펴보고자 한다. 이를 위해 본 장은 크게 데이터 수요 기업과 공급 기업에 대한 사례로 이야기를 나누며, 데이터 수요 기업은 다시 금융권 서비스업과 제조업으로 나누어 이야기를 진행하려 한다.

1. 금융권 기업 조사

금융업은 제조업을 제외한 여타 서비스업 중에서도 가장 활발하게 빅데이터 영역에 진출해 있다. 이는 비단 한국만의 현상은 아니며, 금융업에서는 최근 빅데이터를 마케팅, 금융 기록상의 이상 행동 감지, 신용 변동이나 이와 연관된 소비자의 이상 형태 등을 감지하는 데 활발히 활용하

고 있다. 일례로, 어떠한 카드 사용 패턴을 보이는 사람이 곧 채무 불이행에 빠질 가능성이 높거나, 어떠한 특정 상품을 구매할 가능성이 높은지 등을 빅데이터를 통해 예측하여 이에 맞춘 금융 상품 홍보, 또는 신용한도 조정 및 경고 메시지 발송 등에 적극 활용하고 있다. 따라서 금융업은 다른 어떤 서비스업보다도 빅데이터와 관련된 인력 수요가 많은 집단이라 할 수 있다. 이에 본고에서는 금융업 대상으로 시중 은행과 시중 카드사 한 곳씩을 인터뷰 및 조사하였다.

가. A 은행

A 은행은 조직 내에 다양한 빅데이터 및 데이터 관련 부서가 있으며, 기존의 데이터 부서와는 별도로 빅데이터 관련 부서를 최근 신설하였다. 이와 별개로 본사 내에 데이터 및 빅데이터 관련 조직이 따로 존재하고 있으며, 개별 부서나 조직 간에 주어진 업무는 다르나 중복되는 부분도 일부 있는 것으로 파악된다. 은행의 빅데이터 부서에는 현재 약 20여 명 정도가 근무하고 있으며, 분석가(analyst)를 중심으로 규모를 확충해 가고 있다.

최근 신설된 빅데이터 부서는 통합 빅데이터 플랫폼을 직접 운영하면서 데이터 분석과 관련된 과제나 사업을 수행할 목적으로 신설되었다. 해당 부서는 현재 업무의 성격상 IT 업종의 회사와 금융 회사 내의 연구소 그리고 현직 금융업 종사자의 집합체와 같은 형태로 구성되어 있다. 이 점이 다른 빅데이터 조직과의 차이점으로 해당 부서는 빅데이터를 직접적으로 활용하여 금융 사업에서 바로 적용하거나 활용할 수 있는 모델을 개발하고 기존 사업을 개선할 수 있는 부분을 찾는 등 빅데이터 기법을 응용 및 활용하는 데 주력하고 있다. 빅데이터 관련 별건 조직의 경우 빅데이터 연구를 위한 측면이 강하다면 신설된 본 부서는 빅데이터의 활용과 응용에 방점을 두고 있다 하겠다.

그렇다면 해당 은행에 존재하던 기존의 데이터 관련 부서는 어떠한 일을 수행하는지, 그리고 해당 업무가 변경되었는지도 살펴보아야 할 것이다. 이에 대해서 면담자는 해당 조직의 업무에는 큰 변화가 없었으며, 이

는 빅데이터를 활용할 수 없는 영역, 예를 들어 충분한 양의 구조화된 자료가 확보되지 않거나 특정 집단만을 대상으로 하는 분석 등의 영역에서는 빅데이터의 정확성이 다소 떨어질 수 있기 때문에 기존의 데이터 관련 부서에서는 기존 업무를 기존과 유사한 방법으로 계속 수행 중이라고 하였다. 이는 빅데이터 분석을 모든 분야에 적용할 수 없는 현실에서 기존의 데이터 분석 부서의 업무에 급격한 변화가 생겼을 때 발생할 기업의 대(對) 소비자 대응 변화를 막고, 기존의 데이터 분석과 비교하여 빅데이터 분석의 유용성을 검증하기 위해서도 유용하고 필요한 조치라 하겠다.

다만 업무의 특수성으로 인해 빅데이터 연구에 보다 집중하는 별도의 빅데이터 조직과 비교하여 최근 신설된 A 은행의 빅데이터 조직은 구성원 특성에서 약간의 차이를 보이고 있다. 별도의 빅데이터 조직은 주로 석사나 박사급의 인력 위주로 구성되어 있는 반면, 최근 신설된 빅데이터 부서는 학사와 석사의 비중이 매우 높다. 또한 별도의 조직에서는 빅데이터 관련 근무 및 연구 경력이 있는 경력직의 비중이 상대적으로 높은 반면, 신설된 빅데이터 부서에서는 경력직의 비중도 비교적 낮은 편이다. 이는 빅데이터를 구축하기 위한 플랫폼을 만들고 데이터를 축적, 가공, 관리하며 분석 방법을 연구하는 데에는 보다 높은 수준의 전문 지식이 요구되고 아울러 빅데이터와 관련된 직접적인 소양이 필수적인 반면, 빅데이터를 활용하여 기존의 사업 영역에 응용하는 것은 상대적으로 금융업 및 금융 상품에 대한 이해가 보다 중요할 뿐만 아니라 기존에 고안된 분석 방법을 적용하는 것이므로 빅데이터에 대한 직접적인 소양이 필요치 않은 데 따른다.

해당 신규 빅데이터 부서는 주로 신규 채용과 타 부서에서의 전직 희망자로 인력을 충원하고 있는데, 현재 A 은행 내에는 빅데이터 관련 부서로 이전을 희망하는 부서원이 비교적 많은 편으로, 최근 빅데이터 관련 이슈들이 주목을 받으면서 상대적으로 여타 부서에 비해서 이전 희망자의 수가 높은 편으로 조사되었다. 이에 대해서 빅데이터 관련 부서가 여타 부서와 다른 인력 관리 방식이나 급여 체계 및 휴가 제도 등을 운영하여 부서 이전 희망자가 많은지를 알아보았다. 하지만 최근 주 52시간제의 도입에 맞물려 A 은행 내부에 각 부서별로 근무 시간의 편차는 거의 없

다고 하며, 휴가나 유연근무제도 등의 적용에 있어서도 빅데이터 부서만을 위한 별도의 체계는 없다고 하여, 해당 부서의 인기가 사회적 분위기와 향후의 전망에 크게 영향 받았음을 짐작케 한다. 다만, 별도의 빅데이터 조직이나 신규 빅데이터 부서 내의 분석가들 중 일부는 전문계약직으로 채용하는 경우가 있는데, 이 경우에는 정규직 근로자들과 다른 연봉 체계 및 휴가제도를 적용받는다.

빅데이터 부서로의 이전 희망자들에 대해서 별도로 운영하는 재교육 프로그램에 대해서도 조사해 보았다. A 은행은 부서 이동자를 위한 별도의 교육 프로그램을 가지고 있는 것은 아니며, 다만 전직원들을 대상으로 데이터 분석가 양성을 위한 프로그램을 운영 중이다. 해당 프로그램은 내부 정규직 직원들은 누구나 신청 가능하며, 프로그램별로 조금 상이하지만 대략 2달 정도 별도의 연수 장소에서 교육이 진행된다. 데이터 분석가 양성 프로그램은 통계 분석, 파이썬이나 C 등의 프로그램 언어 코딩, 분석 결과의 시각화 등의 내용을 포함하고 있으며, 주로 파이썬 코딩과 통계 분석 및 시각화를 많이 선호하고 수강한다고 한다.

