

「데이터로 연결하는 사람·세상·미래 Bridging Data」는 디지털 혁신을 이끄는 핵심 자원인 데이터를 잘 활용하기 위해 주요 이슈를 분석하고 이를 해결하기 위한 전문가들의 제안을 담은 한국정보화진흥원(NIA)의 보고서입니다.

NIA의 승인 없이 본 보고서의 무단전재나 복제를 금하며, 인용하실 때는 반드시 NIA, 「데이터로 연결하는 사람·세상·미래 Bridging Data」라고 밝혀 주시기 바랍니다.

보고서 내용에 대한 문의나 제안은 아래 연락처로 해주시기 바랍니다.

■ **기 획** : 한국정보화진흥원(NIA) 정책본부 미래전략센터

Tel. 053-230-1298

■ **작 성** : 제1편. 한성대학교 김상봉 교수

제2편. 건국대학교 장형준 교수

제3편. 경희대학교 김도훈 교수

제4편. 한국고용정보원 김형래 박사, (주) 이앤지테크 정경희 연구원

제5편. 국립경남과학기술대학교 조문증 교수

CONTENTS

제1편 데이터 가치평가의 필요성과 정책적 준비	1
제2편 데이터 공유·활용 플랫폼의 특성 및 사례 분석	65
제3편 데이터기반 경제에서 알고리즘의 역할과 책임성	165
제4편 데이터 혁신과 고용 창출	225
제5편 엣지 컴퓨팅으로 변화되는 미래사회	307

데이터로 연결하는 사람·세상·미래



Bridging DATA

제1편

데이터 가치평가의 필요성과 정책적 준비

NIA 한국정보화진흥원



CONTENTS

서론	5
I. 데이터산업 시장 현황	6
1. 해외	6
2. 국내	11
II. 데이터 가치평가를 위한 기술금융 및 가치평가 사례 분석	14
1. 기술금융	14
2. 기술가치평가	31
III. 데이터 가치평가 방법, 사례 및 정책 제언	53
1. 소프트웨어 가치평가	53
2. 데이터 가치평가 가능 지표와 사례	58
참고문헌	63

서론

본 자료는 해외 및 국내의 데이터산업 현황을 살펴보고, 데이터 가치평가에 대한 사례를 분석하는 것을 목적으로 한다. 이를 위하여 크게 해외의 데이터산업 시장규모 및 국내 시장규모나 증가 추이 등을 살펴볼 필요가 있다. 또한, 데이터 가치평가에서 중요한 부분은 지표이다. 이를 위하여 데이터 가치평가를 포함하는 선행연구인 기술금융과 기술가치평가에 관한 사례 분석이 필요하다.

데이터는 거시적인 특성과 미시적인 특성을 모두 가지고 있다. 데이터는 거시적인 차원의 기술을 통하여 나오는 결과 또는 과정으로 볼 수 있으므로, 기술을 바탕으로 금융이 지원되는 기술금융 현황과 지표를 살펴볼 필요가 있다. 동시에, 데이터는 미시적인 차원과 관련된 기술에서 나오는 결과 또는 과정으로도 볼 수 있으므로, 거시적인 측면보다 범위를 좁혀 기술가치평가 방법론과 지표를 살펴보는 것도 의미가 있다.

거시적 측면인 기술금융과 보다 미시적인 기술가치평가에 대해 살펴볼 필요가 있고, 기술가치평가의 범위를 좁혀, 현재 존재하지 않는 데이터 가치평가와 가장 가까운 소프트웨어 가치평가에 대하여 이해할 필요가 있다. 소프트웨어의 가치평가를 통해 데이터 가치평가에 대한 보다 세분화된 지표를 통해 정보서비스업과 정보데이터(DB) 관리업 등은 현재 데이터 가치평가와 밀접한 관련이 있거나 데이터 가치평가의 세부항목으로 잡힐 수 있는 변수이기 때문이다.



데이터산업 시장 현황

1 » 해외

데이터는 4차 산업혁명에서 중요한 핵심 자원이 되고 있으며, 다양한 응용 분야에서 새로운 가치를 창출하고 있어 시장규모는 점차 확대 중이다. 2017년 기준 세계 데이터 시장규모(IDC, 2017년 3월 기준 예측)에 따르면, 세계 데이터 시장은 2017년 1,508억 달러(12.4%증가)에서 2020년 2,100억 달러로 연평균 성장률은 11.9%로 예측되고 있다. 2017년에 가장 큰 데이터 산업은 은행, 이산형 제조업, 프로세스 제조업, 연방/지방 정부, 전문서비스업이다. 이러한 2017년 상위 5개 산업의 수익은 724억 달러로 예측되며, 2020년까지 1,015억 달러가 투자될 것으로 예측되고 있다.

데이터 산업에서 은행은 연평균 성장률 13.3%로 가장 빠르게 커지고 있으며 건강, 보험, 주식 및 투자 서비스, 통신 등은 연평균 12.8%의 성장률을 나타낼 것으로 예측되고 있다.

2018년 기준 세계 데이터 시장규모(IDC, 2018년 8월 기준 예측)는 2017년 1,660억 달러(11.7% 증가)에서 2022년 2,600억 달러로 연평균 성장률은 11.9%로 예측되고 있다. 2017년에 가장 큰 데이터 산업은 은행, 이산형 제조업, 프로세스 제조업, 연방/지방 정부, 전문서비스업으로, 2017년 상위 5개 산업의 수익은 810억 달러이며, 2022년까지 1,290억 달러가 투자될 것으로 예측되고 있다. 이 중 소매업(연평균 성장률 13.5%), 은행업(연평균 성장률 13.2%), 전문서비스업(12.9%)의 성장률은 매우 높게 나타난다. 서비스 관련 업종의 수익의 연평균 성장률은 13.2%이고 소프트웨어 투자는 2022년까지 900억 달러 이상이 될 것으로 보인다. 가장 빠른 성장률을 보이는 분야로 인지/AI 소프트웨어 플랫폼은 2022년까지 연평균 36.5%, 비관계형 분석 데이터 스토어는 연평균 30.3%가 될 것으

로 보인다.

그러나 해외의 데이터 산업 시장은 정의와 범위에 따라 규모가 다양하다. IDC&The Lison Council(2018)은 데이터 시장의 범위를 ‘데이터를 가공하여 제품 및 서비스로 재 생산되는 디지털 데이터 시장으로 정의되고 있다.

국가별 데이터 산업의 규모 측면에서 미국이 가장 크며, 연평균 성장률도 유럽과 더불어 가장 빠르게 증가하고 있다. 데이터 시장규모는 데이터 관련 기업의 매출을 기반으로 산출되며, 데이터 기업의 매출액은 데이터 관련 제품의 총가치와 같고, 각 국가의 데이터에 대한 총가치는 자국 내 기업이 생산한 서비스의 매출을 기반으로 하고 있다.

미국의 경우, 2014년 1,210억 달러에서 2017년 1,695억 달러로 연평균 성장률 11.9%로 빠르게 성장하고 있다. 특징으로는 대기업 중심의 성장을 보이고 있으며, 다양하지 못한 분야에서 성장을 보이지 못하고 있는 한계점이 존재한다. 2014년 ICT 투자 대비 데이터 기업의 수익률은 10.0% 정도를 나타내고 있다.

EU의 경우, 2014년 592억 달러에서 2017년 757억 달러로 연평균 성장률 8.5%로 빠르게 성장하고 있다. 중소기업과 벤처기업이 주도하고 있으며, 공공데이터 개방 정책으로 민간데이터보다 공공데이터가 높은 비중을 차지한다. 2014년 ICT 투자 대비 데이터 기업의 수익률은 8.7%를 나타낸다.

일본의 경우, 2014년 259억 달러에서 2017년 323억 달러로 연평균 성장률 7.6%로 성장하고 있고 2014년 ICT 투자 대비 데이터 기업의 수익률은 24.0%를 나타내고 있다.

브라질의 경우, 2014년 62억 달러에서 2017년 73억 달러로 연평균 성장률 5.6%로 성장하고 있으며 2014년 ICT 투자 대비 데이터 기업의 수익률은 4.0% 정도이며, 브라질 데이터산업은 빠르게 증가하며 프랑스, 이탈리아와 경쟁할 수 있는 수준이다.

[표 I-1] 2014~2017년 주요국 데이터산업 규모

단위 : 억 달러, %

	2014	2015	2016	2017	CAGR (2014~2017)
미국	1,210	1,345	1,504	1,695	11.9
EU	592	633	696	757	8.5
일본	259	280	297	323	7.6
브라질	62	61	70	73	5.6

자료: IDC&The Lisbon Council(2018), 재산출

주요국 데이터기업을 보면 미국, EU, 일본, 브라질 순으로 나타나고 있다. 여기서 데이터기업의 정의는 데이터 공급, 데이터 생산·유통·활용하는 기업을 말한다. 미국의 데이터 기업은 2014년 277,821개에서 2017년 302,810개로 증가하였으며 EU는 2014년 243,600개에서 276,450개로 증가하였고, 일본은 2014년 95,919개에서 2017년 104,664개, 브라질은 2014년 34,840개에서 36,387개로 증가하였다. 또한, 이러한 데이터 기업 수의 증가는 고용 등 직·간접적으로 경제에 영향을 미치고 있다.

[표 I-2] 2014~2017년 주요국 데이터산업 기업수

단위 : 억 달러, %

	2014	2015	2016	2017	CAGR (2014~2017)
미국	277,821	283,340	289,556	302,810	2.9
EU	243,600	249,100	261,450	276,450	4.3
일본	95,919	99,001	101,612	104,664	3.0
브라질	34,840	35,456	35,979	36,387	1.5

자료: IDC&The Lisbon Council(2018)

이러한 데이터산업의 경제적 효과는 직접 효과와 간접 효과로 구분할 수 있다. 여기서 데이터산업의 직접 효과는 데이터 생산과 관련된 모든 비즈니스 활동에 의해 형성된 효과로, 판매된 데이터 제품 및 서비스의 수익을 기준으로 계산된다. 미국의 직접 효과는 2014년 1,157억 달러에서 2017년 1,324억 달러, EU의 직접 효과는 2014년 569억 달러에서

2017년 757억 달러, 일본의 직접 효과는 2014년 249억 달러에서 2017년 349억 달러, 브라질의 직접 효과는 2014년 62억 달러에서 2017년 74억 달러로 집계되고 있다.

[표 1-3] 2014~2017년 경제적 효과 : 직접 효과

단위 : 억 달러, %

	2014	2015	2016	2017	CAGR (2014~2017)
미국	1,157	1,272	1,264	1,324	4.6
EU	569	606	623	757	10.0
일본	249	267	319	349	11.9
브라질	62	64	72	74	6.1

자료: IDC&The Lisbon Council(2018), 재산출

데이터산업의 간접 효과는 데이터산업이 타 산업에 미치는 영향을 나타내며, 전방 효과와 후방 효과로 구분될 수 있다. 전방 효과는 데이터산업이 생산물을 중간투입물로 사용하는 타 산업을 발전시키는 효과이다. 전방 효과로는 데이터에 기반을 두고 프로세스를 통해 생산을 효율화하는 생산 및 공급 프로세스 최적화, 타 산업에서 데이터 분석을 토대로 마케팅하여 수익을 제고하는 효과, 데이터 기반 조직 구성으로 조직의 생산성을 높이는 효과 등이 있다. 후방 효과는 데이터산업에 속한 기업이 데이터 기반의 상품과 서비스 활동과정에서 타 산업으로부터 자원을 공급받을 때 자원을 제공한 산업에서 발생 되는 효과이다. 미국의 후방 효과는 2014년 53억 달러에서 2017년 90억 달러로, EU의 후방 효과는 2014년 23억 달러에서 2017년 39억 달러로, 일본의 후방 효과는 2014년 10억 달러에서 2017년 15억 달러로 증가하였다.

미국의 경제적 효과는 2014년 1,210억 달러에서 2017년 1,414억 달러로, EU의 경제적 효과는 2014년 592억 달러에서 2017년 796억 달러로, 일본의 경제적 효과는 2014년 259억 달러에서 2017년 363억 달러로 증가, 브라질의 경제적 효과는 2014년 64억 달러에서 2017년 78억 달러로 증가하였다.

[표 I-4] 2014~2017년 직접 효과와 후방 효과를 합한 경제적 효과

단위 : 억 달러, %

	2014	2015	2016	2017	CAGR (2014~2017)
미국	1,210	1,345	1,349	1,414	5.3
EU	592	633	656	796	10.4
일본	259	280	333	363	11.9
브라질	64	67	75	78	6.8

자료: IDC&The Lisbon Council(2018), 재산출

국내총생산(GDP) 대비 경제적 효과를 살펴보면 다음과 같다. GDP 대비 경제적 효과로 보면, 미국의 경제적 효과는 EU의 약 2배이나 성장 속도 측면에서 EU가 더 빠르게 나타난다. GDP대비 경제적 효과 비중을 보면, 미국은 2016년 0.78%에서 2017년 0.81%로 3.8%p 증가하였으며 EU는 2016년 0.42%에서 2017년 0.52%로 23.8%p 증가하였다.

[표 I-5] 2016~2017년 GDP 대비 경제적 효과 비율

단위 : %, %p

	2016	2017	증가율
미국	0.78	0.81	3.8
EU	0.42	0.52	23.8
일본	0.93	0.95	2.2
브라질	0.16	0.16	0.0

자료: IDC&The Lisbon Council(2018), 재산출

2 » 국내

국내 데이터산업 시장규모는 14조 3,047억 원으로 2016년 대비 4.0% 성장하였으며 2010년부터 2017년까지 연간성장률 7.5%를 나타내고 있다. 2010년~2017년 중, 데이터 솔루션은 연간 13.7% 성장하였으며 데이터 서비스 7.0%, 데이터 구축·컨설팅은 6.6% 성장하였다. 비중 측면에서 데이터 서비스, 데이터 구축·컨설팅, 데이터 솔루션 순으로 나타나고 있다.

[표 1-6] 국내 데이터산업 시장규모

단위 : 억 원, %

	2010	2011	2012	2013	2014	2015	2016	2017(E)	CAGR (2010~2017)
데이터 솔루션	6,725	8,717	10,497	10,769	13,619	14,124	15,720	16,536	13.7
비중	7.8	9.2	9.9	9.6	10.9	10.6	11.4	11.6	
데이터 구축·컨설팅	37,407	43,180	47,715	49,985	53,730	55,280	55,850	58,565	6.6
비중	43.3	45.4	45.2	44.2	43.1	41.4	40.6	40.9	
데이터 서비스	42,242	43,218	47,317	52,258	57,329	64,151	65,977	67,946	7.0
비중	48.9	45.4	44.9	46.2	46.0	48.0	48.0	47.5	
전체	86,374	95,115	105,519	113,032	124,678	133,555	137,547	143,047	7.5
비중	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	

자료: 과학기술정보통신부, 2017년 데이터산업 현황조사

각 분야별 시장규모는 데이터 서비스, 데이터 구축·컨설팅, 데이터 솔루션 크기로 나타나며 중분류별 시장규모는 다음과 같다. 데이터 서비스 시장은 2013년 이후 연평균 7.0% 성장하고 있으며, 시장 크기는 정보 제공, 데이터 분석 제공, 데이터 거래 순서로 나타나며, 데이터 거래와 데이터 분석 제공 분야의 성장률이 큰 것으로 나타난다.

[표 1-7] 데이터 서비스 중분류별 시장규모

단위 : 억 원, %

	2013	2014	2015	2016	2017(E)	CAGR (2013~2017)
정보 제공	48,284	51,985	58,171	59,854	61,575	6.3
비중	92.4	90.7	90.7	90.7	90.6	
데이터 분석 제공	2,266	3,325	3,601	3,529	3,662	12.7
비중	4.3	5.8	5.6	5.4	5.4	
데이터 거래	1,709	2,019	2,379	2,594	2,709	12.2
비중	3.3	3.5	3.7	3.9	4.0	
전체	52,258	57,329	64,151	65,977	67,946	6.8
비중	100.0	100.0	100.0	100.0	100.0	

자료: 과학기술정보통신부, 2017년 데이터산업 현황조사

데이터 구축·컨설팅 시장은 2010년 이후 연평균 6.6% 성장하고 있으며, 데이터 구축의 성장률은 연간 6.5%, 데이터 컨설팅 시장은 연간 11.2% 성장하였다. 데이터 구축 시장은 데이터 이행·처리, DB 설계·구축, 기계 처리형 데이터 구축 분야를 포함한다. 데이터 컨설팅 시장은 데이터 설계 컨설팅, 데이터 품질 컨설팅, 데이터 관리 성능개선 컨설팅, 데이터 거버넌스 컨설팅, 데이터 활용 컨설팅 등의 영역을 포함한다.

[표 1-8] 데이터 구축·컨설팅 중분류별 시장규모

단위 : 억 원, %

	2010	2011	2012	2013	2014	2015	2016	2017(E)	CAGR (2010~2017)
데이터 구축	36,610	42,374	46,865	49,021	52,673	54,142	54,571	56,893	6.5
비중	97.9	98.1	98.6	98.1	98.0	97.9	97.7	97.1	
데이터 컨설팅	797	806	650	964	1,057	1,138	1,279	1,672	11.2
비중	2.1	1.9	1.4	1.9	2.0	2.1	2.3	2.9	
전체	37,407	43,180	47,515	49,985	53,730	55,280	55,850	58,565	6.6
비중	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	

자료: 과학기술정보통신부, 2017년 데이터산업 현황조사

데이터 솔루션 시장은 2010년 이후 연평균 13.7% 성장하고 있으며, 시장 크기는 데이터베이스 관리시스템(DBMS), 데이터 관리 등의 순서로 나타나며, 데이터 관리, 데이터 수집, 데이터 설계, DBMS 분야의 성장률이 큰 것으로 나타난다. 데이터 솔루션 시장은 라이선스, 개발, 유지보수로 구분될 수 있다. 2017년 기준, 업종별 데이터 솔루션 매출 비중은 공공(26.7%), 유통·서비스(24.7%), 제조·건설(18.9%), 통신·미디어(15.7%), 금융(10.4%), 의료(3.6%) 순으로 나타난다.

[표 1-9] 데이터 솔루션 중분류별 시장규모

단위 : 억 원, %

	2010	2011	2012	2013	2014	2015	2016	2017(E)	CAGR (2010~ 2017)
DBMS	3,169	4,272	5,301	5,488	5,502	5,727	6,148	6,192	10.0
비중	47.1	49.0	51.0	50.9	40.4	40.5	39.1	37.4	
데이터 관리	1,237	1,516	2,075	2,489	3,446	3,574	4,022	4,417	19.9
비중	18.4	17.4	20.0	23.1	25.3	25.3	25.6	26.7	
데이터 수집	448	570	517	622	1,076	1,115	1,345	1,382	17.5
비중	6.7	6.5	5.0	5.8	7.9	7.9	8.6	8.4	
데이터 분석	1,056	1,274	1,401	1,169	1,121	1,157	1,249	1,328	3.3
비중	15.7	14.6	13.5	10.8	8.2	8.2	7.9	8.0	
데이터 플랫폼					1,389	1,426	1,611	1,821	9.4
비중					10.2	10.1	10.2	11.0	
데이터 품질	686	949	942	855	883	918	1,120	1,127	7.3
비중	10.2	10.9	9.1	7.9	6.5	6.5	7.1	6.8	
데이터 설계	129	136	151	166	202	207	225	269	11.1
비중	1.9	1.6	1.5	1.5	1.5	1.5	1.4	1.6	
전체	6,725	8,717	10,387	10,789	13,619	14,124	15,720	16,536	13.7
비중	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	

자료: 과학기술정보통신부, 2017년 데이터산업 현황조사

II

데이터 가치평가를 위한 기술금융 및 가치평가 사례 분석

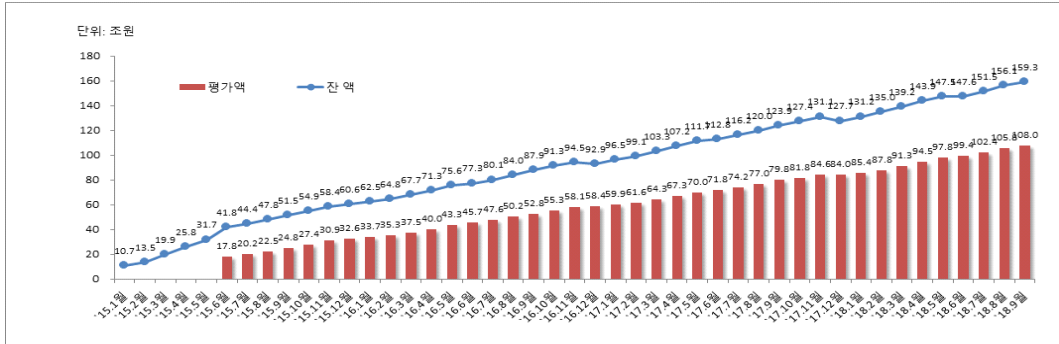
1장에서 전체적인 데이터 시장의 규모를 살펴보았다. 2장에서는 데이터 가치평가를 위하여 기술금융과 기술가치평가 사례를 분석할 필요가 있다. 데이터는 보다 거시적인 차원에서 기술을 통하여 나오는 결과 또는 과정으로 볼 수 있으므로, 기술을 바탕으로 금융이 지원되는 기술금융 현황과 지표를 살펴볼 필요가 있다. 또한, 데이터는 보다 미시적인 차원에서 관련 기술에서 나오는 결과 또는 과정으로 볼 수 있으므로, 거시적인 측면보다 범위를 좁혀 기술가치평가 방법론과 지표를 살펴볼 필요가 있다.

1 » 기술금융

가. 기술금융 현황

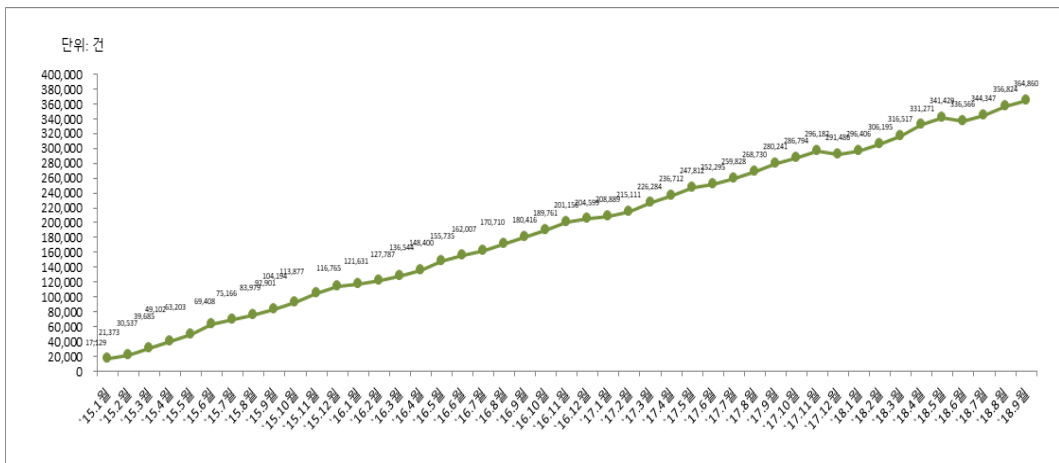
2014년 7월, 금융위원회는 기술신용정보 활용을 위한 금융기관 업무 협약식을 개최하고, 지속적으로 기술금융(기술신용대출) 확장에 주력하였다. 금융위원회, 금융감독원, 전국은행연합회, 18개 은행, 정책금융공사 및 기술보증기금이 참여하여 은행이 기술보증기금의 보증을 받은 기업에 대출을 하거나 정책금융공사의 온렌딩(간접대출)을 이용할 경우 기술신용정보를 활용하기 위하여 협약하였다. 기술신용정보 제공기관(TCB)로 기술보증기금과 한국기업데이터(KED), 나이스평가정보, 이크레더블 4개사로 시작하였으며, 나이스디앤비, SCI평가정보가 포함되었다.

〈그림 II-1〉 기술금융 잔액 및 평가액



자료 : 전국은행연합회

〈그림 II-2〉 기술금융 건수



자료 : 전국은행연합회

2018년 9월 말 현재, 기술금융 건수는 364,860건, 누적 잔액은 159.3조 원으로 나타나고 있어 매우 빠른 속도로 건수와 잔액이 늘고 있다. 기술신용대출 신규평가액은 기술신용대출 공급액 중 신규 및 증액대출을 통해 실제 기업에 자금공급이 증가한 규모이며 기간연장, 단순대환 등은 제외되었다.

은행별 누적 잔액 기준으로, 특수은행에서 기업은행의 잔액이 가장 많고, 시중은행 중 국민은행, 신한은행, 우리은행의 잔액이 많으며, 지방은행 중 부산은행의 잔액이 많이 나타난다.

[표 II-1] 은행별 기술금융 잔액(누적)

구 분		'14.12월	'15.12월	'16.12월	'17.12월	'18.9월
특수 은행	기업	22,165	149,214	268,453	397,863	517,495
	산업	1,753	16,328	30,360	39,641	40,227
	수출입	447	1,315	1,390	1,382	951
	농협	3,840	24,655	44,155	52,072	53,680
	수협	830	1,956	2,769	4,521	5,553
시중 은행	국민	7,464	91,267	129,219	180,367	230,800
	신한	17,360	107,983	145,411	180,771	210,328
	우리	13,123	78,018	110,646	152,326	190,138
	KEB하나	17,709	75,629	100,996	144,835	182,868
	씨티	74	3,610	5,049	6,773	8,479
지방 은행	SC	51	2,715	1,925	1,966	2,289
	경남	887	11,420	23,551	33,766	43,085
	광주	391	3,240	4,358	5,277	7,523
	대구	1,544	16,833	25,473	35,541	45,396
	부산	1,510	20,931	34,015	39,341	52,730
	전북	95	351	354	355	405
	제주	4	375	425	402	581
합 계		89,247	605,838	928,550	1,277,199	1,592,528

자료 : 전국은행연합회

[표 II-2] 은행별 기술금융 건수(누적)

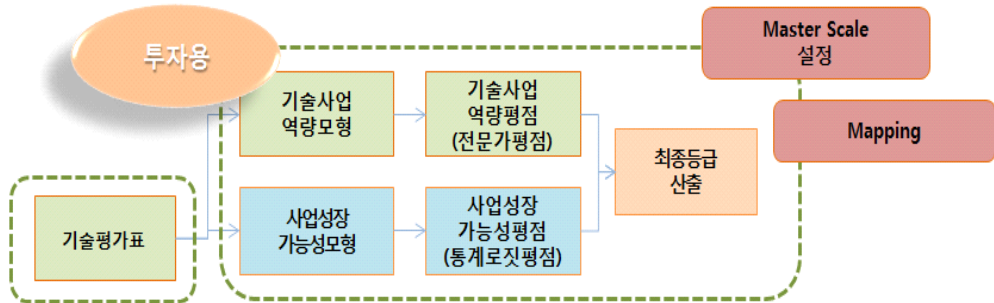
구 분		'14.12월	'15.12월	'16.12월	'17.12월	'18.9월
특수 은행	기업	4,064	25,983	54,417	86,315	113,546
	산업	148	1,191	2,149	2,560	2,383
	수출입	33	79	71	78	46
	농협	692	4,786	9,797	13,677	11,162
	수협	21	164	203	384	625
시중 은행	국민	1,352	20,357	33,887	47,825	61,128
	신한	2,696	22,598	38,293	47,844	56,804
	우리	2,258	15,741	25,268	33,952	41,387
	KEB하나	2,233	12,779	22,050	32,654	41,999
	씨티	46	512	722	998	1,330
지방 은행	SC	19	365	283	185	163
	경남	182	1,917	4,764	7,653	9,771
	광주	77	788	1,176	1,588	2,332
	대구	270	3,037	4,913	7,386	9,244
	부산	244	3,313	6,331	8,111	12,536
	전북	77	173	161	150	174
합 계	제주	1	94	114	126	230
		14,413	113,877	204,599	291,486	364,860

자료 : 전국은행연합회

2016년 4월, 금융위원회는 기술금융 활성화를 위한 투자용 기술금융 평가를 개시하였고, 4월 1일부터 4개 기술신용평가기관에서 일제히 투자용 기술금융 평가를 실시하였다. 유망한 투자대상을 찾는 금융기관이나 투자유치를 원하는 우수 기술기업이 4개 기술신용평가기관에 투자용 기술금융 평가를 요청하면 약 15일 이내에 평가가 완료된다.

투자용 기술금융 평가는 기업의 부실위험 예측에 중심을 둔 기존의 은행권 기술신용평가와 차이가 존재한다. 투자용 기술금융 평가는 기업의 성장 가능성을 예측하는 데 초점을 둔 평가로서, 기술신용평가는 예상 부도율 평가에 초점을 둔 반면 투자용 기술금융 평가는 성장 가능성 평가에 초점을 두고 있다. 투자기관이 실제 투자결정에 이용하는 핵심요소인 기술보호성, 수익성, 기업가정신 등에 대한 투자자 관점의 평가를 강화하였다. 그러나 은행, 벤처캐피탈 등 금융기관이 우수 기술기업을 기술금융 투자 대상으로 발굴하는데 크게 기여할 것으로 기대되었으나, 현실적으로 투자되는 부분은 많지 않다.

〈그림 II-3〉 투자용 기술평가 모형 구조



자료 : 금융위원회(2016.4)

나. 기술금융(기술신용대출) 지표

초기 은행이나 신용평가사들은 기술평가를 할 수 없었으므로 기술보증기금의 KTRS모형의 기술평가표를 기준으로 도입하였으며, 경영주역량, 기술성, 시장성, 사업성 및 수익성을 평가할 수 있도록 33개의 소항목으로 구성되어 있다. 각 소항목에 대하여 평가자가 평가기준에 따라 직접 A~E 등급 중 하나를 입력한다. KTRS-SM(창업 후 5년 이내인 창업 기업용)은 23개 소항목으로 구성되어 있으며 KTRS-BM(창업 후 5년 초과기업으로서 당

기매출액 10억 원 이하인 소규모기업용)는 18개 소항목으로 이루어져 있다.

[표 II-3] 기술보증기금의 기술평가모형 기술평가표

대항목	중항목	세부항목
1. 경영주역량	1.1 기술경험(지식)수준	1.1.1 경영주의 동업종경험수준
		1.1.2 기술경영전략
		1.1.3 경영주의 기술지식수준
		1.1.4 경영주의 기술이해도
	1.2 관리능력	1.2.1 기술인력관리
		1.2.2 기술경영전략
	1.3 경영진인적구성 및 팀웍	1.3.1 경영진의 전문지식수준
		1.3.2 자본참여도
		1.3.3 경영주와의 관계 및 팀웍
2. 기술성	2.1 기술개발추진능력	2.1.1 기술개발전담조직
		2.1.2 기술(디자인)인력
	2.2 기술·연구개발투자현황	2.2.1 기술개발 및 수상(인증)실적
		2.2.2 지식재산권 등 보유현황
		2.2.3 연구개발투자비율
	2.3 기술혁신(선도)성	2.3.1 기술의 차별성/디자인의 우수성
		2.3.2 모방의난이도
		2.3.3 기술의 수명주기 상 위치/트렌드의 부합도
	2.4 기술완성도 및 확장성	2.4.1 기술의 완성도/디자인의 완성도
		2.4.2 기술의 자립도
		2.4.3 기술적 파급효과
3. 시장성	3.1 경쟁상황	3.1.1 시장구조 및 경쟁상황
		3.1.2 목표시장의 규모
		3.1.3 시장의성장성
		3.1.4 법규제 등 제약/장려요인
	3.2 제품의 경쟁력	3.2.1 인지도
		3.2.2 시장점유율
		3.2.3 경쟁제품과의 비교우위성
4. 사업성	4.1 기술의 제품화능력 및 생산능력	4.1.1 생산역량
		4.1.2 재료 및 부품조달 용이성
	4.2 운용능력	4.2.1 자본조달능력
		4.2.2 투자규모의 적정성
	4.3 수익전망	4.3.1 판매계획의 타당성
		4.3.2 판매처의 다양성 및 안정성
		4.3.3 투자대비 회수가능성

최근 민간 TCB사들은 평가지표들을 줄이고 있는 추세이며, 업종별로 다양한 모형 사용하고 있다. 일반모형 평가시, 나이스평가정보 26개, 한국기업데이터 25개, SCI 신용평가 정보 26개, 나이스디앤비 24개 평가 소항목을 사용한다. 각 기업은 산업 분류별로 다양한 모형 세분화된다. 제조업, 전기가공업, 수도/하수 및 폐기물업, 건설업, 도매 및 소매업, 운수 및 창고업, 숙박 및 음식점업, 정보통신업, 금융 및 보험업, 부동산업, 전문과학 및 기술 서비스업, 사업시설관리업, 교육서비스업, 보건업 및 사회복지업, 예술스포츠여가업, 협회 단체 등 기존의 업종 외에 바이오업종, SW/IT업종, 지식서비스업종, 문화컨텐츠 업종 등으로 세분화되고 있다.

다. 투자용 기술평가 모형 평가지표

전체모형을 일반기업, 창업기업, 사업화이전 모형으로 구성된다. 창업기업은 창업 후 5년 이내인 기업이며 일반기업은 창업기업이 아닌 경우이고, 사업화이전은 창업기업 중 제품이 시장 출시 이전 단계인 경우를 나타낸다.

투자용 기술평가모형의 기술평가표는 대출 관점에서 투자지향적으로 지표를 변경하였다. 투자목적에 부합하도록 대출 관점의 KTRS를 미래성장성 관점으로 전환하였으며, 평가 시 중요하게 고려하는 사항을 반영하여 대항목 체계를 구성하였다.

경영역량은 투자에서 강조하는 창업자의 자질이나 역량을 고려하기 위하여 신뢰 역량과 기술역량으로 세분화하였으며, 투자용 평가지표는 경영주의 기술지식 및 경험뿐, 아니라 창의적인 기업가정신 및 신뢰성 평가를 강화한다.

기술성은 기술인프라(기술개발현황, 기술개발능력)와 기술우수성(기술혁신성, 기술완성도 및 확장성) 관점에서 구분하여 각각의 평가항목을 설정하였으며, 투자자가 중시하는 기술의 보호성을 추가하였다.

시장성은 시장여건 및 경쟁력으로 중심으로 평가하며, 사업화단계(경쟁상황)와 개발단계(진입가능성)를 차별화하여 구성되었다. 사업성은 사업추진능력과 향후전망으로 새롭게 구성되었다.

[표 II-4] 투자용 기술평가 모형 세부 평가지표

대항목	중항목	창업기업(사업화이전, 34개)	창업기업(사업화이후, 34개) 일반기업(비창업기업, 33개)
		소항목(평가항목)	
1.경영역량	1.1기업가정신과신뢰	1.1.1 기업가정신	1.1.1 기업가정신
		1.1.2 신뢰성	1.1.2 신뢰성
	1.2최고경영자	1.2.1 동업총경험수준	1.2.1 동업총경험수준
		1.2.2 기술지식수준	1.2.2 기술지식수준
		1.2.3 기술이해도	1.2.3 기술이해도
	1.3경영진	1.3.1 경영진의 전문성	1.3.1 경영진의 전문성
		1.3.2 자본참여도	1.3.2 자본참여도
		1.3.3 경영주와의 팀워크	1.3.3 경영주와의 팀워크
2.기술성	2.1기술개발현황	2.1.1 기술개발 및 수상실적	2.1.1 기술개발 및 수상실적
		2.1.2 지식재산권 등 보유현황	2.1.2 지식재산권 등 보유현황
		2.1.3 연구개발투자	2.1.3 연구개발투자
	2.2기술개발능력	2.2.1 기술개발전담조직	2.2.1 기술개발전담조직
		2.2.2 기술인력	2.2.2 기술인력
		2.2.3 기술인력관리	2.2.3 기술인력관리
	2.3기술혁신성	2.3.1 아이템의 혁신성	2.3.1 기술의 차별성
		2.3.2 기술의 수명주기상 위치	2.3.2 기술의 수명주기상 위치
	2.4기술자립도 및 확장성	2.4.1 기술의 자립도	2.4.1 기술의 자립도
		2.4.2 기술적 파급효과	2.4.2 기술적 파급효과
		2.4.3 기술의 완성도	
	2.5기술보호성	2.5.1 모방의 난이도	2.5.1 모방의 난이도
		2.5.2 기술보호	2.5.2 기술보호
3.시장성	3.1시장현황	3.1.1 시장규모	3.1.1 시장규모
		3.1.2 시장성장성	3.1.2 시장성장성
	3.2경쟁상황	3.2.1 경쟁상황	3.2.1 경쟁상황
		3.2.2 법규제 등 제약/장려요인	3.2.2 법규제 등 제약/장려요인
		3.2.3 시장진입용이성	3.2.3 시장진입용이성(창업기업)
	3.3제품경쟁력		3.3.1 시장점유율
		3.3.1 경쟁제품과 비교우위성	3.3.2 경쟁제품과 비교우위성
			3.3.3 제품인지도
4.사업성	4.1사업능력	4.1.1 생산계획의 타당성	4.1.1 생산능력
		4.1.2 판매계획의 타당성	4.1.2 마케팅능력
		4.1.3 판매처 확보여부	4.1.3 판매처의 다양성 및 안정성
		4.1.4 자본조달능력	4.1.4 자본조달능력
	4.2 향후전망	4.2.1 성장전망	4.2.1 성장전망
		4.2.2 수익전망	4.2.2 수익전망

자료 : 금융위원회(2016.4)

라. 투자용 기술평가모형과 기술보증기금의 KTRS와의 지표 차이

투자용 기술평가모형은 기술보증기금의 KTRS를 기반으로 구축하였으며 이에 따라 먼저 KTRS를 소개하고 이를 활용한 투자용 기술평가모형에 대해 살펴본다.

KTRS와 투자용 기술평가모형 지표 중 경영역량을 비교하면 다음과 같다. 경영역량은 기술보증기금의 경영주 대항목에 1.1 기업가정신과 혁신이라는 중항목을 신설한 것으로 투자자가 중시하는 창업주의 긍정적인 특성을 평가하기 위한 것이다. 기업가정신과 혁신, 최고경영자, 경영진으로 구성되어 있는데, 최고경영자와 경영진은 기존 KTRS 경영주 역량의 중항목으로 구성되었다.

1.1 기술경험(지식)수준 → 1.2 최고경영자

1.3 경영진인적구성 및 팀웍 → 1.3 경영진

기업가정신과 신뢰라는 중항목(1.1)을 새롭게 설정하여 기업가정신과 신뢰성의 평가소항목을 설정하였다. 기업가정신은 KTRS_SM(창업기업)의 특성지표를 적용하였다. 신뢰성은 투자대상기업 경영진의 도덕성 등을 평가하는 항목으로 새롭게 작성한 평가기준이다.

[표 II-5] KTRS와 투자용 기술평가모형 지표 비교 : 경영역량

대항목	중항목 (KTRS)	소항목 (KTRS)	중소항목 (투자용)	소항목 (투자용)	비고
경영주 역량			1.1 기업가정신 과 신뢰성	1.1.1 기업가정신	KTRS_SM적용
				1.1.2 신뢰성	
	1.1 기술경험 수준	1.1.1 경영주의 동업 종경험수준	1.2 최고경영자	1.1.1 동업종경험수준	유지
		1.1.2 기술경영전략			삭제
		1.1.3 경영주의 기술 지식수준		1.1.2 기술지식수준	유지
		1.1.4 경영주의기술 이해도		1.1.3 기술이해도	유지
	1.2 관리능력	1.2.1 기술인력관리	삭제		기술성으로 이동(out)
		1.2.2 기술경영전략			삭제
	1.3 경영진인적 구성 및 팀웍	1.3.1 경영진의 전문 지식수준	1.3 경영진	1.3.1 경영진의전문성	유지
		1.3.2 자본참여도		1.3.2 자본참여도	유지
		1.3.3 경영주와의 관계 및 팀웍		1.3.3 경영주와의 팀워크	유지

KTRS와 투자용 기술평가모형 기술성 지표 비교는 다음과 같다. 기술성은 KTRS의 기술성을 기존 4개 중항목에서 기술보호성(중)을 신설하여 5개 항목으로 재구성되었다. 여기서 기술개발추진역량(중)은 2+1으로, 기존 항목은 기술개발전담조직, 기술(디자인)인력이며, 이동 항목은 기술개발인력관리(경영자 관리능력)에서 이동되었다. 기술의 완성도 및 확장성(중)에서 사업화이후 단계의 경우, 기술의 완전성이 변별력이 없어 적용하지 않는다. 기술의 완성도는 제품화를 만점으로 정의하고 있으며 사업화단계는 모두 만점 처리된다. 사업화이전 단계에서는 매우 중요한 지표이며 지표의 명칭도 기술의 자립도 및 확정성으로 정의되었다.

기술의 혁신성(중)의 경우 사업화이전 단계의 경우, 기술의 차별성을 아이템의 혁신성으로 대체하였으며 이는 사업화이전 단계에 완성되지 않은 기술의 차별성을 평가하는 것보다는 아이템의 혁신성을 평가하는 것이 적절하다는 판단에 근거한다.

기술보호성(중)은 투자자의 요구를 수용하여 신설한 평가요소이다. 기술의 혁신성(KTRS)의 모방난이성을 기술보호성(투자용)으로 이동되었으며 기술 자체 속성으로서 기술보호능력을 의미한다. 기술보호는 해당 기업이 자신의 기술을 보호할 수 있는가를 평가하는 것으로 기술 자체보다는 기술을 보유하고 있는 조직의 기술보호 능력을 평가한다.

[표 II-6] KTRS와 투자용 기술평가모형 지표 비교 : 기술성

대항목	중항목	소항목 (KTRS)	소항목 (투자용)	비고
2.기술성	2.1 기술개발 추진능력	2.1.1 기술개발전담조직	2.1.1 기술개발전담조직	
		2.1.2 기술(디자인)인력	2.1.2 기술(디자인)인력수준	
			2.1.3 기술인력관리	경영주 관리능력(중) 이동(in)
	2.2 기술·연구 개발투자현황	2.2.1 기술개발및수상실적	2.2.1 기술개발및수상실적	
		2.2.2 지식재산권등보유현황	2.2.2 지식재산권등보유현황	
		2.2.3 연구개발투자비율	2.2.3 연구개발투자비율	
	2.3 기술혁신성	2.3.1 기술의차별성/ 디자인의 우수성	2.3.1 기술의차별성 / 디자인의 우수성(일반기업, 창업기업)	사업화이전 /일반, 창업기업 구분

대항목	중항목	소항목 (KTRS)	소항목 (투자용)	비고
			2.3.1 아이템의 혁신성 (사업화이전)	
		2.3.2 모방의난이도		기술보호능력(중) 이동(out)
		2.3.3 기술의수명주기상위치 /트 랜드의부합도	2.3.3 기술의수명주기상위치 /트 랜드의부합도	
	2.4 기술자립도 및 확장성	2.4.1 기술의완성도 /디자인의완 성도	2.4.3 기술의완성도 /디자인의완 성도(사업화이전)	일반, 창업기업 에서는 삭제
		2.4.2 기술의자립도	2.4.1 기술의자립도	
		2.4.3 기술적파급효과	2.4.2 기술적파급효과	
	2.5 기술보호능력		2.5.1 모방의난이도	기술혁신성(중) 이동(in)
			2.5.2 기술보호	신설

KTRS와 투자용 기술평가모형 시장성 지표를 비교하면 다음과 같다.

시장여건 및 경쟁력은 KTRS의 시장성의 3개 중 항목을 그대로 활용하여 구성하였지만 사업화이전 단계에 적용할 모형에서는 일부 수정되었다. 경쟁요인에서 ‘시장진입 용이성’을 추가되었는데 사업화이전 단계에서는 이후 시장진입 가능성을 판단하는 것 중요하기 때문이다. 경쟁력에서는 ‘인지도’를 제외되었는데 사업화이전 단계에 인지도를 파악하는 것이 현실적으로 어렵고 무의미하기 때문이다.

[표 II-7] KTRS와 투자용 기술평가모형 지표 비교 : 시장성

대항목	중항목	소항목 (KTRS)	소항목 (투자용)	비고
3.시장성	3.1 시장현황	3.1.1 목표시장의 규모	3.1.1 목표시장의 규모	
		3.1.2 시장의 성장성	3.1.2 시장의 성장성	
	3.2 경쟁요인	3.2.1 경쟁상황	3.2.1 경쟁상황	
		3.2.2 법·규제 등 제약/ 장려요인	3.2.2 법·규제 등 제약/장려 요인	
			3.2.3 시장진입 용이성 (창업기업, 사업화이전)	KTRS_SM의 시장진입용이성 지표를 추가
	3.3 경쟁력	3.3.1 인지도		사업화이전에서는 적용하지 않음
		3.3.2 시장점유율		
		3.3.3 경쟁제품과의 비교 우위성		

KTRS와 투자용 기술평가모형 사업성 지표를 비교하면 다음과 같다. 사업추진역량은 KTRS의 사업성을 시장개척과 투자역량으로 재구성하였으며 나아가 향후전망을 추가하여 평가자의 전망을 정리하는 항목을 추가하였다. 사업추진능력은 생산능력, 마케팅능력, 판매처의 다양성과 안정성, 자금조달능력으로 구성하였으며 이중 생산능력, 마케팅능력, 판매처의 다양성과 안정성은 사업화이전에 대해서는 R&D평가지표를 적용하였다.