이러한 프로그램을 운영하는 한 이유는 최근 신설된 빅데이터 부서에서도 여전히 신규 채용 등 인력 수급에 있어서 양질의 인력을 충분히 확보하지 못한 측면도 있다. 특히 신규 채용과 경력직 채용에 있어서는 적합한 경력과 학력을 가진 사람을 찾기가 힘든 편이며, 최근 빅데이터 관련 인력 수요가 폭증함에 따라 빅데이터 관련 경력직들을 채용하기에는 노동 시장에 충분한 경력직 인력들이 나오지 않아서라고 한다.

A 은행의 사례를 정리해보면 이러하다. 최근 은행 내에서 빅데이터 관련 조직과 인력 규모를 키우려고 하나, 신규 채용 및 경력직 채용에 있어 어려움을 겪고 있기 때문에 재직 중인 직원을 대상으로 별도의 교육 프로그램을 운영 중이다. 이러한 프로그램은 빅데이터 분석 기법 자체보다는 활용과 응용에 방점을 두고 있으며, 고도의 컴퓨터 관련 기술을 요구하지는 않고 있다. 하지만 이러한 인력들조차도 현재 노동시장에는 충분히 공급되지 않아서 인력 수급에 어려움을 겪고 있다고 할 수 있다.

나. B 카드

B 카드 역시 금융권 회사로서 최근 빅데이터 관련 부서를 신설하였다. 하지만 B 카드도 A 은행과 마찬가지로 빅데이터 부서가 생기기 전에 이미 회사 내에서 기존의 데이터 분석 관련 부서가 별도로 존재하고 있었으며, 현재도 존속 중이다. 기존의 데이터 분석 부서와 빅데이터 분석 부서의 차이점은, 기존의 데이터 관련 부서가 기존의 전통적인 분석 방법에 초점을 두는 반면, 빅데이터 부서는 주로 파이썬 등을 통해 기계학습을 하는 방향으로 주력한다는 점이다. 비록 두 부서가 별도로 존재하고 있고, 사용하는 분석 방법이나 프로그램은 차이가 있으나 중첩된 업무도 많고 현재까지 충원된 빅데이터 부서 인력의 다수가 기존의 데이터 부서에서 이동하기도 하여 외부에서 보기에 두 부서 간에 아주 유의미한 구별이 지는 않은 것이 현실이다.

조금 전에 언급했듯이 두 부서 간에는 인력 이동이 활발한 편이다. 그러나 데이터 부서에서 빅데이터 부서로 이동하는 인력은 다소 있으나 빅데이터 부서에서 데이터 관련 부서로 이동하는 인력의 수는 거의 없다. 기존의 데이터 부서 인력이 빅데이터 부서로 이동할 때는 일정 기간 교육을 받는데, 이는 파이썬과 기계학습에 대한 이론적, 기술적 훈련이 주(主)가 된다. 아울러 기존에 데이터 부서에서 하던 분석 중 일부를 빅데이터 부서에서도 동시에 수행하는데, 분석 결과에 대한 활용 방법 역시 기존과 큰 차이가 없는 것으로 조사되었다. 예를 들어 마케팅 관련 업무에서는 소비자에 대한 정확한 타기팅(targeting)보다는 공격적인 매출 증대를 위해 2중 오류를 피하는 전략을 사용하는 반면, 신용 관련 부서에서는 1중 오류를 피하기 위해 분석에서의 정확성과 보수적인 태도가 요구된다. 특히 신용 관리와 같은 업무는 일반 데이터 분석과 빅데이터 분석 모두에서 신용 위험 주의 경보가 뜬다 하더라도 반드시 전담 인력에 의해서 최종적으로 승인이 나와 실제 신용 조정 및 경고 발송 등이 이루어지도록 내부 절차가 구성되어 있다고 한다.

현재 B 카드의 빅데이터 관련 부서는 약 80여 명 정도가 근무 중이며, 내부 교육 훈련을 위해 배치된 인력이 현재까지 다수를 차지하고 있다.

하지만 2015년 빅데이터 전담 부서가 설치된 이후 새로 들어온 인력들은 신규 입사자가 부서 이전자보다 많은 편이다. 신규 채용 및 부서 이전자들 다수는 경제학, 경영학, 통계학 및 산업공학 전공자들이 많으며, 의외로 컴퓨터공학 전공자들은 많지 않다. 또한 컴퓨터공학 전공자들로서 전문적으로 프로그래밍 언어를 다루는 사람들은 주로 플랫폼 관련 업무들에 배치되며, 실제 분석과 활용에는 여전히 경제학, 경영학 및 통계학 전공자들이 다수 활동하고 있다.

이러한 부서원의 전공 구조는 앞선 A 은행과 유사하며, 그 사유 역시 비슷하다. 기본적으로 B 카드 역시 빅데이터 분석을 실시하는 이유는 기업의 매출 증대 및 경영 고도화로, 이를 위해서는 금융 업무 및 회사의 각종 금융 상품에 대한 전문적인 지식과 금융 시장에 대한 이해가 전제되어야 훨씬 정확하고 수월하기 때문이다. 다만, 이러한 빅데이터 응용 인력들도 기본적으로 파이썬 등의 프로그래밍 언어를 다룰 수 있는 수준의 기본적인 소양은 요구하고 있다.

빅데이터 관련 부서의 인력 관리 방식 역시 B 카드와 A 은행이 비슷한 모습을 보이는 것으로 조사되었다. B 카드 역시 빅데이터 부서만을 위한 별도의 인력 관리 방식이나 급여 및 휴가 체계를 도입하고 있지는 않다. 승진 및 퇴사의 경우에도 여타 부서와 큰 차이가 없는 규정이 적용되거 있으나 빅데이터 관련 인력 부족 문제로 중도 퇴사자나 권고 사직의 비율은 높지 않은 것으로 알려졌다. 다만 빅데이터 관련 부서 인력들은 내부에서 데이터 사이언티스트로 불리면서 별도의 직군으로 처리하고 있으나, 회사의 급여 체계 등은 직군을 불문하고 전 정규직 직원에게 모두 적용되기 때문에 직군별 인력 관리 방식은 동일하다. 따라서 빅데이터 관련 부서로의 이직 희망자들은 빅데이터 분야의 미래 유망성과 별도의 전문 직군으로 분류된다는 점 때문에 끌리는 것으로 보인다.

B 카드의 빅데이터 관련 부서에서는 인력 부족 문제를 체감하고는 있으나 아주 심각한 수준은 아니라고 한다. 인력 부족이 심각한 경우에는 사내 재교육을 통해 부서 이전 희망자들을 통해 충원해오고 있어 심각한 부족 문제를 겪는 것은 아니라고 한다. 다만 별도의 빅데이터 관련 부서를 구성하기 위해서는 그래도 어느 정도 수준의 전문 지식을 갖춘 인력들

을 확보해야 하기 때문에 국내 카드사들 중 독립된 빅데이터 관련 부서나 조직을 갖춘 곳은 많지 않으며, 다수의 중소 카드사들은 기존의 데이터 관련 부서에서 맡는 업무를 확장하는 한편 일부 업무를 전문 업체에 외주를 주어 해결하고 있다.