4.1.1 생산능력→ 4.1.1 생산계획의 타당성

4.1.1 마케팅능력 → 4.1.1 마케팅능력(판매계획의 타당성)

4.1.2 판매처의 다양성과 안정성 → 4.1.2 마케팅능력(판매처 확보여부)

투자규모의 적정성, 투자대비회수 가능성은 평가항목의 속성이 기업수준에 적합하지 않아 이를 제외하였으며 이를 보완하기 위해서 향후 전망을 적용한다. 향후전망은 성장성과 수익성으로 구성하였으며 평가지침은 기술력평가모형(산업기술진흥원)을 참조하여 작성한다.

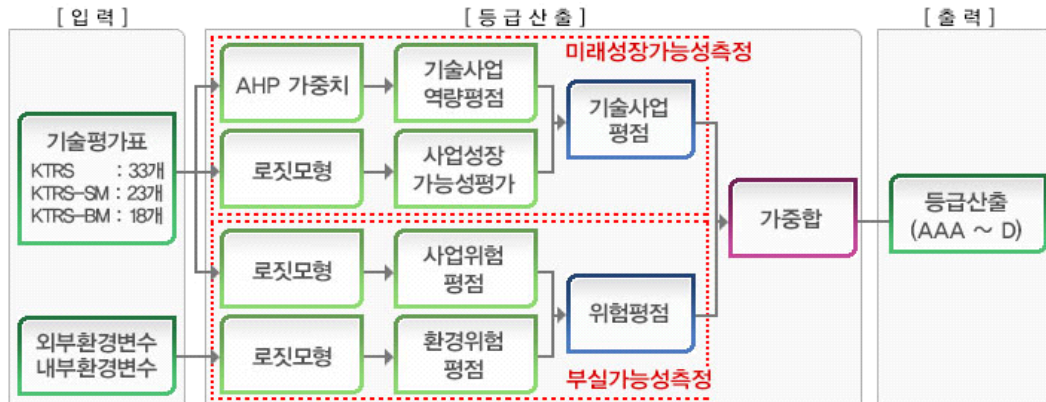
[표 II-8] KTRS와 투자용 기술평가모형 지표 비교 : 사업성

대항목	중항목	소항목 (KTRS)		소항목 (투자용)	비고
4.사업성	4.1 제품화역 량 및 생산 능력	4.1.1 생산역량	4.1 사업추진역량	4.1.1 생산능력 (일반, 창업기업)	사업화이전 단계 R&D 평가지표를 활용
		4.1.2 재료및부품조달의 용이성		4.1.1 생산계획의 타당성 (사업화이전)	
					삭제
	4.2 운영능력	4.2.1 자금조달능력			
				4.1.1 마케팅능력 (일반, 창업기업)	사업화이전 단계 R&D 평가지표를 활용
				4.1.1 판매계획의 타당성 (사업화이전)	
				4.1.2 판매처의 다양성과 안정성(일반, 창업기업)	
				4.1.2 판매처 확보여부 (사업화이전)	
		4.2.2 투자규모의 적정성			삭제
	4.2 수익전망	4.2.1 마케팅능력			이동(4.1.2)
		4.2.2 판매처의 다양성과 안정성			이동(4.1.3)
		4.2.3 투자대비회수가능성			삭제
			4.2 향후전망	4.2.1 성장성	신설
				4.2.2 수익성	신설

마. 기술금융(기술신용대출) 모형

KTRS계열모형의 등급산출은 기업·기술의 미래성장가능성과 부실가능성을 측정하여 기술사업화가능성을 종합적으로 판단하여 이루어진다. 등급의 결정은 미래성장가능성을 측정하는 기술사업 평점과 부실가능성을 측정하는 위험평점의 결합으로 이루어진다. 기술사업 평점은 기술평가표에 따른 평가에 기초하여 기술사업 역량에 대한 평가와 사업성장 가능성에 대한 평가의 결합을 통해서 산출되며, 위험평점은 기술평가표에 따른 평가에 기초하여 사업위험평가와 환경변수에 따른 평가에 기초한 환경위험평가와의 결합을 통해서 산출된다.

〈그림 II-4〉 기술평가등급모형의 등급산출과정



자료: 기술보증기금(2015), 연차보고서

KTRS계열의 등급은 다음과 같은 10등급 체계이다.

[표 II-9] 기술평가모형의 등급정의

	등급정의
AAA	기술사업화역량이 최고수준으로, 합리적으로 예측가능한 장래의 환경변화에도 영향을 받지 않을 만큼 안정적인
AA	기술사업화역량이 매우 우수하며, 사업부실화의 가능성이 매우 낮지만 급격한 환경변화에 다소 영향 받을 수 있음
A	기술사업화역량이 우수하며, 사업부실화의 가능성이 낮지만 급격한 환경변화에 영향 받을 수 있음
BBB	기술사업화역량이 우수하며, 사업부실화의 가능성이 낮지만 환경변화에 따라 영향 받을 수 있음
BB	기술사업화역량이 양호하며, 사업부실화 가능성이 낮지만 환경변화에 따라 상당히 영향 받을 수 있음
B	기술사업화역량이 보통 수준으로 사업부실화의 가능성을 무시할 수 없음
CCC	기술사업화역량이 다소 미흡하여 사업부실화의 가능성이 일부 존재하지만 전반적인 사업추진 타당 시 됨
CC	기술사업화역량이 미흡하여 사업부실화의 가능성이 존재함
C	기술사업화역량이 미흡하여 사업부실화의 가능성이 다소 높음
D	기술사업화역량이 미흡하여 사업부실화의 가능성이 높음

바. 투자용 기술평가 모형

등급정의는 다음과 같다. 투자용 기술평가모형은 기업 보유기술의 사업화를 통한 기업의 미래성장가능성을 평가하는 것으로 향후 시장의 여건, 핵심기술의 우수성, 이를 사업화하여 기업성장으로 이끌어 나갈 기업역량을 평가하여 그 결과를 평가등급으로 제시한다. 투자용 기술평가모형을 10개 등급으로 정의하였으며 그 내용은 다음과 같다.

[표 II-10] 투자용 기술평가모형의 등급정의

평가등급	등급의 정의
1	매우 우수한 기술력과 성장잠재력이 우수한 시장을 바탕으로 미래 성장가능성이 최상위 수준임
2	매우 우수한 기술력과 성장잠재력이 우수한 시장을 바탕으로 미래 성장가능성이 매우 우수한 수준임
3	우수한 기술력과 성장잠재력이 양호한 시장을 바탕으로 미래 성장가능성이 우수한 수준임
4	우수한 기술력과 성장잠재력이 양호한 시장을 바탕으로 미래 성장가능성이 상당히 존재함
5	양호한 기술력과 성장잠재력이 있는 시장을 바탕으로 미래 성장가능성이 존재함
6	양호한 기술력과 성장잠재력이 있는 시장을 바탕으로 미래 성장가능성이 일부 존재함
7	기술력이 다소 부족하고 시장의 성장잠재력도 다소 부족하여 미래 성장가능성이 다소 낮음
8	기술력이 다소 부족하고 시장의 성장잠재력도 다소 부족하여 미래 성장가능성이 낮음
9	기술력이 낮고 시장의 성장잠재력도 낮아 미래 성장가능성이 낮음
10	기술력이 낮고 시장의 성장잠재력도 낮아 미래 성장가능성이 매우 낮음

등급산출 모형은 다음과 같다. 전문가모형과 통계모형을 결합하여 등급을 결정하는 방식으로 모형을 개발한다. 전문가모형은 기술평가표의 모든 평가항목에 대해서 전문가가 생각하는 고성장기업 예측에 대한 기여도를 파악하여 이를 적용하는 방식으로 기술평가표의 모든 항목에 배점이 부여되는 모형으로 미래 지향적이다. 고성장기업은 평가일 이후 3년간 연평균 매출액증가율이 20%이상인 기업을 고성장 기업(“가젤기업”)으로 정의하였다. 통계모형은 기술평가표의 모든 평가항목에 대해서 통계적 분석을 실시하여 고장기업 예측에 유의한 평가항목만을 선별하여 예측에 대한 기여도를 통계적으로 파악하여 적용하는 방식으로 기술평가표의 특정 항목에 대해서만 배점이 부여되는 모형이다.

사. TCB 지표 사례

최근 TCB사들은 소항목을 줄여서 평가하는 경향이 있다. 대부분의 TCB사는 2개 대항목, 8개 중항목은 공통적으로 사용하고 있으며, 25~33개의 소항목을 개별적으로 사용하고 있고, 이를 종합하면 전체 38개 소항목을 사용하고 있다. 전체 소항목은 38개 중, 기보 33개, KED 26개, NICE 25개, 이크레더블 25개, 평가 소항목을 활용하여 기업의 기술력을 평가한다.

[표 II-11] 투자용 기술평가모형의 등급정의

대항목	중항목	소항목	기보	KED	NICE	이크레더블
			33개	26개	25개	25개
기술사업역량	경영주 역량	1. 경영주 동업종경험	O	O	O	O
		2. 경영주 기술지식	O	O	O	O
		3. 경영주 기술이해도	O	X	X	X
		4. 경영주 기술경영능력	O	O	O	O
		5. 경영주 기술경영전략	O	X	X	X
		6. 경영주 기업가정신	O	X	X	X
	경영진 역량	7. 기술인력관리체계	O	X	X	X
		8. 경영진 전문성	O	O	O	O
		9. CTO 기술지식	X	X	O	X
		10. 경영진 자본참여도	O	O	O	O
		11. 경영진 팀워크	O	X	X	X
	기술개발 역량	12. 기술개발조직	O	O	O	O
		13. 기술인력 전문성	O	O	O	O
		14. 기술개발 및 수상실적	O	O	O	O
		15. 지식재산권 현황	O	O	O	O
		16. 연구개발투자비율	O	O	O	O
	제품화 역량	17. 생산역량	O	O	O	O
		18. 품질관리역량	X	X	O	X
		19. 투자규모 적정성	O	X	X	O
		20. 자금조달능력	O	O	O	O

대항목	중항목	소항목	기보	KED	NICE	이크레더블
			33개	26개	25개	25개
	수익 전망	21. 마케팅 역량	O	O	X	O
		22. 판매처 다양성 및 안정성	O	O	O	X
		23. 투자대비 회수가능성	O	O	X	O
		24. 부가가치창출능력	X	X	O	X
기술경쟁력	기술우위성	25. 기술의 차별성	O	O	O	O
		26. 건설공사능력	X	O	X	X
		27. 모방난이도	O	O	O	O
		28. 기술수명주기	O	O	O	O
		29. 기술완성도	O	O	O	O
	시장현황	30. 기술자립/활용성	O	O	O	O
		31. 시장규모/성장성	O	O	O	O
		32. 제약/장려 및 산업여건	O	X	X	X
	시장경쟁력	33. 시장경쟁현황	O	O	O	O
		34. 산업현황	X	O	X	O
		35. 인지도	O	O	X	X
		36. 시장진입 용이성	O	X	X	X
		37. 시장점유율	O	X	O	O
		38. 경쟁제품과의 비교우위성	O	O	O	O

아. 은행 지표 사례

현재 8개 은행이 기술금융을 적용하고 있으며 개별 지표는 다를 수 있으며, 최근 지표를 감소시키는 방향으로 움직이고 있다.

[표 II-12] A은행의 일반모형 지표

대항목	중항목	소항목
1. 기술사업역량	1.1 경영주	1.1.1 경험수준
		1.1.2 기술경영전략
	1.2 관리능력	1.2.1 경영진 전문지식수준
		1.2.2 기술인력관리
	1.3 기술개발능력	1.3.1 기술개발전담조직
		1.3.2 기술인력 전문성
		1.3.3 기술개발 및 인증실적
		1.3.4 지식재산권 보유현황
		1.3.5 연구개발투자
	1.4 제품화역량	1.4.1 생산역량
		1.4.2 자금조달능력
	1.5 수익전망	1.5.1 마케팅역량
		1.5.2 판매처 다양성 및 안정성
		1.5.3 향후수익
2. 기술경쟁력	2.1 기술혁신성	2.1.1 기술차별성
		2.1.2 기술확장성
		2.1.3 기술완성도
		2.1.4 기술자립도
		2.1.5 기술수명주기
		2.1.6 모방난이도
	2.2 제품우위성	2.2.1 기술제품 인지도
		2.2.2 기술제품 경쟁우위성
	2.3 시장현황	2.3.1 시장규모
		2.3.2 시장성장성
		2.3.3 시장구조

2 《 기술가치평가

가. 기술평가의 종류

「기술의 이전 및 사업화 촉진에 관한 법률」제2조 정의 조항에 따르면, ‘기술평가’란 사업화를 통하여 발생할 수 있는 기술의 경제적 가치를 가액(價額)·등급 또는 점수 등으로 표현하는 것이다. ‘기술가치평가’란 기술거래기관 및 평가전문기관 지정요령에 따르면, 기술의 실효성을 제고하고 원활한 기술이전 및 사업화와 기술의 담보제공 등 기술 활용을 촉진하기 위하여 개발된 기술의 가치를 금전적으로 환산하는 것을 말한다.

기술가치평가의 대상 기술은 역시 여러 가지로 정의할 수 있으나 경제적 가치를 실현하는 상품 또는 서비스를 고객에게 제공하기 위해 공정 또는 유통 과정에서 투입되는 과학 또는 공학 기반의 지식으로서, 기계, 문서, 설계도와 같은 유형물에 체화되어 있거나, 특허권이나 실용신안권과 같은 지식재산으로서의 배타적 권리를 인정받고 있거나, 숙련 인력의 기억이나 감각과 같은 무형의 자산으로 체화되어 있는 것이다. 비슷한 속성을 지니지만 공학 또는 과학기반의 지식이 아닌 것인 고객명부, 브랜드 자산과 같은 무형자산은 기술로 분류하지 않는 것이 일반적이다.

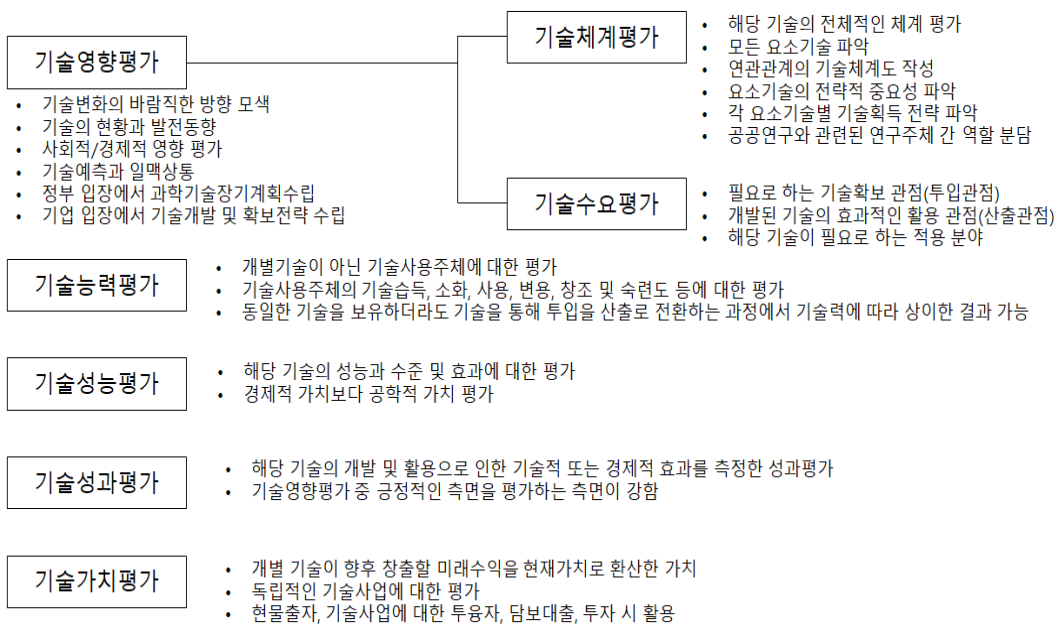
기술가치평가는 평가대상이 경제적 가치를 창출하는 기술자산(technology assets)으로 한정되어야 하기 때문에 기업가치평가, 기술력 평가, 기술영향평가를 포함한 광의의 기술평가와 구분될 수 있다. 기업가치평가(corporate valuation)는 기술자산을 포함하여, 다양한 유형 및 무형의 자산이 종합하여 창출하는 기업의 경제적 가치를 평가한다. 기술력 평가(technology evaluation)는 경제적 가치 창출 이전 단계에 작용하는 기술 개발능력 또는 기술의 우위성을 등급 또는 점수로 평가한다.

기술평가의 한 유형으로, 기술을 활용하는 주체의 인력, 조직, 지원서비스 등을 종합적으로 평가함으로써 그 주체의 기술개발, 흡수, 및 혁신능력을 평가하는 것을 말하며 등급, 점수 등 다양한 형태로 표시될 수 있다. 광의의 기술평가(technology assessment)는 기술영향평가(technology impact assessment)를 포함하며, 그 기술의 적용에서 예상되는

경제적·비경제적 파급효과를 화폐적 또는 비화폐적 측정 수단으로 표현된다.

기술에 대한 광의의 평가 중에서는, 기술 현상의 다양한 속성 중 어디에 초점을 두느냐에 따라 기술영향평가를 기술체계평가(technology system assessment)와 기술수요평가(technology needs assessment)로 구분되고, 기타 기술능력평가(technology capability assessment), 기술성능평가(technology performance assessment) 등이 존재할 수 있으며, 이는 기술가치평가와 구분되는 영역으로 존재한다.

〈그림 II-5〉 기술가치평가와 여타 기술평가 분야와의 관계



자료 : 조경선·임재용(2015), 재구성

산업통상자원부는 「기술의 이전 및 사업화 촉진에 관한 법률」(기술이전법) 제10조 제1항, 제12조의2의 제1항, 제35조 및 동법 시행령 제16조, 제19조의2, 제32조에 따른 기술거래기관, 사업화전문회사 및 기술평가기관 지정요령을 고시하고 있다. 기술의 이전 및 사업화 촉진에 관한 법률(기술이전법)에 따라 전문 인력 확보 등의 일정 기준을 충족하고 정부로부터 별도의 지정을 받아야만 기술가치평가를 수행한다. 기술가치평가기관은 기술이전법 이외에도 기술신용보증법, 발명진흥법 등 다른 법률에서도 지정하고 있는데, 각각

의 법률에서 기술가치평가를 하는 목적에 따라 기술가치평가 기관을 운영하는 형태에는 다소 차이가 있다.

현재는 주무 부처별로 기술가치평가 기관을 지정·운영하고 있으며, 2개 이상의 법률로 중복되어 지정된 예도 있으며 각 법률에 따른 기술가치평가 현황은 다음과 같다.

[표 II-13] 기술가치평가 기관

관련 법률	주무 부처	기술가치평가 기관	기관 수
기술의 이전 및 사업화 촉진에 관한 법률	산업통상 자원부	한국산업기술진흥원, 국방기술품질원, 기술보증기금, 전자부품연구원, 중소기업진흥공단, 한국과학기술정보연구원, 한국발명진흥회, 한국산업은행, 한국보건산업진흥원(보건부), 농업기술실용화재단(농림축산식품부), 한국전자통신연구원, 한국기계연구원 한국생명공학연구원, 한국과학기술연구원, 한국생산기술연구원(미래창조과학부)	15
기술신용보증기금법	중소벤처기업부	기술보증기금	1
벤처기업 육성에 관한 특별조치법	중소벤처기업부	한국산업기술진흥원, 기술신용보증기금, 한국산업기술평가관리원, 한국환경공단, 기술표준원, 한국과학기술연구원, 한국과학기술정보연구원, 정보통신산업진흥원	8
발명진흥법	특허청	한국건설생활환경시험(연), 한국기계전기전자시험(연), 한국산업기술시험원, 한국과학융합시험(연), 농업기술실용화, 한국산업기술진흥원, 기술보증기금, 한국산업은행, 한국과학기술정보연구원, 한국발명진흥회	10
산업기술혁신 촉진법	교육부	「기술의 이전 및 사업화 촉진에 관한 법」에 따른 지정된 기술평가 기관	15

위의 기술평가와 관련된 평가와 기술평가기관을 정리하면 기술평가는 크게 기술가치평가와 기술력평가로 구분이 가능하다. 기술평가 중, 기술력평가는 기술발전이 사회 전반에 미칠 잠재적인 영향을 분석한다. 기술발전이 사회, 경제, 정치제도에 미치는 잠재적인 영향을 체계적으로 분석 및 평가한다. 따라서 기술평가는 정부의 기술프로그램이 가져올 긍정적 부정적 그리고 불확실한 효과에 대하여 정책입안자에게 균형적인 평가를 제공함으로써 기술의 개발을 증진 혹은 발전시킬 것인가에 대한 정책 결정의 지표를 제공하는 것을 목적으로 한다.

기술가치평가는 기술의 경제적 거래가치를 나타내며 보다 거시적인 시각을 가지는 기술력평가와는 달리 개별기술의 가치와 상업성을 판단하는 평가이다.

[표 II-14] 기술평가의 종류

유형	내용
기술력평가	기술평가의 한 유형으로, 기술을 활용하는 주체의 인력, 조직, 지원서비스 등을 종합적으로 평가함으로써 그 주체의 기술개발, 흡수, 및 혁신 능력을 평가하는 것을 말하며 등급, 점수 등 다양한 형태로 표시될 수 있음
기술가치평가	기술평가의 한 유형으로, 사업화하려는 기술이나 사업화된 기술이 그 사업을 통하여 창출하는 경제적 가치를 기술시장에서 일반적으로 인정된 가치평가 원칙과 방법론에 입각하여 평가하는 것을 말함

자료 : 산업통상자원부(2014.6)

나. 기술가치평가 현황

기술가치평가 시장은 공공부문과 민간부문으로 구분할 수 있다. 공공부문의 기술가치평가 수행은 2016년 말 현재 636건으로 나타나며 민간부문의 기술가치평가 수행은 공공부문보다 광범위하게 일어나고 있다.

[표 II-15] 평가기관별 기술가치평가 수행 현황

	~'08년	'09년	'10년	'11년	'12년	'13년	'14년	'15년	'16년
기술보증기금	1,740	384	427	471	368	NA	126	179	110
국방기술품질원	88	37	38	25	153	NA	NA	NA	NA
한국발명진흥회	64	22	37	51	92	NA	197	182	192
농업기술실용화재단	-	-	27	32	47	NA	NA	NA	NA
전자부품연구원	12	30	30	10	17	NA	10	12	10
한국과학기술정보연구원	26	13	10	10	15	NA	NA	NA	NA
한국산업은행	8	4	4	7	8	NA	7	4	6
계	1,938	490	573	606	700	456	493	521	636

자료 : 손수정 외(2014), 지식재산평가센터(2017)

〈그림 II-6〉 공공·민간별 기술가치평가시장 현황(평가건수)



자료 : 한국산업기술진흥원(2011)

다. 기술력평가지표

기술보증기금의 기술력평가지표는 기술사용 주체가 지니고 있는 기술경영 및 연구개발 능력(인적 및 인프라), 보유한 핵심기술의 기술성, 시장성, 사업성 등을 종합적으로 평가하여 평가결과를 등급으로 표시한다. 개별기술의 사업타당성평가 및 가치평가는 개별기술 자체가 지닌 기술성, 권리성, 시장성 및 사업성을 평가한다.

기술성 분석은 평가대상 기술의 정의(발명의 명칭 등), 개요, 기술개발 동향 및 경쟁(신규, 대체) 기술 현황, 경쟁기술 대비 기술수준(우위성, 첨단성, 독창성 등), 기술 활용성 및 파급효과 등에 대해 분석하고 평가한다.

권리성 분석은 기술사업화를 하는 데 있어서 평가대상기술의 특허가 법적 안정성, 권리범위 등의 조사 분석을 통해 시장의 독점적 지위 확보 여부, 기술의 경제적 수명, 경쟁으로부터 사업의 보호 정도를 파악하는 것이다.

시장성 분석은 평가대상기술의 적용제품이 속한 시장에 대하여 시장환경 분석, 시장경쟁 분석 결과에 근거하여 적용제품의 시장경쟁력을 평가하는 것으로서, 평가대상기술이 속해 있는 산업의 특성 및 환경을 살펴보고, 시장구조, 제품 현황, 시장진입장벽, 관련 정책 및 분석을 검토하며, 국내외 시장동향과 업체동향을 조사한 결과에 근거하여 시장경쟁력을 분석한다.

사업성 분석은 기술을 이용하여 사업화를 추진하는 주체의 사업화 기반 역량, 생산 및 영업능력 등 경영요인을 고려하여 평가대상기술을 활용하여 생산되는 제품의 가격 및 품질 경쟁력, 매출전망, 현금흐름 등 사업전망 전반에 관하여 평가하는 절차이다.

[표 II-16] 기술보증기금 기술력평가지표

대항목	평가항목	세부항목
기술성	기술개요	기술개요 및 특징
		기술구성 및 내용
		기술적용 현황
	기술환경 분석	국내외 기술동향 및 업체 현황
		기술의 Life-Cycle
		기술발전방향
	기술 유용성 및 경쟁력 분석	기술 유용성 및 경쟁력 분석
		기술 독창성 및 첨단성 평가
		기존기술 대비 비교우위성
		기술 활용성 및 확장성
		산업적 파급효과
권리성	권리안정성	권리의 안정적 유지가능성 여부
		추가 권리 확보 가능성
	권리범위	관리범위의 광협
		권리의 폭과 깊이
		권리의 차별성
		회피설계 가능성
	사업 연관성	사업과의 관련성
	선행특허	유사특허기술 존재 정도
	권리침해	기존 지식재산권 침해 가능성
시장성	시장개요	기술적용제품 범위
		시장의 정의 및 특성
	시장환경 분석	국내외 시장규모 및 동향
		중장기 성장률
		시장수요 전망
		시장진입가능성
	시장경쟁 분석	적용시장의 경쟁 구조 및 지배유형
		경쟁업체 현황 및 시장점유율
		경쟁 제품과의 비교 분석
		시장지위 확보 가능성

대항목	평가항목	세부항목
사업성	사업화 기반 역량	평가대상 기술사업화 주체가 보유 또는 계획 중인 인적, 물적 사업 기반역량을 종합적으로 파악
	제품경쟁력	평가대상기술 제품의 가격, 품질 등 시장에서 비교 우위를 확보해 주는 경쟁력 속성을 파악
	매출추정 및 수익분석	평가대상기술 제품이 현재 해당시장에서 확보할 수 있는 점유율을 파악하고 미래 당해 시장에서의 매출 확보 가능성 예측

한국산업기술진흥원(KIAT) 기술력 평가지표를 살펴보면 다음과 같다. 한국산업기술진흥원의 기술력평가모형은 특정 기업 내에서 특정기술을 사용 함에 따라 기업의 현재 또는 미래의 현금창출능력을 평가하는 모형이다. 기술력등급모형으로 기술사용 주체가 지니고 있는 기술경영 및 연구개발능력(인적 및 인프라), 보유한 핵심기술의 기술성, 시장성, 사업성 등을 종합적으로 평가하여 평가결과를 등급으로 표시된다. 개별기술의 사업타당성평가 및 가치평가는 개별기술 자체가 지니는 효과를 기술성, 권리성, 시장성 및 사업성을 평가하는 것으로서 기술의 사용주체에 대해 평가하는 것이 아니라는 점에서 기술력등급평가와 구분된다. 기술가치평가는 정량적 평가(평가결과를 화폐금액으로 표시)이고, 기술력등급평가는 정성적 평가를 통해 최종적으로 등급으로 제시한다는 점에서 구분된다.

[표 II-17] 한국산업기술진흥원의 기술력평가등급의 정의

평가등급 (평가점수)	등급의 정의
T1(85≤X≤100)	매우 우수한 기술력을 보유하고 있으며 사업부실화 가능성이 매우 희박함
T2(80≤X<85)	우수한 기술력을 보유하고 있으며 장래의 환경변화에 크게 영향을 받지 않는 수준으로 사업부실화 가능성이 상당히 희박함
T3(77≤X<80)	우수한 기술력을 보유하고 있으며 장래의 환경변화에 다소 영향을 받을 수는 있으나 사업부실화 가능성이 희박함
T4(74≤X<77)	양호한 기술력을 보유하고 있으며 사업부실화 가능성은 낮으나, 장래의 환경변화에 다소 영향을 받을 수 있음
T5(70≤X<74)	기술력이 양호한 수준으로 장래의 환경변화에 따른 사업성 변동가능성이 있으나, 전반적인 사업부실화 가능성이 적음
T6(67≤X<70)	기술력이 보통 수준으로 장래의 환경변화에 따른 사업성 변동가능성이 있으나, 전반적으로 사업추진이 타당시 됨
T7(64≤X<67)	기술력이 보통 수준으로 사업화를 위한 필요사항이 다소 부족하나 전반적인 사업추진이 무난시 됨

평가등급 (평가점수)	등급의 정의
T8(60≤X<64)	기술력이 보통 수준으로 사업화를 위한 필요사항이 부족하나 이에 대한 보완이 이루어질 경우 사업추진이 타당시 됨
T9(55≤X<60)	기술력 및 사업화를 위한 필요사항이 부족한 수준으로 사업성공의 가능성이 미흡한 편임
T10(X<55)	기술력이 취약한 수준으로 사업추진 타당성이 매우 부족함

평가항목으로 대상기업의 기술력을 바탕으로 한 미래 현금창출능력을 평가하기 위해 기술경영 및 연구개발능력, 기술성, 시장성, 사업성 4개 평가항목으로 대항목을 정의하고 있다.

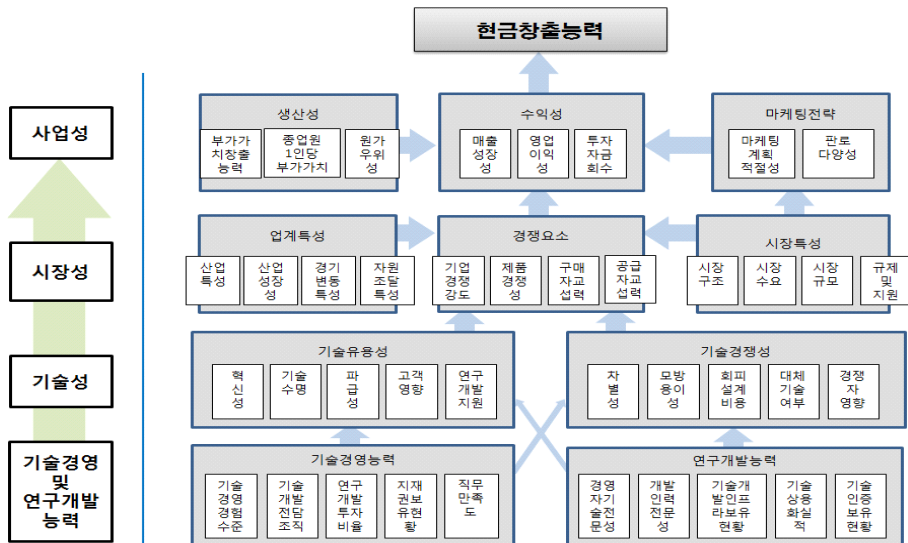
[표 II-18] 한국산업기술진흥원의 기술력평가표

대항목	중항목 (변경후 지표)	세부항목 (변경후 지표)
1. 기술경영 및 연구 개발능력(10)	1-1 기술경영능력 (CEO 평가)	① 기술경영 경험수준
		② 기술개발전담부서
		③ 연구개발투자비율
		④ 지식재산권 보유현황
		⑤ 직무만족도
	1-2 연구개발능력 (인력 및 인프라)	① 기술전문성(CTO)
		② 개발인력의 전문성
		③ 기술개발·인프라보유현황
		④ 기술상용화 실적
		⑤ 기술인증 보유현황
2. 기술성 (10)	2-1 기술 유용성	① 혁신성
		② 기술수명
		③ 파급성
		④ 고객에게 미치는 영향
		⑤ 연구개발지원(기술전망)
	2-2 기술 경쟁성	① 차별성
		② 모방용이성(기술의 난이도)
		③ 회피비용(회피설계비용)
		④ 대체기술 존재 여부
		⑤ 경쟁자에게 미치는 영향

대항목	중항목 (변경후 지표)	세부항목 (변경후 지표)
3. 시장성 (12)	3-1. 업계특성	① 산업특성
		② 산업성장성
		③ 경기변동의 특성
		④ 자원조달의 특성
	3-2. 시장특성	① 시장구조
		② 시장수요
		③ 시장규모
		④ 규제 및 지원
	3-3. 경쟁요소 (기술제품 및 기업)	① 기업간 경쟁강도
		② 제품의 경쟁성
		③ 구매자의 교섭력
		④ 공급자의 교섭력
4. 사업성 (8)	4-1 생산성	① 부가가치 창출능력
		② 종업원 1인당 부가가치
		③ 원가우위성
	4-2 수익성	① 매출성장성
		② 영업이익성
		③ 투자자금회수기간
	4-3 마케팅전략	① 마케팅 계획의 적절성
		② 판로의 다양성 및 구축정도

4개 대항목 간의 관계를 파악하고, 각 대항목 내의 평가항목들이 어떤 의미를 가졌으며 이 평가항목들을 측정함으로써 어떤 효과를 얻을 수 있는지는 알 수 있다.

〈그림 II-7〉 기술력등급평가의 흐름도



특허청 IP담보가치 평가지표는 다음과 같다. 기술성 분석은 기술의 적용제품에 대한 기술개요(기술의 정의 및 개요), 기술개발 동향 및 경쟁 기술 현황, 경쟁기술 대비 기술 수준(차별성, 혁신성 등), 기술 활용성 및 파급효과 등에 대해 분석 및 평가한다. 기술환경 분석과 기술 경쟁성 분석으로 구분된다. 기술환경 분석은 대상기술을 중심으로 경쟁기술, 대체기술, 후방기술(부품, 소재 등) 및 전방기술(조립, 완성, SI 등) 등 주변기술 환경을 조사하는 것이다. 기술경쟁성 분석은 기술환경 분석결과에 근거하여 파악된 각 기술의 특징을 토대로 대상기술의 장단점을 비교 분석하여 기술적 차별성, 혁신성, 파급성 등을 판단한다.

권리성 분석은 평가대상특허의 권리 안정성, 권리보호 가능성, 권리행사 용이성 분석을 통해 평가대상특허 자체의 담보력을 평가한다. 권리 안정성 분석은 선행기술조사를 통해 진보성 및 기타 사유로 인한 무효 가능성이 있는지를 검토한다. 권리보호 가능성 분석은 생산 제품과 특허청구범위의 비교, 경쟁사 제품과 특허청구범위의 비교 등을 통해 특허권의 권리범위가 사업을 보호할 수 있도록 적절하게 작성되어 있는지 검토한다. 권리행사 용이성 분석은 침해 발견과 입증이 용이한 발명인지, 침해 발견과 입증이 용이하도록 청구

범위가 작성되어 있는지, 권리행사에 제한요소가 있는지에 대해 검토한다.

시장성 분석의 사전작업은 대상기술 적용제품(또는 공정), 적용시장에 대한 정의와 범위가 명확하게 설정되어야 하고, 적용제품의 특성(차별성 혹은 우위성)이 요약되어야 한다. 시장성 분석은 대상기술 적용제품이 속한 시장의 환경분석, 경쟁분석 등의 결과에 근거하여 적용제품의 시장경쟁력을 평가하는 것이다. 대상기술이 속하는 산업의 특성 및 환경, 시장구조, 제품 현황, 시장진입장벽, 관련 정책, 국내외 시장동향 및 업체동향을 조사 분석하고, 이를 근거로 시장규모와 시장점유율 추정에 필요한 정보를 수집한다.

IP거래시장성은 현재 해당 기술분야에서 IP거래가 활발한지, 향후 IP거래가 활성화될 가능성이 높은 지 평가하는 것이다. IP거래, 라이선스가 활발히 이루어지고 있는 기술분야인지, 특허관리회사(NPE)가 활발히 매입하는 기술 분야인지, IP분쟁이 자주 발생하는 기술분야인지, 특허출원이 증가추세에 있는 기술분야 인지에 대해 분석한다.

[표 II-19] IP가치평가 세부 평가항목 및 평가지표

평가요소	평가항목		평가지표
합리적 실시료율	기준실시료율		· 신뢰성 있는 실시료율 DB 중에서 선택
	조정지표	기술성	· 차별성 · 혁신성 · 파급성 · 전망성 · 대체성
		시장성	· 수요성 · 수익성 · 생산용이성 · 성장성 · 경쟁성
실시료 기준액	추정 매출액		· 시장점유율에 의한 매출액/수요예측에 의한 매출액 추정/해당기업의 사업계획에 의한 추정 등 다양한 매출추정 방법을 활용
	IP보호 부분의 비중		· 원가비중 또는 전문가 정성평가에 의해 매출 추정이 된 전체 제품 중 IP보호 부분의 비중 산정

평가요소	평가항목		평가지표
IP담보요소	권리성	권리안정성	· 신규성 또는 진보성 무효가능성 · 기타 요인에 의한 무효가능성
		권리보호 가능성	· 해당 제품(서비스) 보호 가능성 · 국내외 IP포트폴리오 구축의 적절성
		권리행사 용이성	· 침해 발견 및 입증 용이성 · 권리행사 제한 가능성
	IP거래시장성		· 분쟁 및 라이선스 활성화도 · 특허출원 활성화도
IP의 수명(n)	기준수명		· 산업부 기술가치평가 모형의 CLT 산출표/IPC 분류별 등록 유지년수 산출표 중 선택
	조정지표	기술요인	· 대체기술 출현 가능성 · 기술적 우월성 · 유사 경쟁기술의 존재(수) · 모방 난이도 · 권리 강도
		시장요인	· 시장집중도(주도기업 존재) · 시장경쟁의 변화 · 시장경쟁강도 · 예상 시장점유율 · 신제품 출현빈도
할인율(r)	상장CAPM 비상장 규모 위험프리미엄 안정성 위험 프리미엄 타인자본비용 자기자본비용		· 산업부 가이드의 산업별 할인율 산출표 이용
	사업화 위험 프리미엄	기술위험	· 기술 우수성 · 기술 경쟁성 · 기술 모방 용이성 · 기술 사업화 환경 · 권리 안정성
		시장 및 사업위험	· 시장 성장성 · 시장 경쟁성 · 시장진입 가능성 · 생산 용이성 · 수익성 및 안정성

자료: 특허청(2013), 지식재산권 담보를 위한 지식재산가치평가 실무 가이드

라. 기관별 기술가치평가 방법

국내에서 공공에 의한 가치평가 모형은 한국발명진흥회(KIPA)의 SMART, 한국특허정보원(KIPI)의 K-PEG, 한국과학기술정보연구원(KISTI)의 Star-value, 한국산업기술진흥원(KIAT)의 NTB(국가기술사업화종합정보망), 기술보증기금 등이 존재한다. 이중 SMART와 K-PEG은 등급평가 기반인 반면 Star-value와 NTB는 가치금액으로 평가결과를 산출한다. NTB의 경우 복합적인 기술평가(기술 사업성, 기업 기술력, 기술가치, R&D 경제성, 시장접근법 가치) 중 한 부분으로서 기술가치를 평가한다. Star-value는 기술이전 및 거래를 위한 가치평가에 치중하여 화폐가치 결과를 산출한다.

[표 II-20] 주요 기관 평가항목 및 모형 특징

	한국발명진흥회	한국특허정보원	KISTI	KIAT
평가항목 특징	· 전문가 평가 기반 평가 모형 사용 · 평가지표 40~50여개 사용	· 전문가 평가가 아닌 국내 특허 중 미국에 출원된 특허를 우수 특허 기준	· 인용특허수명(TCT) DB, 상장/비상장 기업 및 업종별 재무정보 DB, 거래사례 DB, 할인율/로열티/산업기술요소 DB의 독자적 구축 활용	· 업종별 표준재무정보(대, 중, 소 및 중소기업은 업력 구분) DB 및 중소벤처 기업의 업종별 대표기업 재무정보DB · 거래사례 DB, 상장기업과 비상장 기업의 영업이익률로 구한 상관행법 로열티 DB · 인용특허수명(TCT), 할인율, 산업기술요소는 평가기관이 공동 구축 활용
평가	등급평가	등급평가	가치평가	· 기술사업성평가 · 기업기술력평가 · 가치평가 · R&D경제성평가 · 시장접근법 가치평가
평가 모형	다중회귀분석 전문가평가결과를 정답set으로 활용→기계학습화	선형알고리즘으로, 이상치 처리 및 다중공선성 처리가 미흡함	할인현금흐름(DCF) 모형, 로열티절감법 모형 등 전체 5개 모형에 대한 평가모형 KISTI 개발(독자적)	· 할인현금흐름(DCF) 모형 · 로열티공제법 모형

마. 기술가치평가 모형

기업의 가치는 기업이 소유한 영업자산(operating asset)의 시장가치를 의미하며, 영업자산이란 현금을 제외한 그 기업이 가지고 있는 유·무형의 모든 자산을 의미한다. 기업가치는 타인자본가치와 자기자본가치의 합으로서, 기업의 가치를 평가하는 목적은 기업의 주주지분의 가치를 평가하는데 있다. 기술가치평가 방법은 다양하게 존재하는데 일반적으로 시장접근법, 수익접근법, 비용접근법 등으로 구분된다.

수익접근법(Income Approach Method)은 평가대상기술의 경제적 수명 동안 기술사업화로 인하여 발생될 경제적 이익을 추정한 후 할인율을 적용하여 현재가치로 환산하는 방법으로, 수익접근법을 이용하여 가치를 산정하는 경우 4가지 평가요소(기술의 경제적 수명, 현금흐름, 할인율, 기술기여도)를 추정해야 하며 절차는 다음과 같다.

1. 예비평가(screening) : 사업화 가능성을 체크하기 위한 약식평가(평가지표 이용)
2. 사업타당성 평가(assessment) : 정성적 평가로서 기술의 사업추진 타당성을 위한 심층평가이며 기술이 적용될 신규 사업 프로젝트에 대한 평가 : 기술군 특성, 권리성, 시장성, 사업성 평가
3. 가치평가(valuation) : 정량적 평가로써 사업타당성 평가를 전제로 기술의 사업가치 및 기술자산의 가치 측정
 - ① 기술의 경제적 수명 추정 → 현금흐름 산정기간 추정
 - ② 매출액 추정 → 현금흐름 산정
 - ③ 재무 분석 → 매출원가, 판관비용, 운전자본, 자본지출 등 추정
 - ④ 경제적 이익 흐름 → 현금흐름 또는 이익흐름 추정
 - ⑤ 할인율 → 사업위험분석(WACC, CAPM, 경험적 방법 등)
 - ⑥ 사업가치(business value) : 경제적 이익의 현재가치 합계=순현재가치(NPV)
 - ⑦ 기술기여도(tech factor) : 사업가치 중 기술이 공헌한 비율(경제적 이익의 배분)
 - ⑧ 기술가치(tech value) : 사업가치×기술기여도

현금흐름할인법(Discounted Cash Flow method, DCF)에서 가치는 미래 예상 현금흐름을 적절한 할인율로 할인하여 구한 현재가치로 정의한다. 여기서 현금흐름은 총현금유입에서 총현금유출을 뺀 순현금유입의 의미하며 기업이 영업활동을 유지 또는 확대하면서도 자유롭게 사용가능한 현금이다. DCF에 의한 기술의 가치(V_r)은 기술의 경제적 수명기간 동안의 미래 현금흐름을 할인율로 할인한 현재가치의 합이며, 사업 가치에 기술이 기여한 정도인 기술기여도(업종별로 무형자산 중 기술자산이 차지하는 비율)를 곱한 값으로 정의한다.

$$V_r = \sum_{t=1}^n \frac{CF_t}{(1+r)^t} \times \text{기술기여도}$$

여기서, t 는 현금흐름 추정이 이루어지는 기간, n 은 기술의 경제적 수명, CF_t 는 t 기간에서 현금흐름, r 은 할인율, 기술기여도는 사업가치 중 기술이 기여하는 비율이다. 여유현금흐름은 세후 영업이익에 감가상각비 등을 더하고 자본적 지출액과 운전자본 증감액을 차감한 값으로 다음과 같다.

$$\begin{aligned} \text{여유현금흐름} &= \text{영업이익} - \text{법인(소득)세} + \text{감가상각비 등} - \text{자본적 지출액} - \text{운전자본 증감액} \\ \text{영업이익} &= \text{매출액} - \text{매출원가} - \text{판매비와관리비} \end{aligned}$$

정태적 수익접근법은 의사결정의 변동성을 고려하지 않는 것으로, 추정기간 동안 발생가능한 현금흐름의 변동요인(할인율, 변동확률 등)이 고정되어 있다고 가정한다.

동태적 수익접근법은 미래시장 환경 및 의사결정 방향의 변동성을 고려하는 것으로, 사업주체가 창출하는 미래현금흐름에 영향을 미치는 요소를 의사결정나무(decision tree) 형태의 시나리오로 작성하고, 발생가능한 확률을 부여한 후 최적의 시나리오를 선택하여 시나리오에 의한 현금흐름을 추정한다.

기술요소법(Technology Factor Method Approach) : 산업 특성과 개별기술 특성을 동시에 반영하여 기술기여를 산출하는 방법이며 기술기여도(기술요소)는 산업 특성을 반

영하는 산업기수요소(industry factor)와 개별기술의 특성을 평가하는 개별기술강도(technology rating) 곱으로 구성된다.

몬테카를로 시뮬레이션법은 미래현금흐름의 크기에 영향을 미치는 개별 확률변수에 대해 특정의 확률분포를 가정한 상태에서 난수의 발생을 통해 가치평가 결과 및 확률적 기댓값을 계산한다.

25% Rule은 경험측 방법은 과거의 많은 기술거래 협상과정에서 기술도입자(licensee)와 기술판매자(licensor) 간의 합의에 의한 결과로 기술거래에 의한 창출한 이익의 25%는 기술판매자의 몫으로, 75%는 기술도입자의 몫으로 배분하는 것이 합리적이며 기술거래시장에서 관행적으로 이루어져 오고 있으며 이익은 기술 사업에 의해 발생한 세전영업이익이 기준이다.

시장접근법(Market Approach Method)은 평가대상 기술과 동일 또는 유사한 기술이 활성시장에서 거래된 가치를 근거로 기술을 비교분석하여 상대적인 가치를 산정하는 방법으로, 시장접근법을 적용할 경우, 비교대상과 어느 정도 유사성이 있는 지 판단하여 비교대상과 유의한 차이가 있을 때 적절히 차이를 조정해야 하며 절차는 다음과 같다.

1. 시장자료 분석 : 적절한 거래시장을 조사하여 기술의 종류, 기술의 사용현황, 활용산업, 거래일자 등에 있어서 유사하거나 동일한 기술에 대한 매매거래, 목록 등에 대한 자료 수집
2. 사실확인 및 검증 : 자료의 사실과 일치 여부, 실제 기술거래여부 확인 등 신뢰성 검증
3. 비교방법 결정 : 비교대상 분석 항목, 비교분석 방법 적용 등
4. 로열티 산출 : 비교평가 대상기술의 매매 또는 라이선스 거래와 평가대상기술의 거래 조건을 통한 최적의 로열티 산출
5. 조정계수 : 기술의 특성을 반영한 조정계수 적용(기술성, 권리성, 라이선스 계약 우위성, 시장성, 사업성)
6. 기술차지 : 비교거래기술들로부터 산출한 다양한 로열티와 조정계수를 통한 기술가치 추정

거래사례비교법은 거래유형과 자산의 속성이 완벽하게 동질적인 기술자산의 총거래 금액을 쉽게 발견할 수 있다면, 동 금액을 평가대상 자산의 가치로 인정하는 거래사례비교법이 가장 정확한 기술가치평가이다. 그러나 현실적으로 비교용 기술이 대상 기술과 동질적인 경우가 드물고 거래조건도 상이하여, 이러한 조건을 충족하는 거래 사례의 발견이 매우 어려우므로 거래사례비교법을 실무에서 적용하는 경우는 거의 없다.

로열티 공제법과 로열티 참조법은 평가대상기술과 관련된 기술(업종)로부터 합리적인 로열티율을 참조할 수 있는 경우 과거 유사 또는 동일한 기술(업종)의 로열티율을 참조하여 기술의 가치를 추정하는 방법이다.

로열티 공제법은 평가대상이 보유 중인 기술이 만약 제3자로부터 라이선스 되었다고 가정하였을 때 지급되었을 것으로 추정되는 로열티를 기초로 기술의 가치를 추정한다. 로열티 참조법은 평가대상기술이 반드시 제3자로부터 라이선스의 대상이 된다고 가정하지 않더라도, 단지 유사기술의 라이선스 거래에 적용할 로열티율을 참조함으로써 보다 다양한 분야에서 기술의 가치를 추정할 수 있다.

로열티 공제법은 권리확보 및 라이선싱을 전제로 한 지불로열티를 기초로 하지만, 로열티 참조법은 권리유무나 라이선스 유무에 무관하게 유사한 사례의 로열티를 보다 광범위하게 참조하여 기술가치를 산출하는 방법이다.

$$\text{기술가치} = \sum_{t=1}^T \frac{(\text{로열티율} \times \text{조정 계수} \times \text{매출액})(1 - \text{법정세율})}{(1+r)^t}$$

여기서 T 는 기술의 경제적 수명, r 은 할인율

비용접근법은 대체의 경제 원리에 기초를 두고 동일한 경제적 편익을 가지고 있는 기술을 개발하거나 구입하는 원가를 추정하여 가치를 산정하는 방법으로, 비용접근법을 적용할 경우 기술개발비용, 재생산원가, 대체원가 등 상세한 원가정보가 필요하다. 비용접근법은 역사적 원가법, 재생산 비용법, 감가상각된 대체비용법을 사용한다. 과거에 평가대상기

술 개발에 투입된 비용을 평가시점에 재투입한다고 가정할 때 소요될 것으로 예상되는 전체 비용금액에서 평가대상기술의 수요 감소에 기인하는 가치 감소분을 차감하게 된다.

역사적 원가법에서 과거 투입비용 항목은 항목구분 내역을 그대로 인정된다. 과거 투입 비용 금액은 평가대상기술을 재생산하기 위해 평가시점에 제반 비용항목을 재투입한다고 가정하는 경우 소요 예상 금액이다. 단, 특정 비용항목이 과거에는 투입되었으나 평가시점의 연구개발단계에서는 투입되지 않은 상황으로 변경된 경우 평가시점에 대체 투입되는 항목을 기준으로 비용금액을 산정한다.

재생산 비용법은 기존에는 평가시점에 재생산할 경우 투입예상비용이 현재에는 역사적 원가법과 같은 방식으로 적용된다.

감가상각된 대체비용법에서 기존의 물리적 감소는 유형자산의 경우만 인정하고 무형의 기술은 불인정된다. 기능적 진부화는 대체기술의 등장에 기인한 수요 감소로 환산되며 경제적 진부화는 경제적 진부화가 기능적 진부화의 연장선상에서 이루어진다는 가정 하에 기준이 적용된다.

표준적인 비용접근법인 조정재생산 비용접근법은 역사적 원가법, 재생산 비용법, 감가상각된 대체비용법의 취지를 결합하여 단일화한 것으로, ① 재생산 가정 하에 과거의 생산요소 투입항목은 그대로 유지하되, ② 생산요소의 단위비용은 평가시점 기준으로 재조정하고, ③ 최종적으로 합산된 비용가치에 시장수요의 변동에 따른 가치변동분을 가산 또는 차감하는 세 가지 절차로 수행된다.

R&D배수법은 연구개발 투입비용이 기업의 가치와 비례한다는 가정 하에 R&D투입비용, 기술기여도, 조정계수 등을 이용하여 기술가치를 간략하게 산출하는 방법으로, 평가대상기술의 R&D투입비용을 합리적으로 추정이 가능한 경우에 적용된다.

$$\text{기술가치} = (\text{상장기업 기업가치}) / (\text{상장기업 R\&D비용}) \times \text{투입R\&D비용} \times \text{기술기여도} \times \text{비상장조정} \times \text{사업화단계조정}$$

이러한 방법들은 평가목적, 대상기술, 평가 상황 등에 따라 평가방법이 달라질 수 있다.

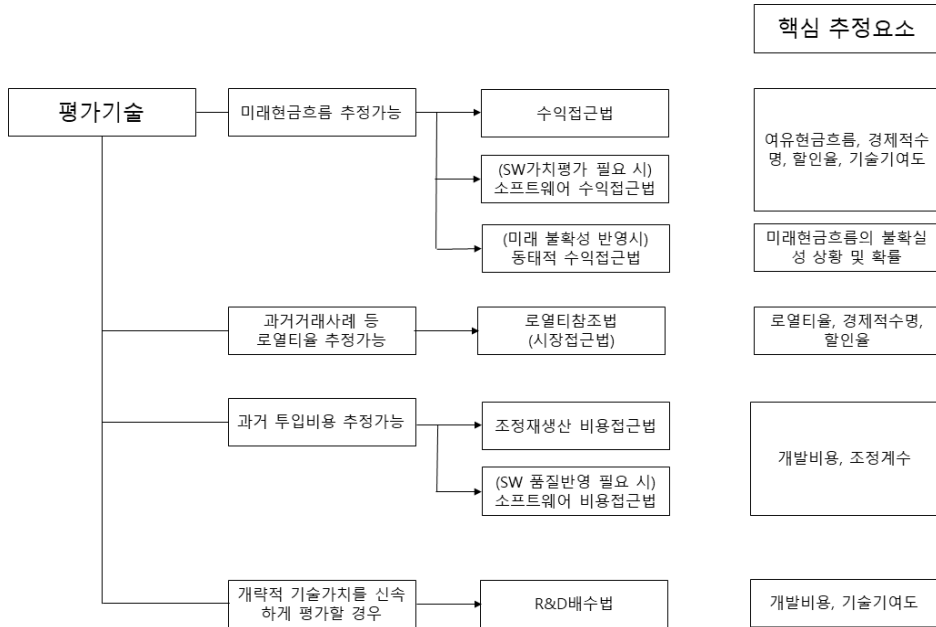
평가시점 현재 동질성 있는 기술에 대한 가격을 시장에서 관찰할 수 있는 경우에 시장접근법을 우선 적용할 수 있다. 평가자는 기술가치평가 실무에서 다양한 상황과 투입변수의 견고성 등을 고려하여 적절한 평가접근법이나 산출방법을 적용해야 한다. 가치산출과 추정요류를 최소화하고 보다 합리적 측정을 위하여 복수의 평가접근법이나 방법을 사용하는 것을 기업회계 기준 등에서 권고하고 있다.

[표 II-21] 기술가치평가 방법론

수익접근법	현금흐름할인법(DCF) 기술요소법(Technology Factor Method Approach) 경험측(25% Rule) 로열티공제법 다기간초과이익법(Multi-Period Earning) Residual Income Method 몬테카를로 시뮬레이션 실물옵션(Real Option) 퍼지수익접근법(Fuzzy-Period Method)
시장접근법	거래사례비교법 산업표준법(Industry Standards) 경매(Auction) 로열티공제법(Relief form Royalty) 로열티참조법(Royalty Comparison Method) 조건부가치평가(contingent Valuation Method) 가상경매시장법(Virtual Auction Market Method)
비용접근법	역사적기원법(Historical Cost Method) 재생산비용법(Reproduction Cost Method) 대체원가(Replacement Cost Method) 조정재생산비용법(Adjusted Reproduction Cost Method) R&D배수법

미래현금흐름 추정가능 여부, 로열티율 추정 가능 여부, 과거 개발비용 추정 가능 여부 및 평가대상기술의 특성, 용도 등을 종합적으로 고려하여 기술가치평가 방법을 선택된다. 기술가치평가 시 발생가능성이 높은 가정이나 조건을 설정하여 가장 효율적이고 효과적으로 평가 대상 기술을 평가할 필요가 있다. 기술가치평가 결과는 상황 변화에 따라 변동될 수 있다.

〈그림 II-8〉 기술가치평가 방법론과 핵심 추정요소



기술가치평가에서 많이 이용되는 접근법은 장점과 단점을 동시에 가지고 있음

[표 II-22] 주요 기술가치평가 접근법의 장단점

종류	평가방법	장점	단점
수익접근법	과거에 평가대상기술 개발에 투입된 비용을 평가시점에 재투입한다고 가정할 때 소요될 것으로 예상되는 전체 비용금액에서 평가대상기술의 수요 감소에 기인하는 가치 감소분을 차감함	· 예상수익을 산정방식의 기준으로 하나, 지적재산권의 실시권의 허여에 의해 실제 실시로 등의 수익이 발생하고 있다면 비교적 용이하게 객관적인 가액의 평가가 가능 · 제반요소가 객관적으로 결정될 수 있다면 이론적으로는 가장 타당한 방법	· 평가요소 추정에 따라 편차가 발생할 수 있음. 평가요소는 일정한 가정에서 추정되는데, 충분한 객관적 근거가 없거나 평가자가 임의로 값을 부여할 경우, 가치평가 결과는 객관성과 합리성이 결여 · 기술가치를 과대하게 산출할 가능성이 존재하며, 이는 추정 매출의 과대평가, 재무비용 적용 오류, 할인율 과소추정에 의해 발생 · 투입정보의 수준에 따라 평가결과가 상당히 달라짐
현금흐름 할인(DCF)	미래유동현금흐름 (FCF)의 현재가치 합	· 이론적으로 합당함 · 실질적 증거들이 법의 신뢰성을 뒷받침함	· 미래현금흐름추정이 어려움 · 비상장 기업의 변수추정이 어려움 · 가중자본비용(WACC)은 일정한 자본

종류	평가방법	장점	단점
			구조와 세율 가정 · 측정값이 잔존가치 성장률에 민감하게 영향을 받음
경제적 부가가치 (EVA)	법인세를 차감한 후의 순영업 이익에서 기업활동의 사용된 총자본에 대한 자본비용을 차감한 값	· 회사가 언제, 그리고 어떻게 (when and how) 가치를 창출하는 지 명시적으로 나타냄	· 가중자본비용(WACC)은 일정한 자본 구조와 세율을 가정 · 측정값이 잔존가치 성장률에 민감하게 영향을 받음
조정현재가치 (APV)	무부채기업(unlevered enterprise value)의 가치에 이자비용절세 효과 및 기타 부채사용효과를 더하는 방법	· 이론적으로 합당함 · 자본구조, 실효세율의 변동이 예상될 때 적합	· 계산이 DCF/EVA 보다 다소 복잡 · DCF처럼 미래 현금흐름의 측정이 어려움
실물옵션 (Real Option)	포기, 축소, 확장, 연기 등의 미래 옵션을 고려하여 불확실성에 따른 투자전략의 유연성을 감안하여 평가하는 방법	· 이론적으로 합당함 · 미래 상황에서의 관리자의 유연성을 고려할 수 있음	· 업계에서 일반적으로 사용되지 않고, 이해가 용이하지 않음 · 옵션 솔루션이 존재하는 모형이 제한적이므로, 실재상황을 단순화 시키는데 어려움 경우가 많음 · Black-Scholes 모형에 한정적으로 사용 (BS모형의 한계점을 가지고 있음)
시장접근법	평가대상 기술과 동일 또는 유사한 기술이 활성시장에서 거래된 가치를 근거로 기술을 비교 분석하여 상대적인 가치를 산정하는 방법으로, 시장접근법을 적용할 경우, 비교대상과 어느 정도 유사성이 있는 지 판단하여 비교대상과 유의한 차이가 있을 때 적절히 차이를 조정	가장 직접적이고 이해가 쉬운 평가방법	특정 지적재산이 그것 자체만으로 거래되는 경우는 드물며 실제로 그러한 거래가 있다고 하더라도 거래조건이 명확하지 않는 경우가 많아 시장접근법을 지적재산에 적용할 때에는 비교가능성의 확보가 무엇보다도 어려움
시장가치 비교법 (Comparables)	평가대상의 재무상황 및 영업 환경이 유사한 기업들을 선택하여 배수를 구하는 방법	· 쉽고 빠르게 적용 가능 · DCF에 비해 미래성과 추정이 간략하고 평가결과에 대한 이해 분석이 용이 · 평가시점의 시장상황 반영	· 비교대상기업선정이 어려움 · 위험, 성장, 현금흐름에 대한 고려가 생략됨 · 현재 시장이 일시적 과대평가 혹은 과소 평가되었을 경우 평가결과 역시 왜곡될 수 있음
비용접근법	과거에 평가대상기술 개발에 투입된 비용을 평가시점에 재투입한다고 가정할 때 소요될 것으로 예상되는 전체 비용금액에서 평가대상기술의 수요 감소에 기인하는 가치 감소분을 차감함		· 비용접근법은 대상자산이 보유하는 미래의 서비스와 동일한 서비스를 얻기 위한 필요 금액을 산출함으로써 자산보유자의 미래이익을 측정하는 방법이며 일반적으로 사용되고 있는 시장접근법과 수익접근법에 비하면 포괄성이 부족 · 가치에 영향을 미치는 극히 중요한 요소가 검토대상에서 제외되어 있으므로 이러한 요소에 대해서는 비용접근법의 기본적인 요소와 별도로 검토하여야 함

바. 사례

IP가치평가 산식과 요소이며 다음의 식과 변수로 정의되며 이후 세부지표는 앞의 특허청 지표를 따른다.

$$V_{IP} = \sum_{t=1}^n \frac{\text{합리적실시료율} \times \text{실시료기준액}}{(1+r)^t} \times \text{IP담보요소}$$

[표 II-23] 현행 IP가치평가 산식 요소

	정의	비고
VIP	IP의 가치	독립된 재산권으로서의 IP의 담보가치
합리적 실시료율	합리적실시료율 (Reasonable Royalty Rate) = 기준실시료율 × 조정지표	동일·유사 기술분야의 실시료율을 기준으로 평가대상 IP의 기술 및 시장 평가를 통해 조정된 실시료율
실시료 기준액	실시료기준액(Royalty Base) = 추정 매출액 × IP보호 부분의 비중	· 추정 매출액에 IP로 보호되는 부분의 비중을 반영한 실시료 산정 기준액 · 추정매출액 = 시장규모 × 시장점유율 · 과거 자료가 존재하지 않을 경우 수요예측에 의한 매출액을 추정할 수도 있음
IP 담보요소	IP담보요소 (IP Collateral Factor) = 권리성(권리안정성, 권리보호가능성, 권리행사 용이성) 및 IP거래시장성이 반영된 조정지표	담보로서 재산권(권리)의 완전성 여부를 측정하기 요소로서, IP자체의 품질 및 IP거래시장의 활성도를 포함
n	IP의 수명 = 기준수명 × 조정지표	· IP가 수익을 창출할 수 있는 기간 · 기준수명은 CLT(Cited patent Life Time) 지수나 IPC(International Patent Classification) 활용 · 수명조정지표는 기술요인 5개와 시장요인 5개의 산술평균
r	할인율(WACC) = 자기자본비용 × 자기자본비율 + 타인자본비용 × 타인자본비율 × (1-법인세율)	· 미래의 현금흐름을 현재가치로 전환하기 위한 환산율 · 중소기업 자기자본비용(K _e)=상장기업 CAPM+사업화 위험프리미엄 +안정성 위험프리미엄+규모프리미엄 · 타인자본비용(K _d)=상장기업 타인자본비용+추가위험 스프레드 · 상장기업 CAPM, 안정성 위험 프리미엄, 규모 프리미엄, 타인자본비용은 제조업 산업별 할인율 산출표를 이용함

자료: 특허청(2013)

III

데이터 가치평가 방법, 사례 및 정책 제언

2장에서 거시적인 기술금융과 보다 미시적인 기술가치평가에 대해 알아보았고, 기술가치평가의 범위를 좁혀, 현재 존재하지 않는 데이터 가치평가와 가장 가까운 소프트웨어 가치평가에 대해 살펴볼 필요가 있다. 소프트웨어의 가치평가를 통해 데이터 가치평가에 대한 보다 세분화된 지표를 살펴볼 필요가 있다. 특히, 정보서비스업과 정보데이터(DB)관리업 등은 현재 데이터 가치평가와 밀접한 관련을 가지고 있거나 데이터 가치평가의 세부항목으로 잡힐 수 있는 변수이다.

1 » 소프트웨어 가치평가

가. 수익접근법

현재까지 데이터 가치평가에 가장 가까운 가치평가는 소프트웨어로 볼 수 있으며 이에 대한 수익접근법은 다음과 같이 평가되고 있다. 소프트웨어는 「소프트웨어산업진흥법」 제2조1항에 따라 “컴퓨터·통신·자동화 등의 장비와 그 주변장치에 대하여 명령·제어·입력·처리·저장·출력·상호 작용이 가능하도록 지시·명령(음성이나 영상정보를 포함한다)의 집합과 이를 작성하기 위하여 사용된 기술서, 기타 관련 자료”를 말한다.

기술의 경제적 수명기간 동안 평가대상기술이 활용되어 발생하는 미래 여유 현금 흐름의 현재가치 합계액에 기술기여도를 곱하여 기술가치를 산출하는 프로세스는 앞의 수익접근법과 동일하다.

[표 III-1] 소프트웨어 사업분야 분류 및 해당 업종

	SW 사업분야	업종코드	
		표준재무정보	한국은행
패키지 소프트웨어	패키지-ERM Application	5822	582
	패키지-Application		
	패키지-Information & Data Management Software		
	패키지-Application Development & Deployment	6201	62
	패키지-Security		
	패키지-System Infrastructure		
비 패키지 소프트웨어	IT서비스	6202	62
	임베디드	5822	5822
	정보서비스	63	63

자료: 기술보증기금(2018)

기술의 경제적 수명에서 패키지 소프트웨어의 경제적 수명은 소프트웨어의 출시년도로부터 단종년도에 이르는 기간으로 정의된다. 즉, 패키지 소프트웨어의 수명은 소프트웨어 단종년도에서 출시년도를 차감한다.

패키지 소프트웨어의 경제적 수명을 분석하기 위하여 한국소프트웨어산업협회(KOSA)에 등록된 회원사가 출시한 소프트웨어를 대상으로 출시년도, 생존여부 또는 단종년도 정보를 추적 조사하여 4분위수를 계산한다.

비 패키지 소프트웨어의 수명은 각각의 산업특성을 반영하여 산출한다. 임베디드 소프트웨어의 경우, 해당 소프트웨어가 장착되는 장치의 기술순환주기(TCT) 지수를 바탕으로 기술수명 영향요인 평가를 적용하여 기술의 경제적 수명을 결정한다. IT서비스의 경제적 수명은 IT시스템 구축년도부터 동 시스템을 보다 개선된 시스템으로 재구축하는 연도까지의 기간으로 정의한다. IT서비스의 수명은 개선된 대체시스템으로 재구축하는 연도에서 IT시스템 구축년도를 빼면 된다. 정보서비스 사업의 경제적 수명은 한국기업데이터(KED) 정보 보유기업 중 정보서비스업(J63)으로 분류되는 기업 중 설립년도와 폐업년도 사이의 경과기간, 설립년도 이후 생존 연령자료를 바탕으로 산출한다.

[표 III-2] 소프트웨어 종류별 기술순환주기 지수

	1분위	중간값	3분위	산출평균
기업자원관리(ERM)	5	3	11	7.5
애플리케이션, 단말기용	4	7	11	6.9
정보데이터(DB) 관리	3	5	10	5.7
애플리케이션 개발도구	5	7	11	6.6
보안	5	7	11	6.9
시스템기반	3	7	12	6.7
IT서비스	9	10	12	10.11
임베디드	TCT법에 의한 관련 기계장치 기술의 경제적 수명			
정보서비스	5	8	13	8.3

자료: 기술보증기금(2018)

여유현금흐름 추정은 앞의 수익접근법 방법에 따라 추정하며, 표준재무정보 자료에 의한 추정방법을 활용하는 경우, 아래의 재무 정보를 활용한다.

[표 III-3] 소프트웨어 사업분야별 운전자본증감액 산출용 참고자료

	분야	매출채권회 전율	재고자산 회전율	매입채무 회전율
패키지 소프트웨어	패키지-ERM Application	4.8	31.5	17.8
	패키지-Application	5.0	68.9	10.6
	패키지-Information & Data Management Software	3.9	84.2	16.5
	패키지-Application Development & Deployment	3.4	32.4	26.4
	패키지-Security	4.4	19.7	12.4
	패키지-System Infrastructure	2.4	19.4	5.7
비 패키지 소프트웨어	IT서비스	5.7	84.9	9.5
	임베디드	4.8	26.5	15.3
	정보서비스	7.3	615.3	36.2

자료: 기술보증기금(2018)

[표 III-4] 소프트웨어 사업분야별 표준재무비율

	분야	매출 원가율	판관비율	법인(소득)세/ 법인(소득)세 차감전 순이익률비율	감가 상각비 비율	자본적 지출 비율	운전 바본 증감액 비율	영업 이익률
패키지 소프트웨어	패키지-ERM Application	62.14	29.16	3.60	1.91	8.27	2.79	8.70
	패키지-Application	69.66	20.58	5.16	1.38	3.16	1.15	9.75
	패키지-Information & Data Management Software	57.45	34.96	7.26	2.15	5.35	1.37	7.60
	패키지-Application Development & Deployment	53.31	42.98	5.74	5.44	5.44	-1.87	5.71
	패키지-Security	45.36	44.18	6.34	2.66	2.66	2.07	10.46
	패키지-System Infrastructure	63.38	32.94	0.61	2.99	2.99	-0.34	3.68
비 패키지 소프트웨어	IT서비스	74.36	18.42	7.00	0.83	0.83	1.84	6.82
	임베디드	52.85	38.87	5.63	2.44	2.44	2.70	8.27
	정보서비스	21.32	41.55	28.78	2.74	2.74	1.02	37.13

자료: 기술보증기금(2018)

할인율 및 기술기여도에서 할인율 및 기술기여도도 수익접근법을 따르되, 할인율 추정 산식에 사용되는 업종별 자본비용 및 자본구성비율, 산업기술요소는 소프트웨어 업종(J582, J62, J63)에 대해 제공된 데이터를 활용한다.

[표 III-5] 비상장 소프트웨어 업종별 할인율 산출 지침표

산업 분류 코드	자기자본 비중(%)	자기자본비용(%)				타인자본 비중(%)	타인자본비용(%)			
		대	중	소	창업		대	중	소	창업
582	58.57	9.19	10.1	11.16	12.63	41.43	3.71	6.31	7.36	8.72
62	60.36	8.94	9.71	10.61	11.85	39.64	4.3	6.91	7.95	9.32
63	59.18	8.64	9.41	10.32	11.57	40.82	3.22	5.82	6.87	8.24

자료: 기술보증기금(2018)

[표 III-6] 소프트웨어 업종별 산업기술요소

산업분류코드		산업기술요소(%)
582	소프트웨어개발 공급업	67.75
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	64.07
63	정보서비스업	71.11

자료: 기술보증기금(2018)

나. 비용접근법

소프트웨어 가치평가 방법은 특성상 상용화를 목적하는 일부를 제외하고는 시장사례 수집이나 현금흐름의 추정이 현실적으로 불가능한 경우가 많아, 비용접근법을 중용하여 적용 가능한 평가방법이 필요하다. 평가대상 소프트웨어 품질 특성을 가치평가 결과에 반영할 필요가 있다고 판단되고, 조정재생산비용의 진부화율 등을 적용하는 것이 불합리하다고 판단되는 경우에 활용한다.

평가대상 소프트웨어 개발에 투입된 재생산비용에 소프트웨어 품질에 따른 조정계수를 곱하여 산출한다. 즉, 소프트웨어 기술가치는 재생산비용×조정계수이며 조정계수=산출 소프트웨어품질점수/60이다.

재생산비용 추정방법은 조정재생산비용법과 동일한 가치로 추정한다. 조정계수는 소프트웨어의 품질에 따라 기술가치에 영향을 주는 것으로 소프트웨어 품질평가표의 점수로 산출된다. 소프트웨어 품질평가표는 기능성, 사용성, 초기부담비용, 효율성, 신뢰성 등 내적요인과 이식성, 호환성, 유지보수성 등 외적요인으로 구성된 평가표로 소프트웨어 분야 별로 차별화된 가중치를 부여한다.

소프트웨어품질점수(Software Quality Score, SQS)는 ISO/IEC9126의 소프트웨어품질표준을 바탕으로 실무적 특성을 고려하여 개발된 것으로 전통정건 방법(Man-Month, Constructive Cost Model, Function Point)과 차이가 있다. ISO/IEC9126-Software Engineering Production Quality의 SQS는 기능성(Functionality), 신뢰성(Reliability), 사용성(Usability), 효율성(Efficiency), 유지보수성(Maintainability), 이식성(Portability)의 6대 지표로 구성된다.

[표 III-7] 소프트웨어 품질평가표 및 분야별 가중치

	평가항목	IT서비스	임베디드	기업지원 관리	어플리 케이션	데이터 관리	어플리케이션 개발도구	보안	시스템기반
내적 요인	기능성	8.02	9.08	8.33	8.1	9.15	13.66	13.94	4.06
	사용성	7.41	7.8	8.33	12.42	5.52	9.07	10.5	3.78
	초기부담비용	6.48	8.23	8.33	4.91	6.42	5.47	5.52	4.06
	효율성	18.79	25.97	16.68	15.23	18.49	23.74	23.75	13.01
	신뢰성	9.3	15.59	8.33	9.33	10.42	14.73	12.96	8.43
외적 요인	이식성	13	11.11	10	16.67	13	10.92	11.11	27.5
	호환성	20.63	11.11	20	16.67	20.63	13.75	11.11	21.83
	유지보수성	16.37	11.11	20	16.67	16.37	8.66	11.11	17.33

자료: 기술보증기금(2018)

2 데이터 가치평가 가능 지표와 사례

가. 데이터 가치평가 사례

한국정보화진흥원의 「공공데이터 개방품질 평가모형」(2016), 한국데이터진흥원의 「데이터 품질평가 모형」(2002), 「데이터 품질관리 성숙 모형」(2006) 등이 이전 연구가 될 수 있다. 한국데이터진흥원의 모형은 개방데이터를 대상으로 유효성, 정확성, 일관성, 완전성 등 4개 영역 14개 진단항목을 포함하고 있다. 한국데이터진흥원의 모형은 데이터 값, 관리 프로세서, 품질관리 성숙수준 등 3가지로 구성되고 있다.

한국정보화진흥원의 「공공데이터 개방품질 평가모형」(2016)은 4개 영역, 14개 진단항목으로, 품질 오류율 기준으로 5등급 평가를 하고 있다. 대상은 개방데이터(파일데이터, 오픈 API) 관련 원천데이터이며 방법은 데이터 오류진단 프로파일링 기준으로 하고 있다.

유효성 : 여부, 수량, 금액, 율, 날짜, 코드, 번호

정확성 : 번호, 명칭

일관성 : 시간순서, 컬럼논리관계, 계산 및 집계

완전성 : 필수값, 관계·키

프랑스 데이터 중개 기업인 DAWEX는 데이터 등급을 평가하고 있다. 10개의 지표로 운영되고 있으며 10개의 지표 내에서 각 세부지표가 존재하고 있다.

기업신뢰도 : 신용도, 평판 및 규모, 지속가능성, 전문성

희귀성 : 대체성

사실성 : 정확성, 일관성

규모 : 대표/타당성, 수집기간

최신성 : 갱신주기, 등록주기

정형화 : 표준화, 구조화

문서화 : 명확성, 설명력

완전성 : 다양성, 세분화

지속성 : 가용성, 내구성

가치성 : 활용성, 유용성

[표 III-8] DAWEX 데이터 등급 지표

주요지표	세부지표	설명
기업신뢰도 (Market Position)	신용도 평판 및 규모 지속가능성 전문성	데이터 공급 업체의 평판과 지위
희귀성 (Rarities)	대체성	데이터를 대체할 수 있는 비슷한 데이터의 존재
사실성 (Truth)	정확성 일관성	데이터의 값과 실제 값의 일치 여부
규모 (Volume)	대표/타당성 수집 기간	데이터 샘플의 대표성과 수집 기간
최신성 (Frequency)	갱신주기 등록주기	업데이트 갱신 주기
정형화 (Structured)	표준화 구조화	데이터를 정형화하기 위한 추가 노력/데이터의 정돈된 수준

주요지표	세부지표	설명
문서화 (Documented)	명확성 설명력	데이터에 대한 명확하고 자세한 설명
완전성 (Completeness)	다양성 세분화	운영 시점에서 확보 가능한 데이터(데이터 운영 시점이 실시간인 경우)
지속성 (Longevity)	가용성 내구성	지속적으로 확보가능한 데이터
의존도 (Dependency)	활용성 유용성	데이터로 인해 발생하는 의사결정의 중요성

자료: DAWEX 홈페이지

Cai and Zhu(2018)의 연구에서는 데이터 질적인 측면을 강조하고 있다. 데이터의 질적인 측면은 다음과 같이 5개 지표, 14개 지표로 구성되고 있다.

[표 III-9] Cai and Zhu의 가치평가를 위한 지표

지표	세부지표
유용성(Availability)	접근성(Accessibility) 적시성(Timeliness) 권한부여(Authorization)
사용성(Usability)	정의/문서화(Definition/Documentation) 신용성(Credibility) 메타데이터(MetaData)
신뢰성(Reliability)	정확성(Accuracy) 무결성(Integrity) 일관성(Consistency) 완전성(Completeness) 감사가능성(Auditability)
적절성(Relevance)	적합성(Fitness)
데이터 발표의 질 (Presentation Quality)	가독성(Readability) 구조(Structure)

자료: Cai and Zhu(2018)

나. 데이터 가치평가 가능 지표

기술금융 지표, 기술가치평가모형 지표, 한국정보화진흥원 모형의 지표, 한국데이터진흥원 모형의 지표, DAWEX 지표를 이용하면 데이터 가치평가에 대한 새로운 모형과 지표가 도입될 수 있다. 기술금융 지표와 기술가치평가모형, 소프트웨어 가치평가의 관련 지표를 맵핑하면 된다. 기존 한국정보화진흥원 모형 지표, 한국데이터진흥원 모형 지표를 맵핑한다. 또한, DAWEX 지표 등과 비교 맵핑한다.

거시적인 기술금융 지표 중, 데이터 가치평가에 이용될 수 있는 지표는 경영주역량, 기술성, 시장성, 사업성에서 일부 지표가 이용될 수 있다.

경영주역량 : 데이터 경험 수준, 관리능력 등

기술성 : 데이터개발 추진능력, 데이터 완성도 및 확장성 등

시장성 : 세부 시장 경쟁상황 등

사업성 : 운용능력 및 수익성 등

미시적인 기술가치평가 지표 중, 데이터 가치평가에 이용될 수 있는 지표는 기술성, 시장성, 권리성 등에서 일부 지표가 이용될 수 있다.

기술성 : 차별성, 혁신성, 파급성, 전망성, 대체성, 기술적 우월성, 모방난이도 등

시장성 : 수요성, 수익성, 생산용이성, 성장성, 경쟁성, 시장집중도, 시장점유율 등

권리성 : 권리안정성, 권리보호 가능성, 권리행사 용이성 등

거래시장성, 데이터 수명 등

소프트웨어 가치평가 지표 중, 이용될 수 있는 지표는 내적요인과 외적요인으로 구분된다.

내적요인 : 가능성, 사용성, 초기부담비용, 효율성, 신뢰성

외적요인 : 이식성, 호환성, 유지보수성

사례의 DAWEX 지표나 Cai and Zhu 등에서 이용가능한 지표는 다음과 같다.

기업신뢰도 : 신용도, 평판 및 규모, 지속가능성, 전문성

회귀성 : 대체성

사실성 : 정확성, 일관성

규모 : 대표/타당성, 수집기간

최신성 : 갱신주기, 등록주기

정형화 : 표준화, 구조화

문서화 : 명확성, 설명력

완전성 : 다양성, 세분화

지속성 : 가용성, 내구성

가치성 : 활용성, 유용성

데이터 가치평가의 지표는 관점을 어떻게 가져가는 지가 매우 중요하며, 시장 확대를 위하여 공급자 관점의 지표를 먼저 산출하고, 수요자와 플랫폼 관점의 지표를 거시적인 측면부터 미시적인 측면까지 가능한 지표를 산출해야 한다. 데이터 인프라와 같은 거시적 측면과 개별 데이터의 가치 자체에 대한 평가와 미시적인 데이터가 혼합되어야 한다. 이를 위하여 기술금융 지표, 기술가치평가 지표, 소프트웨어 가치평가 지표, 사례 지표 등을 인용할 필요가 있다.

더불어 공급자 관점의 데이터 비용과 데이터 품질, 수요자 관점의 데이터 가치와 서비스 품질, 플랫폼 관점의 시스템 품질 등도 기준이 될 수 있다.

공급자 관점의 가능 지표 : 정확성, 일관성, 최신성, 완전성, 유용성, 조직성 등

수요자 관점의 가능 지표 : 유용성 등

플랫폼 관점의 가능 지표 : 접근성, 적시성, 보안성 등

데이터 가치평가를 통한 시장 확대를 위하여, 먼저 공급자 관점의 지표를 다수 개발하여야 하며, 이후 수요자 관점과 플랫폼 관점의 지표를 개발할 필요가 있다. 각 지표 그룹에 대한 세부지표 맵핑, 정의, 평가대상, 측정방안, 가중치(중요도)가 산출되어야 한다.

참 | 고 | 문 | 헌

- Cai, L. and Zhu, Y.(2015), The Challenges of Data Quality and Data Quality Assessment in the Big Data Era, *Data Science Journal* 14(2), 1-10.
- Heizenberg, J., Lo, T., Zaidi, E., and Duncan, A. D.(2018), *Market Guide for Data and Analytics Service Providers*, Gartner.
- IDC and The Lisbon Council(April 2018), *How the Power of Data Will Drive EU Economy*, The European Data Market Monitoring Tool Report.
- International Data Corporation(July 2017), *Worldwide Semiannual Big Data and Analytics Spending Guide*.
- International Data Corporation(August 2018), *Worldwide Semiannual Big Data and Analytics Spending Guide*.
- Knight, S., and Burn, J.(2005), Developing a Framework for Assessing Information Quality on the World Wide Web, *Information Science Journal* 18, 159-171.
- Li, J. Z. and Liu, X. M.(2013), An Important Aspect of Big Data: Data Usability, *Journal of Computer Research and Development* 50(6), 1147-1162.
- Nature(2008), *Big Data*, Retrieved November 5, 2013 from the World Wide Web, <http://www.nature.com/news/specials/bigdata/index.html>
- Newman, D.(2011), *How to Plan, Participate and Prosper in the Data Economy*, Gartner.
- Science(2011), *Special Online Collection: Dealing with Data*, Retrieved November 5, 2013 from the World Wide, Web: <http://www.sciencemag.org/site/special/data/>
- Wikibon(2018), *Big Data Analytics Trends and Forecast 2018*.
- 4차산업혁명위원회(2018), 데이터 산업 활성화 전략.
- 과학기술정보통신부(2017), 2017년 데이터산업 현황조사.
- 과학기술정책연구원(2014), 기술가치평가 기반 국가R&D사업의 성과평가 및 기술료 연계 가능성 탐색연구, 정책연구 2014-24.
- 금융위원회(2018), 데이터를 가장 안전하게 잘 쓰는 나라를 만들겠습니다, 보도자료.
- 기술보증기금(2018), 기술가치평가 실무 매뉴얼.

- 김은미·김상봉·홍우형(2016), “기술금융이 은행의 성과에 미치는 영향 분석”, 신용카드리뷰, 10(2), 3-20.
- 김상봉·김시연(2014), “신용카드지수 산출 방법론에 관한 연구”, 8(2), 131-152.
- 김상봉·조경준·김시연·김재진·심기준(2017), “기술평가모형 개발 방법론에 관한 연구”, 신용카드리뷰, 11(1), 41-67.
- 박규만·김상봉(2018), “기술 사업성 및 위험 요인을 고려한 기술 가치 평가모형에 관한 실증 연구”, 신용카드리뷰, 12(3), 45-58.
- 배효성·엄성희(2016), “기술가치평가에 대한 향후 개선 방안에 관한 연구”, 신부동산학연구, 26(4), 131-146.
- 산업통상자원부(2014), 기술평가기준 운영지침.
- 이철평·김상봉(2016), IP금융의 투자로의 전환 필요성 연구, 연구용역보고서.
- 조경선·임재용(2014), 지식재산금융과 가치평가 실무.
- 중소기업중앙회, 중소기업 금융 애로실태 조사보고서, 각년도.
- 한국데이터진흥원(2017), 2017 데이터산업 백서.
- 한국데이터진흥원(2018), 2018년 데이터산업 현황조사.
- 한국발명진흥회(2017), 지식재산평가센터 안내자료.
- 한국산업기술진흥원(2011), 기술평가시장 현황분석 및 활성화 방안, 연구용역보고서.
- 한국정보통신진흥협회(2014), 데이터 거래소 설립전략, 빅데이터 조찬간담회, 빅데이터연합회.

데이터로 연결하는 사람·세상·미래



Bridging DATA

제2편

데이터 공유·활용 플랫폼의
특성 및 사례 분석

NIA 한국정보화진흥원



CONTENTS

I. 서론	69
II. 데이터 공유 플랫폼 사례	71
1. 개인정보 공유 플랫폼 DECODE (DEcentralized Citizen-Owned Data Ecosystems)	71
2. MS, 구글, 트위터, 페이스북 데이터 전송 프로젝트 (DTP)	94
3. SSAP2(Smart Ship Application Platform2) 프로젝트	108
4. 건설 데이터 공유 플랫폼 Trimble connect[19]	118
5. 블록체인 공공 네트워크 기반의 인스턴트 개인인증 플랫폼[25]	126
6. 전자상거래 물품 통관 데이터 공유 [27]	148
III. 정책적 제언	160
참고문헌	162



서론

4차 산업 혁명시대에서 데이터는 산업발전과 새로운 가치 창출을 위한 핵심 동력이며, 이에 따라 각국 정부 및 기업은 양질의 데이터를 수집하고 제공하기 위한 정책 수립과 데이터 플랫폼 구축을 위해 노력하고 있다.

영국은 CKAN 오픈 데이터 관리 플랫폼을 활용한 공공데이터 포털을 구축하여 경제, 환경, 교육, 교통 등의 데이터 서비스를 운영하고 있고, 미국은 CKAN과 Wordpress를 통합한 형태의 포털을 통해 농업, 기후, 교육, 에너지, 금융 등의 데이터를 서비스 하고 있다. 일본 역시 공공데이터는 국민에게 공유해야 하는 재산이라는 인식에 입각하여 재무행정, 경제, 사법, 안전, 교통 등의 데이터를 제공하고 있으며, 한국 또한 오픈 API 서비스를 제공하여 교육, 국토관리, 공공행정, 재정금융, 산업고용 등 국가 중점데이터를 국민들에게 제공하고 있다[1].

기업의 경우 기존에 자신이 가진 데이터와 타사의 데이터를 결합해 새로운 시장을 창출하고, 수익을 극대화하기 위해 기업 간 데이터 공유를 위한 협업을 시작했다. 예를 들어 병원의 의료비 청구서, 건강진단 등의 정보를 활용하게 되면, 기업은 새로운 건강 서비스 개발이 가능하게 된다. 해외 주요국은 산업 내에서 자체적으로 데이터 공유를 위한 플랫폼을 개발하고 상호간의 약속을 통해 데이터를 교환·활용하고 있다.

데이터 플랫폼은 센서, 장비, IoT기기, 사람으로부터 생성되는 데이터를 수집하고, 수집된 데이터를 여러 사람, 기업 및 기관에 제공함으로써 필요한 정보를 쉽고 빠르게 얻을 수도 있고 데이터 분석을 통하여 장비의 고장진단이나 예지보전 같은 서비스도 개발할 수 있다.

하지만 데이터의 가치가 높아질수록 이 데이터에 불법적으로 접근해 탈취하려는 시도들

도 많아지고 있다. 특히 데이터의 활용범위가 민간 영역까지 광범위하게 확대되면서 시민들의 개인 정보 침해 및 해킹에 대한 우려가 확산되고 있다.

최근 이러한 개인정보 침해와 해킹 문제를 정부의 정책과 블록체인을 활용한 보안기술을 통해 해결하려 하고 있다. EU의 경우 새로 개정된 개인정보보호법(GDPR)을 2018년 5월부터 시행하여 정보주체인 개인의 권리를 강화하고 기업의 책임을 강화하였고 법령 위반 시 과징금등 행정처분을 부과하고 있다[2].

또한 기술적 측면에서 중앙 집중화된 현재의 데이터베이스는 해킹에 취약할 수밖에 없다. 블록체인은 중앙에서 데이터를 저장하는 데이터베이스와 정 반대로 모든 데이터를 네트워크상의 노드에 블록 형태로 저장한다. 따라서 해커가 데이터를 해킹하기 위해선 51% 이상의 익명의 블록을 해킹해야 하고, 해킹 속도 또한 새로운 블록이 생성되는 속도보다 빨라야하기 때문에 현존하는 컴퓨터로 해킹이 불가능하다.

이처럼 세계 각국은 국민의 삶의 질을 높이고 산업발전과 새로운 가치 창출을 위해 양질의 공공 및 민간 데이터 제공을 위한 정책수립과 안전한 데이터 플랫폼 구축에 노력을 아끼지 않고 있다.

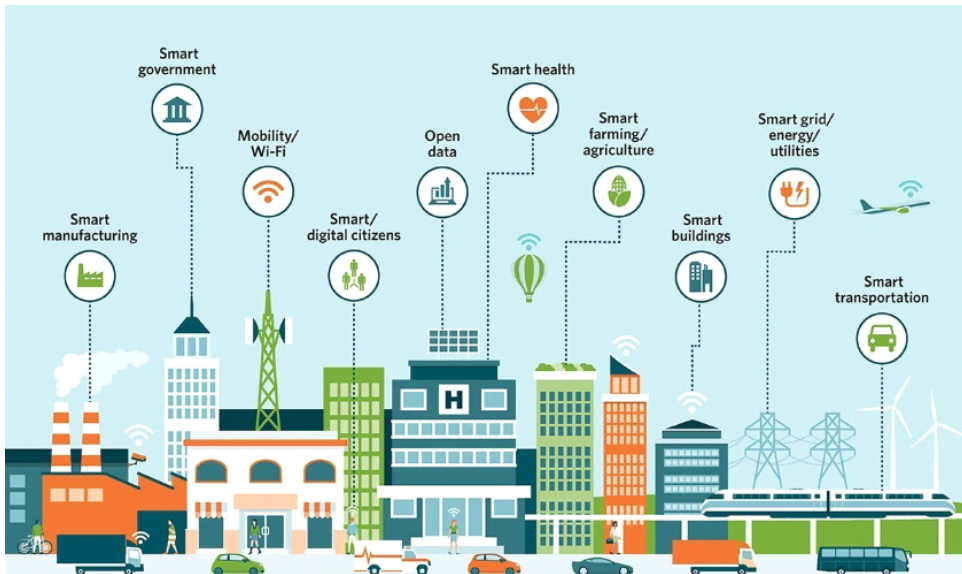
우리 정부 역시 4차 산업을 선도하기 위한 제도, 기술, 산업 기반 등 데이터 전반에 걸친 활성화 정책을 구상해야 하고, 개인정보 및 기업중요정보에 대한 보호 방안도 마련해야 할 것이다.