B 카드에서 주로 채용하는 빅데이터 인력들의 전공은 통계학과 경제학이라고 한다. 단순한 공학 기술자들은 빅데이터 플랫폼 구축 및 관리 등에만 필요하며 신용카드 시장에 대한 이해가 결여되어 이에 소요되는 재교육 시간 및 비용이 크기 때문에 막상 크게 선호되는 것은 아니라고 한다. 하지만 현재 국내에서 통계학과 경제학 전공자들 중 빅데이터 관련 숙련이나 기술을 갖춘 인력이 많지 않기 때문에 신규 채용 시 적합한 인재를 찾기 힘든 상황이라고 한다.

신규 입사자들 다수는 최소 학사 이상의 학력이 요구되며, 최근의 신규 입사자 다수는 석사 이상의 학력을 소지하고 있다고 한다. 경력직 채용의 경우 소규모 업체에서 경력을 쌓은 인력들 위주로 채용 중이라고 한다. 하지만 앞선 A 은행에서 살펴보았듯이 충분한 인적 자본과 숙련, 기술 및 경험을 갖춘 경력직들은 워낙 인기가 많아서 시장에서 찾기 어렵기 때문에 B 카드 정도의 대기업에서 근무한 경력이 있는 경력직들은 사실상 찾기 힘들다고 한다. 대규모 업체 간에 임금 구조 및 업무 성격과 양에 있어서 차이가 있다면 경력직 이직이 활발하겠으나, 아직은 대규모 업체 간에는 그러한 차이가 거의 없는 편이라고 한다. 다만, 소규모 업체들은 상대적으로 임금에서 대기업보다 낮은 수준을 제공하기 때문에 소규모 업체에서 경력을 쌓은 일부 빅데이터 관련 인력들을 채용하는 경우는 있다고 한다.

2. 제조업 기업 조사(외국계 기업 C)

본 연구를 위하여 자동화 관련 신기술 도입에 적극적이고 고도의 스마트공장을 도입하고 있다는 외국계 기업 C의 국내 사업본부를 대상으로 인터뷰 조사를 실시하였다. 해당 부서는 각종 건설 및 에너지 관련 사업을 진행하고 있는 곳이며, 본사는 이 외에도 각종 전기 및 전자, 산업 및

의료 장비 등을 생산, 판매하는 곳이다.

업체 C의 중견 간부는 해당 업체에서 아직 구체적으로 빅데이터 분석이나 데이터 분석을 활용하는 경우는 없다고 답하였다. 에너지 관련 공사 등에서 빅데이터 관련 분석이나 기존의 데이터 분석이 필요할 수는 있으나, 많은 경우에 주로 발주처에서 이러한 분석을 마치고 설계상의 반영 요구 사항에 이미 삽입하는 경우가 많아서 별도의 분석이 필요하지는 않다고 한다. 즉, 각종 시공 및 시설 설치에 있어서 필요한 여러 사항에 대해서는 이미 수요처에서 충분한 분석을 하여 발주 시 요구 사항에 반영하고 있다는 것이다. 업체 C가 자체 분석을 통해 발주처의 요구에 대해서 수정을 요청할 수는 있으나, 일반적으로는 발주처의 특성이 다양하며, 이러한 특성을 일일이 파악하기 위한 조사를 실시하는 것은 힘들 뿐만 아니라, 이 경우 업체의 특성을 모두 반영하지 못하기 때문에 굳이 실시하지는 않는 것이다.

다만 설계 및 시공 초기 단계에서 CAD를 활용하는 등 신기술을 도입하여 자동화 비율을 높이는 것에 대해서는 기업 차원에서 대응하고 있으나, 이는 신기술의 수용이지 이 과정에서 데이터 활용이나 빅데이터 분석이 크게 개입되지는 않는다. 해당 업체는 사내 그룹웨어 역시 제작 및 유지·보수에 있어서 본사와 외주업체에 크게 의존하며, 한국 지사 내에 별도로 데이터 분석팀이나 빅데이터 관련 부서를 두고 있지 않다.

이에 대해서 건설 현장에서 발생하는 각종 재해나 여러 특성에 대해서 데이터를 축적한 후 이를 활용할 여지가 있지 않은지 알아보았으나, 각 시공마다 위치, 지리적 특성, 공사 규모, 계절적 특성과 공사 당시의 가용한 기술 수준 등이 상이하기 때문에 일괄적으로 데이터베이스화하기 힘들며, 해당 사항들을 모두 데이터베이스화하는 것이 오히려 추가적인 부담이 될 수 있다고 언급하였다. 또한 이는 대부분의 건설사와 각종 시공사들에도 적용되는 것이라 하느냐, 제조업 및 건설업 내에서 산업별 데이터의 활용도상의 차이는 크지 않은 것으로 보인다.

일부 의료기기 및 산업용 기기 판매 부서에서 기존의 고객 수요에 대한 데이터를 축적하여 활용할 수는 있겠으나 이에 대해서 별도의 분석팀이나 조직을 가지고 있지는 않으며, 필요한 경우 주로 외부 업체에게 용

역을 맡겨 의뢰하는 것이 일반적이라 한다. 산업용 기기나 건설의 경우 표준화된 제품이 대량 판매되는 것이 아니라 발주처의 요구에 맞추어 소량만이 생산되거나 제공되며, 이후 생산하는 제품은 또 다른 수요처에 따라 다르게 제작되기 때문에 데이터베이스를 구축하기도 어렵고, 구축된 데이터가 광범위하게 여타 사업에서도 활용되기 쉽지 않은 것으로 보인다. 또한 데이터 분석 관련 업무가 상시적으로 있는 것이 아니라서 이러한 수요가 발생했을 때는 해외 본사나 외부 업체에 의뢰하는 것이 보다 효율적이라고 답변하였다.

건설 현장에서 자재의 관리 및 수요에 대해서 데이터베이스화하는 것은 비교적 쉽기 때문에 전사적 자원 관리 시스템(ERP)을 도입하여 활용하는 것은 가능하지 않냐는 물음에 해당 시스템은 이미 도입되어 있으나, 역시 발주처별로 상이한 자재가 소요되고 공사 및 제품의 특성에 따라 요구되는 자재 및 원재료도 다르기 때문에 시스템을 통해서는 자재 및 원재료의 출납 현황을 체계적으로 관리하는 데에 중점을 맞추고 있다고 응답하였다. 즉, 자재 등에 대해서도 데이터베이스 구축은 가능하나 이것이 여타 발주 사업에도 그대로 적용하기 어렵기 때문에 기업 내부의 여러 전산 시스템들은 업무의 효율적인 진행에 초점을 맞추어 설계되어 있으며, 이것이 추후의 조직 및 사업의 개선이나 합리화에 활용되기 위한 목적으로 구축되지는 않아서 전용(轉用)하기 어렵다고 하였다.

이를 통해 현재 제조업 및 건설업 내에서는, 일부 업종을 제외하고, 특히 B2C 업체가 아닌 B2B 업체들의 경우에는 데이터 관련 서비스의 수요가 크지 않으며, 이로 인해 전담 부서나 조직을 설치할 필요성을 느끼지 못하고 있는 것으로 보인다. 한편으로 데이터 관련 인력에 대한 수요 급증으로 인력을 구하기 힘든 노동시장의 현실을 고려한다면, 수요에 맞추어 외부에 의뢰 및 처리하는 것이 반드시 비합리적이거나 근시안적이라고 할 수도 없다.