II

데이터 공유 플랫폼 사례

1 » 개인정보 공유 플랫폼 DECODE(DEcentralized Citizen-Owned Data Ecosystems)

〈그림 1〉 스마트 시티



출처 : Margaret Rouse, 2017. "Smartcity". *IoT Agenda*, July

(1) 배경 및 문제점:

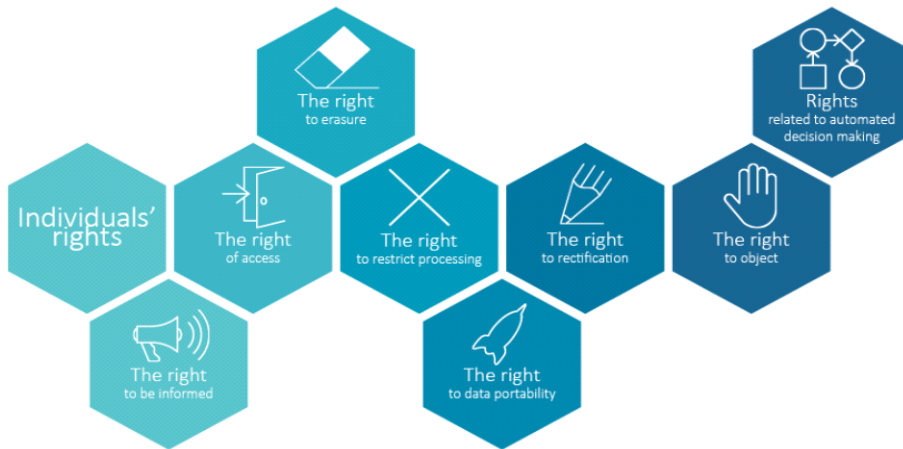
현재 암스테르담, 뉴욕, 시애틀, 샌프란시스코, 시드니, 바르셀로나 등 여러 도시에서 스마트 시티의 기반을 마련하기 위한 노력을 기울이고 있다. 도시의 규모가 커질수록 여론 조사, 범죄 모니터, 교통량조사, 도시계획 수립등과 같은 정책을 수립하는데 있어 더 많은 데이터가 필요하기 때문에, 각 도시는 사물인터넷, 인공지능과 같은 새로운 기술과 빅데이

터를 도시 운영에 통합 시키고 있는 추세이다. 하지만 초기 스마트 도시는 실제 사람들의 요구에 대응하기 보다는 새로운 기술의 도입에만 너무 초점을 맞추고 있었다. 즉, 시민들을 도시 기반 시설 개발에 기여하는 능동적인 존재가 아닌 단순한 수동적 센서같은 도구로 인식 하는 문제가 있었다[3].

개인 데이터를 이용한 공유경제에서 도시는 중요한 역할을 하고 있다. 시 정부의 경우 폐기물관리, 대중교통, 치안, 응급상황 대처 등 시의 모든 업무에 대해 데이터에 입각한 접근 방법이 필요한 상황이다. 시 정부는 시민들로부터 더 많은 요구를 받고 있으나, 시 예산은 그에 맞춰 증가하지 않는 현실에 직면해 있다. 이러한 맥락에서 볼 때, 시 정부가 이용할 수 있는 많은 양의 데이터에서 더 많은 가치를 찾아내는 것이 이 문제를 해결 할 수 있는 방법 중에 하나로 생각된다. 또한 민간 부문의 경우 수집한 데이터를 바탕으로 Uber, Airbnb와 같은 공유 플랫폼에서부터 지역서비스를 제공하는 업체까지 데이터 수집을 통하여 다양한 방면으로 수익모델을 창출 할 수 있다. 그러나 개인 데이터가 점점 경제 분야와 일상생활에서 중요한 요소가 되면서 개인정보 보호 및 데이터 거버넌스와 관련된 문제 또한 발생하고 있다.

스마트 시티의 데이터 활용은 시민들에게 효율성, 편의성, 맞춤 서비스 같은 장점을 제공 해주지만, 개인의 사생활을 침해할 수 있다. 도시들은 연결되고, 데이터 중심적이기를 원 하지만, 무의식적으로 시민들에 대한 대규모 감시에 노출되게 된다. 또한 사람들은 자신의 개인 데이터가 어떻게 수집되고 사용되는지에 대해 알기 어렵고, 정책담당자가 시민들의 동의를 통하여 데이터를 수집하는 시스템도 갖춰지지 않았다. 이러한 문제점 때문에 시민의 개인정보를 보호하고 안전하게 활용하기 위한 시도들이 나타나고 있다. 유럽의 경우 시민의 개인 정보 보호를 위해 2016년도에 유럽의회에서 GDPR(General Data Protection Regulation)을 발표하였다. GDPR은 사용자가 본인의 데이터 처리 관련사항을 제공받을 권리, 열람 요청 권리, 정정 요청 권리, 삭제요청권리, 처리 제한 요청 권리, 데이터 이동권리, 처리 거부 요청 권리, 개인정보의 자동 프로파일링 및 활용에 대한 결정 권리를 법령화한 것으로, 유럽 내 모든 기업 및 EU주민의 개인정보를 다루는 전세계 모든 기업에선 GDPR을 준수하여야 한다[2].

〈그림 2〉 EU - GDPR에 적용된 개인권리 범위



출처 : Pramod Chandrayan, 2018. "GDPR Compliance : It's Time To Protect Your Users Data". *Hackernoon*,

또한 세계 여러 도시들은 스마트도시를 도입에 앞서 개인정보를 보호하는 시도를 하고 있다. 몇 가지 사례를 보면, 뉴욕시는 개인정보에 대한 책임 있는 사용과 취급을 촉진하는 다양한 이니셔티브를 추구하고 있다. 예를 들면 도시 전체의 공공 공간에 IoT 장치를 배치하기 위한 프라이버시 표준을 설정하는 일련의 사물 인터넷 지침을 만들고, 공공부문의 의사결정 시스템 사용을 감시하기 위한 테스크 포스(TF)의 설립을 의무화하는 법도 도입했다.

네덜란드 암스테르담의 경우 암스테르담 경제 위원회는 특정 기관이 시민의 데이터를 보다 책임 있는 방식으로 사용하는 것을 돕기 위해서 6가지 원칙으로 이루어진 TADA 선언문[4]을 발표하기도 하고, 불필요한 개인 데이터 수집을 최소화하면서 사람들이 익명으로 지역 전자정부 서비스에 접근할 수 있도록 허용하는 DECODE 프로젝트를 시범운영 중이다[5]. DECODE 프로젝트는 범국가 차원에서, 산업체, 학교, 연구기관 간의 활발한 교류를 지원하는 유럽의 연구 혁신 프로그램인 Horizon 2020중 하나로, 온라인에서 생성된 개인 데이터의 저장, 관리 및 사용에 대해 개개인이 쉽고 효율적으로 제어할 수 있도록 하는 프로젝트이다[6]. 이 프로젝트는 바르셀로나에서도 동일하게 시범 운용되고 있으며, 시민들은 시의회의 e-참여 시스템을 사용할 때 익명성을 유지하면서 자신이 공유하고자

하는 정보를 선별적으로 선택할 수 있는 시스템을 개발했다[7].

(2) 목적

디지털 데이터는 제한 없이 복제될 수 있고 공유 될수록 더더욱 가치를 갖는다. 특히 개인데이터의 공유를 통해 얻게되는 경제적, 사회적 가치는 매우 높다. 개인의 데이터를 활용하기 위해선 공유, 집계 및 분석을 해야 하지만 이 과정엔 여러 위험이 따른다. 민감한 개인정보가 원치 않게 노출될 수도 있고, 개인정보를 악용하는 사례 또한 발생할 수도 있다. 예를 들면 페이스북과 구글, 아마존, 마이크로 소프트 등 대형 인터넷 기업들은 40여 가지의 개인정보를 수집하고 있으며 이중엔 종교, 정치성향, 인종 등 민감한 정보 또한 포함되어 문제가 되고 있다.

[표 1] 기술기업 수집/보유 개인정보(음영은 민감정보)

	Google	Facebook	Apple	Twitter	Amazon	Microsoft
이름	○	○	○		○	○
성별	○	○				○
전화번호	○	○	○	○	○	○
이메일	○	○	○	○	○	○
위치	○	○	○	○	○	○
결혼여부		○				
소득수준		○				
교육수준	○	○				
인종	○	○				
종교		○				
주소		○	○			○
얼굴인식		○				○
정치성향		○				
신용카드정보		○	○	○	○	○
신분증					○	
IP주소	○	○	○	○	○	○
이메일내용	○					○
연락처	○	○		○		○

	Google	Facebook	Apple	Twitter	Amazon	Microsoft
전화기록			○			○
채팅내용		○	○	○	○	○
일정관리	○	○				○
검색기록	○	○			○	○
시청동영상	○			○	○	○
웹 방문기록	○				○	○
브라우저 정보	○	○			○	○
업로드 비디오	○	○				○
업로드 사진	○	○		○	○	○
게시물 상태	○	○		○	○	
좋아요 기록		○		○		○
보관된 문서	○					○
사용기기	○	○	○	○	○	○

출처 : "The Data Big Tech Companies Have On You". *SECURITYBARON*, September

따라서 사용자가 공유한 데이터가 의도한 사람에게만 사용된다는 확신을 줄 수 있는 안전한 데이터 공유 환경을 구축하기 위해서 DECODE는 사람들이 데이터에 대해 세분화된 사용 조건을 설정할 수 있는 도구를 개발하고 있다. DECODE는 다음과 같은 3가지 주요 목표를 가지고 있다[8].

① 사용자에게 모든 권한을 부여하는 유연한 규칙 설정

현재 사람들이 자신의 데이터를 통제하고 공유 할 수 있는 기술적이고 법적인 규범이 부족한 상황이다. 사람들이 공공의 이익을 위해 개인 데이터를 공유하거나 익명화된 공개 데이터 사용을 가능토록 하기 위해, 속성기반 자격 증명과 분산원장 기술을 결합하여 개인 데이터에 특정 권한을 설정 할 수 있는 유연한 스마트 규칙 시스템을 구축한다.

② 개인데이터의 가치를 극대화하고 신뢰할 수 있는 플랫폼 개발

개인데이터들은 단독으로써 보다, 다른 개인데이터들과의 통합 및 대조를 통하여 가치가 실현된다. 이처럼 데이터를 수집하고 분석 할 수 있는 플랫폼구축을 위해서 DECODE는 GebiedOnline과 같은 플랫폼 협동조합과 Decidim(바르셀로나의 디지털 포털)과 협

력하여 DECODE참가자들이 특정 목적을 위해 개인 데이터를 공유 할 수 있도록 한다.

③ 시범도시에서의 커뮤니티 구축

DECODE는 스페인 바르셀로나와 네덜란드 암스테르담과 제휴되어 있으며, 두 도시 전체에 걸쳐 디지털 공동 공유 환경 구축을 위한 테스트를 진행한다. 이 같은 시범적용을 통해서 데이터 공유에 기반을 둔 경제 모델을 구축 한다.

(3) 기술 상세

DECODE는 분산신뢰(Decentralized trust), 분산 원장, 영-지식 증명(Zero- Knowledge Proof), 속성 기반 암호화(ABC), 암호화된 자격검증 (Cryptographically verifiable entitlement), 스마트 규칙언어, 검증가능하고 통제된 실행 환경을 기반으로 만들어 졌다. 각각의 내용은 다음과 같이 요약될 수 있다.

분산신뢰(Decentralized trust)는 시스템이 소수에 의해 통제 되어선 안 된다는 개념으로써 DECODE의 분산신뢰 시스템 모델 구축에 기반이 된다. 분산원장은 당사자가 그 안에 기록된 어떤 사실을 입증할 수 있도록 하는 기술로써 중앙 집중적인 관리자나 데이터 저장소가 존재하지 않아도 기능을 수행할 수 있다 이러한 특성 때문에 기록에 대한 복원력이 좋으며 조작에 대해서 안전하다. 영지식 증명(Zero-knowledge proof)은 거래 자체의 데이터를 공개적으로 기록할 필요 없이 거래의 체결 유무에 대한 증거만을 원장에 기록할 수 있도록 해준다[9]. 속성기반암호화 및 암호화된 자격검증은 참여자가 자격증명을 할 때 다른 식별 정보를 요구자에게 공개하지 않고 필요한 정보만 증명할 수 있는 메커니즘이다. 예를 들면 참가자는 출생정보, 주민등록번호 같은 민감한 정보를 노출 하지 않고 특정 도시의 거주 여부를 암호화된 방식으로 증명 할 수 있다. 따라서 이 메커니즘은 상호 작용의 진위성을 보증하며 개인의 익명성을 보장하는데 사용될 수 있다. 스마트 규칙언어는 사용자 친화성을 중점으로 시민들에게 하여금 이러한 자격을 표현하고 이해할 수 있는 수단을 제공한다.

DECODE의 구성개체(Entity)는 개인시민권자와 단체(공공기관, 기업, NGO 및 시민

사회단체)로써 구성되며 참가자, 운영자, 속성 검증자(Attribute Verifier), 노드호스트의 역할을 수행한다. 참가자는 DECODE의 최종 소비자를 말한다. 참가자는 데이터를 저장하고 애플리케이션과 상호작용을 하는 역할을 맡는다. 운영자는 실행되는 애플리케이션을 개발하고 유지관리 한다. 속성 검증자의 경우 특정 속성과 관련된 주장을 검증할 수 있는 역할이며, 주장에 대한 암호학적 증거를 제공한다. 노드호스트는 네트워크의 기본 인프라를 운영한다. 일반적으로 네트워크 내에서 노드를 호스팅하고 실행하며, 데이터 서비스 또는 온라인 지갑 서비스와 같은 특수한 서비스를 실행 할 수도 있다.

DECODE 내의 모든 상호 작용은 공개키 혹은 개인키로 암호화되어 계정에 연결된다. 공개키로 암호화 되어 전송된 데이터를 계정 관리자는 개인키를 이용하여 복호화하여 확인할 수 있다. 모든 응용 프로그램은 참여자가 접근/조작하려는 속성에 대해 투명해야 하며 이 속성 집합을 프로필이라고 한다. 속성은 애플리케이션에 의해서만 생성되며 애플리케이션과 1대1로만 생성된다. 예를 들면 각각의 애플리케이션에서 ‘이름’ 이라는 속성을 생성해도 생성된 속성은 별개의 속성 인스턴스를 갖는다. 특징을 중심으로 구조화 된 모델을 만들면 스마트 규칙과 같은 상위 구조를 개발에 용의하며, 암호화 및 개인정보보호에 대한 전문 지식이 없어도 개인정보를 제어하는 응용프로그램을 쉽게 제작할 수 있다. 때문에 DECODE 개발자는 친숙하고 편리한 환경에서 개발할 수 있다.

① 속성(Attribute)

개인정보 보호를 위하여 메타데이터를 속성에 첨부하는 개념적 모델을 구현하였다. 속성은 일반적으로 다음과 같은 구조로 구성된다.

『ATTRIBUTE=(SUBJECT PREDICATE OBJECT) 』

〈그림 3〉 JSON포맷의 일반적인 속성구조 예시

locality	:	[	"decode-account:543232",	"schema:addressLocality",	"Paris"]
~		~	~	~	~
ATTRIBUTE		SUBJECT	PREDICATE	OBJECT	

출처 : George Danezis and etc., "First Version of DECODE Architecture", DECODE project, Oct. 2017.

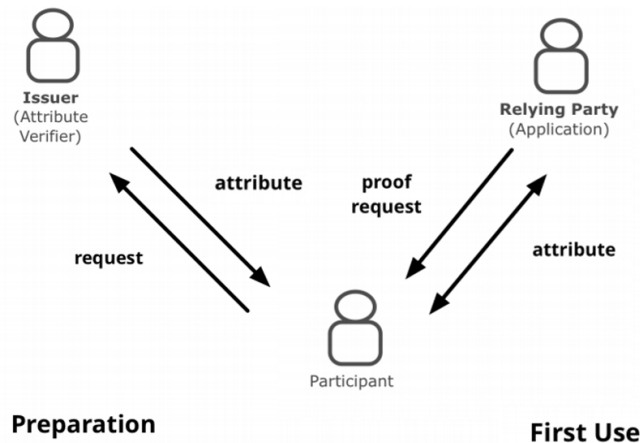
DECODE프로젝트에서 속성은 위와 같은 구조에서 출처(PROVENANCE)와 범위(SCOPE)가 추가되어 구성된다.

『 ATTRIBUTE=(SUBJECT PREDICATE OBJECT PROVENANCE SCOPE) 』

PROVENANCE는 속성의 출처와 속성이 생성된 애플리케이션에 대한 정보 등을 나타낸다. 위 예시에선 파리지 당국에서 발급한 계정의 개인키와 연관된 속성 기반 자격증명이 이에 해당된다. SCOPE는 속성의 소유자와 애플리케이션간에 체결된 정보의 유지 범위와 관련이 있다.

DECODE참가자들의 개인정보는 속성으로써 저장이 된다. 당국에서 발급한 여권이나 주민등록증과 같은 신원증명 문서를 DECODE계정에 등록하면, 등록된 데이터는 속성으로 분류되어 개인키 등의 암호화 방법을 통해 저장된다. 이 속성이 저장될 때 PROVENANCE와 URN에 대한 내용이 기록되기 때문에 기관을 통해 공식적 검증된 속성으로써 사용 될 수 있다. 예를 들어 온라인 주류 구매 애플리케이션을 사용하여 주류를 구매할 때, 필요한 정보는 18세 이상이라는 정보이다. DECODE환경에선 구체적인 생년월일을 요구하지 않고 18세 이상이라는 속성만으로 인증이 가능하다. 이와 같은 암호화된 속성을 기반으로 증명하는 속성기반 자격증명은 애플리케이션과의 상호작용에서 꼭 필요한 최소한의 정보만 제공할 수 있기 때문에 개인데이터를 보호할 수 있는 이점을 갖는다. 속성기반자격증명은 속성 검증자(Issuer), 참가자, 애플리케이션(Relying Party) 간의 상호작용을 통하여 외부에서 생성된 속성에 대해서도 신뢰 메커니즘을 구현할 수 있다.

〈그림 4〉 외부에서 생성된 속성의 검증과정



출처 : George Danezis and etc., "First Version of DECODE Architecture", DECODE project, Oct. 2017.

위의 그림에서 속성 검증자는 참가자가 요청한 속성을 PROVENANCE를 포함하여 개인키를 통해 암호화한 토큰형태로 제공한다. 참가자는 이 토큰을 애플리케이션에게 전달하여 속성을 증명하고, 애플리케이션은 속성의 PROVENANCE를 통해 공식적으로 검증된 속성이란 것을 확인 할 수 있다. 앞선 온라인 주류구매의 예시에 대입하여 보면, 참가자는 시당국에 18세 이상임을 증명할 수 있는 속성을 요청하고, 시당국은 시에서 발급한 PROVENANCE가 포함된 속성을 암호화 하여 참가자에게 전달한다. 참가자는 주류구매 애플리케이션을 이용하기 위해서 이 속성데이터를 애플리케이션에 전달하고, 애플리케이션은 전달받은 속성데이터를 통해서 참가자가 18세 이상임을 확인하고, PROVENANCE를 통해서 이 속성이 신뢰할 수 있는 기관에서 생성됐음을 확인한다.

② 자격(Entitlements)

자격은 데이터 소유자가 데이터 공개 및 사용에 대한 권한을 설정하는 것이다. DECODE에서 사용권한은 정책에 의해서 정의되어 있으며 암호화를 적용하여 구현한다. DECODE의 구조는 분산되어있기 때문에 자격정책을 관리하기 위해서 다음과 같은 문제를 고려해야한다. 첫 번째는, 데이터가 저장되는 곳이라면 어디서든 데이터를 공유하거나 복제하더라도 사용 권한 정책이 일관적이어야 한다. 두 번째로 고려해야 할 문제는 분산된

자격 정책이 데이터를 관리하는데 있어서 정책을 모으고 이해하는데 어려움을 준다는 점이다. 세 번째 문제는 자격정책의 공식적인 표현은 다양한 표현 모델을 가져야 한단 점이다. 마지막으로 기본 속성이 다른 속성으로 대체 될 때 자격을 어떻게 적용할지이다. 예를 들면 주소라는 속성이 이사로 인해 다른 주소로 바뀔 경우에 자격이 자동으로 새 주소에도 적용돼야 하는가와 같은 문제이다. DECODE는 이러한 문제를 스마트 규칙언어를 통하여 기준을 규정하는 방식으로 해결하였다. 특정 속성을 “민감한, 개인, 공용” 과 같은 그룹으로 분류하여 계층적 사용권한 시스템을 구축하며, 모든 사용자격을 각각의 속성에 세분화하여 지정하였다.

[표 2] DECODE에서의 데이터 접근 단계

접근 단계	설명
owner-only	비공개 설정
can-discover	데이터 항목에 속성에 대한 값이 있는 건 알 수 있지만, 어떤 값인지에 대해선 알 수 없다.
can-access	데이터 항목에 속성에 대한 값이 있는지 알 수 있고, 그 값 또한 알 수 있다.

대부분의 경우 참가자는 자격을 직접 생성하진 않는다. 참가자가 DECODE 애플리케이션과 상호작용할 때, 애플리케이션은 참가자에게 자격부여를 요청하고 참가자는 OAuth와 같은 인증방식을 통해서 권한 부여를 수락하는 방식으로 자격부여가 이뤄진다.

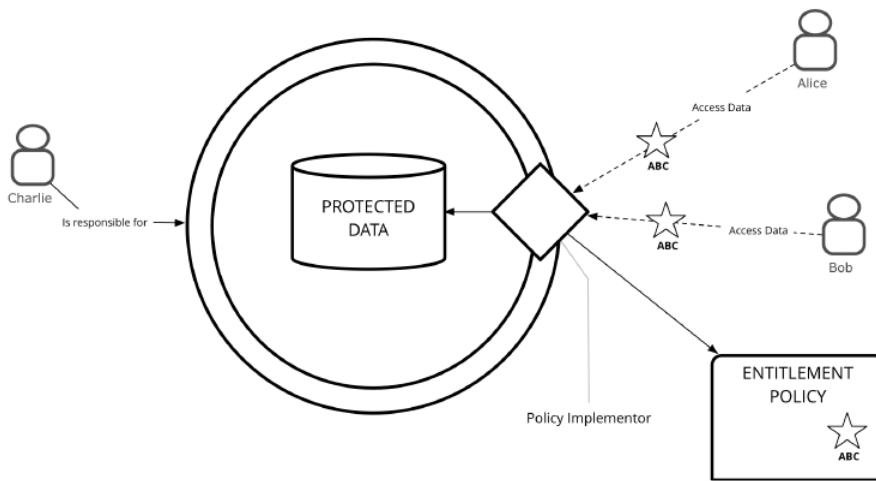
DECODE는 분산적인 시스템 모델을 구축하기 위해서 자격증명의 발행과 검증을 포함하여 데이터와 접근 제어 메커니즘이 모두 분산 될 수 있도록 하였다. DECODE는 분산원장기술과 속성 자격증명을 기반하며 현재는 중간단계로써 기존 시스템에 이 기술을 적용하여 시스템의 접근 제어 메커니즘을 개선하였다. 다음은 두 가지 예시를 통해서 DECODE와 기존 시스템의 통합과정을 설명한다.

(i) 속성기반 자격 증명(ABC)을 권한부여 메커니즘으로 사용

이 예시에선 중앙 엔티티에 모든 사람들의 데이터가 저장되고, 접근제어를 통해 이를 보호하는 전통적인 아키텍처를 활용한다. Json Web Token(JWT)규격 기반에서

DECODE는 승인정보시스템과 통합되어 사용될 수 있다. 이 경우 애플리케이션은 DECODE에 저장된 자격 정책을 참조하여 사용자가 제공한 속성 기반 자격증명을 수락하고 검증할 수 있다. 이처럼 DECODE를 통해서 암호화 보안시스템과 기존시스템을 쉽게 통합하여 사용할 수 있다.

〈그림 5〉 속성기반 자격 증명을 인증 메커니즘으로 사용 예



출처: George Danezis and etc., "First Version of DECODE Architecture", DECODE project, Oct. 2017.

(ii) 분산 기반 암호화를 접근 제어 메커니즘으로 사용

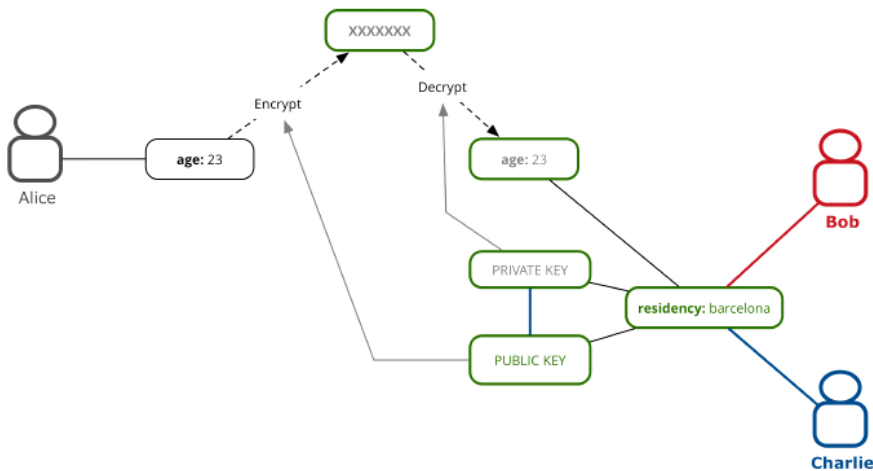
보다 분산적인 접근법으로 모델을 확장해 나가면서, DECODE는 데이터가 IPFS (Interplanetary File System)와 같은 대규모 분산 파일 공유시스템에서도 저장될 수 있도록 사용권한을 가진 모든 사용자의 공용키로 데이터를 암호화하였다.

아직 개발 중에 있지만, 분산기반암호화를 이용하면 사용자 그룹이 암호화된 데이터 소스에 대해 접근을 허용할 경우 발생하는 오버헤드(특정목표를 달성하기 위해 간접적으로 요구되는 시간, 메모리, 대역폭과 같은 컴퓨터 자원)를 상당히 줄일 수 있다.

이는 각각의 참가자가 고유한 키를 갖고 있는 것처럼 생각할 수 있으며, 암호화된 데이터는 어떤 키로도 복호화가 가능하다. 예를 들면 찰리는 거주중인 A아파트의 월별 에너지

사용량 데이터를 같은 A아파트의 주민들과 공유하려한다. 찰리는 속성기반암호화를 사용하여 이 데이터를 A아파트의 구성원이라는 속성의 공개키를 사용하여 암호화 할 수 있고 A아파트 구성원 이라면 자신이 가진 키를 이용하여 데이터를 복호화 할 수 있다. 이처럼 사용자는 제 3자의 도움 없이 데이터에 대한 접근을 직접 제어 할 수 있다.

〈그림 6〉 속성기반 암호화



출처 : George Danezis and etc., "First Version of DECODE Architecture", DECODE project, Oct. 2017.

속성기반암호화를 이용하면 사용자는 더 복잡한 접근조건 또한 설정 할 수 있다. 예를 들어 Alice는 18세 이상이며, 같은 아파트 거주민을 위한 파티초대장을 속성기반암호화를 통해 보내려 한다. Alice는 A아파트 구성원이라는 속성의 공개키로 초대장을 암호화 하고, 18세 이상이라는 속성을 갖는 사람만 복호화 할 수 있도록 한 초대장을 배포한다. 배포된 초대장은 A아파트 구성원중 개인지갑에 18세 이상의 속성을 갖는 사람만이 내용을 해독할 수 있다.

③ 분산원장(Distributed ledger)

DECODE내에서 원장은 무결성 및 가용성이라는 두 가지 특징을 갖으며, 속성기반자격 증명과 속성기반암호화와 함께 어떤 개인 데이터도 원장에 저장하지 않는 다는 핵심 원칙을 통해서 개인정보를 보호한다. 이를 위한 암호화 기술로는 영-지식증명(Zero-

knowledge proofs)이 적용된다. 이는 참여자가 거래를 수행하고 거래 내에서 사용된 데이터의 노출 없이 거래의 입증 가능한 증거만 기록 할 수 있게 한다. 이 과정을 통해 생성된 증명서엔 개인데이터가 없기 때문에 많은 수의 노드에도 안전하게 저장할 수 있다. 또 복수의 노드가 트랜잭션을 검증하고, 증명서에 대한 신뢰를 구축 할 수 있다.

DECODE는 데이터 스토리지, 검증 노드, 지갑, 원장 구현, 스마트 규칙엔진, 운영체제 아키텍처, 하드웨어 허브, 키 및 스마트카드, IoT 연결과 같은 항목으로 구성된다.

① 데이터 스토리지

DECODE 시스템 내에는 속성, 원장 트랜잭션(Ledger Transaction), IoT 데이터 스트림의 3가지 타입의 데이터가 저장된다.

속성은 속성 저장소에 IPFS와 같은 P2P프로토콜을 사용하여 분산 저장되며, 사용자가 원하는 경우 암호화된 형식으로 사용자의 지갑(wallet)에 로컬 저장도 가능하다. 분산 저장의 장점은 백업과 복구가 간편하다는 점, 여러 기기에서 동일한 접근이 가능하다는 점, 지갑자체에 저장기능이 필요 없다는 점이다.

원장 트랜잭션(Ledger transaction)은 원장 노드 시스템에 저장되며, 원장에는 개인정보를 제외한 암호화된 데이터만 저장된다.

IoT 데이터 스트림에서 IoT데이터는 많은 용량의 시계열 데이터를 포함할 가능성이 있어 특수한 경우의 데이터로 정의한다. AWS IoT와 같은 기존의 데이터 저장소에 대한 사용 권한 및 접근 제어 메커니즘으로 DECODE를 통합 할 수 있는 도구를 제공함으로써 AWS IOT와 같은 기존 IoT 데이터 저장소 또한 활용할 수 있다.

② 검증 노드

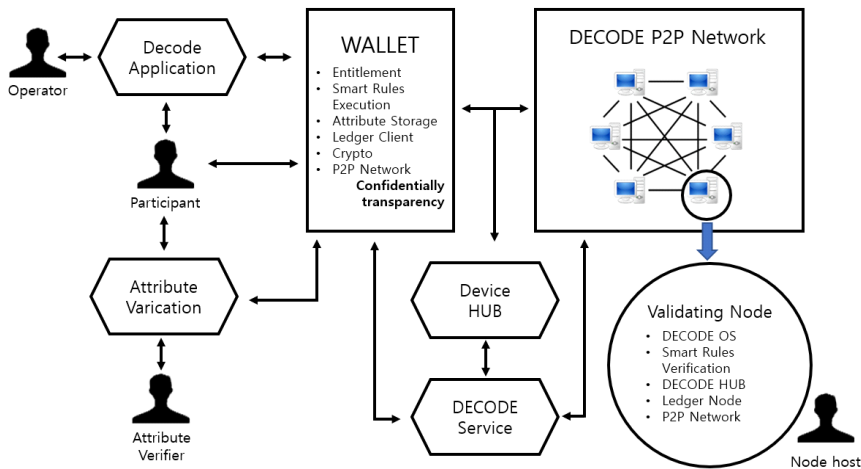
네트워크의 신뢰도와 복원력은 검증 노드의 분산 네트워크를 통해 제공된다. DECODE의 주요 특징은 계약의 실행을 논리 증명과 분리한 점이다. DECODE OS에 기반을 두어 검증노드를 구성하였고 각 노드에는 DECODE네트워크에 참여하는데 필요한 노드와 기타 라이브러리 및 소프트웨어가 포함된다. DECODE는 모든 참가자가 검증노드를 호스트하지 않아도 DECODE Wallet을 통하여 참여할 수 있다. 반면 검증노드를 유지하기 위해선

하드웨어 구입비용, 전기료 같은 비용이 발생한다. 다른 블록체인 플랫폼에선 검증 노드에 금전적인 보상을 제공하지만, DECODE의 경우 공익을 위해 작동할 수 있는 공유 모델을 구축하려 한다.

③ 지갑

지갑은 사용자가 DECODE와 상호작용하기 위해 필요한 최소의 구성요소이다. 모든 참가자는 자신의 지갑을 갖고 있으며, 각각의 지갑은 암호화 자료(개인키), 개인키에 연결된 속성기반 자격증명을 안전하게 저장한다. 스마트 규칙을 통해 DECODE 트랜잭션을 실행하고, 결과 확인을 위해 원장(Ledger)을 제출하고, 참가자의 속성 저장 및 암호화 기능을 수행하며, 애플리케이션과의 상호작용을 기반으로 자격정책을 정의 및 게시할 수 있다. 또한 참가자에게 속성, 자격 및 애플리케이션을 관리할 수 있는 그래픽 인터페이스와 같은 기능을 제공한다.

〈그림 7〉 DECODE 개요



출처 : George Danezis and etc., "First Version of DECODE Architecture", DECODE project, Oct. 2017.

④ 원장 구현

DECODE는 개인정보보호와 확장성을 고려하여 설계한 체인스페이스(<http://chainspace.io>)를 통하여 원장을 구현하였다. 체인스페이스에선 트랜잭션 실행과 검증을

구분하는 확장성이 뛰어난 BFT(Byzantine Fault Tolerance)합의 알고리즘방식이 적용되었다.

⑤ 스마트 규칙 엔진 (Smart rules engine)

DECODE에서 스마트 규칙은 유한 온톨로지로 참조되고 분산 컴퓨팅 클러스터에 의해 실행되는 의미론적 모델이다.

⑥ 운영체제 아키텍처

DECODE OS는 DECODE 프로젝트를 위해 설계, 개발 및 배치된 모든 소프트웨어를 실행하는 기본 운영체제이다. 이 운영체제는 데비안 GNU+리눅스에 기반을 두고 있다.

⑦ 하드웨어 허브

하드웨어 허브는 지갑 또는 유효한 노드중 하나를 실행할 수 있는 기본 컴퓨팅 인프라를 말한다. 예를 들어 물리적 서버 또는 PC, AWS, Google Cloud와 같은 가상머신, 제한된 암호와 코드를 실행하는 스마트카드, 모바일기기 등이 여기에 해당된다.

⑦ 키 및 스마트카드

기존의 PIN코드 나 비밀번호 같은 인증요인은 보안에 취약한 문제가 있다. 해킹을 방지하기 위하여 2차,3차 인증수단으로써 Nitrokey, U2F zero, FST-01과 같은 장치를 사용한다.

⑧ IoT연결

DECODE프로젝트의 주요한 사용사례중 하나로써 개인 IoT데이터를 안전하게 관리하는 도구를 제공하는 것이다.

(4) 현황 및 구축 예

DECODE프로젝트는 유럽의 글로벌 경쟁력을 확보하기 위한 연구 및 혁신 프로그램으로 유럽집행위원회의 재정지원을 통해 진행 되고 있다. 2017년부터 시범운영이 적용되었고, 2019년 까지 암스테르담과 바르셀로나에서 운영 될 예정이다. 프로젝트 참여자는 자

신의 IoT장치와 그들의 데이터를 관리하고 제어 할 수 있다. 현재 프로젝트가 진행 상황과 참여자에 대한 구체적인 내용이 홈페이지를 통하여 추가로 제공 될 예정이다.

시범 운영 중인 바르셀로나의 경우 DECODE를 이용한 탄원서 및 온라인 투표와 같은 정책 참여 시스템과 IoT데이터 활용 등을 중점적으로 테스트한다.

① IoT 데이터 활용

대규모 광장과 같이 소음공해가 큰 이슈가 되는 지역은 소음을 측정하는 사물인터넷(IoT) 기기를 설치하는 비율이 다른 지역에 비해 높다. 수집한 소음공해 데이터와 지역 거주민의 건강 데이터(수면 패턴)를 공유하면, 두 데이터의 상관관계를 분석 할 수 있다. 구체적으로 지역 거주민들은 웨어러블 IoT장치를 착용하고 정기적으로 수면, 혈압, 심박수와 같은 개인 건강 데이터를 측정하고 저장한다. 수집한 개인 데이터와 도시의 환경오염도(소음포함) 공공 데이터를 결합하면 시민의 생활패턴과 소음공해에 대한 상관관계에 대한 정확한 분석이 가능하다.

② 정책참여 예시

시민들이 시의 의사결정에 좀 더 직접적인 참여할 수 있는 방법으로써, 시민들이 청원서에 서명함으로써 제안된 정책을 지지할 수 있도록 지원하는 사례이다. 현 단계에서 실제 투표에 적용하는 것은 법적 제약이 있어 DECODE의 범위를 벗어난 부분이 있지만, 암호화 사용의 한 사례로써 의미를 찾을 수 있다. 투표자는 속성기반 자격증명을 통해서 투표자격을 증명하며, 익명성을 보장받으며 투표에 참여할 수 있다. 투표 데이터는 블록에 등록되기 때문에 위변조가 불가능하며, 투표자 또한 이중투표가 불가능하다. 또한 DECODE를 통해서 단순한 지지율뿐만 아니라 데이터 공개 범위설정을 통하여 연령범위와 같은 집계통계데이터 또한 수집할 수 있다.

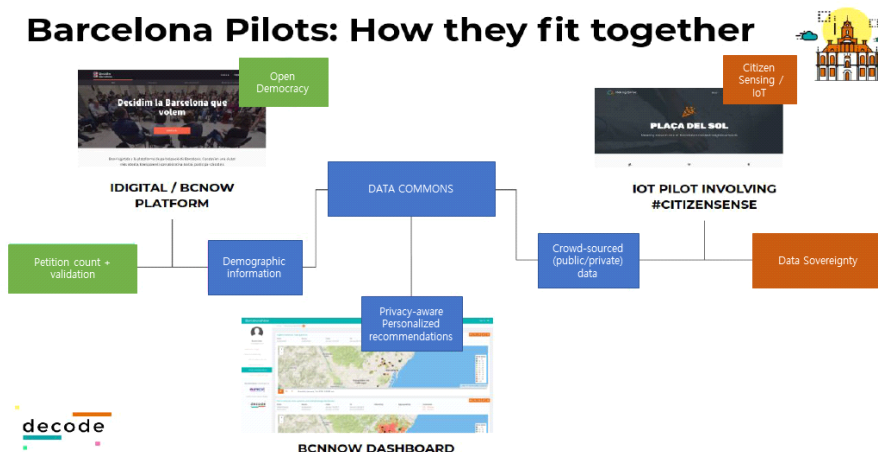
③ 바르셀로나의 BCNNow 와 CityOS

• BCNNow

데이터에 대한 통제권과 소유권을 시민들에게 돌려주기 위한 목표를 달성하기 위해선 실제 현장에서 테스트 할 수 있는 기반시설의 마련이 필요하다.

시의회가 참여하여 암호화, 개인정보 보호 기술, 데이터 고유 라이선스 및 데이터 관리의 사회적 모델로 생성된 자산을 이용하여 현장 테스트를 할 수 있지만, 효율적인 테스트를 위해 시민소유의 기반시설과 공공 행정간 통합된 하이브리드 형태가 필요하였다. 따라서 DECODE참여 도시 중 하나인 바르셀로나에선 BCNNow라는 인프라를 구축하였다. 바르셀로나에선 이 인프라를 통해서 DECODE참여자를 통해 생성된 데이터가 시민과 공공 심의를 충족시키는 장소가 되는 것을 목표로 한다. BCNNow는 바르셀로나의 DCODE사용자가 시각화를 생성하고 공개적으로 사용가능한 데이터, 개인소유의 데이터 및 군중 출처의 데이터를 기반으로 개인화 된 보기 및 정보를 얻는 데 사용되는 도구이다. 바르셀로나에서 이 인프라는 바르셀로나 시의회가 주도하여 데이터 관리에 대한 새로운 접근 방식부터 개발된다. 바르셀로나는 디지털 전환을 위한 윤리적인 표준으로써, 설계 및 데이터 윤리에 의한 데이터주권, 프라이버시 및 보안을 의무화 하는 새로운 데이터 지침을 발행하였다. 이는 이 기준이 정부의 계약 및 조달 과정에 조항으로 통합된다는 것을 의미한다. 시는 또한 데이터를 공공 인프라, 물, 전기, 도로 및 공기와 같은 새로운 메타 유틸리티로 정의한다. 데이터는 시민이 통제해야 하며 공공 가치를 제공하는 미래데이터 기반 서비스를 구축할 수 있는 모든 시민, 지역 기업 및 기타조직에 대한 적절한 개인정보 보호, 데이터 사용권한 및 윤리표준을 사용하여 액세스 할 수 있어야 한다.[10]

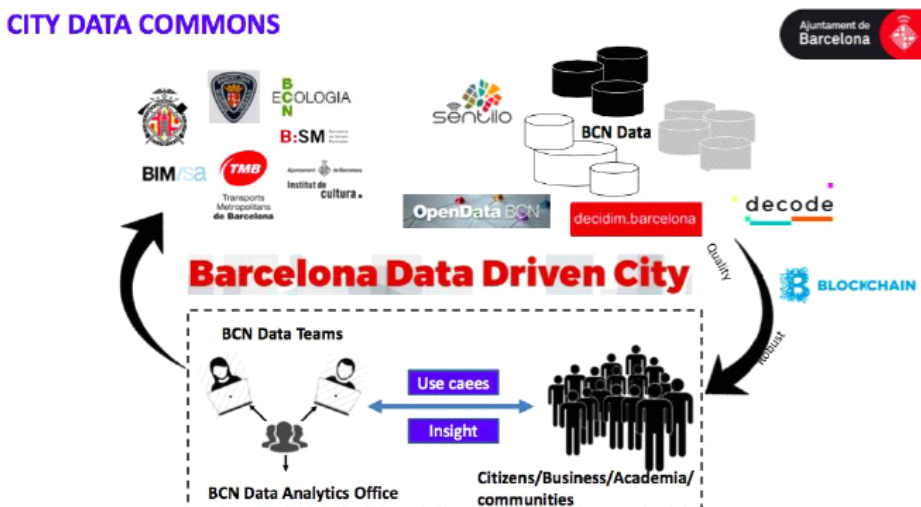
〈그림 8〉 Schema of relations between BCN Pilots and BCN Now



출처 : Oleguer Sagarra and etc., "CityOS connection", DECODE project, Jul. 2018.

BCNNow는 바르셀로나 DECODE사용자가 향후 제공 가능한 도시 공공 데이터를 수집하고 정의하기 위해 데이터 권한을 시각화 할 수 있는 플랫폼이다. 이 결과물은 바르셀로나 시의회가 WP4의 프레임워크에서 개발된 다양한 DECODE 노드를 호스팅 할 수 있도록 개발 중인 작업과 통합된다. 이 작업의 비전은 DECODE를 바르셀로나 도시데이터 인프라의 중요한 요소로 만드는 것이며, 이를 통해 도시가 분산되고 개인정보를 보호하며 권리를 유지하는 공공 도시 데이터를 개발할 수 있도록 한다.

〈그림 9〉 Barcelona Data Commons Framework



출처 : Oleguer Sagarra and etc., "CityOS connection", DECODE project, Jul. 2018.

DECODE는 시민들로 하여금 자신의 데이터를 통제 할 수 있도록 하고, 시청과 공유하고자 하는 데이터를 '공공재'로 결정할 수 있도록 한다. 동시에 DECODE자격 부여와 데이터 통신 사용허가 프레임워크를 통해 소유권을 유지하고 사용조건을 설정할 수 있도록 한다. 이와 같은 방식으로 바르셀로나는 도시 데이터로부터 창출되는 부가가치를 더 잘 분배할 수 있다.

• CityOS

CityOS는 바르셀로나가 교통, 에너지 효율, 인구, 기상데이터 등과 같은 도시기능과

관련된 모든 데이터를 실시간으로 파악하여 중요한 이벤트를 포함한 모든 종류의 상황을 예측하는 빅데이터 기술 플랫폼이다. CityOS는 BCNNow 및 DECODE와 같은 플랫폼과 통합되어 도시 데이터를 수집하고 정보 접근관리 운영체제를 제어하는 시티 플랫폼으로써의 기능을 한다. CityOS의 주요 목표는 다음과 같다.

- ① 도시와 관련된 데이터(시립/비시립 데이터베이스, 센서 시스템, 소셜 네트워크 등)를 상호 연관시키고 통합한다.
- ② 시에서 제공하는 다양한 서비스가 수평적으로 관리(부서 간)되고 수직으로 관리(글로벌 감독센터를 통해) 될 수 있도록 한다.
- ③ 도시서비스의 효율성과 효능을 높이고 예측가능성을 향상시켜 비상사태를 예측한다.
- ④ 의사 결정에 실시간 지원, 시민의 삶의질과 데이터 통제권을 향상시킨다.
- ⑤ 단일 창 액세스 및 내부데이터 관리를 사용하는 오픈소스 빅 데이터기술 기반 인프라

현재 CityOS는 개발 중이며 대부분의 서비스가 최종단계에 접어들고 있다.

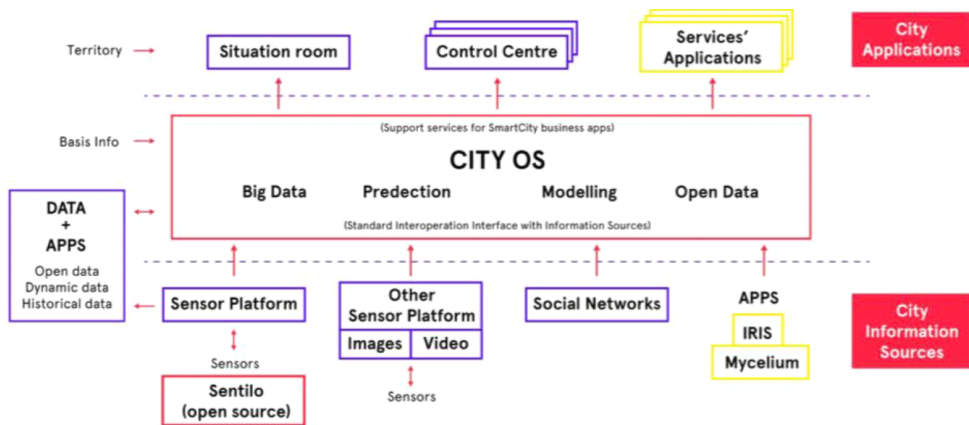
〈그림 10〉 CityOS, a graphical overview of its functions



BCNNow는 CityOS와 공개 API를 통해 연결된다. 이용 할 수 있는 CityOS데이터와 관련하여, 시의회 기관이 직접 관리하지는 않지만, 시의회 데이터와 외부 기관의 데이터를

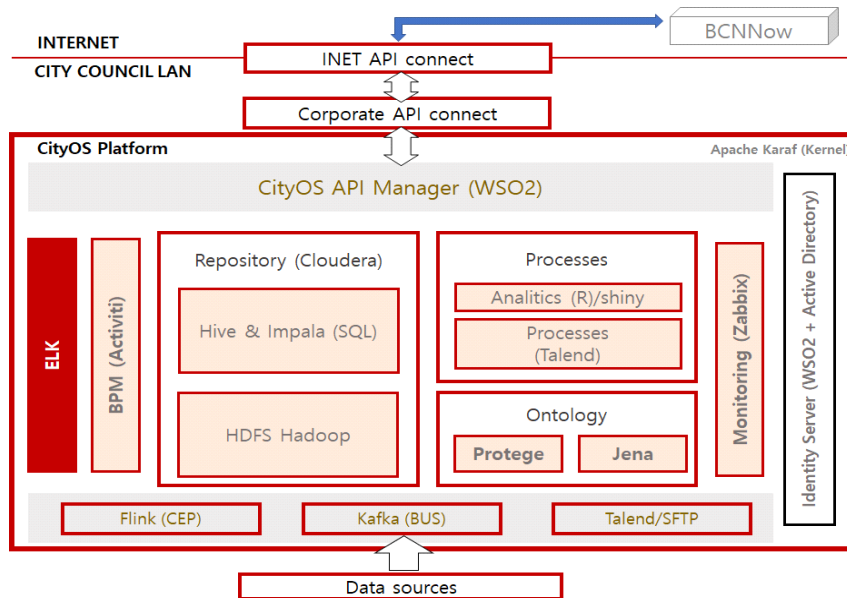
사용하여 도시 관리 정보를 확인 할 수 있다. 이용 가능한 데이터는 계약의 판결, 보조금, 도시행동계획사업, 지구, Open Data BCN (여러 공공 데이터셋으로 도시의 개방형 데이터 인프라), Sentilo(도시의 센서 플랫폼), IRIS(시민의 요구와 요구사항을 저장하는 플랫폼), ASIA(모든 공공활동의 안전을 저장하고 도시건물의 사용을 기록하는 플랫폼), Decidim Barcelona, 외부기관(기타 공용센서 플랫폼, 비디오 플랫폼) 등이 있다.

〈그림 11〉 바르셀로나 도시플랫폼 통합

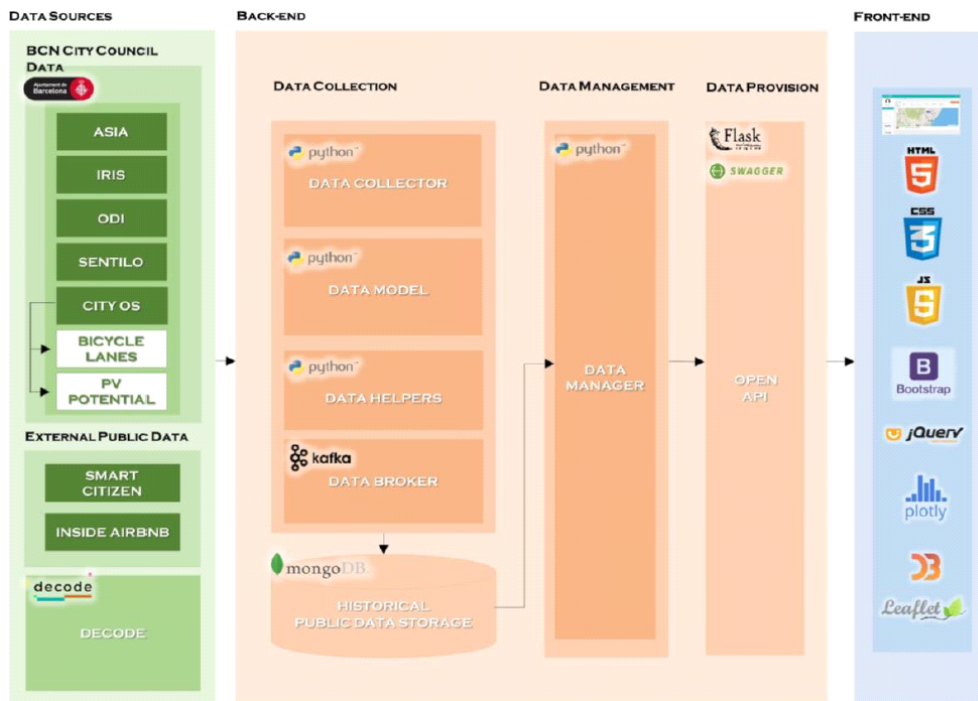


BCNNow인프라로 CityOS플랫폼에 연결하기 위해, CityOS는 WFS (Web Feature Service)게시자에게 접근 할 수 있는 임시 API를 제공한다. BCNNow 인프라에서 CityOS플랫폼으로 요청이 수행되면 INET API Connect는 청원을 받아 이를 기업 API Connect로 재전송한 다음 같은 방법으로 CityOS API Manager로 전송한다.

〈그림 12〉 CityOS인프라 및 BCNNow와 연결



〈그림 13〉 Workflow for the connection of CityOS with BCNNOW infra.



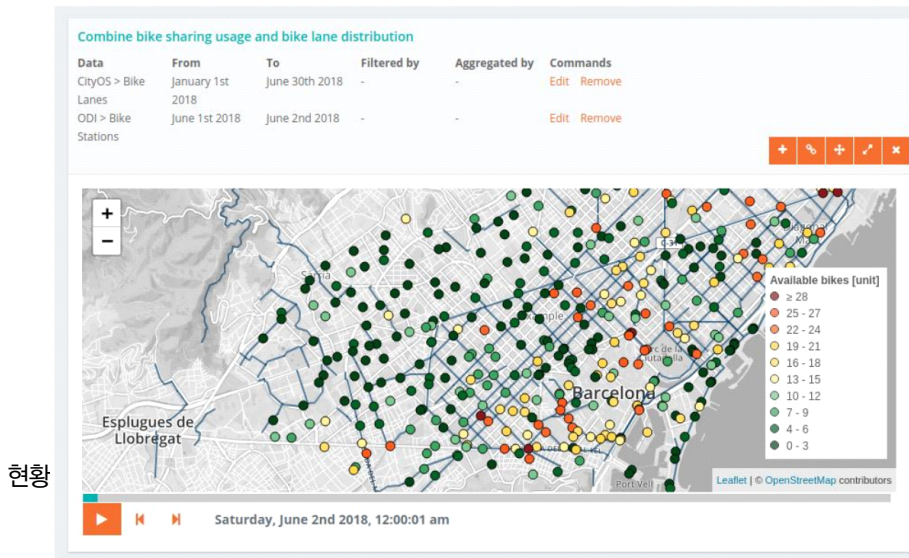
실제 데이터 표기를 보면 CityOS Data Collector는 “Institut Municipal d’Informatica (IM) del Ajuntament de Barcelona”에서 제공하는 임시 API를 통해 “cityos:ptt_carril_bici”, “city:potencial_fotovoltaiac”와 같이 특성을 갖는 WFS(Web Feature Service)형태로 BCNNow 인프라와 연결된다. WFS “cityos:ptt_carril_bici”는 자전거 전용 도로와 연관되며, WFS “city: potencial_fotovoltaiac”는 바르셀로나 지방의 태양광 발전과 관련이 있다.

〈그림 14〉 BCNNow에 표기되는 WFS “cityos:ptt_carril_bici”



위 그림은 BCNNow의 데시보드에 WFS “cityos:ptt_carril_bici”을 이용하여 바르셀로나 도시 지도위에 자전거 도로를 표기한 그림이다. 이 같은 시각화 된 데이터를 바탕으로 바르셀로나의 자전거 대여 사업인 Bicing(Bicycle Borrowing in Barcelona)에서 운용중인 자전거에 대한 상태 정보를 아래와 같이 확인 할 수 있다.

〈그림 15〉 BCNNow에 표기되는 Bicing 자전거

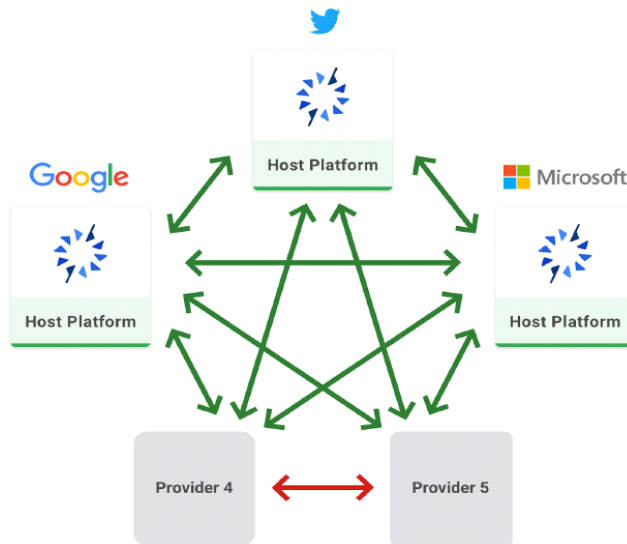


(5) 결론

DECODE 프로젝트는 최신 암호화 기술을 도입하고, 분산원장의 개념을 발전시켜 시민으로 하여금 자신의 데이터에 접근할 수 있는 사람 혹은 기관을 설정할 수 있을 뿐 아니라 무슨 목적으로 어떻게 사용하는지도 모니터링 할 수 있는 플랫폼을 개발했다. DECODE 프로젝트는 공공 데이터 구축이라는 목적으로 이러한 기술을 도입해 개인 데이터의 중앙집중화, 데이터 권한 남용이라는 기존의 개인 데이터에 대한 잘못된 관리에서 벗어나 공공의 이익을 위한 안전하고 자발적인 개인 정보 공유에 한발 다가섰다. 현재 진행형인 DECODE프로젝트는 백서(White paper) 및 진행과정이 웹사이트를 통해 공개되며, 트위터([https://twitter.com/ decodeproject](https://twitter.com/decodeproject))를 통해서도 확인 할 수 있다.

2 》 MS, 구글, 트위터, 페이스북 데이터 전송 프로젝트 (DTP)

〈그림 16〉 데이터 공유



(1) 배경 및 문제점

2017년 7월20일 미국의 거대 IT 기업인 페이스북, 구글, 마이크로소프트, 트위터는 이용자가 자신의 데이터를 쉽게 다른 인터넷 서비스로 옮길 수 있는 데이터 전송 플랫폼(DTP: Data Transfer Project) 개발에 합의했다[11].

최근 페이스북은 이용자의 친구관계, 팔로잉 관계 등 이용자 네트워크 구조에 대한 정보를 제공하는 소셜 그래프 API의 루프홀(loophole:빠져나갈구멍)을 악용한 캠브리지 애널리티카 스캔들을 통해 5천만명에 이르는 이용자 정보가 외부로 유출 되었고[12], 구글 또한 자사와 협력하는 수백만명의 외부 소프트웨어 개발자에게 지메일 사용자의 이메일을 열람할 수 있도록 권한을 부여했다는 논란에 휩싸였다[13].

이에 따라 미국 정부는 실리콘 밸리 기업들이 이용자의 수많은 데이터를 손쉽게 수집,

공유, 유출하는 것에 대한 규제를 강화하는 움직임을 보이고 있으며, 현재 세계 각국에서 해당 기업에 대한 조사 및 손해배상이 잇따르고 있다.

이런 배경에서 실리콘밸리 IT기업들의 데이터 전송 플랫폼 개발은 개인 정보에 대한 주체가 개인에게 있으며, 선택권이 없었던 이용자들에게 새로운 서비스를 선택할 기회를 제공하고 기업간에는 혁신과 경쟁을 가능하게 한다는데 큰 의미가 있다.

DTP는 오픈 소스 플랫폼으로써 누구든지 깃허브(Github)에서 소스코드를 다운 받아 개발에 참여할 수 있으며, 자신이 서비스하는 플랫폼에 적용할 수 있다[14].

(2) 목적

DTP를 사용하면 사용자는 특정 인터넷 서비스를 탈퇴하더라도 그 서비스에서 내가 생성한 데이터(사진, 영상, 음악 재생 리스트, 친구목록 등)를 다른 플랫폼으로 쉽게 옮길 수 있게 된다. DTP 백서에는 여러 유스케이스를 통해서 DTP 플랫폼을 통해 사용자가 할 수 있는 여러 예시를 다음과 같이 설명한다[15].

- ① 사용자는 인터넷에서 발견한 멋진 포토북 형식의 사진 인쇄 서비스를 사용하고 싶지만 모든 사진이 SNS에 저장되어있어서 다운 받고 업로드 하는 과정이 불편했다. DTP를 이용하면 불편한 다운로드/업로드 없이 사진 인쇄 서비스를 제공하는 웹 사이트 또는 앱을 방문하고 기존 SNS에서 사진 가져오기 기능을 통해 사진첩 서비스로 직접 사진을 이전할 수 있게 된다.
- ② 사용자는 자신이 이용하는 음악 서비스의 변경된 개인정보 보호 정책에 동의하지 않아 탈퇴를 했고 서비스는 즉시 중지 되었다. 다른 음악 서비스에 가입은 하지 않았지만, 기존에 만들어놨던 음악 재생 목록은 가져오고 싶다. DTP 플랫폼을 사용하면 기존 서비스 업체의 내보내기 기능을 통해 재생 목록 사본을 클라우드에 저장할 수 있고, 추후 새로운 음악 서비스를 결정한 후 재생 목록을 새 제공자에게 가져올 수 있다.
- ③ 회사는 곧 사업이 중단되는 기존 공급자로부터 자신의 데이터를 가져오려는 고객의

요청을 받고 있으나, 기존 공급자는 고객 데이터 전송에 대한 기술적인 지원이 어렵다고 한다. 이때 회사는 기존 공급자의 API용 DTP Adapter를 개발하여 사용자로 하여금 회사의 서비스로 데이터를 전송할 수 있게 하며, 동일한 데이터 유형을 처리하는 다른 회사에도 도움을 줄 수 있다.

- ④ 한 재택 근무자는 인터넷이 느린 환경에서 새 건축 프로젝트에 대한 도면 및 그래픽 작업을 설계담당자와 함께 진행했다. 프로젝트가 끝나면 둘 다 모든 파일을 개인 공유 스토리지 시스템에서 고객의 클라우드 스토리지로 전송하려고 하지만, 낮은 인터넷 대역폭 때문에 시간도 많이 걸리고 요금도 비싸서 걱정이다. 이럴때 클라우드 스토리지 DTP 유저 인터페이스(UI)에 접속하여 수백개의 대용량 파일을 대역폭을 그대로 유지하면서 직접 이동시킬 수 있다.
- ⑤ 슈퍼마켓 협회는 고객으로 하여금 자신의 구매 기록을 한 매장에서 다른 매장으로 옮길 수 있게 하고 싶습니다. 이를 통해 매장 간에 구매 습관을 기반으로 한 쿠폰을 발행하고 싶습니다. 협회는 DTP의 산업 특화 호스트 플랫폼을 통해 이 서비스를 제공할 수 있다.

(3) 기술 상세

1) Architecture

DTP 플랫폼은 서비스 공급자가 표준을 채택하고 사용하기 쉽도록 설계되었으며, 다음과 같은 제약조건을 고려하였다.

① 기존 표준 사용

OAuth 및 REST와 같은 기존 표준을 지원함으로써 DTP 서비스를 이용하는데 필요한 기본 작업이 최소화 된다. 새로운 표준이 개발되는 경우 내부 검토를 통해서 채택된 후 DTP 플랫폼에 통합된다.

② 서비스 추가에 필요한 작업 최소화

DTP는 타 서비스 공급자가 자신의 핵심 인프라를 변경하지 않고도 DTP서비스를 추가

할 수 있도록 설계되었다.

서비스 공급자는 어댑터를 구축하고 기존 API 및 인증매커니즘을 활용하여 가져오기 및 내보내기 기능을 사용할 수 있다.

③ 표준 배포 지원

DTP는 누구나 어디서든 서비스를 호스팅할 수 있도록 유연하게 설계되었다.

2) 시스템 구성요소(System components)

DTP는 데이터 전송방법에 대한 표준 형식인 데이터모델, 고유한 데이터 및 인증형식을 시스템에서 사용할 수 있는 형식으로 변환하는 어댑터, 작업관리 라이브러리와 같은 3가지 주요 요소들로 구성되어 있다.

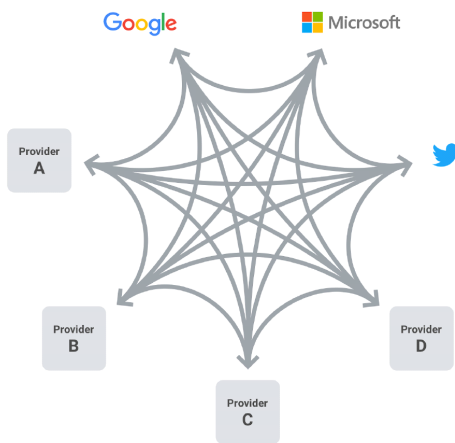
① 데이터모델 : 파일 전송을 위한 데이터 모델은 2파트로 구성된다. (i)파일유형 (ii) 수신 받는 공급자 측에서 데이터를 정확하게 가져 오는데 필요한 추가 메타데이터.

- 예를 들어 사진파일의 경우 파일형식은 JPEG와 같은 표준형식이고, 메타데이터에는 제목, 설명, 앨범 등의 정보가 포함된다.
- 데이터 모델은 일반적으로 업종별로 함께 클러스터링 되어 있으며 이를 버티컬 (Vertical)이라고 한다.
- 공급자는 하나 이상의 버티컬에 데이터를 보유할 수 있으며, 예를 들어 버티컬은 사진, 이메일, 연락처 또는 음악과 같다.
- 각 버티컬에는 관련 파일 유형을 완벽하게 전송할 수 있는 자체 데이터 모델 집합이 있다. 예를 들어, 음악 버티컬은 음악, 재생 목록 및 뮤직비디오 데이터 모델을 가질 수 있다.
- 이상적으로 버티컬은 잘 정의되고 광범위하게 적용된 소수의 데이터 모델을 보유하게 된다.
- DTP의 한 가지 목표는 어떤 조직으로 하여금 시스템에서 공통 데이터 모델을 사용하도록 권장하는 것이며, 이 공통 데이터 모델의 사용은 시스템을 처음 설

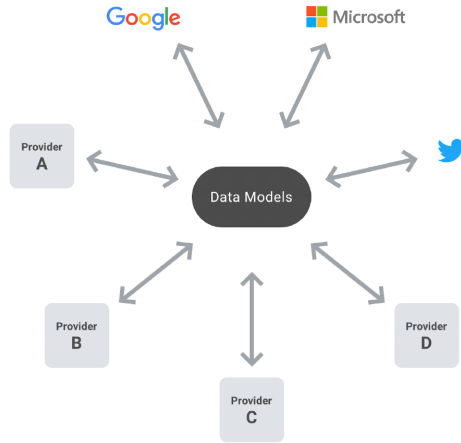
제하거나 업데이트를 제공시 데이터를 주고 받는 경우에 발생한다.

- 공통 데이터 모델을 사용하면 회사가 API를 유지 관리하고 업데이트해야 하는 필요성이 크게 줄어든다.

〈그림 17〉 DTP가 없는 환경의 데이터 교환
각 공급자는 데이터 교환을 위해 다른 모든 공급자의 전용 API를 각각 개발하고 유지관리 해야 함



〈그림 18〉 DTP를 이용한 데이터 교환
각 공급자는 표준 형식을 기반으로 하는 DTP 데이터 모델기반 API 하나만 개발하고 유지 관리함



출처 : Google, "Data Transfer Project Overview and Fundamentals", Jul. 2018

- ② 어댑터 : 어댑터는 (i)데이터 어댑터와 (ii)인증 어댑터로 나뉜다. 이 두가지 어댑터는 공급자의 핵심 인프라 외부에 존재하며 공급자가 작성할 수도 있고 공급자와 데이터 전송을 원하는 제3자가 작성할 수도 있다.

• 데이터 어댑터(Data Adapters)

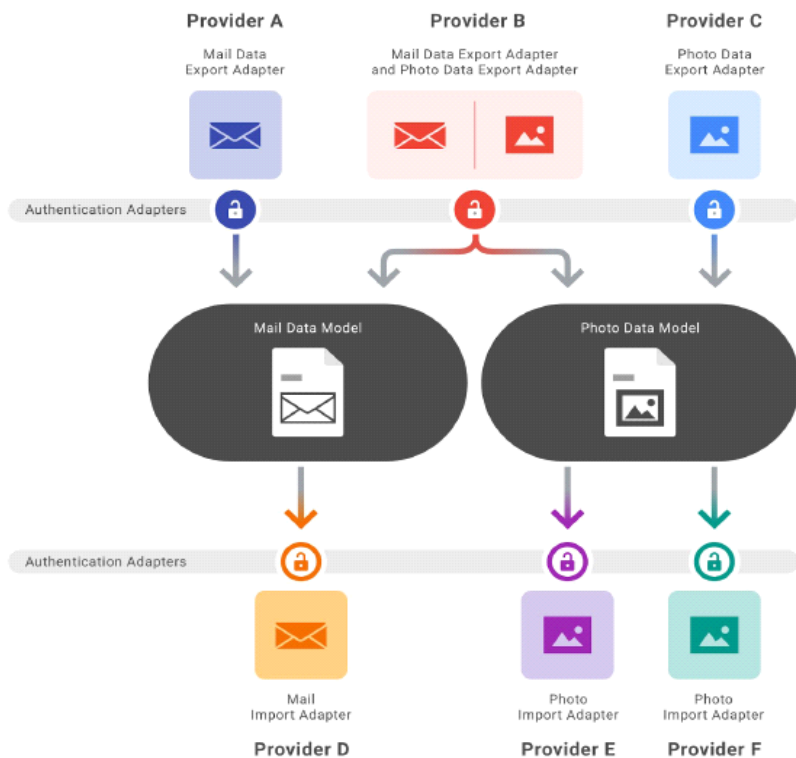
- 데이터 어댑터는 특정 제공자의 API를 DTP에서 사용되는 데이터 모델로 변환하며, 제공자의 API를 데이터 모델로 변환하는 내보내기 도구와 데이터 모델에서 제공자의 API로 변환하는 가져오기 도구로 구성되어 한 쌍으로 제공된다.

• 인증 어댑터(Authentication Adapters)

- 인증 어댑터는 데이터를 다른 외부 공급자로 전송하기 전에 사용자가 계정을

인증 할 수 있도록 해준다. OAuth 방식은 대부분의 공급자가 선택할 수 있지만, DTP는 인증 유형에 상관없이 소프트웨어 프로그램에 사용 가능하도록 설계되었다.

〈그림 19〉 데이터 어댑터와 인증 어댑터 사용 예시



출처: Google, "Data Transfer Project Overview and Fundamentals", Jul. 2018

③ 작업관리

작업관리 라이브러리는 두 가지 관련 어댑터간의 호출, 보안 데이터 스토리지, 반복실행 로직, 속도 제한, 페이지 배치 관리, 오류 처리 및 개별알림과 같은 백그라운드 작업을 처리한다.

DTP는 두 공급자 간에 데이터 전송을 위해 어댑터를 활용하는 방법에 대한 참조 구현으로써 작업 관리 라이브러리 모음을 개발했다.

또한 공급자는 DTP의 데이터 모델과 어댑터를 사용하는 작업관리 라이브러리를 자체 개발할 수도 있다.

작업관리 라이브러리는 일반 클라우드 인터페이스의 최상위단에 구현되어있어서 호스트 플랫폼은 로컬 시스템, 사내망 시스템, 클라우드 플랫폼 등 어떤 환경에서도 실행이 가능하다.

클라우드 인터페이스는 현존하는 모든 밴더의 클라우드 플랫폼에서 구현될 수 있도록 높은 수준의 추상화만을 사용하는 것을 목표로 한다.

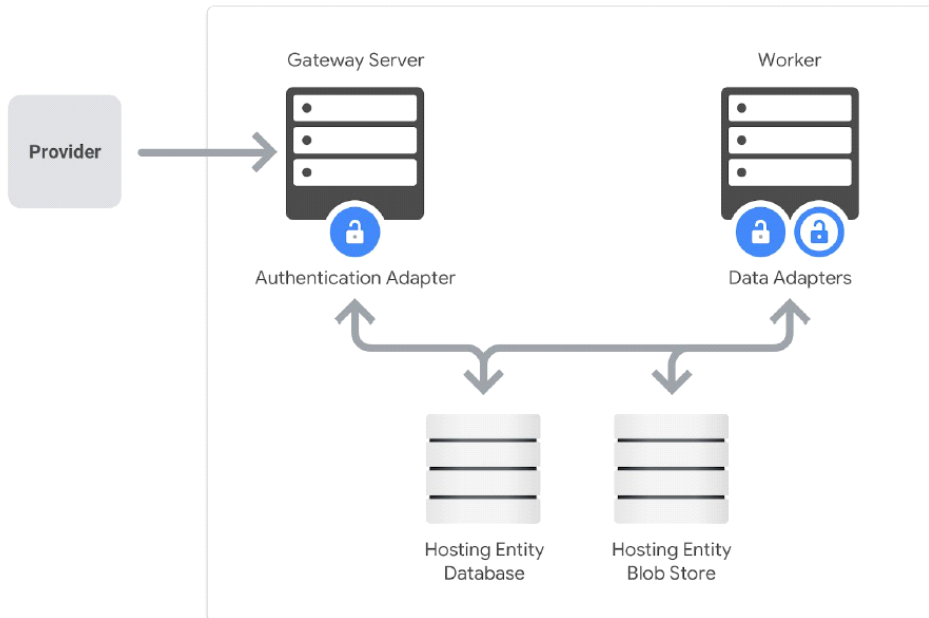
- Worker

- Worker는 작업관리 라이브러리를 사용하여 어댑터를 실행한다.
- Worker는 데이터 전송이 시작될 때 만들어지고 해당 전송이 완료되면 종료되는 독립 가상 머신이다.
- Worker는 생성될 때 임시 키를 생성하고, 종료될 때 해당 키는 삭제된다.

- Securely Storing Data

- DTP를 통해서 전송되는 데이터의 보안은 매우 중요하다.
- DTP를 디자인하는데 있어서 가장 중요한 목표는 호스팅 주체가 데이터를 전송할때나 그렇지 않을때에도 사용자의 데이터에 접근할 수 없도록 하는 것이다.
- DTP 보안은 업계 best practices를 따르며, 위에서 설명한 Worker가 생성한 임시 키로 암호화 된 모든 데이터를 저장하고 전송계층 보안도 요구한다.
- DTP 시스템 디자인은 관리자가 사용자의 암호화 키에 접근할 수 없게 하여 사용자의 데이터를 보호한다.

〈그림 20〉 DTP시스템 구성요소간 상호작용 개요



게이트웨이 서버는 인증 어댑터를 통해 사용자가 데이터를 내보내고 가져올 수 있도록 권한을 부여하고, 암호화된 자격증명 및 전송요청에 대한 메타데이터를 데이터베이스에 저장함.

작업자 프로세스는 특정 전송 요청에 할당되고 작업관리 라이브러리를 호출하여 내보내기 및 가져오기 작업을 조정 및 실행하며, 내보내기 및 가져오기 실행시 임시적으로 blob 저장소에 데이터를 암호화하여 저장함.

출처 : Google, "Data Transfer Project Overview and Fundamentals", Jul. 2018

④ 시스템 액세스(System Access)

각 공급자는 자신의 시스템에 저장된 데이터에 대한 액세스 권한을 가진 사람에 대한 완전한 제어권을 갖는다. Hosting entity가 DTP 호스트 플랫폼을 실행하면 Hosting entity는 각 공급자에게 데이터를 주고 받을 수 있는 키를 요청한다. 이때, DTP는 공급자 간의 데이터 접근 권한에 대해 관여하지 않는다. 이렇게 하면 API할당량을 공급자가 자체적으로 관리할 수 있기 때문에 트래픽 급증 및 공급자간의 부정적 영향을 최소화할 수 있다. 누구나 DTP에 어댑터를 제공할 수 있지만 각 Hosting entity는 상호작용할 공급자를 결정한다. 각 Hosting entity는 API키를 요청할 공급자를 결정하고 각 공급자는 API키를 부여할 Hosting entity를 선택한다. 이때, 각 Hosting entity는 공급자의 개인정보 보호 의무를 고려해야 하며 사용자가 데이터를 적절히 사용하는지 확인해야 한다.

3) 배포(Deployment)

① 호스트 플랫폼 배포(Deploying a host platform)

Hosting entity가 DTP 인스턴스를 배포하기 위해서는 자신의 필요에 맞게 DTP를 커스터마이징 하는 몇 단계 작업이 필요하다.

• API키 취득(Acquire API keys)

- Hosting entity는 데이터를 주고 받을 수 있는 모든 데이터 공급자의 API 키를 취득해야 한다. API 키를 통해 Hosting entity는 상호작용하고자하는 데이터 공급자를 결정할 수 있고, 각 공급자는 원하는 경우 다른 공급자 혹은 Hosting entity가 키를 제공하는 것을 제어할 수 있다. 또한, 호스팅 제공업자는 공급자가 제공한 API 키를 사용하는데 필요한 이용 약관에 동의해야 할 수도 있다.

• 배포(Deploy)

- DTP는 Docker 이미지를 사용하여 실행되도록 제작 되었다. DTP의 오픈소스 코드 저장소에는 다양한 클라우드 공급자 또는 로컬 시스템에 DTP를 배포하는 스크립트 소스코드를 제공한다.

• 사용자 인터페이스 생성(Create a User Interface(UI))

- 배포된 DTP는 전송작업을 관리하기 위한 http 프로토콜 기반의 Restful API를 제공한다. Hosting entity는 사용자가 DTP를 쉽게 사용할 수 있도록 DTP API위에 UI를 개발한다.

② 배포 모델(Deployment model)

DTP의 호스트 플랫폼을 배포하는 방법은 분산식, 중앙집중식, 자체관리식의 3가지가 있다. 각각의 방식은 장단점이 있으며, 모든 유스케이스에 맞는 솔루션은 존재하지 않는다.

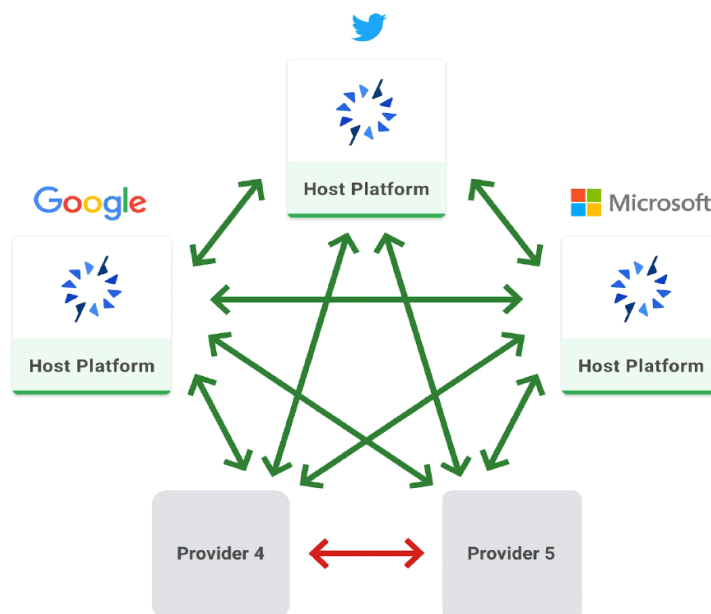
• 분산식(Distributed)

- 분산 호스팅 모델에서 각 공급자는 Hosting entity가 되어 호스트 플랫폼을 실

행한다. 사용자가 특정 공급자와 데이터를 주고 받기를 원하면 사용자는 전송을 시작한 공급자의 호스트 플랫폼을 통해서 데이터 전송을 시작한다. 예를 들어, Google과 Microsoft는 모두 호스트 플랫폼을 실행하고 있다. 사용자가 Microsoft UI로 이동하여 Google에서 가져오기를 시작하면 Microsoft에서 실행하는 호스트 플랫폼을 사용하게 된다. 분산형 호스팅의 장점은 데이터가 제3자의 관여 없이 오직 원본 공급자와 데이터를 받는 공급자 사이에서만 처리된다는 점이다. 단점은 호스트 플랫폼을 실행하는 공급자로 전송이 제한된다는 점이다. 또한 모든 공급업체가 별도의 호스트 플랫폼을 유지 관리해야 하기 때문에 과도한 유지관리 비용이 발생할 수 있다.

〈그림 21〉 분산 배포 모델

Distributed DTP Environment



이 분산 모델에서 Microsoft, Google, Twitter는 각각 Hosting entity이며 DTP 어댑터 및 데이터 모델을 이용하여 서로 간에 데이터를 전송할 수 있다. 공급자4 와 5는 DTP 어댑터를 가지고 있지만 호스트 플랫폼을 실행하지 않는다. 사용자는 Google, Microsoft 및 Twitter에서 공급자4나 공급자5로 데이터를 전송할 수 있다. 그러나 공급자 4와5는 호스트 플랫폼을 실행하고 있지 않으므로 사용자는 공급자 4와 5 사이에서 직접 데이터를 전송할 수 없다.

출처 : Google, "Data Transfer Project Overview and Fundamentals", Jul. 2018

- 중앙집중식(Centralized)

- 중앙집중식 시스템은 독립적인 제3의 기관이 Hosting entity가 되어 호스트 플랫폼을 구성하고 모든 공급자에 대한 요청을 처리하는 방식이다. 이 방식의 장점은 호스트 플랫폼을 실행하기 위한 비용과 전문지식이 부족한 중소기업이 쉽게 DTP에 참여할 수 있다는 것이다. 중앙집중화 시스템을 이용하기 위해서 중소기업은 단지 어댑터를 개발하고 유지관리만 하면 된다. 중앙집중식 모델의 단점은 독립적인 제3의 기관의 신뢰성이 검증되지 않은 경우 사용자의 데이터에 접근해 무단으로 개인정보를 획득할 수 있다는 점이다. 따라서 독립적인 제3의 기관에 대한 검증이 필요하고, 고도로 숙련된 데이터 보호 기술이 뒷받침되어야 한다. 또한 독립 제3의 기관은 하드웨어와 네트워크 비용이 발생하기 때문에 시스템을 유지할 수 있는 구체적인 사업 계획이 동반되어야 한다.

- 자체관리(Self-managed)

- 사용자는 자신의 컴퓨터 또는 개인 클라우드에서 DTP 복사본을 다운로드하고 실행할 수 있다. 자체 관리의 장점은 사용자가 전송 환경을 완전히 제어할 수 있다는 것이다. 자체관리 방식을 사용하면 공급자가 호스트 플랫폼 실행을 중지해도 사용자는 해당 공급자와 데이터를 주고 받을 수 있다. 이 방식의 단점은 호스트 플랫폼을 실행하는 것이 개인이 실행하기에는 복잡한 작업이고 자원 및 비용이 발생한다는 점이다.

4) 보안 및 개인 정보 보호(Security & Privacy)

사용자 데이터의 보안 및 개인 정보 보호는 DTP의 기본 원칙이다. 그러나 DTP 데이터 전송에 관련하여 여러 당사자(사용자, Hosting entity, 공급자 및 제공자)가 있기 때문에 어떤 사람이나 기관도 전체 시스템의 보안이나 개인정보에 대한 완벽한 보장을 약속할 수 없다. 대신에 책임은 모든 참가자들 사이에서 공유된다. 자세한 설명은 아래 구체적으로 설명한다.

① 데이터 최소화(Data minimization)

공급자간 데이터 전송시 최소한의 데이터만 전송해야 한다. 데이터를 제공받는 기관은 자신의 서비스를 제공하는데 필요한 최소한의 개인 데이터 집합만을 처리하고 유지해야 함을 의미한다. 데이터를 보내는 기관은 제공받는 기관이 필요로 하는 모든 정보를 제공해야 하지만, 그 외에 어떤 데이터도 제공해서는 안된다.

② 사용자 알림(User notification)

Hosting entity는 사용자의 데이터가 전송이 시작되면 사용자로 하여금 전송 사실을 알 수 있는 알림 기능을 호스트 플랫폼에 구성해야 한다. 파일 전송시 보내는 원본 계정과 받는 대상 계정의 사용자는 동일인이어야 한다. 따라서 사용자 알림 기능은 원본 계정 및 대상 계정 모두에게 전송 요청 경고 알림을 보내야 한다. Hosting entity는 전송되는 데이터의 민감도에 따라 전송 시작 지연시간을 설정함으로써 사용자가 알림을 받은 후 전송을 취소할 수 있도록 한다.

③ 전송 횟수/빈도 제한(Rate Limiting)

공급자 및 Hosting entity는 사용자의 전송횟수와 빈도를 제한해야 한다. 이 방법은 사용자의 계정이 해킹당한 경우 시스템에 영향을 제한하는데 도움이 된다. 시스템 이용의 편의성과 보안은 트레이드 오프 관계에 있으며, 모든 경우에 맞는 한 가지 제한 기준은 없다는 것을 알 수 있다. 대신 공급자 및 hosting entity는 적절한 제한 기준을 결정하기 위해서 알려진 해킹 공격뿐 아니라 데이터의 민감도를 고려해야 한다.

④ 토큰 해지(Token revocation)

전송이 완료되면 DTP는 전송에 사용된 인증 토큰을 폐지하려고 시도한다. 공급자는 API가 토큰 취소를 지원하는지 확인해야 한다. 이 기능은 보안 매커니즘 중 하나가 손상되어 사용자 인증 토큰이 유출되었을 경우 그 유효기간이 전송기간으로 제한되도록 하여 시스템의 상위 계층을 보호할 수 있도록 한다.

⑤ 인증 토큰의 최소 범위(Minimal scopes for Auth Tokens)

공급자는 인증 토큰에 대한 세분화된 범위를 제공해야 한다. 이는 두가지 이점을 제공한다. 첫째, 정확하게 어떤 데이터가 이동될지 투명성을 제공한다. 둘째, 토큰이 어떤 방식으로든 유출 되면 침침 방어 매커니즘을 통해 최소한의 권한만 갖도록 한다. 최소한 읽기 권한은 모두에게 제공되므로 보내는 공급자에게 쓰기/수정/삭제 권한은 제공되지 않는다.

⑥ 데이터 보존(Data retention)

DTP는 전송 작업이 이뤄지는 시간 동안에만 데이터를 저장한다. 또한 시스템을 통해서 전송되는 모든 데이터는 대기 상태 및 전송 상태 모두 암호화되어 처리 된다. 특히, 대기 상태의 모든 데이터는 단일 사용자의 작업만을 처리하기 위해 할당된 임시 가상 메모리에 사용자별 세션키로 암호화 되어 저장된다.

⑦ 남용(Abuse)

공급자 및 Hosting entity(호스팅 사업자, 공급자와 별도인 경우)는 API에 강력한 남용 방지 기능을 내장해야 한다. DTP는 단일 전송 제한 시간을 초과하는 사용자 데이터를 보유하지 않으며, 해킹 공격에 대비해 여러개의 호스팅 엔티티가 활용될 수 있기 때문에 공급자는 악의적인 시스템 접근을 탐지하고 대응할 수 있는 최상의 도구를 갖추고 있다. 공급자는 API 키에 대한 접근이 가능한 Hosting entity를 신중히 제어해야 한다. 공급자는 인증 토큰과 관련하여 강력한 보안을 유지해야 한다. 액세스 권한을 부여하기 전에 재 인증을 요구하거나 보안 질문에 대한 답변을 요구하는 경우가 여기에 해당한다.

[표 3] 책임 공유 테이블: 보안 및 개인정보보호

Task	User	Provider-exporter	Provider-Importer	Hosting Entity	DTP System
데이터 최소화	전송할 데이터 선택	내보낼 데이터 세부 제어 제공	서비스에 불필요한 데이터 삭제	적절한 전송 대상만 구성	N/A
사용자 알림	전송 알림 수신 및 검토	N/A	N/A	메일 송신자 및 지연 정책 구성	알림 송신, 취소 가능한 딜레이 제공
전송 회수/빈도 제한	N/A	시행	N/A	남용을 막기 위한 합당한 제한 설정	공급자별 전송회수/빈도 제한 지원
토큰 해지	공급자가 토큰 자동폐기를 지원하지 않을시 수동 해지	토큰 해지 제공	토큰 해지 제공	N/A	(공급자 지원시) 사용 후 인증토큰 해지
인증 토큰의 최소범위	적절한 요청 범위 확인	세분화된 범위 구현	세분화된 범위 구현	N/A	각 전송에 대한 최소 범위 요청
데이터 보존	데이터 전송이 삭제됨을 의미하지 않으므로, 사용자는 원본 데이터를 삭제	사기 및 남용을 방지하는데 필요한 데이터만 저장	개인정보 취급 방침에 따라 가져온 데이터만 보관. 사기 및 악용 방지에 필요한 메타데이터 저장	식별가능한 정보를 보유하지 않도록 시스템 구성	전송후 데이터 파기
남용	계정 자격 증명 보호(강력한 암호, 이중 인증 등)	API에 사기 및 남용 방지 기능 구현	API에 사기 및 남용 방지 기능 구현	UI에 사기 및 남용 방지 기능 구현	임시키를 사용하여 전송/대기중인 데이터 암호화, 각 전송에 독립/전용 가상 머신 사용

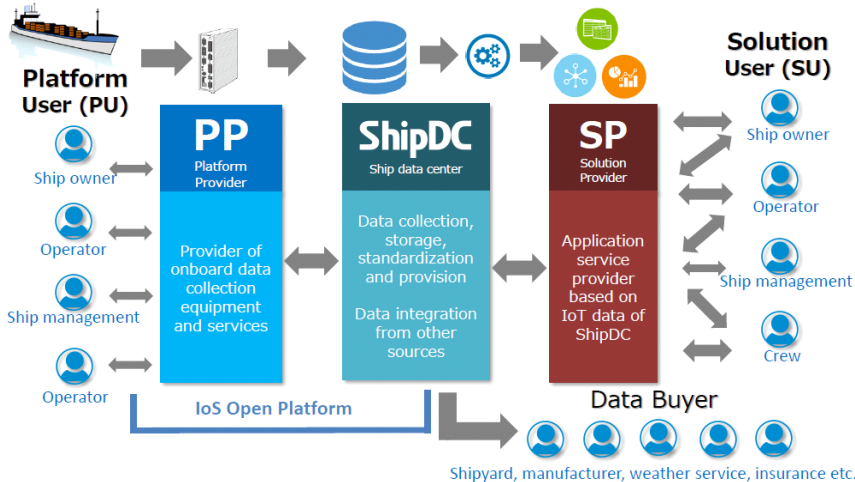
출처 : Google, "Data Transfer Project Overview and Fundamentals", Jul. 2018

(4) 결론

DTP 프로젝트는 표면적으로는 SNS상의 개인정보를 다른 서비스로 복사할 수 있는 편리함을 목적으로 한 것으로 보이지만, 개인정보의 주체가 개인에게 있고 사용하는 서비스에서 자신의 개인 정보를 언제든지 주체적으로 변경, 삭제할 수 있다는 개인정보보호에 대한 새로운 인식이 내재되어 있다. 다시 말해 업계는 그동안 개인정보를 손쉽게 수집하고 이용하여 경제적 이득을 취했던 그동안의 관행에서 벗어나 이제 새로운 개인정보 패러다임에 적응해야 한다. 새로운 서비스를 개발하는 신생 기업들은 보다 더 쉽게 사용자를 확보할 수 있고, 기존 거대 IT기업의 데이터를 활용하는 다양한 서비스를 개발할 수 있어 신규 시장 진입에도 도움이 되고 IT시장 전체가 더 활성화 될 전망이다.