제2절 데이터 서비스 공급 기업 사례 조사

1. 빅데이터 공급 기업 조사

가. 중소 빅데이터 업체 D

본 사례의 기업 D는 조사 및 분석 업체로서 일부 부서에서 빅데이터 관련 기술을 이용하여 자연어 처리를 대행해 주거나 혹은 각종 온라인 상의 정보를 수집 및 분석하고, 아울러 빅데이터 관련 위탁 조사 업무를 수행 및 검증하는 업무를 한다. 1999년에 설립된 기업 D는 이미 2010년대 중반부터 각종 빅데이터 기술을 도입하기 위한 전담부서 설치와 인력 충원을 시작한 중소·중견 기업 중 가장 업력(業曆)이 오래된 업체이다.

해당 업체의 빅데이터 인력 규모는 전체 약 20여 명 정도이며, 이 중 전문적인 빅데이터 및 데이터 분석 관련 기술 인력은 5명 이내 정도이고, 다수는 앞서 살펴본 금융권과 비슷하게 빅데이터 관련 응용 및 해석 업무를 수행 중이다. 해석 관련 업무에 종사하는 근로자의 다수가 통계학이나 상경 계열을 전공하였으며, 기술직 인력은 컴퓨터공학을 전공하였다.

기업 D는 신규 채용자에 대해서는 정규 채용 절차에 따라 매년 수 회 정도의 공식 채용 절차를 진행하고 있으며, 경력직의 경우 수시로 채용하고 있다. 하지만 중소기업의 특성상 높은 연봉을 주지 못하는 점 등으로 인해 빅데이터 관련 코딩 등의 기술과 결과물에 대한 해석 능력을 모두 갖춘 인력을 채용하는 데에는 난항을 겪고 있으며, 두 업무에 대해서 각각의 능력을 갖춘 인재들을 채용한 후 분리된 업무를 수행하고 있다. 하지만 이러한 신규 채용 및 상시 경력직 채용을 시도함에도 여전히 인력난에 시달리고 있어 빅데이터 관련 인력의 부족을 해당 사례에서도 절감할 수 있었다.

기업 D는 중소·중견 기업 중에서는 이직률이 비교적 낮은 편인 빅데이터 서비스 회사이며, 특히나 기술직 인력이 응용 인력보다 근속기간이

더 긴 편인 것이 특징이다. 회사 내에서 해당 데이터 및 빅데이터 관련 부서와 여타 부서 간의 인력을 따로 관리하고 있지 않아 급여, 휴가 및 승진 체계에서 차등이 존재하지는 않는다. 다만 빅데이터나 데이터 관련 부서보다 그렇지 않은 부서의 규모가 더 크기 때문에 고위급 임원의 분포상 데이터 관련 부서 출신의 인력 규모가 상대적으로 적은 편이긴 하다.

해당 빅데이터 부서들이 하는 역할은 크게 세 가지로 나눌 수 있다. 첫 번째는 직접적으로 빅데이터 분석 관련 서비스를 제공하는 것이다. 하지만 이것은 직접적인 빅데이터 기술 개발 등은 아니며, 빅데이터를 활용하여 신제품에 대한 소비자들의 반응을 조사하거나 시장의 수요를 파악하는 것, 혹은 인물 개인이나 기업과 같은 조직의 평판 및 신뢰도를 분석하는 등의 역할을 하며, 주된 분석 대상 역시 온라인에서 수집되는 자료에 한정되어 있다. 즉, 해당 업무는 직접적으로 빅데이터 관련 기술을 개발하거나 연구하기보다는 키워드 중심으로 온라인상에 존재하는 다양한 정보를 모아 단순한 분석을 제공하는 것이다. 이 경우 일부 빅데이터 전문가들은 이를 빅데이터 분석이 아니라고도 하며, 사안에 따라서는 데이터 자체의 크기도 작기 때문에 빅데이터 분석이라 부르기 힘든 경우도 있으나, 온라인상의 많은 데이터를 쉽게 처리해서 보거나 분석하기 편하게 제공한다는 측면에서 빅데이터 분석이라 부르고 있다.

두 번째 업무는 자연어 처리와 기계학습 등 보다 직접적으로 빅데이터 관련 기술을 활용하는 분석이다. 앞선 빅데이터를 활용한 신뢰도 및 평판 분석이 긍정/부정의 단순한 이분법적인 평판에 대한 분석이라면, 최근의 기계학습은 보다 긴 자연어 자체를 수치화하여 그대로 분석함으로써 단순한 단어 및 표현의 등장 빈도나 위치 등에 기반한 단순 분석 이상을 제공하는데, 해당 서비스는 이러한 영역에 속한다. 즉, 앞선 첫 번째 업무가 온라인상의 자료를 수집하여 기초통계를 제공하고 향후의 분석에서 활용할 수 있는 방향의 데이터를 구축하는 데 초점이 맞추어져 있다면 본 업무는 빅데이터 기술을 활용하여 이를 비정형 데이터로 구축하거나 비정형 데이터를 정형 데이터로 변환하고 구축하는 데 초점을 두고 있는 것이다.

첫 번째 및 두 번째 업무와 관련된 특징 중 하나는 일부 업무의 경우 외주를 주기도 한다는 점이다. 빅데이터 관련 기술적 인력이 많지 않으며,

학력 수준도 매우 높지 않기 때문에 의뢰가 들어오는 모든 위탁 사업에 대해서 자체 인력만으로 처리가 되지 않으므로 해당 업무들은 외주 처리한다는 것이다. 이는 앞서 살펴본 B 카드사의 경우에도 언급했듯이, 일부 카드사들은 빅데이터 관련 업무의 일부를 자체에서 소화하되 몇몇 업무를 외주에 맡기는 것과 유사하다 하겠다.

세 번째는 여타 부서에서 조사된 결과에 대해서 이상치를 검증하거나, 결과가 정확한지를 검사하는 역할이다. 조사의 규모와 범위 및 대상이 커지면 이 역시 빅데이터와 비슷한 수준이 되는데, 이에 대해서 결과치의 이상이나 결과의 불일치성 등을 점검하는 것이다. 이는 앞서 살펴본 금융권 사례에서 등장하는 기존의 데이터 분석 부서에서 수행하던 역할과 비슷하게 기존의 데이터 분석이 적용되는 분야이다.

기업 D는 이러한 빅데이터 부서의 업무 처리를 위해서 신규 인력 채용과 경력직 스카우트를 모두 활용하고 있으나, 앞선 금융권과의 차이점은 내부 여타 부서에서 빅데이터 부서로의 전직이 많지 않다는 점이다. 현재 빅데이터 부서에 근무하는 인력들과 여타 조사 부서들에 근무하는 인력들의 학력 수준이나 경력상의 특징, 그리고 전공은 대체로 비슷하나 비슷한 부서 내에서 이동하는 경우는 있어도 데이터 분석 관련 부서와 그렇지 않은 부서 간의 인력 이동은 많지 않은 것으로 조사되었다. 이는 앞서 살펴본 두 금융권 회사들이 대기업이기 때문에 자체적으로 내부 교육 프로그램을 진행할 수 있었으나 기업 D는 전체 직원이 100여 명 내외의 중견 기업이기 때문에 그러한 프로그램을 운영하는 것이 사실상 힘들다. 그러므로 한 번 숙련된 업무 이외의 다른 업무를 익히기가 어려우며, 만일 빅데이터 관련 기술과 지식을 습득한 경우 반드시 내부에서의 부서 이동이 아니라 외부로의 이직도 가능하기 때문에 내부에서의 인력 이동이 많지 않은 것으로 보인다.

다만 희망자에 한하여 정부에서 실시하는 교육에 파견을 보내고 교육 비용을 지불하거나, 혹은 대학원에 진학하여 빅데이터 관련 교육을 이수하는 것을 장려하는데, 대학원의 경우 학비를 보조하는 정도는 근속 기간 및 직급에 따라 다르게 책정하여 30% 이하부터 최고 100%까지 지원하고 있다. 한편 빅데이터에 한정되지는 않지만 퇴근 후 및 휴일에 실력 양성

을 위한 사교육을 받는데 대해서도 지원이 이루어지고 있으며, 이에 따라 일부 인력들이 빅데이터 관련 교육을 받기도 한다.

이는 새로운 기술에 적합한 인력 양성 시 기업 내부에서 교육 및 훈련 프로그램을 진행하면 내부 직원들에 대한 재교육과 신기술 대응이 가능하지만 그렇지 않은 중소기업들에서는 이러한 재교육 프로그램을 운영하지 못함에 따라 기존 인력들이 신기술로부터 소외될 수 있음을 시사한다. 또한 내부 재교육 프로그램의 존재 여부가 직원들의 신기술 습득 수준 및 내부 인재 양성에도 차이를 주고 있음도 알려주고 있다.

2. 일반 데이터 공급 기업 조사

가. 중견 데이터 업체 E

본 사례의 기업 E는 B2C(Business-to-Consumer)가 아니라 B2B(Business-to-Business) 업체이다. 해당 업체는 1990년대 중반에 설립되어 국내와 동남아시아에서 사업을 진행하고 있다. 2000년대 초반 사업 부문을 대대적으로 조정하여 모바일 프로그램 및 소프트웨어에 대한 B2B 판매를 제외한 나머지 부분을 매각하고 조달청 입찰 분야도 정리한 후 현재는 일부 직접 판매 회사들의 판매 시스템을 공급, 유지 및 보수하는 업무 위주로 사업을 조정하였다. 기업 E는 여러 데이터 산업 종사 기업 중에서는 비교적 규모도 크고 건실한 편이며, 영업이익 등의 재무지표 역시 좋은 편이다. 사업 구조조정 이후 매출의 반 이상이 직접 판매 회사 두 곳의 사용자 인터페이스(User Interface)를 제공 및 관리하는 데에서 창출되어 여타 기업들과 비교하면 매출이 안정적이고 사업의 불확실성이 적다고 할 수 있다. 이렇게 고정 수입이 반 이상을 차지하기 때문에 프로젝트 단위로 경영 성과의 부침이 큰 중소기업 데이터 서비스 제공 업체들과 비교하여 고용 안정성 및 급여 지급의 안정성이 높은 편이며, 영세 업체들에서는 흔한 편인 임금 체불이나 미지급 사태는 발생한 적이 없다.

기업 E는 현재 약 100여 명의 동남아시아 현지 인력을 포함하여 국내에서 약 40여 명이 근무하고 있으며, 국내 인력 중 반 정도는 본사에서 근무

무하고 나머지 반은 프로젝트의 수행을 위해 발주처에 파견되어 있거나 장기 거래처의 A/S를 위해 상주하고 있다. 국내에서 선발하는 인력은 주로 산학협력 제휴를 맺은 지방 대학교나 수도권 중위권 대학교에서 충원하며, 동남아시아 인력은 한국계 기업으로서 현지 기업보다 높은 연봉을 제시하기 때문에 해당국의 최정상급 대학 졸업생들로 채워져 있다.

데이터 서비스 제공 업체로서 전체 인력의 약 80% 정도가 프로그램 개발자이며, 영업과 경영지원 인력이 나머지를 구성하고 있다. 특히 거래처 및 발주처에 파견된 인력은 거의 다수가 프로그래머로 구성되어 있으며, 영업과 경영지원 인력은 주로 본사에서 활동하고 있다. 동남아시아 파견 국내 인력은 소수의 경영 지원 관련 인력을 제외한다면 다수가 시니어급 개발자로 현지 인력에게 소프트웨어나 시스템 개발 관련 노하우와 관리 방식을 교육하고, 진행 중인 프로젝트 전반의 관리와 제품 설계를 담당하고 있다.

기업 E의 내부 조직은 프로그래머 직군과 나머지 영업 및 경영지원 직군이 있는데, 영업 및 경영지원 직군의 경우 소속 부서가 일정하게 변하지 않는 반면, 개발 부서의 경우 프로젝트에 맞추어 부서 간 인력 이동이 잦은 편이다. 또한 영업 및 경영지원 직군은 근무 시간에 편차가 크지 않은 반면, 프로그래머 직군은 프로젝트의 진행 과정에 따라 근무 시간의 편차가 크게 나타나고 있다. 이것이 반드시 단점은 아닌 것이 일부 단기 프로젝트의 경우, 3~4개월 정도의 장시간 근무 이후 2주 정도의 휴가가 주어지는데 이 경우 해당 휴가의 활용도 및 만족도가 낮지 않은 것으로 조사되었다. 직군에 따른 근무 형태뿐만 아니라 인력 특성도 매우 상이한데, 프로그래머 직군의 경우 주로 컴퓨터공학 전공자들이며 나머지 직군들은 대체로 인문·상경 계열 출신들이 많다.

직군 간의 다른 특성으로 이직에 관한 것을 들 수 있다. 프로그래머 직군의 경우 일부 핵심 인력을 제외하면 대체로 3년마다 지속적으로 이직을 하는데, 이직을 해야 연봉이 많이 상승하기 때문이다. 반면 비개발자 직군의 경우 40대 중반 이전에는 비슷한 주기로 계속해서 이직을 하지만 그 이후에는 한 군데에서 근무하는 경향을 보인다. 다만 개발자 직군의 경우 워낙 시장의 수요가 많기 때문에 일단 어느 정도의 숙련과 기술이

확보되면 나이가 많더라도 충분히 이직의 기회가 주어져 이직을 하는 경우가 많다.

내부 인력이 이직을 하는 경우 다시 신규 및 경력직 채용을 하고 해당 선발 인력에 대해서 추가로 교육을 진행해야 해서 적어도 핵심 인력에 대해서는 회사에서 특별히 관리를 하여 이직을 막으려 하고 있다. 그렇다고 연봉 등에 대해서 차등지급하는 것은 법률 위반 소지가 있으므로, 핵심 개발 인력의 경우 청년층은 청년내일채움공제 가입을 권유하고 중년층의 경우 재직자내일채움공제 가입 혜택을 제공한다. 기업 E의 경우 핵심 인력을 관리하기 위한 특별한 방법으로 청년내일채움공제와 재직자내일채움공제의 근로자 납입분에 대해서도 회사가 대납하여 공제 가입 기간 동안 해당 가입자들의 가치분 소득이 하락하는 것을 막고 있다. 즉, 다수의 가입자들이 공제에 가입함으로써 기업 가입분과 정부 지원분뿐만 아니라 근로자 기여분까지 회사에서 지급받아 내일채움공제 수령액 전액이 추가적인 임금 상승으로 이어지는 구조이다. 이러한 공제 가입이 이직을 방지하는 효과는 뚜렷하여 적어도 핵심 인력 내에서 40대 중반 이전에 이직하는 프로그래머의 비율은 높지 않은 것으로 나타났다. 다만 핵심 인력과 비핵심 인력 간의 위화감 조성을 방지하기 위해서 내일채움공제 가입에 대해서는 개별 면담을 통해 진행하며, 여타 근로자들에게 발설하지 않고 비밀을 유지할 것을 근로자들에게 요구하고 있다.