3 SSAP2(Smart Ship Application Platform2) 프로젝트

〈그림 22〉 선박 데이터 플랫폼



출처 : JSMEA, "The role of Open Platform in the Coming Era of Ship Digitalization", Apr. 2018

(1) 배경 및 문제점

세계 조선 해운 산업은 최근 몇 년간의 신조선 발주량 감소 및 선복량 과잉으로 인해 산업 전체가 장기 불황 상태에 접어들었다. 또한 글로벌 주요 선사는 저운임으로 막대한 손실을 감수하면서, 선복량을 늘리는 등 치킨게임을 주도하며 경쟁사를 도태시키는 전략을 수행하고 있다[16]. 최근 4차 산업혁명 및 ICT융합의 기술 발전에 따라 세계 각국은 스마트 선박 개발을 통한 해운업의 운영비 절감, 선박 효율 향상 등 위기 극복을 위한 다양한 시도가 이뤄지고 있다.

현재 선박에서는 최적트림, weather routing, 엔진모니터링 등 선박 응용서비스를 위해 선박내의 단위 시스템(전자해도, 블랙박스(VDR), 엔진 데이터 로거 등)에 직접 인터페이스하여 데이터를 취득함으로써 응용 프로그램이 추가될때마다 불필요한 데이터 케이블링 작업이 동반되며 중복된 데이터 전송으로 인해 단위시스템에 부하가 발생한다. 또한 육상 데이터 전송시 중복 데이터 전송으로 인한 네트워크 비용의 증가도 발생하며 데이터

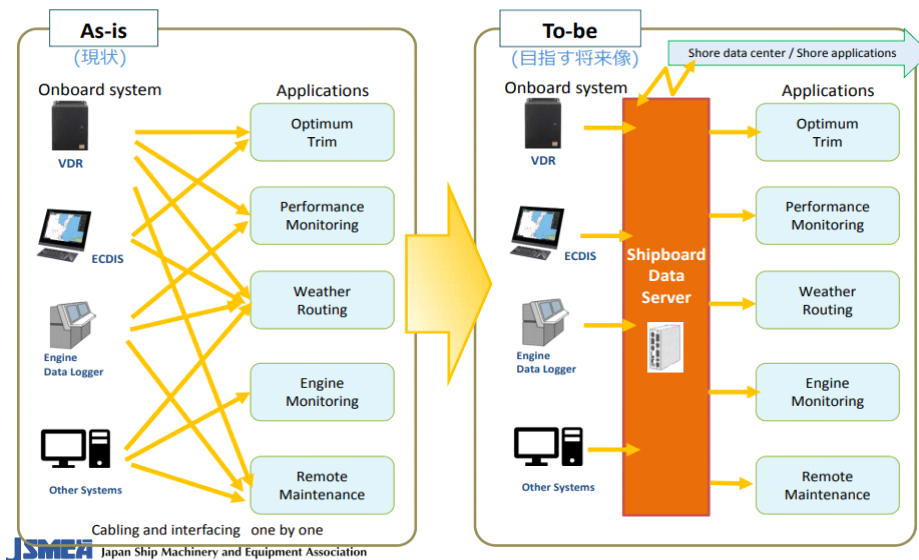
의 주체가 불분명하여 공유가 제대로 이뤄지지 않고 있다.

일본에서는 2012년 12월부터 2015년 3월까지 36개 기관이 선박 마스터 데이터베이스 중심으로 하는 스마트십 플랫폼(SSAP1) 개발 프로젝트를 진행했다. 그 후속 과제로 2015년 8월부터 선주, 선사, 조선소, 기자재 업체, 연구소 등 48개 기관이 프로젝트 그룹을 생성하여 Ship data center(ShipDC)를 중심으로 하는 선박-육상간 데이터 표준화 및 데이터 플랫폼(SSAP2) 개발을 시작했다.

(2) 프로젝트의 목적

이 프로젝트의 목적은 선박, 육상 관제센터에서 운용하는 응용 서비스(weather routing, 최적 트림, 성능 모니터링, 엔진 모니터링, 원격 유지보수 등)에 필요한 데이터를 이해당사자(선주, 선사, 조선소, 기자재 업체, 보험회사 등)에게 쉽고 빠르게 제공함으로써 IoT 응용 서비스 개발을 지원하여 선박의 안전과 에너지 효율 향상에 기여하는데 있다[17].

〈그림 23〉 개발 전 후 비교



출처 : JSMEA, "The role of Open Platform in the Coming Era of Ship Digitalization", Apr. 2018

기존에는 선박의 설계와 건조 단계에서만 엔지니어링 기술과 시뮬레이션 도구가 사용되었다. 또한 설계자와 엔지니어는 제조가능성, 유용성, 유지보수성 등 선박의 수명주기에 걸친 주요 성능을 설계 단계에서만 고려하고 선박의 운영시에는 기술 지원이 없다. 선박에 IoT기술이 접목되면 선박의 수명주기 전반에 엔지니어링 기술과 시뮬레이션 도구가 사용될 수 있고, 선박 운영에 필요한 새로운 서비스를 창출 할 수 있으며 새로운 선박 디자인에 대한 피드백을 받을수도 있다. IoT와 빅데이터를 활용한 응용 서비스는 다음과 같다.

역할	기능	응용 서비스
선주	기술 경영	· 안전 운항 · 상태 모니터링 및 유지보수 · 환경 규제 준수 · 선체 및 프로펠러 청소 · 부품 재장착 및 개조
	신조선	· 설계 최적화
해운사	운항	· 에너지 절약 운항 · 안전 운항 · 일정 관리
	선단 계획	· 선단 배정 · 서비스 기획 · 용선
조선소	운항중인 선박	· 선박 성능 분석
	신조선	· 설계 최적화
장비 업체	유지보수	· 원격 모니터링 및 진단

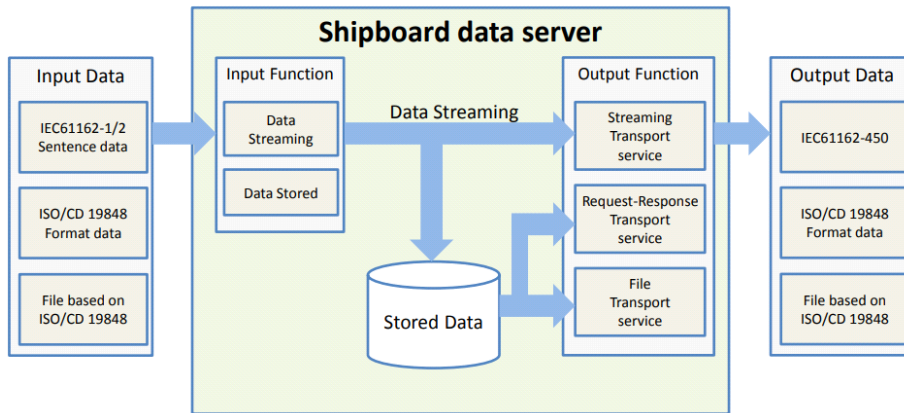
(3) 기술 상세

1) 데이터 표준화

① ISO 19847-Shipboard data servers to share field data on the sea

선박 데이터 서버에 대해 정의한 국제표준인 ISO 19847은 선박 장비 및 시스템에서 데이터를 저장하고 저장된 데이터를 육상으로 전달하는 데이터 서버에 대한 성능, 기능, 서비스 및 안전 요건에 대한 방법을 정의한다.

〈그림 24〉 ISO19847 표준에 기술된 선박 데이터 입출력 개념 모델



출처 : JSMEA, "Activities of Smart Ship Application Platform 2 Project(SSAP2)", Sep. 2017

선박 데이터 서버는 입력, 출력, 데이터베이스로 구성되어있고 데이터 입출력시 이더넷 인터페이스를 사용하여 데이터 전송을 위해 요청-응답 데이터 전송 서비스, 실시간 데이터 및 파일 전송 서비스 기능을 제공한다.

입력은 ISO 19848, IEC61162-1/2 표준에 정의된 형식의 데이터를 수신하여 출력으로 내보내고 데이터베이스에 저장한다.

출력은 실시간 전송, 요청-응답 전송 및 파일 전송 서비스로 구성되어 있으며 사용자는 MQTT(Message Queueing Telemetry), REST API 및 FTP(File Transport Protocol) 프로토콜을 통해 데이터를 수신받는다.

② ISO 19848-Standard data for machinery and equipment part of ship

기존 IEC61162 표준은 항법장치의 데이터 교환 표준으로 기계장비, 안전관련 장비 및 선체 등 선상 시스템의 데이터는 표준화 되지 않았다[18].

선박의 장비 및 센서에서 생성되는 모든 디지털 데이터를 교환하기 위해서 ISO19848은 센서ID, 데이터 모델 및 데이터 포맷을 정의한다.

〈그림 25〉 ISO19848 표준에 따른 선박 장비 명명 규칙 예

Name Object	Description	Example	Entries
System	Discrete equipment, systems, and units.	MainEngine , DieselGeneratorSet	39
Component	Part of the System above or detail description of the system.	Cylinder TurboCharger	34
Content	Contents that flow through pipelines or parts in the above component.	CoolingFreshWater HeavyFuelOil	tbd
Position	Position that indicates relative positions in System and Components.	Inlet Outlet	33
Item	Indicates temperatures, pressures and other measurement objects.	Temp Press	58
Suffix (Option)	Only if Data Channels cannot be specified with the Elements shown above (from "System" to "Item"), suffix may be added after Local ID to identify the Data Channels. Suffix may consists of and DataChannelType as shown below. Suffix: ../<modifier>/<DataChannelType>	Modifier is typically used to describe correction parameter such as temperature for fuel oil viscosity, like "at15degC" .	

LocalID = /jsmea_mac/MainEngine/Cylinder/CoolingFreshWater/Inlet/Temp

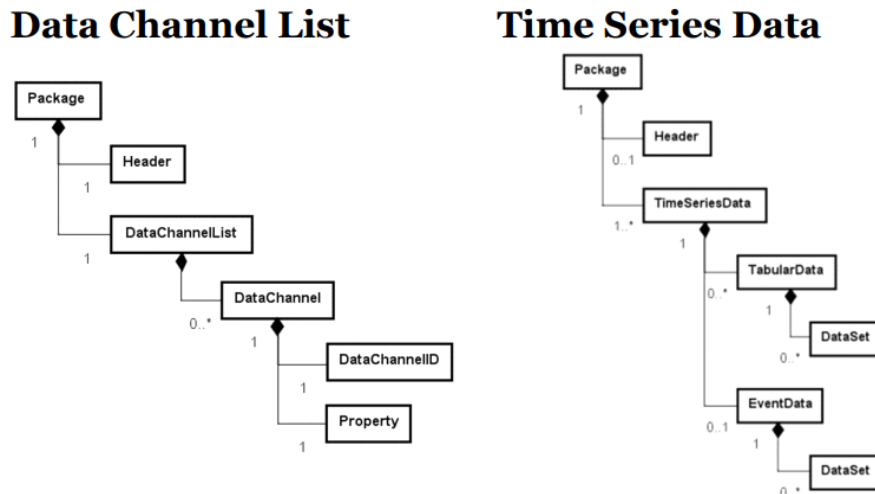
출처 : JSMEA, "Introduction of ISO/FDIS 19848 Standard data for shipboard machinery and equipment", Apr. 2018

ISO19848표준은 2종류의 데이터 구조와 파일 포맷을 정의한다. (i) 필수 메타데이터를 포함한 데이터 채널 리스트. (ii)측정 데이터 저장을 위한 시계열 데이터 포맷

시계열 데이터 포맷은 최소한의 메타데이터 정보만 포함하도록 간단히 설계 되었다.

데이터 채널은 채널ID와 채널 속성으로 구성 된다. 채널 ID는 고유값을 가지며 논리 데이터 채널을 식별할 때 사용한다. 채널 ID는 세 가지 유형이 있다. (i)선박에 사용된 고유 식별자인 로컬ID (ii)선박ID(예, IMO번호) 및 로컬ID로 구성된 유니버설ID (iii)단축ID

〈그림 26〉 ISO19848 표준에 정의된 데이터 구조



출처 : JSMEA, "Introduction of ISO/FDIS 19848 Standard data for shipboard machinery and equipment", Apr. 2018

신뢰할 수 있는 데이터 교환을 위해 데이터는 XML(Extensible Markup Language) 및 XML스키마를 사용하도록 규정한다.

또한 편의성과 효율성을 위해서 JSON 및 CSV형식의 데이터 구조도 정의한다.

표준에 따라 인코딩된 데이터는 보안상의 이유로 암호화할 수 있으며, HTTP (Hyper Text Transfer Protocol) 방식으로 컴퓨터 응용 프로그램간에 교환된다. 또한 MQTT(Message Queue Telemetry Transport) 방식으로 전자메일로 전송할 수 있다.

2) 선박 데이터 센터(Ship Data Center)

① 선박 표준 플랫폼의 장점

응용 서비스 제공자는 선주로부터 제공받은 선박 운용 데이터를 이용하여 쉽게 다양한 서비스를 개발하여 선주에게 제공가능

선주는 데이터 수집 플랫폼을 공유함으로써 선내 데이터 수집 비용 및 응용서비스 사용료 경감

조선소/장비 제조업체는 자신이 제작한 장비의 생애주기 데이터 취득을 통해 새로운 서비스 가능

표준화된 데이터 포맷과 프로토콜을 통해 더욱 다양한 IoT 응용 서비스 개발이 가능

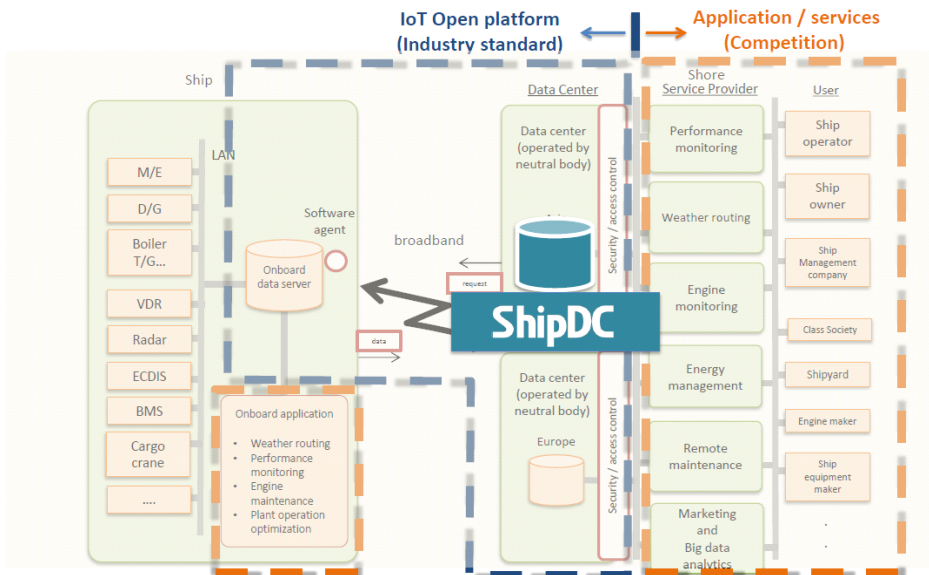
② 선박-육상간 데이터 교환을 위한 선박 오픈 플랫폼

선박 오픈 플랫폼은 선내 데이터 서버와 육상 데이터 센터로 구성된다.

선내 데이터 서버는 엔진, 레이더, 전자해도 등 선내의 항법장치 및 기계장치로부터 데이터를 수집해 항로최적화, 성능모니터링, 엔진 유지보수 등 선내 응용 어플리케이션에 데이터를 제공한다. 또한 소프트웨어 에이전트를 통해 수집된 선내 데이터를 통합해 주기적으로 거점별 육상 데이터 센터로 전송한다.

육상 데이터 센터에 축적된 데이터는 보안, 접근제어를 통해 여러 이해당사자(선주, 조선소, 장비업체, 해운사 등)에 제공되며, 각 이해당사자는 이 데이터를 활용해 필요한 서비스를 개발, 운영한다.

〈그림 27〉 SSAP 오픈 데이터 플랫폼 구성도



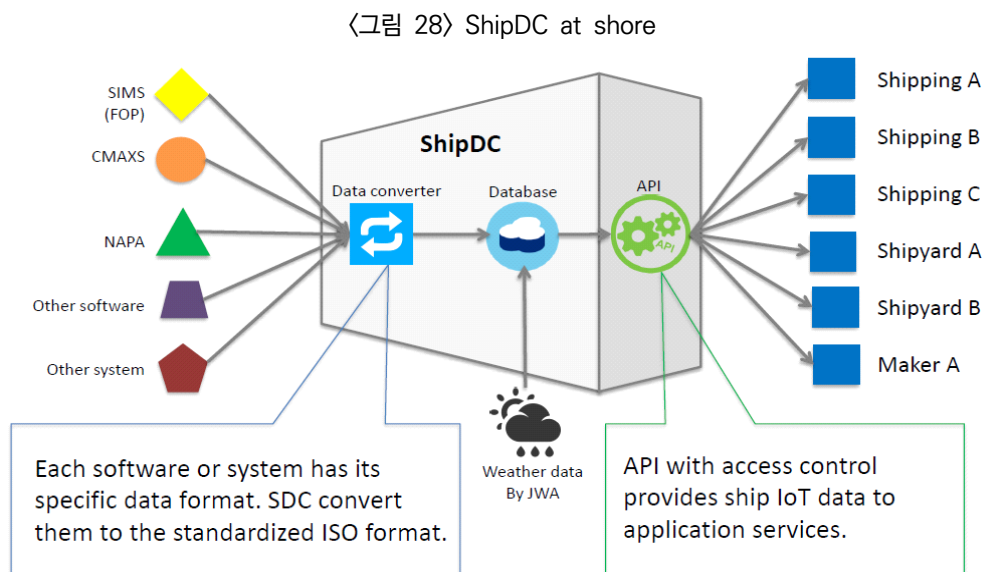
출처 : JSMEA, "Activities of Smart Ship Application Platform 2 Project(SSAP2)", Sep. 2017

③ 선박 데이터 센터(shipDC-ship data center at shore)

선박 데이터 센터는 선박 데이터를 안전하고 편리하게 사용할 수 있는 플랫폼을 제공한다. 플랫폼을 통해 육상에서 선박 데이터의 활용이 쉬워지고, 선박 IoT어플리케이션 서비스의 개발 및 운영이 편리해진다.

선박에 기 적용된 여러 소프트웨어, 시스템이 가지고 있는 독자적인 데이터 포맷은 데이터 변환기를 통해서 ISO19848 표준 포맷으로 변경된다.

선박 데이터 센터에서 제공하는 API를 통해 각 기관은 접근 가능한 데이터를 수집하여 자신의 어플리케이션 서비스에 활용할 수 있다.



출처 : JSMEA, "Activities of Smart Ship Application Platform 2 Project(SSAP2)", Sep. 2017

④ 선박 인터넷 오픈 플랫폼(Internet of Ship(IoS) open platform)

플랫폼 공급자는 선박 내의 센서, 장비 데이터를 수집하여 선주, 선원 등 선내 플랫폼 사용자에게 안전하고 편리한 서비스를 제공한다.

ShipDC는 선박 플랫폼의 데이터를 표준화된 포맷으로 변환하여 저장하고, 선박 외부의 데이터 소스로부터 데이터를 수집하여 통합한다.

육상 솔루션 공급자는 여러 선박플랫폼으로부터 수집된 데이터를 이용하여 응용 소프트웨어 서비스를 개발하고 이해당사자에게 제공한다.

또한, 선박의 데이터를 이용해 부가가치를 창출하고 싶은 이해당사자들은 비용을 지불하고 ShipDC로부터 데이터를 구매할 수도 있다.

(4) 현황 및 향후계획

일본은 SSAP프로젝트를 통해 선박 데이터 센터(ShipDC), 선내플랫폼 및 IoT를 위한 데이터 표준을 마련하였다. 이러한 기반을 바탕으로 CBM(Condition based maintenance), PHM(Prognostics and health monitoring) 등 선박의 누적된 데이터를 바탕으로 고장을 예측하고, 현 상태를 모니터링하는 신규 서비스를 2020년까지 개발하겠다는 계획을 가지고 있다.

선단 운영에 있어서 빅데이터와 IoT기술을 활용하여 선박 성능 최적화를 통해 비용을 절감하고, 더 나아가 무인 선박의 전초 단계인 자동화를 2020년까지 완성하라는 계획을 가지고 있다.

〈그림 29〉 SSAP 로드맵

Topic		2016	2017	2018	2019	2020
Application	R&D projects (e.g. I-Shipping in Japan and autonomous ships in Europe)			R&D projects (navigation, engine, hull, cargo, operation and etc.)		
	CBM ^{*1} & PHM ^{*2} services & class inspections		Condition based maintenance and PHM for main engines, machineries and equipment.		Condition-based class inspection services ?	
	Big data and IoT utilization in fleet operation		Integration of ship with shore operation system. Optimization, automation and simulation technologies.			
Platform	SSAP2 and standardization (ISO DIS 19847/19848)	ISO CD/DIS 19847/19848		(FDIS)		
	Ship data center and IoT (Internet of Ship)	SSAP 2	Preparation for IoT program		SSAP 3 ?	Ship Data Center in operation
Regulatory	Cyber safety and cyber resilient ship		Cyber safety (BIMCO guideline, IMO MSC guideline, Class guidelines)			
	EU MRV ^{*3} and IMO DCS ^{*4}				IMO DCS	EU MRV
	e-Navigation and autonomous ship projects	Model development	Standardization		Implementation	Operation
			Remote control/autonomous ship projects in Norway, Finland and other nations			

^{*1} CBM: Condition-Based Maintenance, ^{*2} PHM: Prognostics and Health Monitoring
^{*3} MRV: Monitoring Reporting and Verification, ^{*4} DCS: Data Collection System

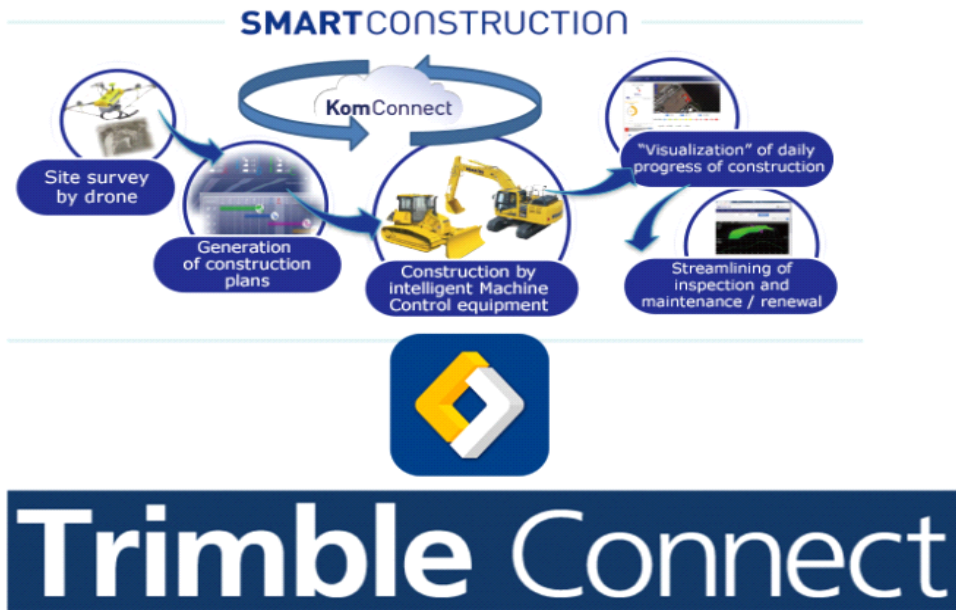
출처: JSMEA, "Activities of Smart Ship Application Platform 2 Project(SSAP2)", Sep. 2017

(5) 결론

선박 해양 사고는 국내의 경우 2013년 1,093건에서 2017년 2,582건으로 증가 추세에 있으며 인명피해도 꾸준히 증가하고 있다. 사고의 원인 중 하나는 선박의 항해시 선원의 견시에 크게 의존하기 때문이다. 3교대 근무를 법제화 하고 있지만 최근 전세계적으로 3D업종을 기피하는 추세와 해운사의 원가절감을 이유로 선원이 줄어 개개인의 업무량이 증가했다. 이러한 문제를 해결하는 방법중 하나는 선박의 업무를 육상에서 보조하는 것이다. 이를 위해서 선박의 각종 센서, 장비의 데이터를 통합하여 육상 데이터 센터로 전송하고, 육상에서는 이를 실시간으로 분석해 선박 및 육상 관계자에게 전달하여 관련 이해당사자로 하여금 빠른 의사결정을 하도록 보조해주는 플랫폼이 필요하다. 일본의 경우 선급이 중심이 되어 조선소, 해운사, 선주의 참여 아래 데이터 표준 개발, 선박 데이터 플랫폼 구축, 육상 데이터 센터 구축 등 가장 앞서 기술 개발을 하고 있다. 이를 통해 더 많은 선박, 육상 서비스를 개발할 수 있는 새로운 시장을 만들 수 있고, 더 안전하게 항해할 수 있고, 해운사는 선박 운영 비용을 줄일 수 있다.

4 » 건설 데이터 공유 플랫폼 Trimble connect[19]

〈그림 30〉 'Trimble Connect' 와 Komconnect 협업



출처 : Julkaistu, 2017. "Trimble ja Komatsu". *koneporssi*, August

(1) 배경 및 문제점

2017년 7월 Trimble사와 Komatsu사는 각 회사의 건설 소프트웨어 사용자들이 서로 건설현장 데이터를 쉽게 교환 할 수 있도록 하는 플랫폼 개발 협업을 발표했다[20]. 이를 통해서 사용자는 다양한 제조사의 토목공사 장비를 보다 쉽게 관리할 수 있고 전체 현장 운영 상황을 한눈에 모니터링 할 수 있게 되었다.

기존에는 Trimble의 현장 관리 프로그램과 Komatsu의 ICT 플랫폼을 각각 실행하여 모니터링 해야 했고, 각 프로그램에서 생성된 데이터가 연동이 되지 않아 한 프로그램의 결과 데이터를 다른 프로그램에 새로 작성해야 하는 번거로움이 있었다.

이번 협업의 주요 내용은 Trimble사와 Komatsu사는 각사의 플랫폼간의 호환이 가능

하게 해주는 API를 개발하는 것이다. 클라우드 기반 Trimble connect 플랫폼은 이 API를 통해서 Komatsu사의 디지털 건설 데이터를 Trimble connect환경에서 사용 할 수 있도록 해준다.

Trimble connected site의 솔루션은 시공업체, 건설장비의 제조사 및 모델에 관계없이 계약자에게 생산성 검토, 자재량, 이동량 과 차량 및 자산 관리정보를 제공하여 현장의 모든 정보를 한 개의 화면에서 확인하고 처리할 수 있는 플랫폼이다.

Komatsu사의 komconnect는 드론, ICT가 적용된 중장비를 활용하여 시공 효율성을 높이는 스마트 건설 플랫폼이다. 드론을 통한 측량은 기존의 방식에 비해서 시간과 인원을 절약할 수 있고 ICT가 적용된 중장비는 미숙련 작업자도 쉽게 중장비를 조작할 수 있도록 해준다.

스마트 건설의 과정은 다음과 같다.

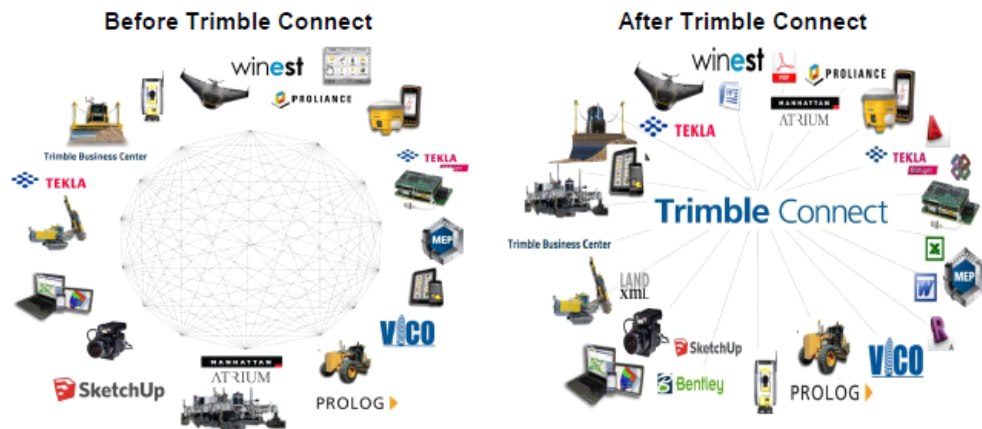
측량기사는 드론을 이용하여 건설현장의 측량데이터를 계측한다. 이 후 설계자는 3D측량 데이터를 바탕으로 시공계획을 수립한다. 시공사는 시공계획을 바탕으로 ICT가 적용된 중장비를 이용하여 시공을 진행한다. 시공이 진행되는 동안 측량기사는 드론을 이용하여 실시간으로 현장을 다시 측량하여 측량한 데이터와 시공계획을 실시간으로 비교할 수 있어 오차가 발생한 경우 빠르게 수정이 가능해진다.

(2) 프로젝트의 목적

Trimble connect는 프로젝트의 생애주기(설계, 구축, 운영단계)에 걸쳐 데이터를 통합하고 원활하게 교환함으로써, 건축 및 인프라스트럭처(infrastructure) 프로젝트의 효율성을 향상시키고 비용을 절감한다.

Trimble connect는 소프트웨어 및 하드웨어 공급업체가 직접 연결 할 수 있는 오픈 클라우드 기반의 협업 플랫폼을 구축하였다. 이를 통해 Trimble connect는 프로젝트 관계자가 파일을 관리하는데 드는 노력과 비용을 최소화 하여 관계자들이 프로젝트에만 집중할 수 있도록 해준다.

〈그림 31〉 Trimble Connect

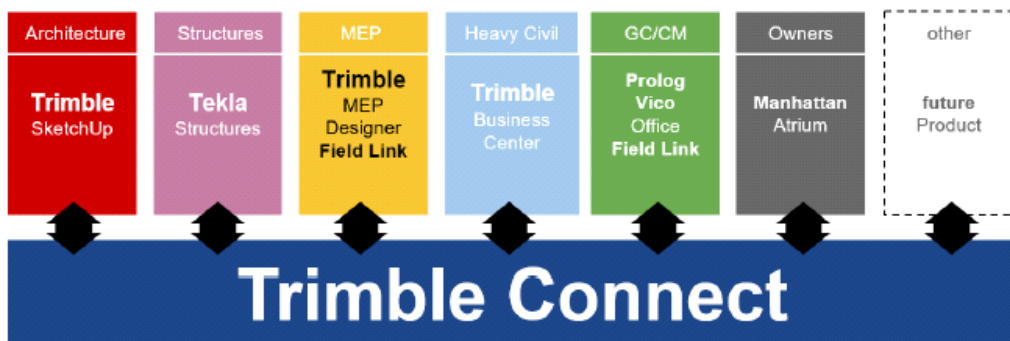


출처 : Trimble, "Trimble connect overview", Available at <https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

(3) 상세 기술

사용자는 Trimble connect의 데스크탑 파일 동기화 도구를 사용하여 다른 소프트웨어 및 하드웨어 솔루션과 파일을 쉽게 교환할 수 있다[21]. 다음은 Trimble connect와 통합된 Trimble family의 소프트웨어 및 하드웨어 솔루션이다.

〈그림 32〉 Trimble connect에 통합된 소프트웨어 및 하드웨어 솔루션



출처 : Trimble, "Trimble connect overview", Available at <https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

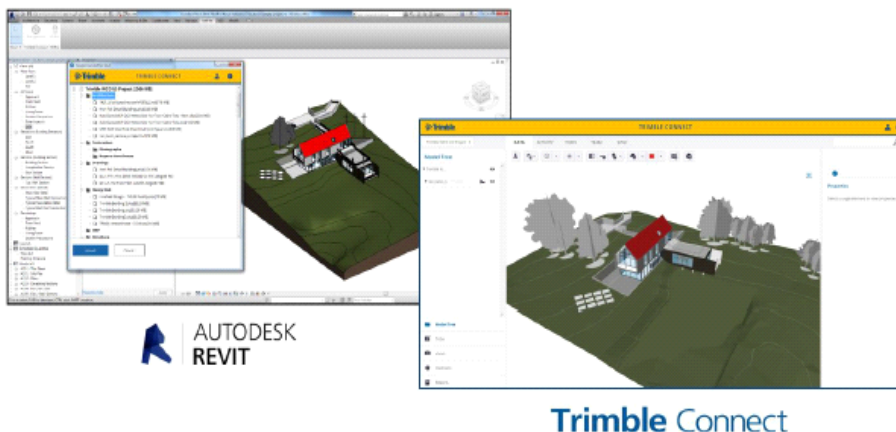
〈그림 33〉 Trimble connect가 지원하는 파일 포맷

분류	파일 포맷	비고
2D	PDF, DWG	PDF, DWG
3D	IFC (2x3), IFCZIP, SKP (2017 and below)	IFC, SKETCHUP
	3DM (Rhino 5.0 and below)	
	DWG (AutoCAD 2013 and below)	REVIT, DWG
	Revit (2017 and below)	
	LandXML, DXF, DGN, OBJ, FBX, 3DS, 3DM, STP, STEP, IGES, IGS, STL, SLDASM	TEKLA, RHINO
문서	DOC, DOCX, XLS, XLSX, RTF, TXT, PPT, PPTX	Word, Excel, PowerPoint, PDF
이미지	JPG, JPEG, GIF, PNG, TIF, TIFF, BMP	JPG, PNG, TIF, BMP
공간정보	SHP, KML, VCE, VCA, JXL, JOB, KMZ, GDB, Geospatial Zip files	(Beta Ver.)

출처 : 트림블커넥트, "지원파일포맷", <http://blog.naver.com/trimbledbo/221065385752>

한 예로, Trimble은 Autodesk사의 Revit 소프트웨어를 위한 플러그인을 생성하였다. Revit 소프트웨어 사용자는 플러그인을 이용하여 Trimble connect에 Revit 모델을 게시할 수 있다. 게시된 모델은 Revit 소프트웨어를 갖고 있지 않은 프로젝트 관계자도 관리 및 조회할 수 있고 코멘트를 달 수 있다. 이러한 소프트웨어 및 하드웨어 솔루션의 통합은 다양한 작업을 간소화하고 최신 정보 공유를 신속하게 한다.

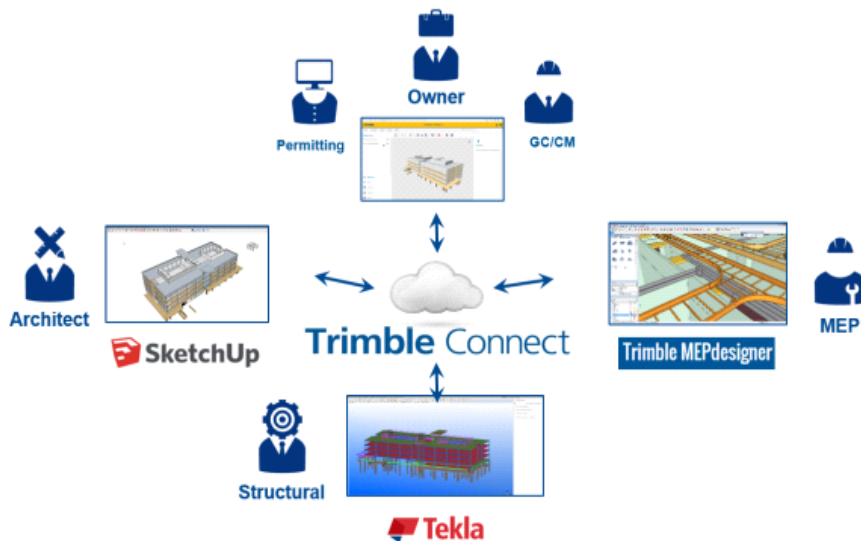
〈그림 34〉 플러그인을 이용한 revit모델공유



출처 : Trimble, "Trimble connect overview", Available at <https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

또 한 예로, 설계작업에서 SketchUp을 사용 하는 설계자는 파일을 내보내고 수동으로 업로드하지 않고도 자신의 3D모델을 직접 클라우드에 게시한다. 소유주는 클라우드 웹에서 업로드 된 모델을 검토하고 코멘트를 추가 할 수 있다. 설계자는 코멘트를 실시간으로 검토 할 수 있다. 구조엔지니어는 Tekla structure에서 설계자의 모델을 참조할 수 있다. 또한 설계를 조정하기 위해 자신의 구조 모델을 클라우드에 게시하여 설계자가 참조할 수 있도록 한다.

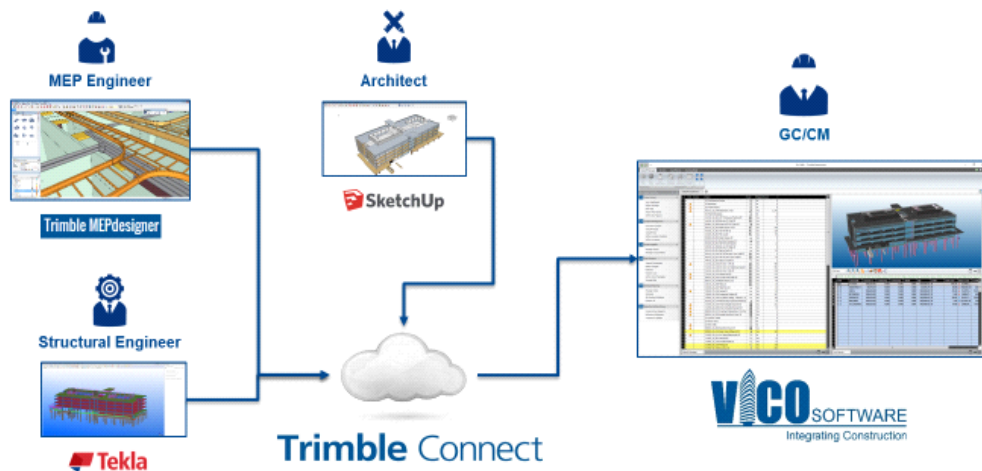
〈그림 35〉 Trimble connect를 이용한 설계 데이터 공유



출처 : Trimble, "Trimble connect overview", Available at
<https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

건축가, 구조엔지니어, MEP 설계자는 Trimble connect에 모델을 쉽게 게시할 수 있다. 이러한 모델을 Trimble의 VICO Office 소프트웨어와 결합하여 견적과 4D일정을 생성 할 수 있다. 새 버전의 디자인이 클라우드에 게시될때, VICO Office는 평가자에게 새 디자인에 따라 기존 견적 및 일정을 비교, 조정할수 있도록 보여준다.

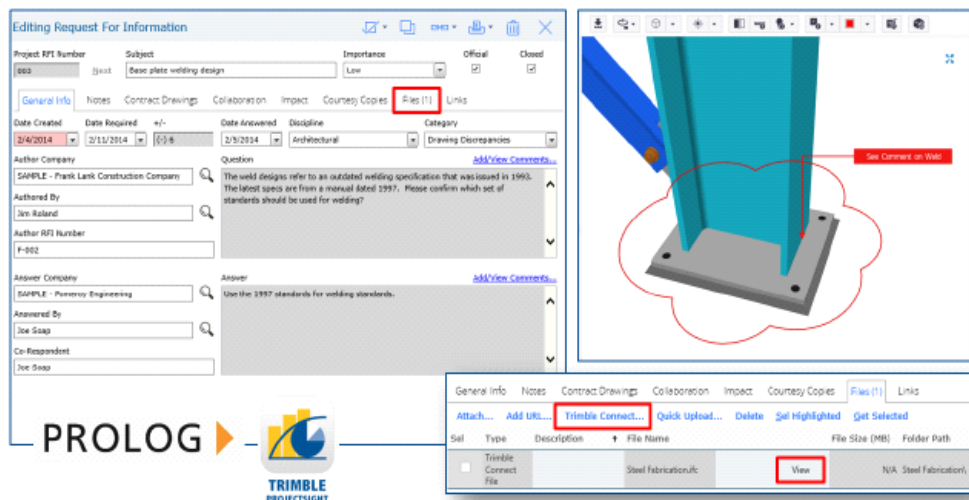
〈그림 36〉 설계데이터와 Vico office연동



출처 : Trimble, "Trimble connect overview", Available at
<https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

Trimble connect에 업로드 된 3D모델은 자료 요청서, 주문 변경서, 자산 또는 Prolog 및 Projectsight 소프트웨어에 첨부 할 수 있다.

〈그림 37〉 3D모델을 첨부한 자료요청서

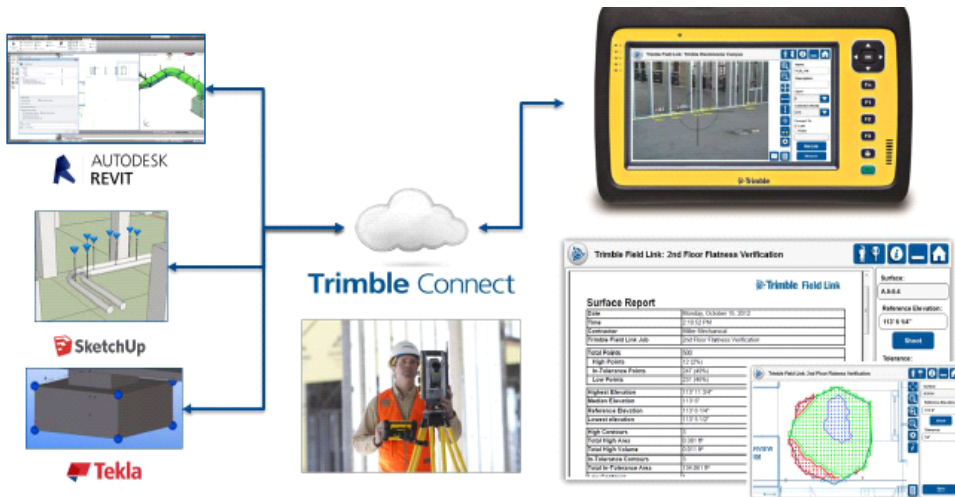


출처 : Trimble, "Trimble connect overview", Available at
<https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

위 그림처럼 주문 변경서에 3D모델이 첨부되면, 강철 받침판과 관련된 요청 내용을 보다 쉽게 이해할 수 있다.

건설작업에서는 디지털 3D빌딩정보 모델로 부터 주요 레이아웃 포인트를 추출 할 수 있다. 이 포인트 데이터는 클라우드 플랫폼을 통하여 현장의 RTS(Robotic Total Stations; 측량장비)로 전달된다. 이를 통해서 기존에 2D건축 도면의 레이아웃 정보를 수동으로 넘겨주고 측량을 위해 2인이상이 필요하던 과정을 제거함으로써 반복적 오류나 시간소모를 줄일 수 있다. 또한 측량기사는 모바일 태블릿에 설치된 Trimble Field Link를 이용하여 도면 요약 보고서와 현장의 불일치 보고서를 다운로드 할 수 있다. 이 보고서들은 클라우드에 다시 게시되어 프로젝트 관리자와 팀원들은 실시간으로 클라우드 웹을 통해 보고서 검토를 할 수 있다.

〈그림 38〉 Trimble Connect



출처 : Trimble, "Trimble connect overview", Available at
<https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>

Trimble connect는 클라우드를 통해 프로젝트 정보를 안전하게 저장, 관리, 전송한다. 프로젝트 정보는 북미, 유럽, 아시아의 주요 프로젝트 호스팅사이트를 통해 전세계에 호스팅되고 백업된다. 폴더 및 사용자 권한을 이용하여 관리자는 프로젝트를 온전하게 유지

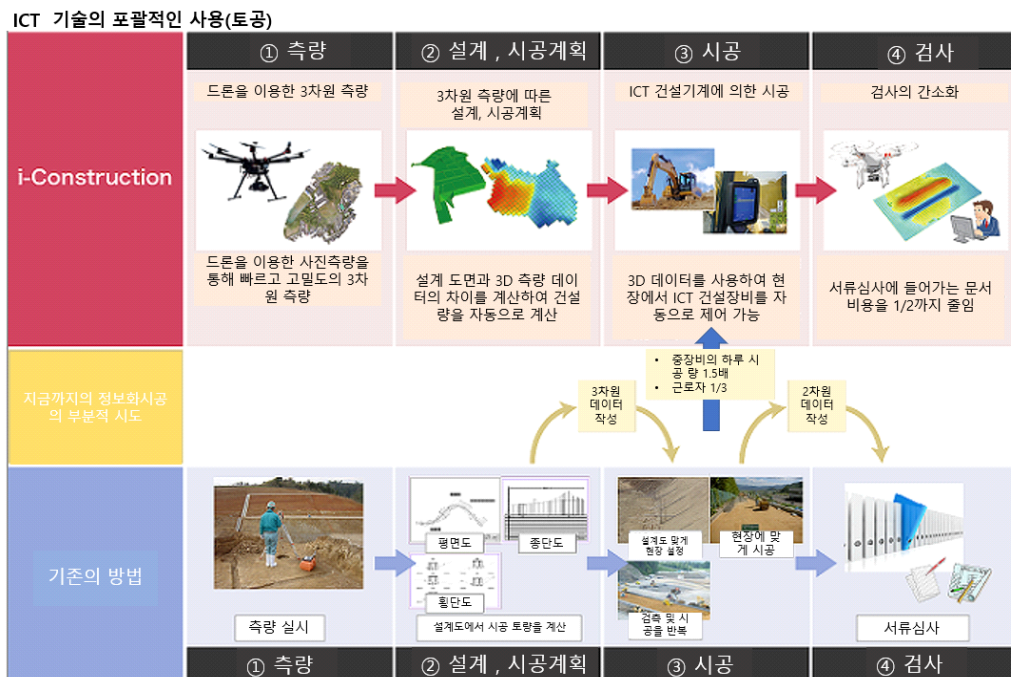
관리 할 수 있다. Trimble connect는 사용자가 현재 편집중인 파일을 잠글 수 있도록 Check in, Check out 기능을 제공한다.