한편 기업 E는 프로그래머 직군의 다수가 청년층이거나 젊은 층이어서 이들을 대상으로 한 맞춤형 복지 서비스를 제공하고 있다. 특히 원두와 원두 그라인더를 구비한 것이 사내 만족도가 높는데, 이로써 사내 근로자들이 질 좋은 커피를 마시기 위해 외부에 나가지 않아도 되고 사내에서 진행되는 회의의 수준을 높여 보다 많은 거래처와의 회의가 사내에서 진행될 수 있도록 돕고 있다. 실제 경영진 측에서는 원두와 원두 그라인더의 구비 이후 근로자들의 평균 외출 시간이 감소하여 실질 근로 시간이 증가함에 따라 해당 서비스 제공에 대한 비용을 충분히 상쇄할 정도의 이익이 있었다고 내부적으로 진단하였다. 이와 함께 사내 휴게실과 회의실에 다과를 비치함으로써 아침 식사를 하지 못하고 출근하는 직원들이 사내에서 일하면서 아침 식사를 해결하거나, 혹은 바쁜 시기에 사내에서 간

단하게 끼니를 해결할 수 있도록 하였다. 이에 따라 직원들이 회식 및 야근 이후에도 정시 출근하는 비율이 증가하는 등 해당 비용을 상회하는 경영상의 이득이 있는 것으로 내부 평가 결과가 나타났다.

기업 E 역시 2~3년 단위의 직원 이직으로 인해 매년 꾸준한 인력 수요가 있으며, 이를 위해 신규 인력과 경력직 모두 채용 중이다. 특히 신규 채용에서의 원활한 인력 수급을 위해 일학습 병행제를 제도 초창기부터 도입하여 시행 중이며, 현재까지 근로자들과 경영진 모두 만족도가 높은 편이라고 한다. 일학습 병행제의 효과 중 하나로는 특정 대학의 전공교수 한 명과 긴밀하게 협조하여 프로그램을 운영함에 따라 해당 교수 아래의 학생들이 기업 E의 업무에 보다 친숙하게 잘 적응할 수 있는 방향으로 교육이 이루어지고, 이로써 신규 채용자라 하더라도 업무 적응 속도가 빠르며 업무 처리 능력도 좋아진 점을 들 수 있다. 해당 교수와 학교의 경우 기업 E와의 일학습 병행제 운영으로 재학생의 취업률을 높이는 데 효과를 보고 있으며, 이들이 취업을 걱정하지 않고 안정적으로 기술 습득 및 학업에 매진하도록 하는 효과를 보고 있어 만족도가 높아 해당 프로그램이 장기간 지속되고 있다고 한다.

나. 중소 데이터 업체 F

기업 F 역시 주로 B2B 사업을 하는 업체로서 1990년대 후반에 설립되어 기업에 여러 품질 관리 솔루션 프로그램을 개발 및 제공하는 일을 하고 있다. 현재 기업의 인력 규모는 대략 30~40여 명 정도이며, 소수의 경영 지원 인력을 제외한 다수가 프로그래머 직군에 해당한다.

이들이 기업에 제공하는 제품은 전사적 기업 관리 시스템, 생산 과정 관리 시스템, 통계적 공정 관리 시스템이나 품질 관리 시스템 등으로, 기업의 수요에 맞추어 다양한 종류의 프로그램을 제공하고 있다. 해당 기업은 홍보에도 적극 나서서 다양한 제품 전시 박람회 등에 참여하고 있으며, 인력 채용 사이트에도 꾸준히 인력 공고를 내면서 기업 홍보도 겸하고 있다.

해당 기업 역시 다수의 프로그래머 직군들이 프로젝트 단위로 조직되

어 근무하여, 특정 시기에는 근로 시간이 길어지지만 장기 프로젝트가 종료되면 소수의 사후 지원 인력을 제외하고는 긴 휴식을 갖거나 짧은 시간만 근로를 하는 형태를 띠고 있다.

역시 다수의 인력이 컴퓨터공학 전공자이며, 근로자들의 평균 연령대는 40 전후로 상당히 젊은 편이다. 그러나 한편으로는 중소기업이기 때문에 대기업 및 중견 기업과 비교하여 연봉이나 처우가 열악하여 이직이 잦은 편이라고 한다. 이들 기업 역시 핵심 인력의 이직을 막기 위한 복안을 고민 중이나 아직은 중소기업으로서 제공할 수 있는 여력이 많지 않기 때문에 특별한 유인책을 제공하고 있지는 못하다고 한다. 다만 중년 이상의 경력직은 곧 중견 관리자급이 되기 때문에 이들의 경우 상대적으로 오랜 기간 일하기는 하나, 역시 보다 높은 연봉을 제외받으면 이직하는 경우가 상당수라고 한다.

기업 F와 기업 E의 가장 큰 차이는 프로젝트의 성격에 따른 외주 여부이다. 기업 E는 매출도 상대적으로 안정되고 규모도 중견기업이기 때문에 일부 프로젝트의 경우 내부 인력이 바쁘거나 전담 인력이 없을 때는 보다 작은 규모의 기업체에게도 하청이나 위탁을 주는 반면 기업 F의 경우 재하청의 여력이 많지 않다는 점이다. 이에 따라 기업에서 제공하는 제품들을 어느 정도 표준화해 놓고 새로운 시스템이나 프로그램에 대한 발주에 웬만해서는 참여하지 않으려 하는 등 여러가지 노력을 기울이고 있으나, 발주된 프로젝트에 대해서는 최대한 외주를 맡기지 않으려 한다. 대신 기업 F 역시 매출 안정 등을 위해서 사업을 특화하고 있으며, 이를 위해 사출 및 프레스, 그리고 조립 가공품 산업에 대한 시스템에 특화하여 해당 기업들로부터 많은 주문을 받고 있다.

기업 F의 신규채용 입사자들 다수는 청년이며 컴퓨터공학 전공자이거나 수도권의 주요 대학 출신들이 아니고, 지방대 출신이 많은 편이다. 하지만 대기업 G의 인증 시스템을 통과한 후 해당 대기업 하청 업체들과의 거래가 늘면서 상대적으로 매출이 안정되고 있고, 이에 따라, 여전히 높기는 하지만, 이후 이직률이 조금 낮아진 편이다. 기업 측은 이를 바탕으로 핵심 인력에 대한 관리에 더 많은 자원을 투자하기 위해 여러 방안을 고민 중이라는 것이 중견 간부의 응답이었다.

제3절 소 결

본 장에서는 실제 데이터 서비스 수요 기업들과 공급 기업들에 대해서 방문 및 인터뷰 조사를 진행하여 데이터 산업 관련 노동시장의 특성과 인력 관리 방식을 보다 심층적으로 파악해 보았다. 이를 통해 우리는 데이터 서비스 수요 기업과 공급 기업들의 몇 가지 특성을 파악할 수 있었다.