(4) 현황 및 결론

Trimble사는 일본정부에서 추진하는 I-Construction 사업에 맞추어 Hitachi사, Komatsu사의 스마트 건축 플랫폼과의 API 개발을 진행하고 있으며 이와 같은 협업을 통해서 일본의 i-construction 프로젝트 이니셔티브에도 기여하고 있다. [22][23]

i-construction은 일본 국토교통성에서 2016년부터 진행한 생산성 혁명 프로젝트로써 일본의 건설 인력의 감소에 대응하기 위한 건설현장의 생산성 향상 프로젝트이다[24]. 측량, 시공, 검사 부분에서 ICT기술을 적용하였으며, 2025년까지 건설현장의 생산성을 20%향상을 목표로 진행 중에 있다.

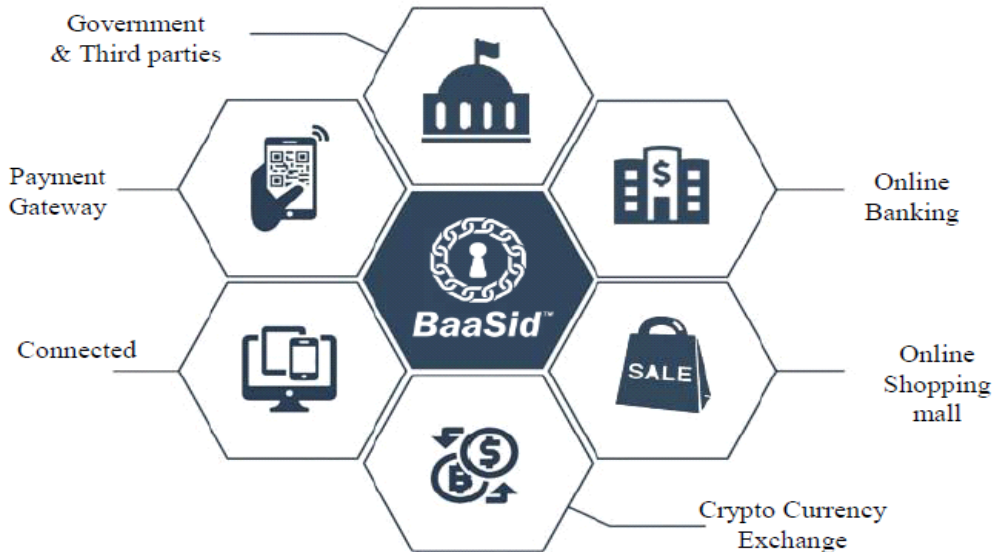
〈그림 39〉 일본 I-Construction



출처 : I-Construction. Available at <http://www.mlit.go.jp/tec/i-construction/index.html>

5 》블록체인 공공 네트워크 기반의 인스턴트 개인인증 플랫폼[25]

〈그림 40〉 인증을 위한 서비스로서의 블록체인



출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

(1) 배경 및 문제점

1968년 미국 국방 프로젝트를 위해 만든 컴퓨터 통신망(ARPANet)은 1990년대 Microsoft사의 주도로 웹 기반의 PC어플리케이션을 탄생시켰고 주요 인터넷 서비스들이 출시되는 web 1.0시대를 만들었다[26].

2000년대 중반에 이르러 아이폰의 iOS와 구글의 Android의 등장은 사용자를 Mobile로 이동하게 만들었고 초기의 정보 검색에서부터 의사소통, 교육, 언론, 광고, 오락 및 여가생활에 이르기까지 사회 전반으로 그 기능이 확대되어 다양한 정치적, 경제적, 문화적, 사회적 변화를 창출하고 있는 등 삶의 모든 부분에 영향을 미치고 있다.

인터넷 서비스의 발전에 따라 개인화된 서비스도 가능하게 되었고, 중앙화된 데이터베이스에 저장된 개인의 다양한 정보는 기업의 수익 창출에 큰 기여를 하게 된다. 즉, 정보가

돈이 되는 세상이 온 것이다. 이에 따라 서비스 공급자는 이러한 소중한 정보자산을 지키기 위해 많은 노력을 기울여 왔다.

그러나 지난 수십년간 개인의 다양한 정보는 해킹, 분실, 변조 등의 방식으로 유출되었고, 특히나 중앙화된 집단(정부, 은행, 기관, 기업 등)의 데이터베이스에 수없이 많은 해킹 사례가 발생하였다.

이에 따라 기존 중앙화된 방식의 보안, 인증의 한계를 극복할 수 있는 기술이 필요하게 되었다. 최근 그에 대한 대안으로 최근 블록체인 기술이 관심을 받고 있다.

(2) 프로젝트의 목적

지금까지의 인터넷 상의 개인 인증 구조와는 정반대의 방식으로 블록체인과 비블록체인 간의 연계, 인증을 위한 원본의 부존재, 제3의 인증기관의 공인인증서 등이 필요없이 순수한 참여자(유저 혹은 Node)의 공공 네트워크만을 통한 탈 중앙화 데이터베이스 개발이 목표이다.

- 개인 민감정보(이름, 전화번호, 아이디, 생체정보 등)를 잘 개 쪼개(Split)서 이를 수백, 수천 개의 노드에 분리, 분산 저장하고 원본은 삭제한다.
- 인증, 로그인시 '생체 인증키(지문 등)'를 통해 공공네트워크 인증으로부터 자신의 쪼개진 수천 개의 암호화된 개인 정보 조각(Split ID)들을 불러들여 복호화, 조합을 통해 '1회성 인스턴트 인증(Instant Access)'하고 인증이 끝나는 순간 즉시 폐기하며 사본은 삭제한다.

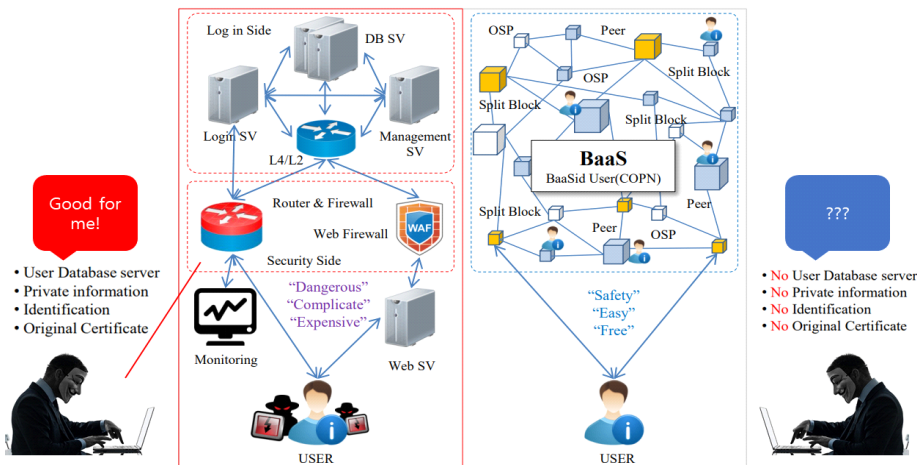
(3) 기술 상세

인터넷상의 모든 서비스는 수 많은 개인 정보가 데이터베이스에 저장되고 사용되며 이를 기반으로 로그인, 정보 확인, 결제 등 기타 모든 사용자의 요청을 공급자와 제3의 인증기관이 인증한다. 이때 중심이 되는 것이 데이터베이스인데, 데이터베이스를 지켜내기 위해 다양한 보안 시스템을 구축, 운영하고 있다.

또한 각국 정부는 개인정보 보호법령을 제정하고 개인정보 유출과 도용, 거래를 막기 위해 개인, 인터넷 서비스 기업, 금융, 기관, 단체 등에 많은 노력을 요구하고 있다. 하지만 중앙화된 데이터베이스는 결국 보안 취약점을 노린 해커의 대상이 되고 한계점에 다다른 상황이다.

블록체인 기술 역시 거래 기록을 동일하게 저장하여 위·변조에 대한 강한 보안성을 제공하지만, 이것은 모두 같은 데이터를 가지고 있어 유저의 민감한 개인 정보와 중요정보, 인증키 등을 저장하기에는 무리가 있다. Baasid 플랫폼은 참여자 개인인증에 필요한 정보는 인증시 대조할 원본조차 그 어떤 서버나 유저 디바이스에도 저장하고 있지 않아 근본적으로 해킹 시도를 차단한다. 오직 공공의 참여자들(사용자들)의 노드에 조각으로 분리 분산 저장되어있을 뿐이다.

〈그림 41〉 중앙화DB시스템과 탈중앙화 블록체인 시스템 비교



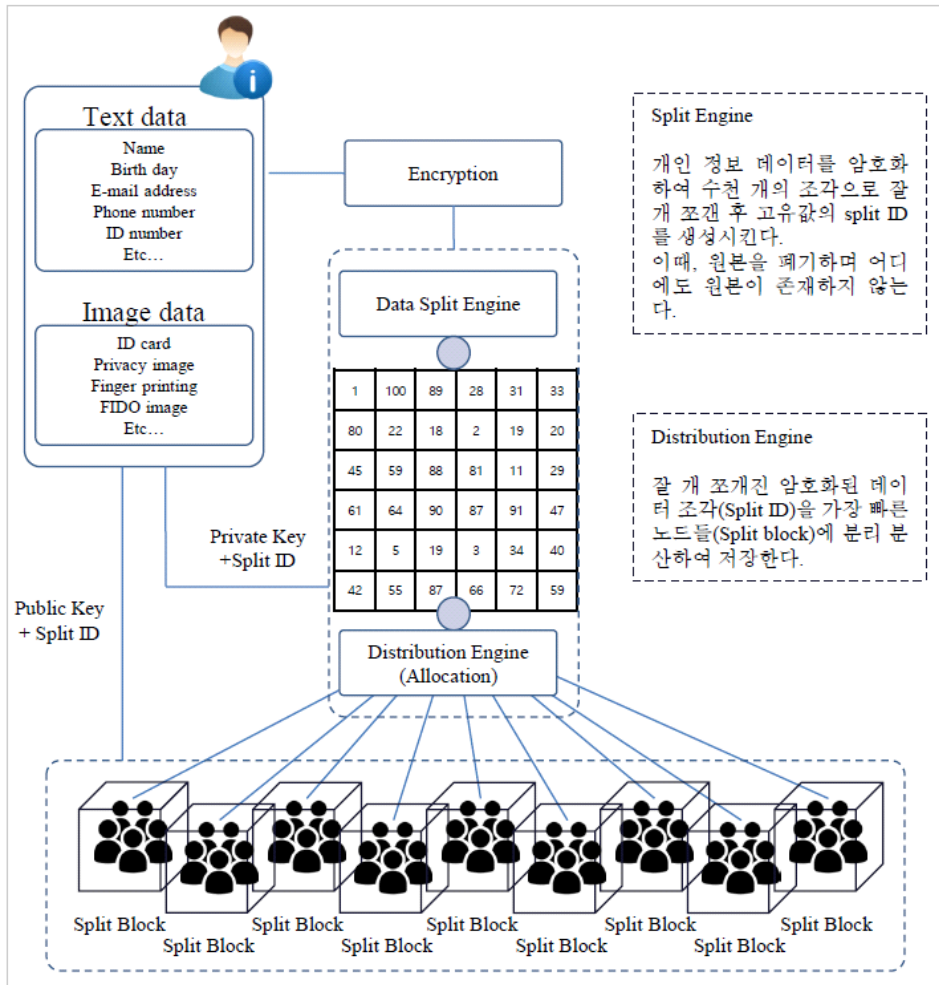
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

① Split & Distribute Engine

Split Engine은 개인 정보 데이터를 암호화 하여 수천 개의 조각으로 잘게 쪼갬 후 고유 값의 split ID를 생성시킨다.

Distribution Engine은 잘게 쪼개진 암호화된 데이터 조각인 Split ID를 가장 빠른 노드(Split block)에 분산하여 저장한다.

〈그림 42〉 BaaSId 인증 통합 서비스



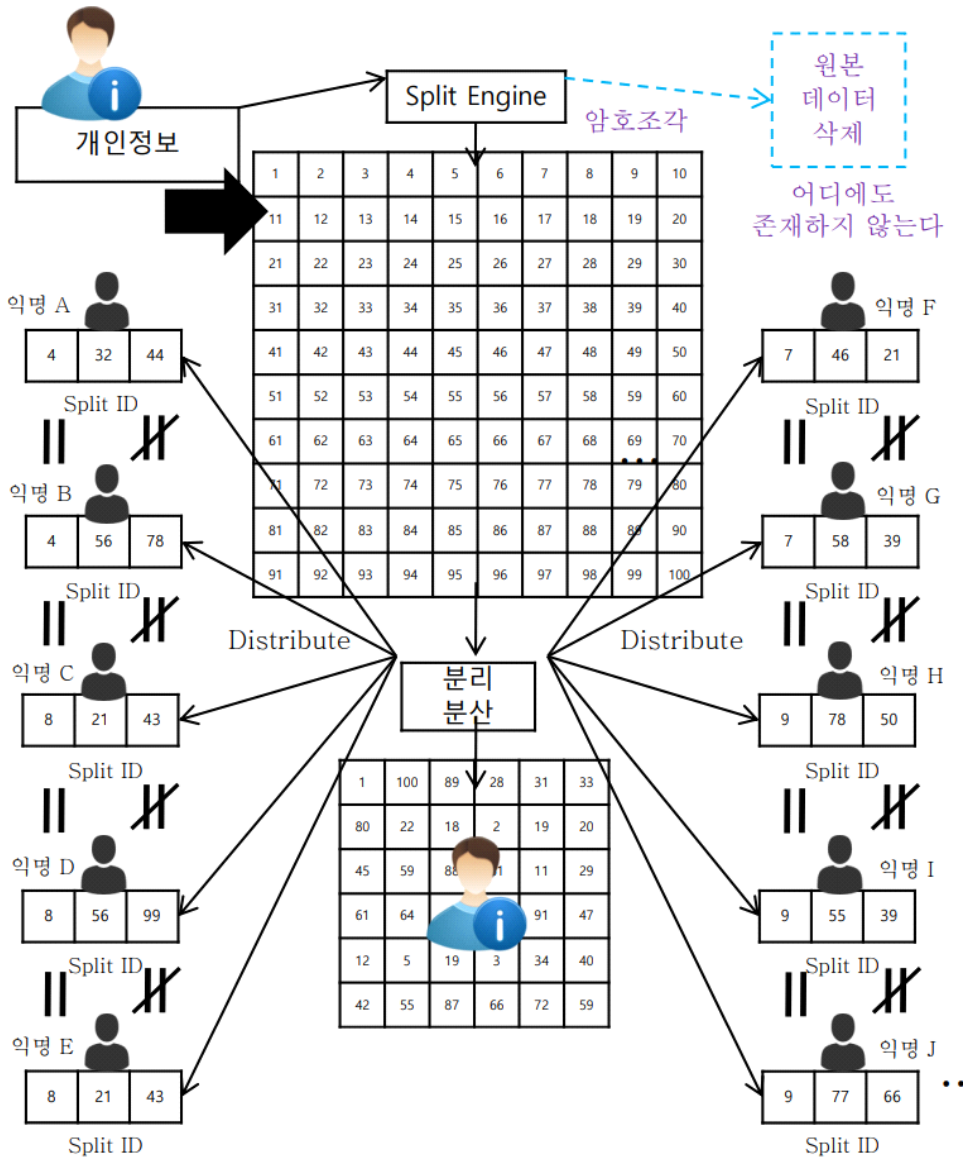
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

② Split ID

참여자의 개인 정보는 Split Engine에 의해 수천 개의 조각으로 분리되며 조각난 모든 Split ID는 암호화되어 수천 명의 Split Block 안에서 서로 다른 Split ID가 저장된다.

이는 P2P의 가장 큰 장점인 분리된 데이터의 할당 값과 분산된 조각들을 합리적으로 파악한 후 가장 빠른 노드와 데이터를 조합하여 분산된 조각들을 하나의 파일처럼 다운받거나 복구할 수 있도록 해주는 프로세스와 유사하다.

〈그림 43〉 Split ID의 생성과 분리 분산



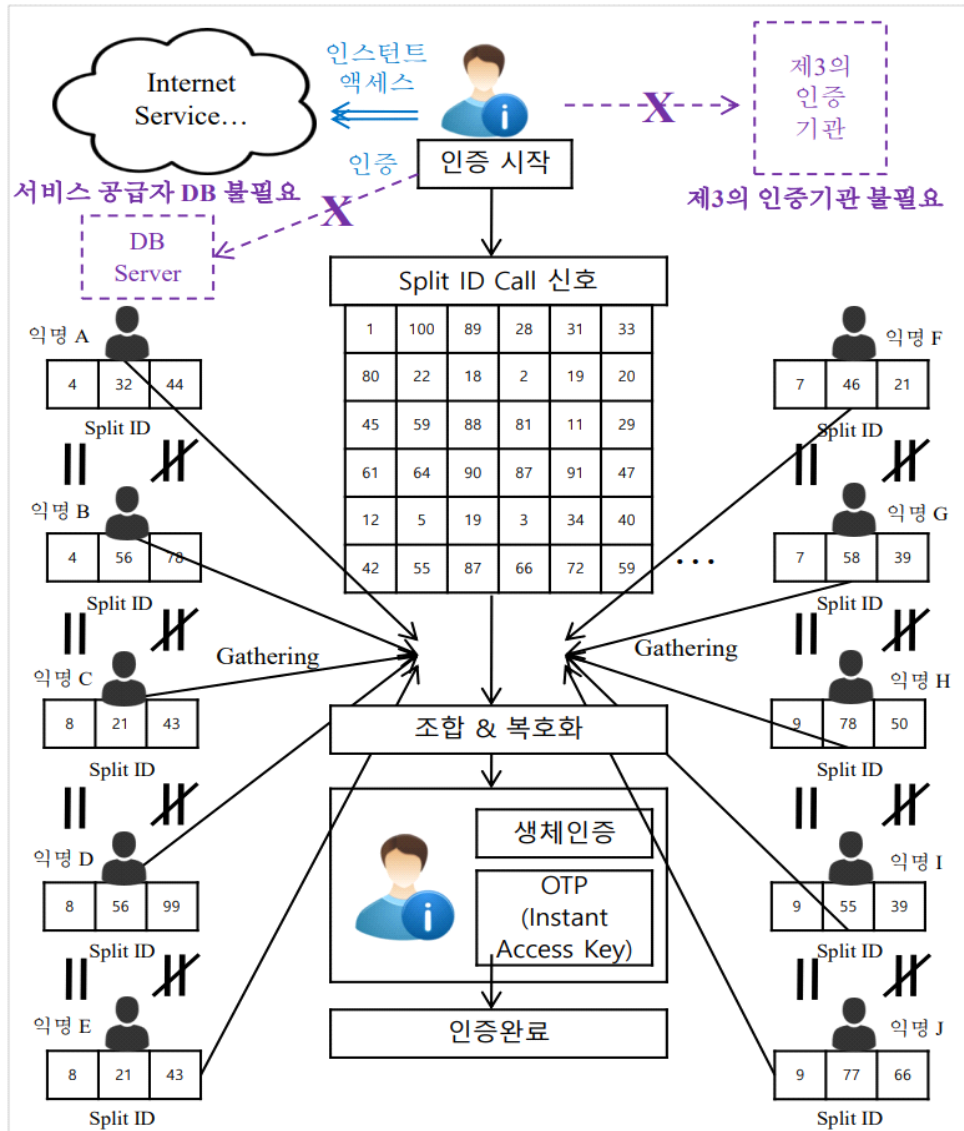
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

③ Combination Engine 복호화 및 인증단계

참여자들의 Split ID들은 Public Key와 함께 각각의 고유한 암호코드로 유지되며 이러한 고유 Split data들은 지속적으로 안전하게 유지된다.

특정 인터넷 사이트에 접속할 경우 일시적인 1회성 인스턴트 인증을 서로 부여해주면서 인증을 증명하고 참여한다.

〈그림 44〉 Split ID의 참여 및 증명



출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

④ One pass ; 복잡한 인증을 하나로

BaaSid는 불특정 다수의 참여자들의 동의로 모든 인증을 허가하고 증명한다.

BaaS API는 공급자들에게 개인정보 보호와 관리책임을 줄여주며, 복잡한 개인정보 정책, 시행 및 단계에서 공급자들을 해방할 수 있는 다양한 서비스를 제공한다.

〈그림 45〉 BaaSid One pass



출처 : BaaSID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

⑤ BaaS Certification API Service

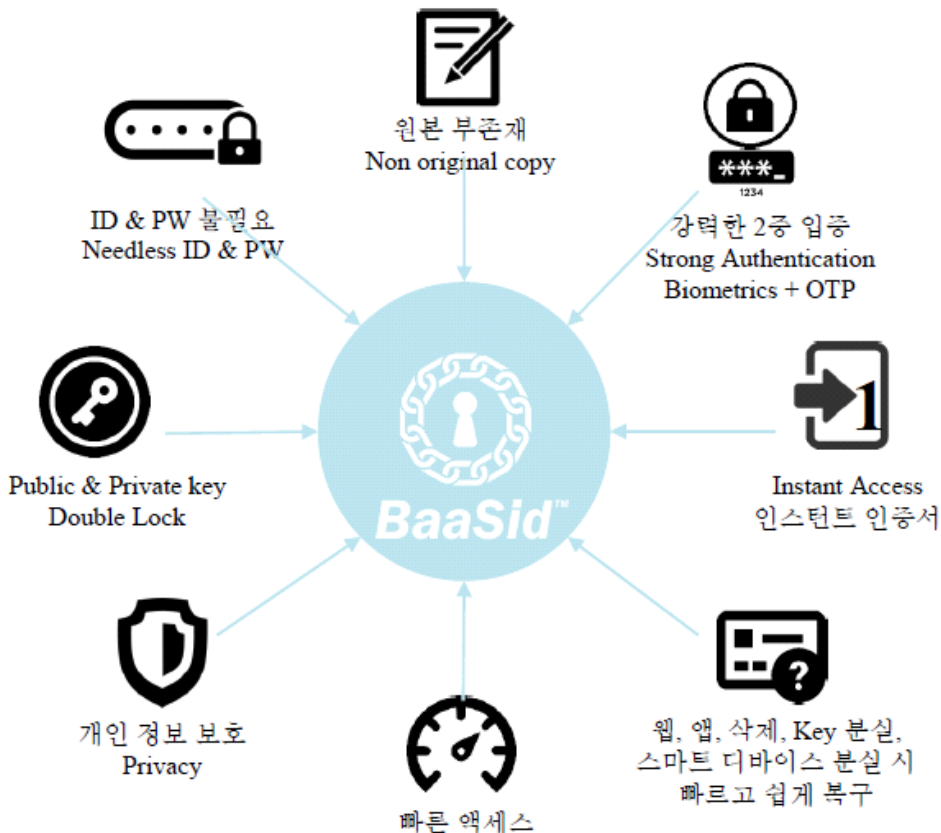
BaaSid는 개인의 Privacy보호를 위해 불특정 다수의 노드에 각각 서로 다른 고유 값을 갖는 작은 조각들을 임의의 분리 분산을 통해 저장한다.

BAAS토큰 거래, 로그인, 각종 인증 등을 통해 증명과 신뢰원장을 생성하고 보존한다.

이와 같은 방식은 POA(Proof of Access)라는 증명방식으로 사용자의 인터넷 서비스 로그인이나 인터넷상의 주요활동 등을 통해 주기적으로 활용도를 증명하는 적극적인 생태계를 말한다.

BaaS 기반의 공공네트워크인증(COPN : Certification of public Network) API는 사용자들의 적극적이고 자연스러운 활동을 기반으로 증명하고 지속적으로 연결된다.

〈그림 46〉 BaaSId API Service



출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

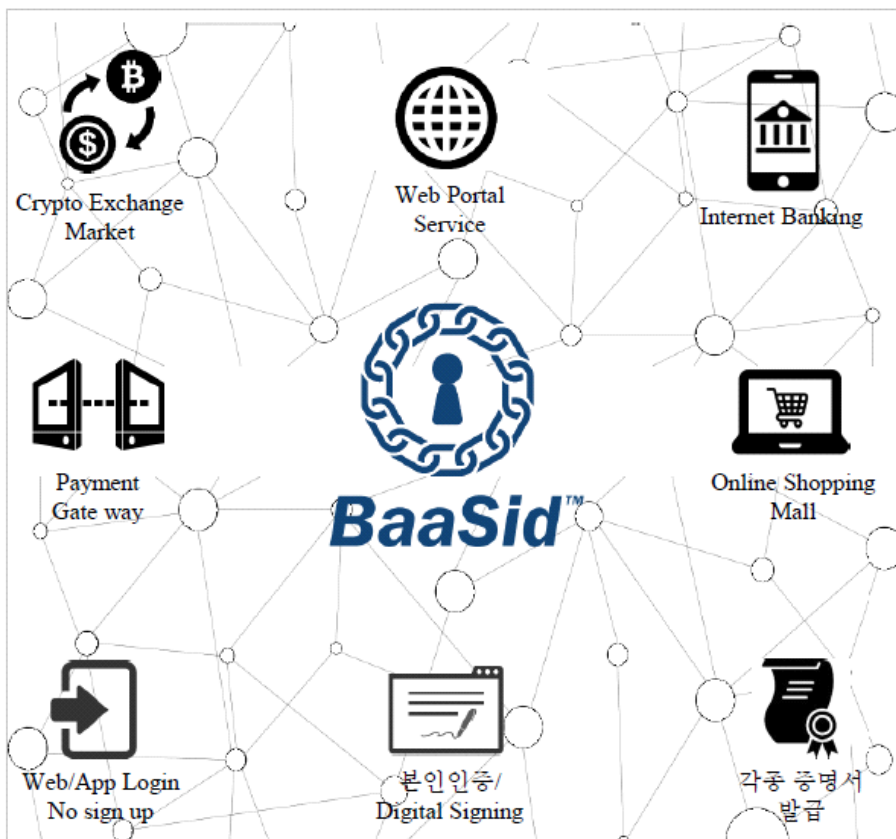
⑥ DB Governance

BaaSID는 공급자가 블록체인, 웹 또는 APP을 기반으로 하는 서비스인지 구분하지 않는다. 이는 모든 서비스에는 사용자를 필요로 하며 사용자는 공급자에게 어떤 형태로든 이익을 가져다주기 때문이다.

BaaSID의 참여자들은 적극적이고 잠재력 있는 고객이며, 서로를 신뢰하고 증명해주는 로그인 및 인증을 위한 데이터베이스 그 자체이다.

BaaSID는 이처럼 준비된 잠재적인 고객 인프라인 동시에 사용자이며 공공 인증을 다함께 처리하는 인증 주체의 거버넌스(governance) 이다.

〈그림 47〉 BaaSID API Service



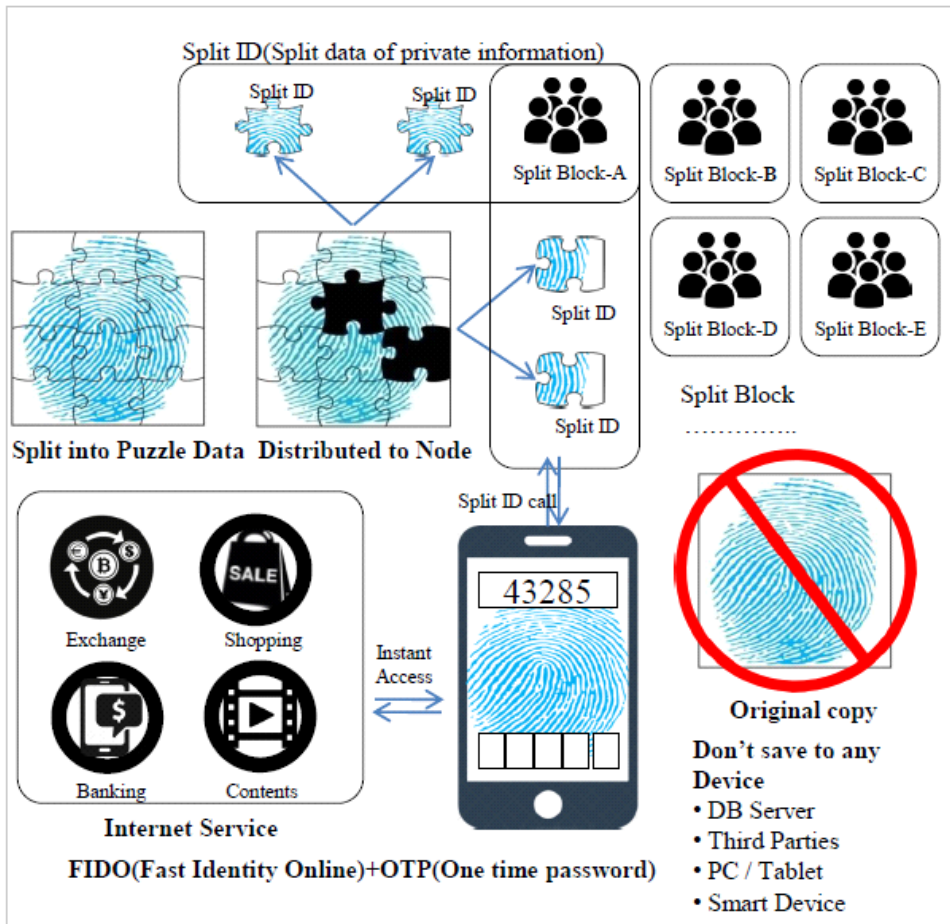
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

⑦ Split & Distribute Data

COPN(Certification of Public Network)의 모든 참여자의 개인정보는 수천 개의 조각으로 쪼개진다. 쪼개진 개인정보는 임의의 같은 Split Block에 참여된 불특정 다수에게 각각 다른 암호화 된 데이터 조각들로 나누어 저장된다. 이렇게 잘게 쪼개진 개인정보 조각을 Split ID라 한다.

블록체인 기반으로 분산된 Split block의 노드 숫자와 Split ID의 용량, 분리 및 분산을 최적화 하여 실시간 인증속도를 높인다.

〈그림 48〉 Instant Access 인증

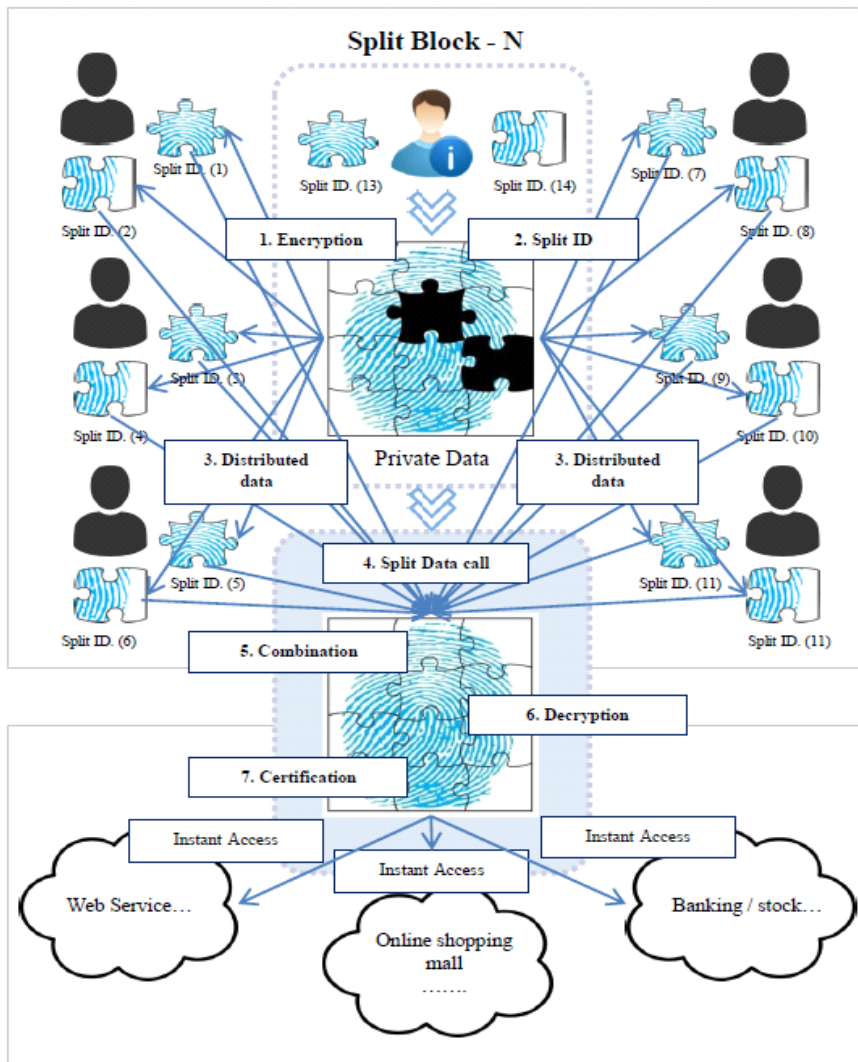


출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

⑧ 공공 네트워크인증(COPN)을 바탕으로 하는 사용자 중심의 인증

인터넷 서비스에서 로그인, 회원가입, 쇼핑, 송금, 금융자산 이동 등의 중요 행위를 할 때 Split ID가 저장된 Split block에서 다시 일시적으로 결합하여 생체 인증키(지문, 홍채, 목소리), 패스워드 또는 OTP를 사용한 인스턴트 인증을 한 후 원본은 즉시 폐기된다.

〈그림 49〉 개인을 중심으로 하는 일시적 중앙화



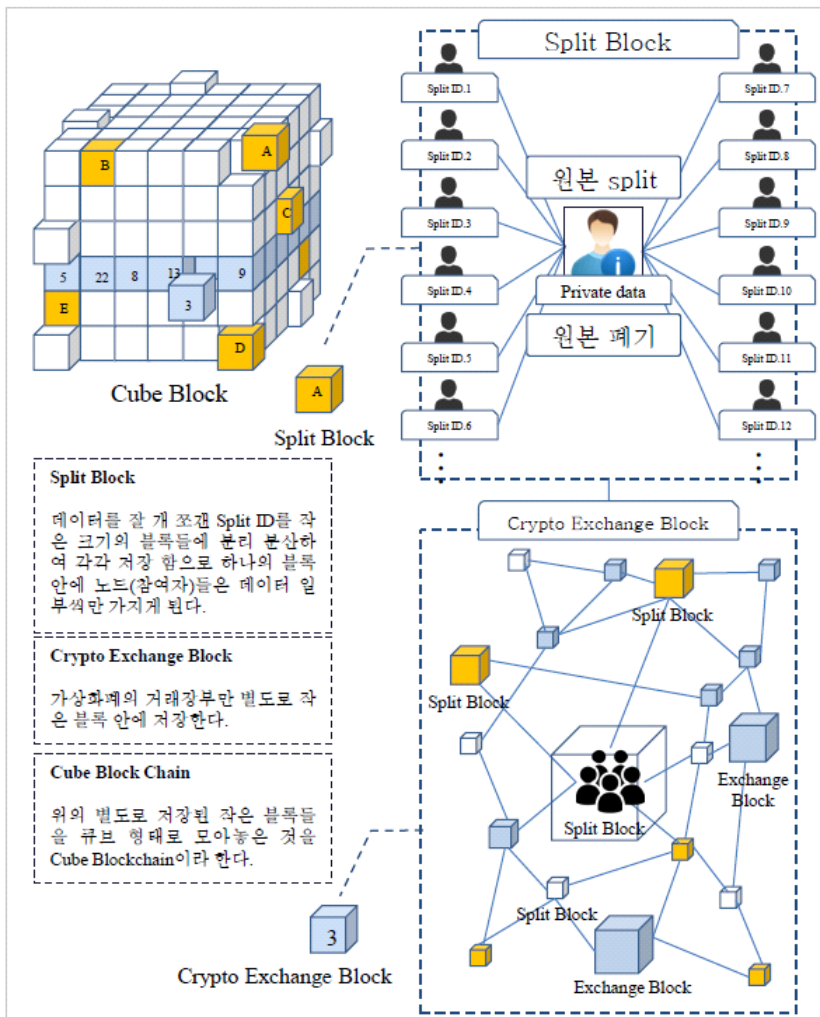
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

⑨ Split Block과 Crypto Exchange Block

고유의 Split ID는 참여자들 각각의 독창성을 인정하듯 서로 다른 암호화된 데이터를 말한다. 이 Split ID를 보유한 하나의 작은 참여자 그룹의 블록을 Split Block이라 한다.

참여자들이 BAAS 토큰을 거래하는 모든 공공 거래장부는 Crypto Exchange Block에 저장된다.

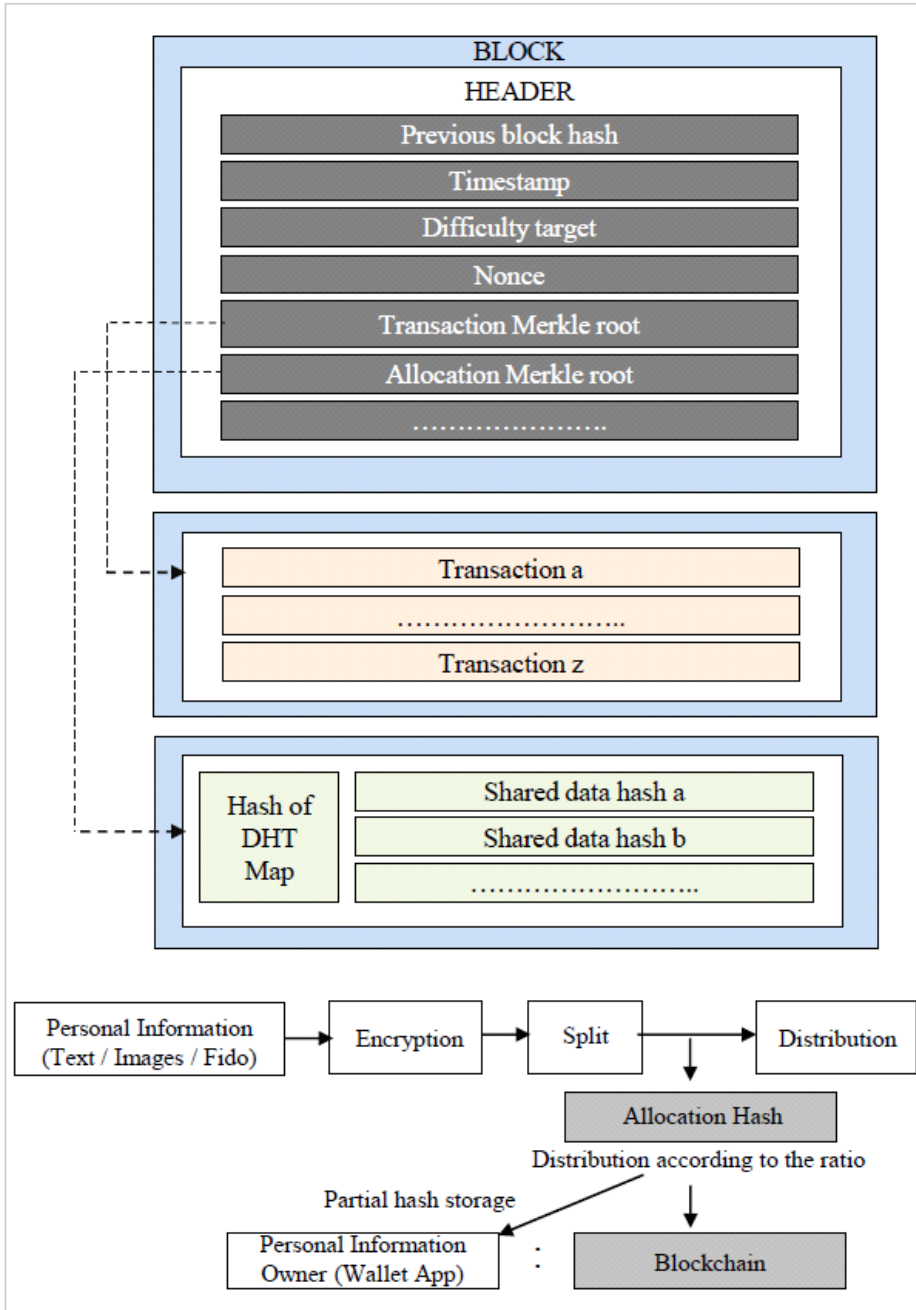
〈그림 50〉 Split Block & Crypto Exchange Block



출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

⑩ Split & Distribute Data의 무결성(integrity) 증명

〈그림 51〉 BaaSid Block 구조



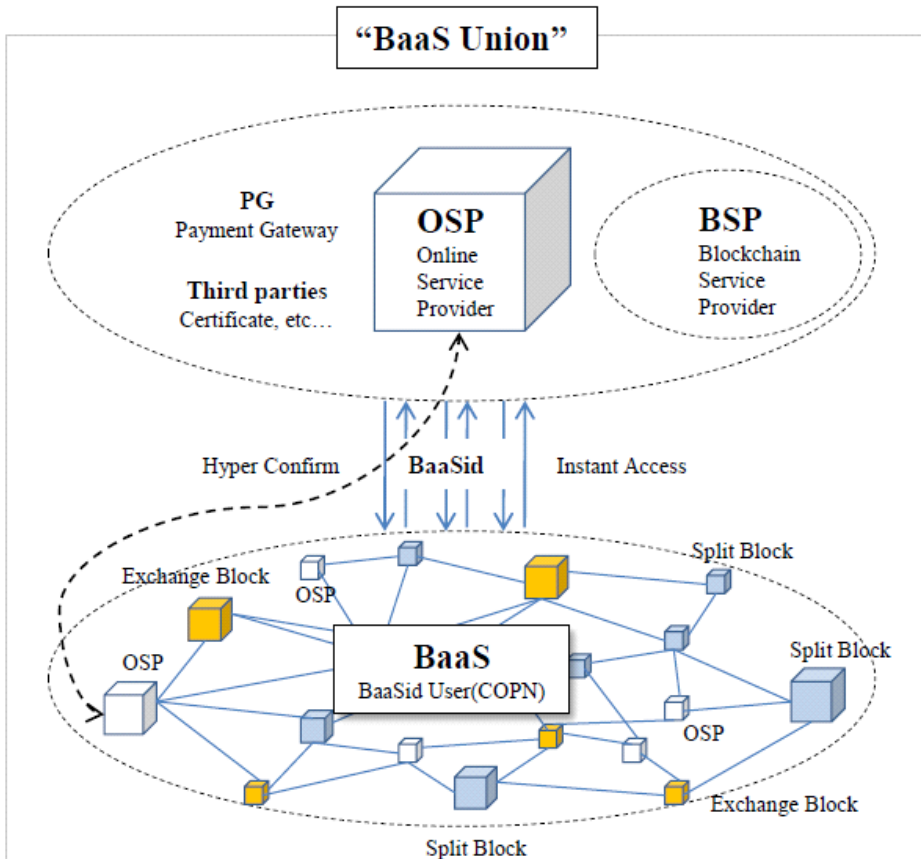
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

⑪ BaaS Union

BaaS를 기반으로 기존 온라인 서비스 공급자(OSP), 블록체인 서비스 사업자(BSP), 그리고 COPN에 참여한 모든 사용자들은 서로 연결되고 신뢰할 수 있는 거대한 협력 공동체인 BaaS Union을 구축함으로써 서로의 신뢰와 균형을 유지한다.

BaaS에 참여하는 온라인 서비스 공급자, 블록체인 서비스 사업자, 일반 참여자들이 모두가 자신의 역할, 책임, 활동, 공동체에 대한 연동을 통해 서로 보상을 주고받을 수 있도록 되어있다.

〈그림 52〉 BaaS Union을 통한 서로 간의 신뢰, 책임과 보상



출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

⑫ 참여자 및 온라인 서비스 공급자의 Baasid를 통한 이점

i. 막대한 데이터베이스와 보안시스템의 구축비용과 유지비용의 절감

BaaSid의 COPN API는 모든 인터넷 서비스 공급자에게 제공된다.

인터넷 서비스 공급자는 자사의 서비스를 위해 데이터베이스 구축, 운영 및 보안시스템을 구축하는 비용을 줄일 수 있다.