첫 번째로 데이터 서비스 수요 기업의 경우, 데이터 관련 서비스이건 빅데이터 관련 서비스이건 모두 응용 분야에 집중하여 수요가 발생하고 있으며, 해당 서비스의 기초 기술에 대한 인력 수요는 비교적 적은 편이라는 점이다. 이는 본 장에서 살펴본 서비스업 기업들 모두 실제 신규 채용 인력들의 전공이 상경계열이라는 점에서도 나타난다. 일부 통계 전공자들도 선호하나 실제 컴퓨터공학 전공자들의 경우는 해당 산업에 대한 이해 부족으로 채용 이후 오히려 더 많은 재교육 시간과 비용이 들어가기 때문에 선호하지 않는 편이다. 즉, 즉각적으로 현장에서 활용 가능한 기술이나 분석을 수행할 수 있는 능력을 갖춘 응용 분야에 대한 인력 수요가 많다는 것이다.

두 번째로 컴퓨터공학 전공자들의 경우 많은 인력 수요로 이직이 잦으나, 데이터 분석 기술을 갖춘 응용 분석가들의 경우 현재 노동시장에서 찾아보기 힘들다는 점이 있다. 즉, 기초기술을 가진 공학자 및 프로그래머들은 경력직 채용 시장에서 쉽게 찾아볼 수 있으나, 빅데이터 관련 기초 기술을 가지고 있으면서 특정 산업에서 바로 응용할 수 있는 분석 능력과 소양도 갖춘 사람들은 현재 노동시장에서 찾아보기 힘들다. 따라서 이들은 신규 채용 및 사내 직원 재교육을 통해 선발되고 있다. 이런 까닭에 앞으로 시장 수요에 부응하는 인재를 공급하기 위해서는 단순히 프로그래밍을 할 줄 아는 공학자보다는 각 전공별로 프로그래밍을 활용할 수 있는 졸업생을 양성하여 다양한 분야에서 발생하는 프로그래밍 응용 관련 수요를 충당할 수 있어야 할 것이다.

세 번째로, 많은 데이터 서비스 공급 기업들이 영세하다는 점을 들 수

있다. 이들 기업은 영세하기 때문에 핵심 인력을 관리할 만한 여력을 갖추지 못하여 재직자들이 잦은 이직을 하게 되고 이로 인해 기업의 성과나 생산성이 떨어지는 악순환이 반복될 가능성이 있다. 이것이 문제가 되는 것은 상경계열 졸업생의 경우 프로그래밍 소양을 갖추게 되면 여러 금융권 회사들로부터 채용되어 상대적으로 괜찮은 수준의 연봉과 근로 환경을 맞이하게 되나, 컴퓨터공학자들이나 프로그래머들의 경우 이러한 기회가 적고 오히려 계속해서 이직을 통해 본인의 연봉과 처우를 개선해 나가야 하기 때문이다. 따라서 빅데이터 및 데이터 분석과 관리에 대한 기초 기술과 소양을 가진 인재 양성이 장기적으로 지체되고, 응용 능력만 장려되는 문제가 발생할 수 있다.

마지막으로 사내 재교육 프로그램 및 재직자 대상 재교육 프로그램의 성공 가능성을 언급해야겠다. 금융권 기업들은 급증하는 인력 수요를 전부 다 신규 채용 및 경력직 채용으로 충당치 못함에 따라 사내에서 기존의 재직자들을 대상으로 재교육 및 훈련 프로그램을 운영하고 있으며, 이는 비단 데이터 및 빅데이터 관련 부서 이동자들이나 이동 희망자들만을 대상으로 하고 있지도 않다. 이로 인해 사원 다수가 빅데이터와 분석 기법에 친숙해짐에 따라 장기적으로는 빅데이터 구축에 용이한 방향으로 업무가 이루어지며, 이에 따라 양질의 데이터가 구축되어 기업의 성과가 장기적으로 개선될 여지가 있다. 반면, 같은 서비스업 기업이라도 사내 교육 프로그램이 없는 회사의 경우 빅데이터 및 데이터 분석 관련 인력을 주로 외부에서만 충당해야 하는 부담이 있으며, 이는 장기적으로 빅데이터 관련 부서와 여타 일반 부서 간의 인식 차이를 유발하고 교류를 줄이게 될 여지를 낳는다. 만일 기업체에서 이러한 재교육 및 훈련 프로그램을 제공하기 힘든 경우 정부나 공공기관 및 유관 단체에서 해당 프로그램을 제공할 수도 있을 것이다. 그러나 이 경우 과연 이러한 표준화된 교육 프로그램이 프로그램 신청자 개인에게 얼마나 도움이 될 것이며 향후 응용 가능할 것인가의 문제와 언제 시간을 내서 이러한 프로그램을 이수할 것인가 하는 문제 등이 예상되므로 선불리 양적인 확대를 추구하는 정책을 펼치기보다는 시장의 수요에 맞추어 신중하게 결정할 필요가 있을 것이다.

제 6 장

결 론

데이터 산업은 그 자체가 창출하는 부가가치도 작지 않지만 기존의 다른 산업 및 서비스와 연계되어 창출하는 연계적 부가가치가 매우 큰 특징을 가지고 있는 산업이다. 기존의 많은 연구들이 데이터 산업의 발전에 따라 창출되는 본연의, 혹은 연계된 부가가치의 크기에 주목하거나, 새롭게 등장하는 산업 및 사업 분야에만 초점을 맞추는 경향이 있다. 다른 한편으로는 기술적 관점에서 최근 부상하고 있는 다양한 데이터 관련 기술과 활용도에 대한 연구들도 폭넓게 이루어지고 있다. 하지만 기존에도 있었거니와, 최근 새로운 기술의 등장으로 성장하고 있는 데이터 산업에서의 노동시장의 모습은 상대적으로 주목받지 못한 상황이다.

본 연구에서는 기술적인 측면이나 사회적인 각도에서 접근한 데이터 산업에 대한 여타의 연구와는 달리 데이터 산업의 노동시장에 대해서 수요 측면과 공급 측면으로 나누어 다각도로 살펴보았다. 데이터 산업은 그 특성상 기존의 사회 내 여러 분야와 접목되기 쉽다. 이는 데이터 산업 자체 혹은 데이터 서비스 공급 측면에서 인력 수요가 증가한다는 차원을 넘어서는 일이다. 기존 산업에서 이미 널리 알려진 여러 사업 분야에 데이터 분석 기술 등이 접목됨으로써 산업의 생산성을 높이거나, 사업의 모습을 일신하는 형태로 응용되는 것이다. 따라서 이러한 데이터 관련 응용 분야의 인력 수요 역시 크게 증가하고 있으며, 다만 시장에서 요구하는 학력 및 경력을 갖춘 인력이 원활하게 공급되지는 못하고 있는 실정이다. 이는 2장과 3장, 그리고 4장과 5장에서 지속적으로 확인되고 있다. 2장에

서는 데이터 산업 분야 충고용 인력의 10% 이상에 대한 초과 수요가 존재하며 다수가 응용 분야임이 확인되었고, 4장에서는 채용 공고에 대한 지원자가 있음에도 적합한 능력이나 경력을 가지고 있지 않아 채용치 못한다는 기업이 채용 부족 문제에서 가장 큰 비율을 차지했다. 한편 이는 데이터 공급 및 수요 기업 모두에서 해당됨이 5장에서 확인되다. 따라서 기초 기술과 응용 기술 양 분야에서 최신 기술을 습득한 인력을 양성하여 시장에 공급할 수 있는 방안을 모색할 필요가 있다. 이러한 인력 수요가 충족된다면 해당 사업체들은 데이터 분석을 기반으로 보다 나은 재화와 용역을 공급함에 따라 사회 전반의 생산성이 올라가고 국제경쟁력 또한 제고될 수 있어 더 많은 부가가치가 창출될 수 있기 때문이다.

이러한 인력 수급 불균형의 문제는 특히나 빅데이터 분야에서 더욱 크게 나타났는데, 빅데이터 관련 분석 능력을 갖춘 응용 분석가에 대한 수요가 많이 있으나 한국은 미국과 비교하여 해당 기술의 수준도 낮을 뿐만 아니라 관련 인력도 부족한 실정이다. 다만 응용 분야에 대해서 인력 수요가 집중되어 있고, 기초 프로그래밍 및 기초적인 빅데이터 기술 분야보다 응용 분야의 보수와 처우가 좋다는 점은 장기적으로 기초 기술에 대한 등한시로 이어질 가능성을 내포하는 대목이다. 따라서 응용 분야의 인력을 공급하되, 빅데이터 관련 분석 등에서의 기초 기술 능력을 갖춘 인재들을 육성하고 이들이 보다 높은 수준의 기술을 연구할 수 있는 환경을 제공하기 위한 지원 방안을 고민해 봐야 할 것이다.

아울러 사례 조사를 통해 빅데이터 관련 인력이나 데이터 관련 전문 인력의 초과 수요에 직면한 대기업들이 내부 직원들을 대상으로 재교육 및 훈련 프로그램을 운영하여 인력을 충당하고 있으며, 해당 프로그램에 대한 일반 직원들의 참여 욕구가 크다는 점이 확인되었다. 이러한 교육 프로그램은 데이터 관련 인력 충원 문제를 해결할 수 있는 방안일 뿐만 아니라 데이터에 대한 근로자 전반의 이해 수준을 높임에 따라 양질의 데이터가 원활하게 축적되고 사용될 가능성을 높여준다는 측면에서 긍정적이라 할 수 있다. 하지만 많은 영세 중소·중견 기업들의 경우 장시간 근로로 인해 이러한 교육 프로그램을 받을 만한 시간을 찾지 못할 뿐만 아니라 이러한 훈련 프로그램을 운영할 수 있는 여유 역시 가지고 있지 못

한 실정이다. 또한 잦은 이직과 열악한 처우 등 데이터 산업 종사 기업의 다수를 차지하는 30인 미만의 영세 업체에서 관찰된 여러 현상들은 여타 산업 내 영세 기업의 문제와 크게 다르지 않은 것이 확인되었다. 이는 단순히 노동 공급 측면에만 초점을 맞추어서는 데이터 산업에서의 노동시장의 미스매치 문제를 해결하기 힘들다는 점을 암시한다. 따라서 빅데이터 등의 첨단 분야에 대한 인력 공급만으로는 입사 이후 잦은 이직과 낮은 근속 기간으로 인해 인적 자본이 축적되지 않는 문제를 해결하기 힘들 것이다. 그러므로 노동 수요 측면에서 영세한 데이터 기업체들이 경쟁력을 가지고 성장하여 중견 기업으로 클 수 있도록 데이터 산업에서도 기존의 중소 영세 제조업체들을 대상으로 하는 육성책과 지원책을 시행함으로써 기업들의 스케일 업과 재직자 처우 개선을 유도할 필요가 있다. 아울러 기존에 노동 공급 측면에만 맞추어져 있던 정책 방향을 노동 수요 측면으로도 배분하여 재직자들의 근속 기간을 늘리고 이직률을 낮추기 위한 방안을 모색해야 할 것이다.

참고문헌

- 과학기술정보통신부·한국데이터진흥원(2018), 「2017 데이터 산업 현황조사」, 2018년 4월.
- 권배근(2015), 「기계학습 기반의 표면결함 검사에 대한 연구」, 부산대학교 기계공학부 제어자동시스템 전공, 박사학위논문.
- 박훤일(2017), 「개인정보의 현지화에 관한 연구」, 『경희법학』, 52(4), pp. 129~164.
- 서동혁(2018), 『4차 산업혁명 관련 산업분류체계 개선 및 지수개발』, 산업연구원.
- 이규엽 외(2018), 『국경 간 데이터 이동에 관한 국제적 논의 동향과 대응 방안』, 대외경제정책연구원.
- 정지원(2019), 『인천광역시 데이터 산업 활성화 방안』, 인천연구원.
- 정혁·정용찬·정진한·고동환·유선실·정부연·김민식·이경남·오정숙·이은민·나상우·김대건·이선희(2018), 「ICT 산업 중장기 전망(2018~2022)」, 『방송통신정책연구』, 정보통신정책연구원.
- 통계청 국가통계 포털.
 _____(2017), 「한국표준산업분류(10차)」
- 한국데이터진흥원(2016), 『데이터 전문인력 양성 방안 연구 결과 보고서』.
 _____(2018), 『2018 데이터 산업 백서』.
 _____(2018), 『데이터 산업 현황조사』.
- 4차산업혁명위원회(2018), 『데이터 산업 활성화 전략 - I-KOREA 4.0 데이터 분야 계획, I-DATA』.
- 한국은행(2019), 『2015년 산업연관표』.
- 중소기업중앙회(2018), 『2017년 기준 중소기업실태조사 보고서』.
 _____(2019), 『2018년 기준 중소기업실태조사 보고서』.

- Bauer, M., H. Lee-Makiyama, E. Marel, and B. Vershelde(2014), “The Costs of Data Localization : Friendly Fire on Economic Recovery,” ECIPE OCCASIONAL PAPER·No.3, European Centre for International Political Economy(ECIP).
- Bureau of Labor Statistics(2018), “Occupational Employment Statistics”.
- EMC Corporation(2012), “IDC’s Digital Universe Study : Big Data, Bigger Digital Shadows and Biggest Growth in the Far East”.
- Laney, Doug(2001), “3D data management : Controlling data volume, velocity and variety”, META Group Research Note. 6(70).
- Newman, David(2011), “How to Plan, Participate and Prosper in the Data Economy”, Gartner 보고서.
- NIST Big Data Interoperability Framework(2015), “NIST Big Data Interoperability Framework”, NIST Special Publication 1500.
- Statista(2019), <http://www.statista.com>(접속일자 : 2019. 9. 24).
- The International Data Corporation(IDC) and The Lisbon Council(2016), “First Report on Facts and Figures : Updating the European Data Market Monitoring Tool”.
- _____ (2018), “The European Data Market and Monitoring Tool Report”.

◆ 執筆陣

- 방형준(한국노동연구원 부연구위원)
- 손연정(한국노동연구원 부연구위원)
- 노세리(한국노동연구원 부연구위원)

데이터 산업의 노동시장 분석

- 발행연월일 | 2019년 12월 26일 인쇄
2019년 12월 30일 발행
- 발행인 | 배규식
- 발행처 | **한국노동연구원**
30147 세종특별자치시 시청대로 370
세종국책연구단지 경제정책동
☎ 대표 (044) 287-6080 Fax (044) 287-6089
- 조판·인쇄 | 미래기획 (044) 866-6331
- 등록일자 | 1988년 9월 13일
- 등록번호 | 제13-155호

© 한국노동연구원 2018 정가 7,000원

ISBN 979-11-260-0366-2