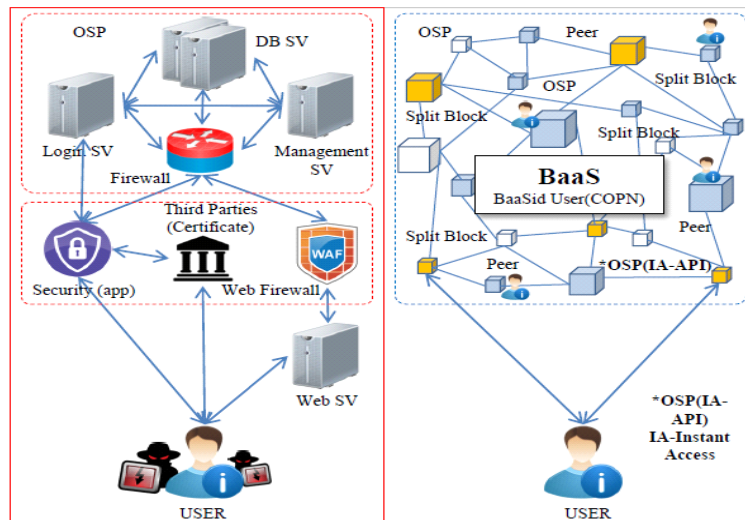
ii. 개인정보 유출에 대한 법률적 리스크 제거

BaaSid의 COPN을 적용하는 모든 인터넷 서비스 공급자는 별도의 데이터베이스 구축을 통해 개인정보를 저장하지 않기 때문에 해킹시도와 유출에 대한 법률적인 리스크가 존재 하지 않는다.

iii. 회원가입의 절차나 인증단계가 간편한 Instant Access API

BaaSid의 참여자는 본인의 생체정보(지문, 홍채)인증을 암호화>조각화>호출>조합>복호화>생체인증 순으로 이뤄지는 일시적인 중앙화를 통해서 1회성의 접근 권한(Instant Access)을 갖는다.

〈그림 53〉 일반 로그인/인증과 Instant Access 비교



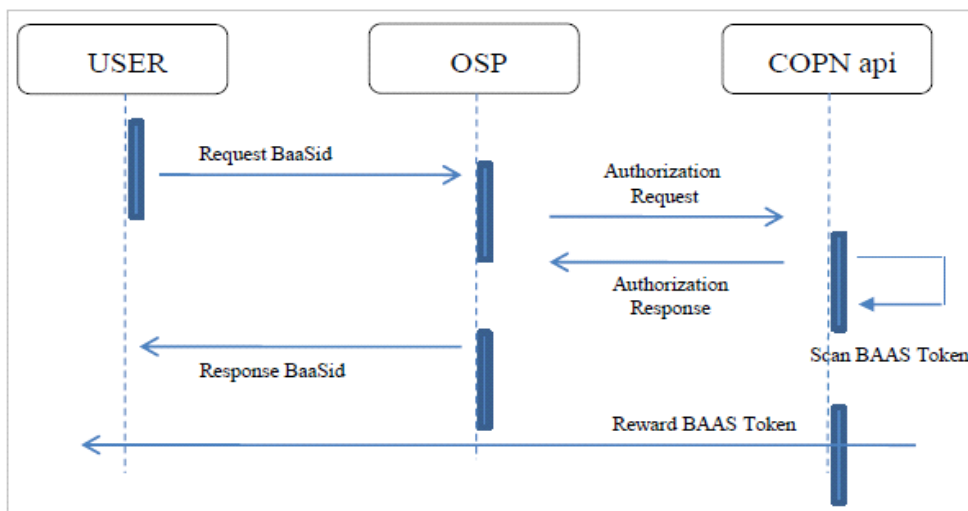
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

유저 스스로가 회원가입이나 기타 인증단계 없이 바로 서비스에 Access할 수 있다.

서비스 공급자는 회원단계에서 이탈하는 고객을 확보 할 수 있고, 로그인 및 인증을 위한 데이터베이스를 구축하는 번거로움을 줄일 수 있다.

BaaSID의 COPN을 기반으로 IA-API(Instant Access API)를 2019년 1분기에 개발을 완료하여 배포할 예정이다.

〈그림 54〉 Instant Access API 구성도



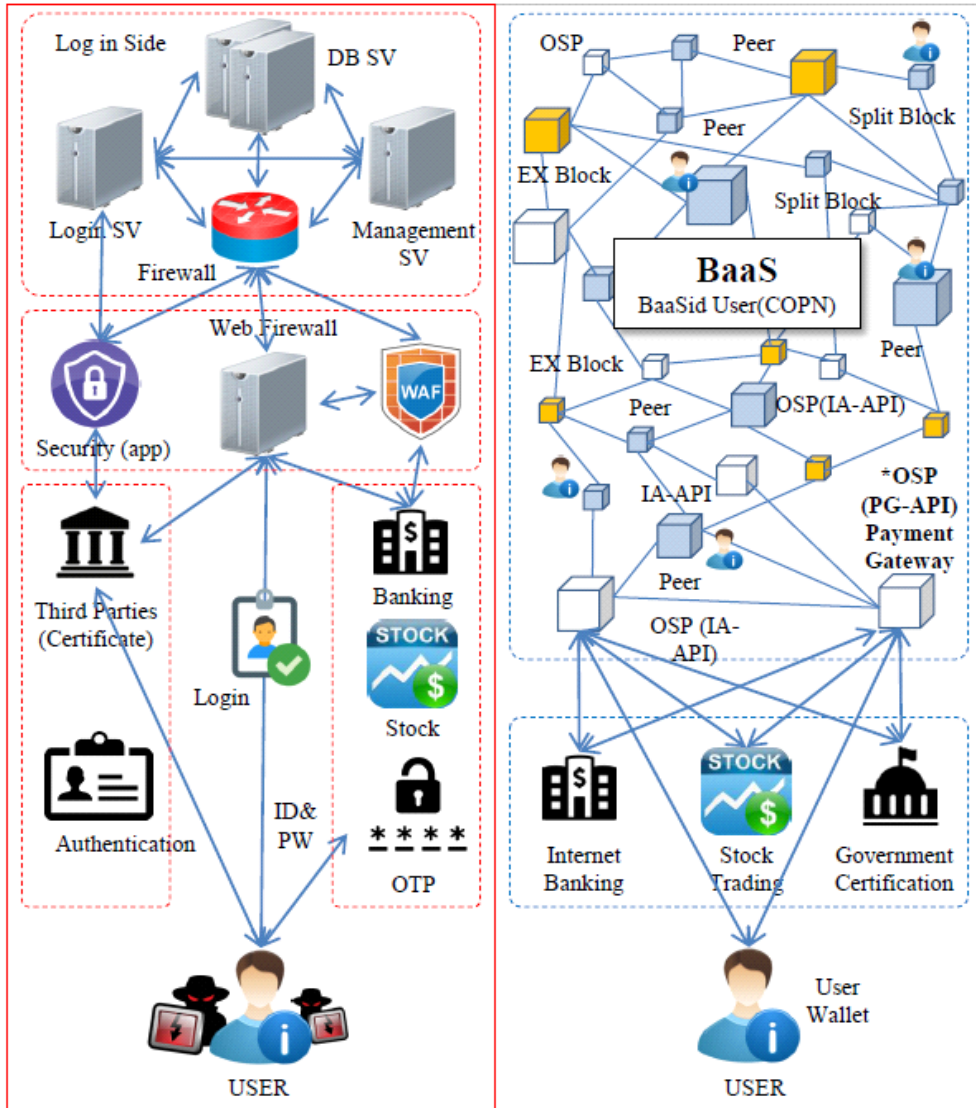
출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

iv. 빠르고 쉬운 중요 인증 Hyper Confirm API

모든 은행, 금융, 증권사, 온라인쇼핑, 기타 콘텐츠의 유료 결제에는 개인 정보와 관련된 제 3인증기관의 인증절차, PG(Payment Gateway)사 연동 등의 다양하고 복잡한 단계를 요구한다.

BaaSID는 COPN인프라를 통해 복잡하고 비효율적인 현재의 중요 인증 단계와 절차를 간편하게 해결 할 수 있는 HC-API(Hyper Confirm API)를 개발하여 제공할 예정이다.

〈그림 55〉 일반 인터넷 결제 인증과 Hyper Confirm 비교

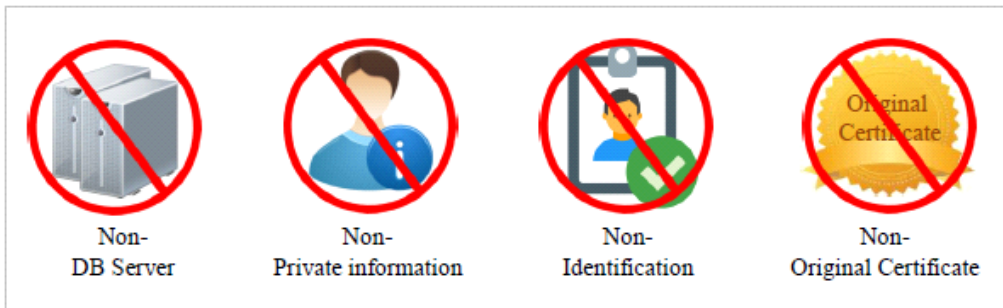


출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

⑬ BaaS 사용자 데이터베이스

BaaS의 목표는 인터넷상의 모든 서비스 공급자의 데이터베이스, 국가의 인증기관, 제 3의 인증기관, 사용자의 모든 저장장치 등에도 사용자의 개인정보와 어떤 원본도 저장하지 않고 인터넷 서비스에 Instant Access가 가능하게 만드는 것이다.

〈그림 56〉 BasSid에서 존재하지 않는 4가지



출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

i. Non-existence of User Database

- 인터넷 서비스 공급자가 운영하는 데이터베이스가 존재하지 않아 탈중앙화를 실현 할 수 있다.

ii. Non-existence of Private information

- 자신의 이름, 전화번호, 이메일 주소, 통장번호, 신용카드번호 등의 민감한 개인 정보를 잘게 쪼개어 가장 최적화된 노드 그룹에서 분리 및 분산하여 저장함으로 개인 정보의 원본을 남기지 않는다.

iii. Non-existence of Identification

- 자신임을 확인하는 정보, 인증 정보 등이 어디에도 저장되지 않음으로 해킹이나 분실의 위험에서도 안전하다.

iv. Non-existence of Original Certificate

- 제3의 인증기관의 서버, 유저의 모든 디바이스(스마트 폰, PC, 태블릿 PC)에 어떠한 인증서나 원본데이터가 존재 하지 않는다.

⑭ 거래 검증 주체 (Transaction validator)

BaaSid의 모든 블록 그룹(Split Block, Exchange Block)의 로그인, 인증 등의 모든 인터넷 활동 내역과 거래장부들의 거래검증은 마이너와 별도로 참여자가 포함된다.

BaaSID는 기본적으로 이더리움의 마이닝 정책을 따르나 PoA와 PoH라는 별도의 증명 방식을 가지고 있다.

i. PoW

- PoW는 이더리움 마이닝을 그대로 이용하여 블록 생성의 보상으로 BAAS토큰을 지급한다. 또한 거래 수수료를 PoW마이너에게 별도로 지급하며 PoS방식으로 변경될 예정이다.

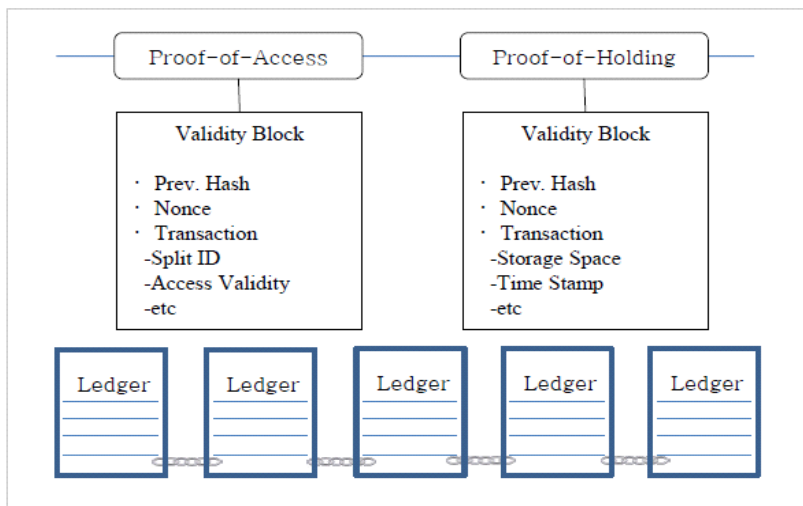
ii. PoA(Proof of Access)

- 온라인 서비스 제공자나 블록체인 서비스 제공자는 참여자의 Instance Access가 이루어지면, 이에 대한 수수료 일부를 BaaSID의 데이터보관 제공자(PoH 마이너)에게 지급한다.

iii. PoH

- COPN의 IPFS(Interplanetary file system)에서 Split ID관리를 위한 별도의 블록을 생성하며 블록이 신규 생성 될 때 BAAS토큰을 지급한다.

〈그림 57〉 PoA 와 PoH



출처 : BaasID, “블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼”, Apr. 2017

⑮ BaaS

BaaS 블록체인은 인터넷상에서 개인정보를 저장하여 로그인, 결제, 기타 모든 인증 절차를 필요로 하는 공급자들이 별도의 데이터베이스를 구축하거나 운영하지 않고도 사용자들 간의 BaaS Blockchain Network를 통해 신뢰 높은 인증을 간편하게 제공할 수 있다.

모든 개인들이 연결된 BaaS Blockchain의 네트워크는 하나의 거대한 인증시스템을 의미한다.

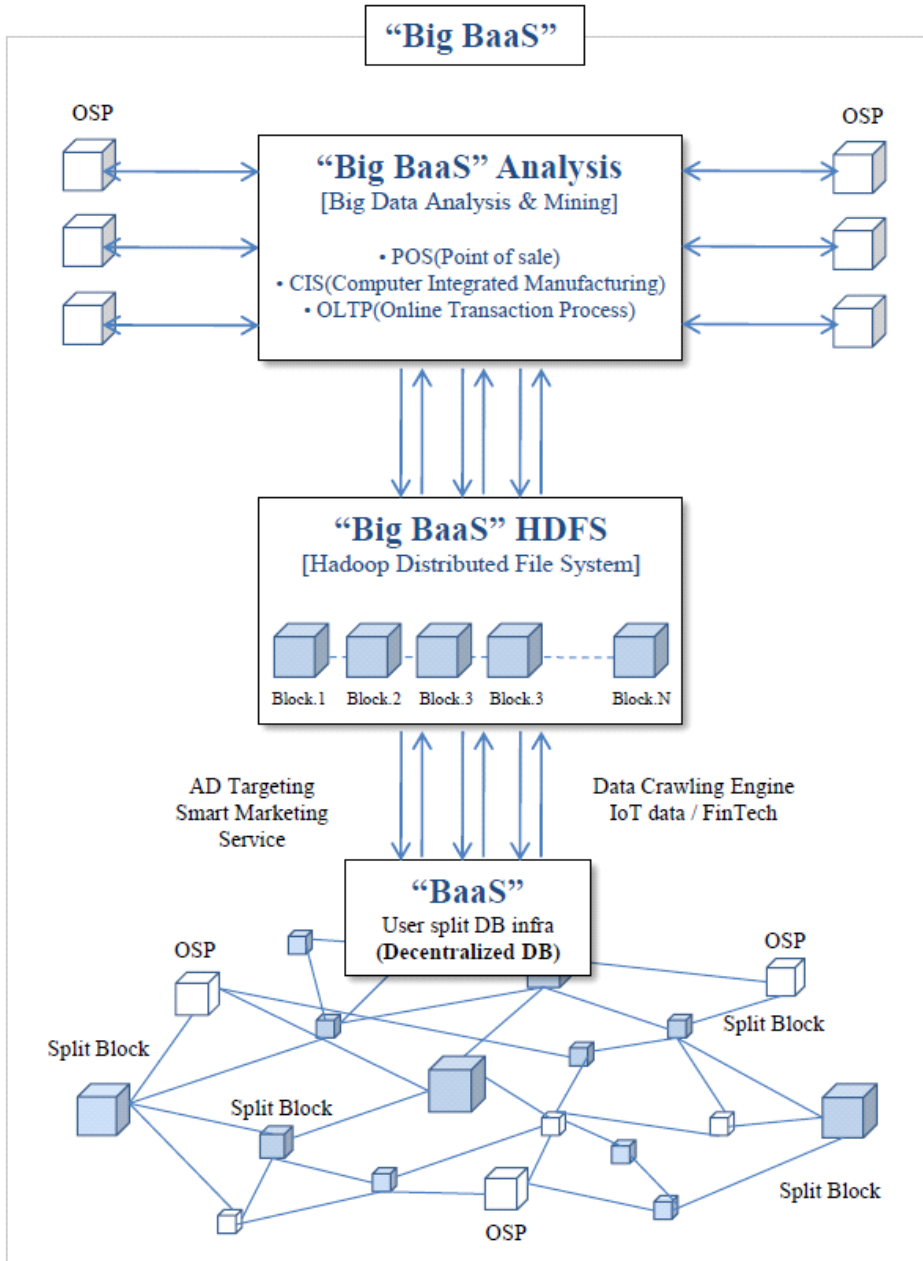
기업 및 기관은 별도의 중앙화 된 데이터베이스, 인증의 권한 및 책임에서 벗어나 BaaS Blockchain 기반의 공공네트워크 인증을 통해 사용자의 인증활동을 진행 시킬 수 있다.

⑯ Big Data Service BigBaaS

BaaS를 사용하는 인터넷 서비스 공급자들은 자체적인 데이터베이스를 보유하지 않으므로 독자적인 각종 프로모션, 이벤트 등 DB가 필요한 서비스를 할 수 없게 된다. 때문에 BaaS에서 개발 중인 BigBaaS는 참여자들의 인스턴트 인증 정보를 모은 다양한 행동 데이터들과 인터넷 서비스에 한번이라도 인증한 참여자들에게 이메일 및 각종 알림 서비스를 인터넷 서비스 공급자들에게 제공한다.

BaaS 참여자(Node 혹은 Device)가 인증한 웹 사이트가 무엇인지, 무엇을 봤는지, 얼마나 머물렀는지, 무엇을 샀는지 등을 익명성과 개인의 프라이버시를 안전하게 보장하면서 기록할 수 있고 이를 바탕으로 명확한 타겟 마케팅을 할 수 있다.

〈그림 58〉 BigBaaS의 HDFS와 분석 시스템



출처 : BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017

(4) 결론

현재 인터넷상의 서비스에서 데이터베이스를 바탕으로 하는 개인인증구조는 보안에 취약하다. 서비스 업자들은 중앙화 된 데이터베이스의 보안을 위해서 막대한 비용을 투자하여 보안설계를 하지만 항상 해킹의 위험에 노출되어있다. 또한 보안을 강화하기 위해 사용되는 2차로그인, OTP나 공인증서 같은 추가 수단은 사용자로 하여금 불편함을 느끼게 한다. BaaSid는 블록체인 기술을 이용하여 개인정보를 잘게 쪼개어 나누고 나뉜 데이터를 임의의 블록에 저장한다. 또한 인증과정에서도 나뉜 데이터를 임시로 결합하여 사용하고 제거함으로 데이터베이스를 필요로 하지 않는다. 해커가 데이터를 해킹하기 위해선 51% 이상의 익명의 블록을 해킹해야 하고, 해킹 속도 또한 새로운 블록이 생성되는 속도보다 빨라야하기 때문에 현존하는 컴퓨터로 해킹이 불가능하다. 따라서 블록체인 기술을 이용한 개인인증은 사용자에게는 편의성과 개인정보 보호를, 서비스 공급자에게는 데이터베이스를 구축하고 유지하는데 드는 비용을 절감시켜 줄 수 있다. 또한 익명의 사용자의 소비패턴 등을 이용하여 마케팅에 이용할 수 있다. 이를 통해 인터넷 환경을 간편하고 안전하게 만들 수 있다. 최근 로봇 청소기, 냉장고, 홈 CCTV등 사물인터넷을 활용한 스마트 홈 기기에서 개인정보 침해되는 사례가 자주 발생하고 있다. 이처럼 4차 산업중 하나인 IoT기술을 이용하는 데 있어 개인정보 보호문제는 해결해야 하는 문제이다. 때문에 블록체인기술을 이용한 개인정보 보호를 통해 안전하고 활성화된 IoT사회를 구현할 수 있다.

6 전자상거래 물품 통관 데이터 공유 [27]

〈그림 59〉 블록체인 기술을 이용한 수출통관 물류 서비스



• 관세청 전자통관시스템(UNI-PASS)으로 실시간 연계되어 수출신고 심사 수행

출처 : 관세청, “18년도 ICT기반 공공서비스 촉진사업 제안요청서”, Apr. 2018

(1) 배경 및 문제점

현재 해외 전자상거래 물품 통관 과정은 몇 가지 불편함이 있다. 사례를 통해 보면 운송업체의 경우 전자상거래업체로부터 전달 받은 주문내역을 근거로 통관을 진행할 때, 일일이 문서파일을 직접 확인하여 진행하며 누락항목 발생 시 전자상거래업체에 다시 자료를 요청함으로써 통관신고가 지연되며, 인력 또한 많이 투입된다. 또한 소비자의 경우 화물의 위치를 파악에 같은 내용을 세관, 운송업체등 여러 곳에 문의해야한다. 또한 현행구조에선 일반 문서파일(엑셀 등)로 전달되는 가격, 품명, 수하인 성명, 주소, 연락처 정보 등 총 28개의 통관정보는 위변조에 쉽게 노출되어 엄격한 통관 질서 확립에 장애물로 작용한다. 이와 같은 통관정보는 각 단계별 발생하는 정보간 연계성이 떨어져 활용이 어렵다. 최근

5년간 매년 2~30% 폭증하는 물량('13년 1,116만건 → '15년 1,584만건 → '17년 2,359만건)에 대하여 전자상거래업체, 운송업체 등 사업자간 정보전달 및 통관목록 작성이 수작업으로 진행되어 인력 낭비, 업무처리 지연 등 비효율이 존재한다.[28] 따라서 이 같은 문제를 해결하기 위해서 과기정통부와 관세청은 블록체인 기반 전자상거래 물품 개인통관 시범서비스 시스템을 구축하고 있다. 블록체인 기술을 통해 통관과 관련한 참여기업 및 기관 간에 실시간으로 데이터를 연동 및 공유함으로써 통관 과정에서의 위·변조 우려 등을 해결 할 예정이다.

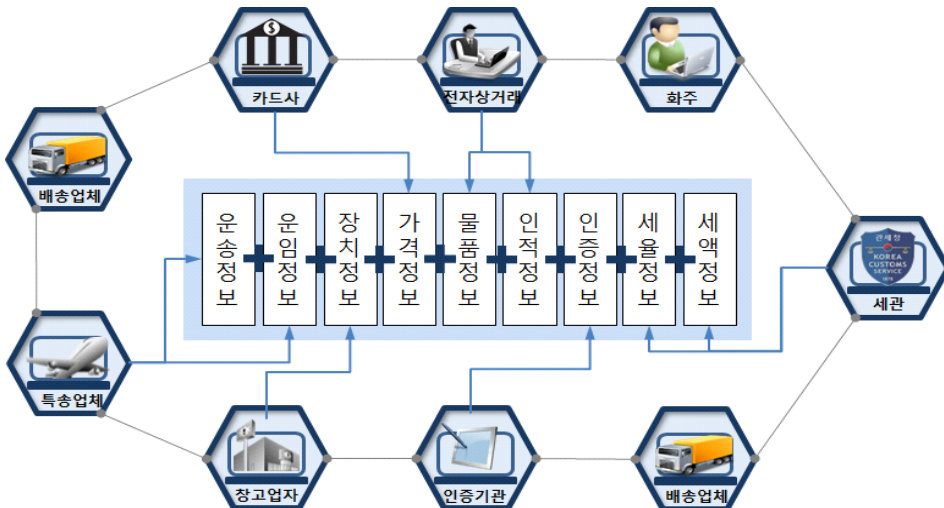
(2) 프로젝트의 목적

과기정통부와 관세청이 함께 주관하는 전자상거래물품 개인통관 시범서비스는 해외 직구물품의 통관과정에 블록체인 기반 분산형 시스템을 적용하여 비효율적인 업무절차를 감소시키는 목적을 갖는다.

〈그림 60〉 현행 통관시스템 : 신고서(서류)는 전자상거래 업체가 제공하는 정보에 근거하여 수기로 작성되므로 위변조에 취약하다



〈그림 61〉 블록체인 기반 통관시스템 개념도 : 각 참여자의 생산정보를 기반으로 신고서의 내용이 순차적으로 채워지므로, 특정인의 의지에 의한 위변조가 불가능하다



출처 : 인터넷진흥원, “지능형 개인통관 서비스 플랫폼 구축 시범사업 상세제안 요청서”, Mar. 2018

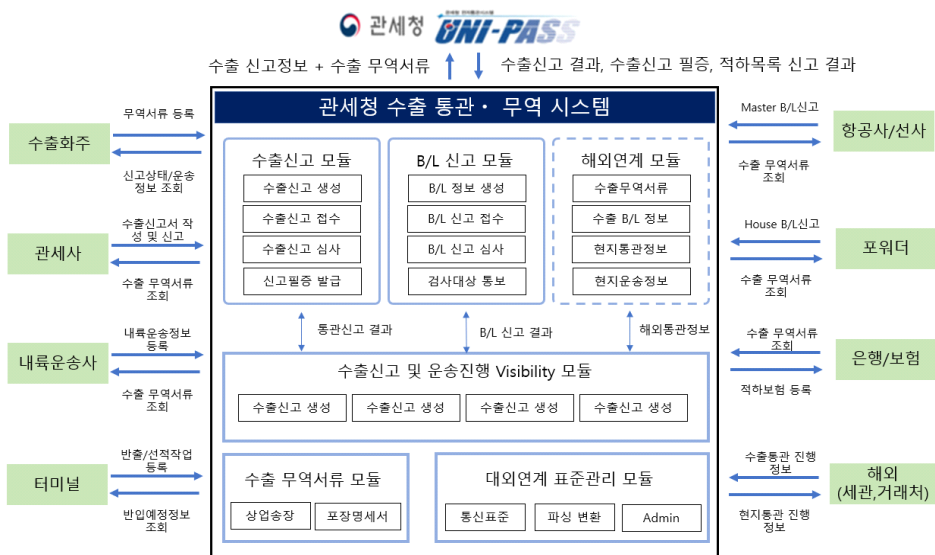
블록체인에 취합되는 정보의 신뢰성과 신속성을 높여 해외 발송부터 국내배송까지 모든 정보를 관리하는 전자상거래 통관 통합 서비스로 확대 개편할 계획이다. 블록체인 기술을 전자 상거래 물품 통관 시스템에 도입하여 신뢰할 수 있는 물품 정보 관리 체계를 구성하고 자정능력을 갖춘 물품관리와 체계적인 관리 체계를 유지한다. 전자상거래업체, 운송업체, 관세청 등을 연계하는 블록체인 기반 개인통관 서비스 플랫폼을 구축하여, 전자상거래 물품 통관시 일반 문서파일을 이용한 수작업 없이 편리하고 정확하게 신고자료를 작성하고, 빠르게 통관할 수 있도록 한다. 또 전자상거래 구매정보와 운송업체 운송정보로 자동으로 작성된 통관목록을 통해서 해외 발송부터 국내 도착까지 화물의 위치를 실시간으로 제공한다[29].

블록체인 기반 전자상거래물품 개인통관 시스템은 전자 상거래 업체의 물품 주문정보와 운송업체의 운송정보를 블록체인을 통해 실시간으로 상호 공유하여 28개 통관정보를 자동 취합하여 정리 할 수 있다. 이로 인해 업체의 관련업무 처리가 자동화되어 불필요한 업무량 감소시키고, 전체 통관시간이 최소 반나절 이상 단축된다. 또한, 관세청은 물품주문과 운

송에 관한 원천정보를 실시간으로 확인 할 수 있게 되어 전자상거래업체 등의 세금탈루 방지는 물론 불법물품 반입을 효율적으로 차단할 수 있으며, 사용자들은 온라인 포탈을 통해 자신의 화물 위치정보 및 세관 신고정보를 조회하여 실시간 화물확인과 허위신고로 인한 피해를 방지 할 수 있다.

(3) 상세 기술

〈그림 62〉 목표시스템 개념도



출처: 관세청, "18년도 ICT기반 공공서비스 촉진사업 제안요청서", Apr. 2018

1) 수출 통관 · 물류서비스 개발

블록체인 기반의 통관 물류 업무절차를 도출하며, 수출 무역서류 데이터를 기반으로 수출신고서를 생성한 후 관세청에 신고하는 기능을 제공한다. 또한 신고 된 건에 대한 선별, 심사 및 수출신고필증 발급기능, B/L을 생성하거나 연계하여 관세청에 신고하는 기능과 신고 된 건에 대한 심사, 검사대상 통보기능을 제공한다. 향후 사업 확대 시 국가 간의 통관정보 및 무역서류를 상호 제공하여 국가 간 무역의 신속하고 안전한 통관 체계를 구축 할 예정이다.

2) 수출 무역서류 / Visibility서비스 개발

수출신고의 무역서류, 수출신고 및 B/L의 신고 상태정보, 화물의 내륙운송과 국제운송 정보를 블록으로 생성하고 조회기능을 제공한다. 또한 대외연계 시 양단간의 통신표준과 문서 스펙별 파싱 규칙 등에 대한 관리자 기능을 제공한다.

3) 시스템 구축내용

수출통관·물류 응용S/W 및 블록체인 기반의 시스템 환경을 다음과 같이 구축한다.

구분	구축내용	비고
H/W*	서버(WEB, WAS, DB, 응용)	서버 1대 메모리(64GB) 이상 증설
S/W	OS, WEB, WAS, DBMS, 서버보안, 블록체인S/W**	

* 관세청이 보유하고 있는 장비(모델명 : IBM x3650 M4) 재사용

* 블록체인S/W는 블록체인 서버 내 최대 10개까지 구성

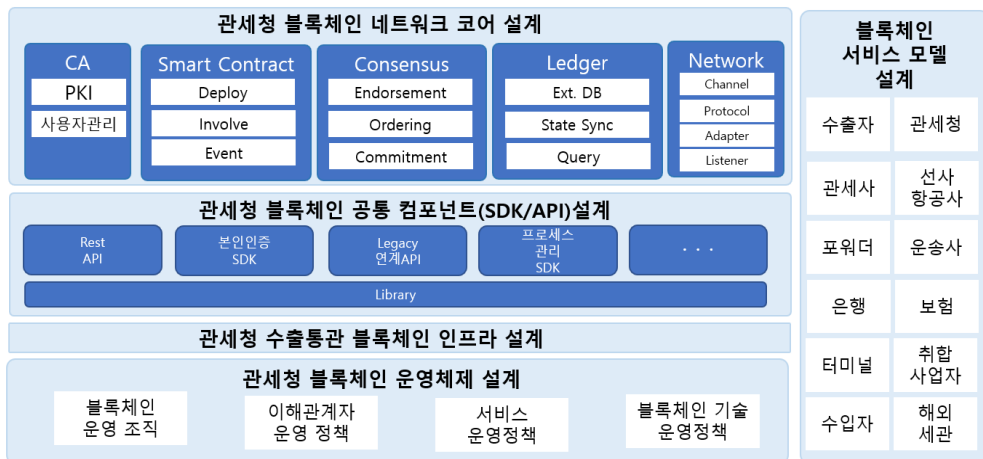
〈블록체인 기반의 수입통관물류 시스템 구축내용: 블록체인 노드는 참여 기관/
기업을 대상으로 최대 50개로 구성함〉

출처 : 인터넷진흥원, “지능형 개인통관 서비스 플랫폼 구축 시범사업 상세제안 요청서”, Mar. 2018

4) 관세청 수출 통관 블록체인 플랫폼

관세청 블록체인기반 수출통관 플랫폼은 물류주체별 역할 및 업무설계, 블록체인 공통 컴포넌트 설계, 블록체인 네트워크 코어 설계, 블록체인 인프라 설계, 블록체인 운영체제 설계 모듈로 구분하여 구축 예정이다.

〈그림 63〉 관세청 수출통관 블록체인 플랫폼 설계



출처 : 관세청, “18년도 ICT기반 공공서비스 촉진사업 제안요청서”, Apr. 2018

① 블록체인 네트워크 코어

블록체인 네트워크 코어는 다음과 같은 내용들을 고려하여 설계된다. 인증서, 사용자, Smart Contract, 합의 알고리즘, 분산원장, 네트워크 관리 기능 설계. 또한 블록체인 노드 구성을 위한 이해관계자(노드, 사용자)를 정의하고 이해관계자간 업무 및 사용 문서 분석을 통하여 블록체인 채널을 정의한다. 블록체인 내에서 블록 송신을 위해 어댑터를 구현하고, 신규블록 수신을 위한 기능을 구현한다. 이외에도 노드 운영과 유지보수 방안 등을 수립한다.

② 블록체인 공통 컴포넌트 (SDK / API)

설계된 블록체인 서비스들의 기능 리스트를 통해, 효율적인 개발과 운영을 위한 공통기능을 도출한다. 이렇게 도출된 각 공통 기능 요구사항을 정의하고 이에 대한 개발계획을 수립한다. 또한 대외 업체 및 관세청 내부 시스템(UNI-PASS)와의 연계를 위한 API를 설계하고, 개발된 공통 컴포넌트의 운영 안정성과 유지보수 방안을 수립한다.

③ 관세청 수출통관 블록체인 인프라

수출통관 시범사업을 위해서 블록체인 물리적 노드는 5개를 운용하며 물류 수행사 50개

노드를 논리적으로 구성한다. 향후 확장되는 이해관계자들과의 인프라 구조 설계와 영향도 파악 및 블록체인에 쌓이는 블록사이즈에 따른 인프라 확장 계획을 고려하여 관세청 인프라 환경에 맞는 블록체인 하드웨어 구성 방안을 설계한다.

④ 블록체인 운영체제

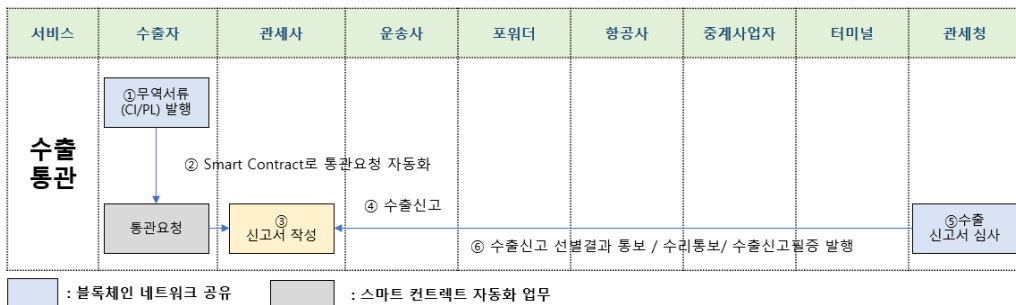
관세청 블록체인 운영 모델과 최적의 운영 프로세스 설계한다. 또한 블록체인 네트워크에 참여하는 이해관계자들의 블록체인 노드 운영 정책, 블록체인 서비스 운영 정책 및 기술 운영 정책을 설계한다.

4) 블록체인기반 서비스 모델

① 수출신고 블록체인 서비스

수출자가 무역서류를 블록체인 네트워크를 통하여 관세사에게 통관을 요청(통관요청)하면 관세사는 블록체인상에 공유된 무역서류를 기반으로 수출신고서 블록을 작성하여 관세청에 전중 및 심사를 요청(수출신고)한다.

〈그림 64〉 수출신고 블록체인 서비스



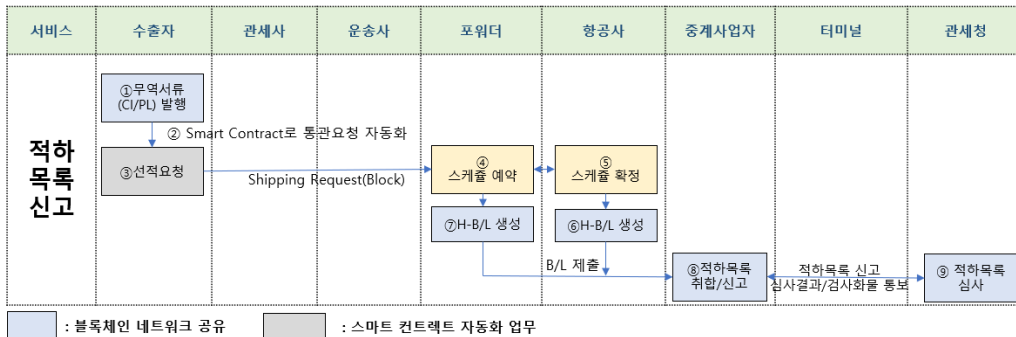
출처: 관세청, "18년도 ICT기반 공공서비스 촉진사업 제안요청서", Apr. 2018

② 수출적하목록 신고 블록체인 서비스

수출자가 블록체인 네트워크(스마트 컨트랙트)를 통하여 포워더에게 선적을 요청하면, 포워더는 블록체인상에 공유된 무역서류를 기반으로 스케줄 예약, H-B/L 블록을 작성하여 부킹 및 B/L을 제출한다. 항공사는 M-B/L을 작성 후 제출하여 취합사업자를 통하여

적하목록을 취합하여 관세청에 적하목록을 신고한다.

〈그림 65〉 수출적하목록 신고 블록체인 서비스

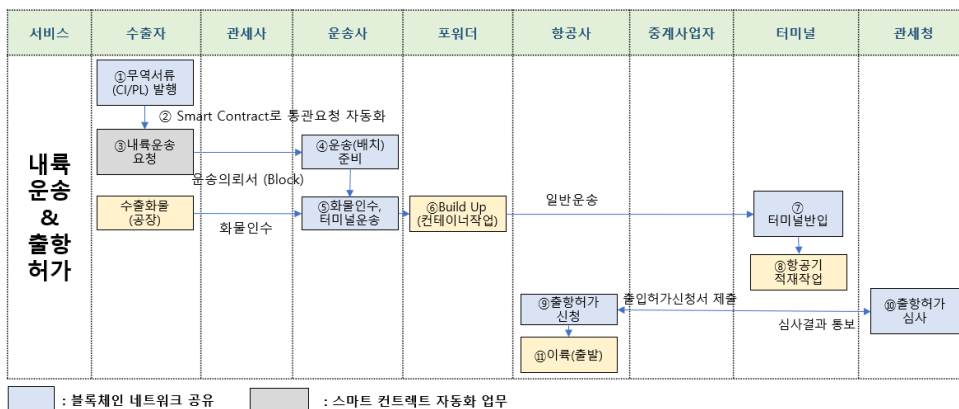


출처 : 관세청, “18년도 ICT기반 공공서비스 촉진사업 제안요청서”, Apr. 2018

③ 내륙운송 및 출항 허가 블록체인 서비스

수출자가 블록체인 네트워크를 통하여 운송사에게 내륙운송을 요청(화물인수요청)하면 운송사는 블록체인상에 공유된 무역 서류를 기반으로 운송차량 배차, 운송정보에 대한 진행상황을 제공한다. 터미널에서는 화물의 반입준비 및 반입처리를 블록으로 생성하며, 항공사는 출항허가 신고를 블록으로 생성 신고한다.

〈그림 66〉 내륙운송 및 출항허가 블록체인 서비스



출처 : 관세청, “18년도 ICT기반 공공서비스 촉진사업 제안요청서”, Apr. 2018

6) 블록체인 기술 적용 주요내용

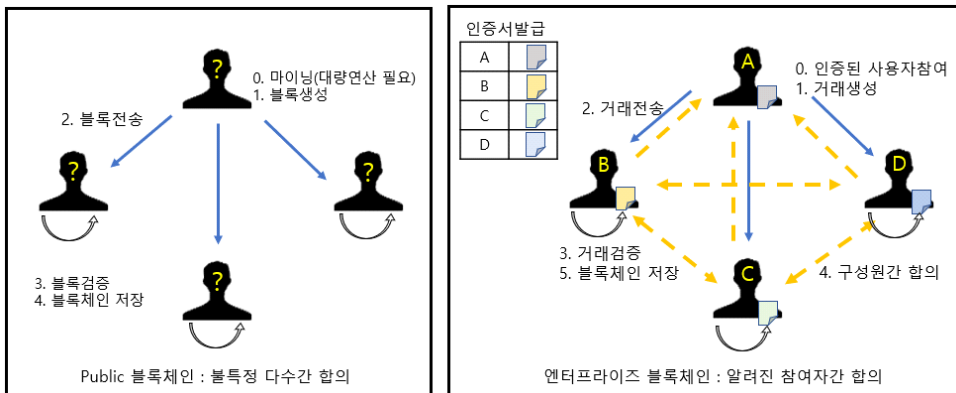
① 전자문서 위변조 방지체계 강화

분산 공유 기술인 블록체인을 활용하여 유통되는 물류서류에 대한 투명성과 위변조 방지체계를 강화한다. 실제로 수출신고를 위해 연간 유통되는 문서 건 수는 '17년 기준 7.2백만여 건으로 EDI기반의 문서유통 및 포털을 활용한 수기입력이 주를 이루며, 전자문서 유통 및 데이터 입력과정에서 무결성 확보가 매우 중요하다. 따라서 블록체인 기술의 특성중 하나인 블록체인에 기록된 정보는 변경이 불가능한 특성을 활용하여 사용자가 송신한 전자문서에 대한 위변조 여부 판단할 수 있다.

② 최적의 분산 합의 알고리즘 적용

공공 블록체인은 불특정 다수 참여자 간의 합의를 위해 복잡한 수학적 계산을 통한 채굴 과정이 수반되어 많은 에너지와 시간 소요된다. 이러한 단점을 해결하기 위해 참여 기업 및 기관을 대상으로 엔터프라이즈 블록체인 플랫폼 기반의 최적의 분산합의 알고리즘을 적용하였다. 이를 통해 수출신고, B/L신고, 정보연계 및 진행정보를 고속처리하고 즉각적 확정성을 제공한다.

〈그림 67〉 Public 블록체인 과 엔터프라이즈 블록체인 비교

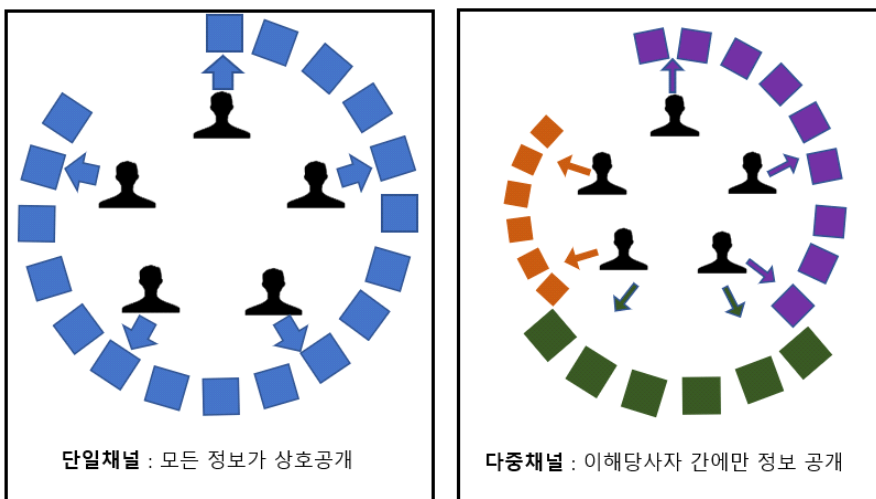


출처 : 관세청, "18년도 ICT기반 공공서비스 촉진사업 제안요청서", Apr. 2018

③ 채널 설계를 통한 접근 권한 관리

전자문서에 포함된 기밀정보 및 무역/물류 진행정보는 이해관계자 간 공유가 되어야 하며 타사에 공개되는 것을 방지하기 위해 Private 전송 기능을 지원 한다. 전체 공유가 필요한 정보의 경우에는 Public 전송 기능을 통해 공유할 수 있도록 한다. 또한 기업/기관 간 거래관계에 따른 프로세스를 분석하여 채널을 설계하고 관련 정보 접근 권한을 관리할 수 있도록 한다.

〈그림 68〉 채널설계를 통한 접근 권한 관리

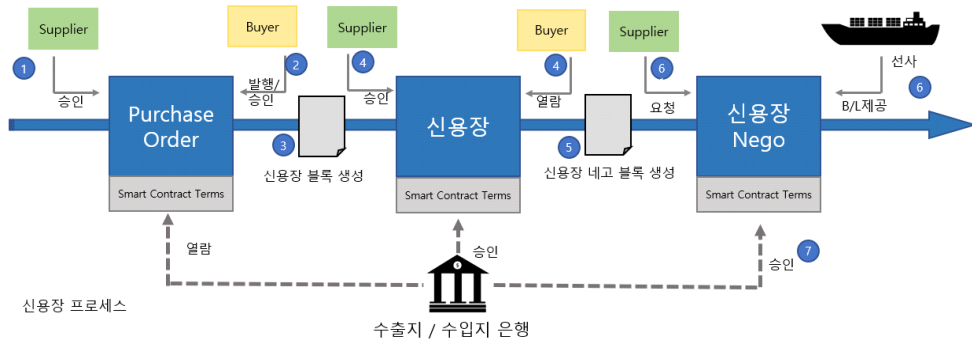


출처 : 관세청, “18년도 ICT기반 공공서비스 촉진사업 제안요청서”, Apr. 2018

④ 스마트 컨트랙트(Smart Contract) 기반 업무 자동화

스마트 컨트랙트를(Smart Contract) 활용한 다양한 처리 조건을 구성하여 정형화된 이벤트(수출신고 승인, B/L접수 등)를 자동으로 수행하여 업무 자동화를 지원하고, 이벤트를 형성하는 주요 조건 및 권한에 따라 거래 참여자의 전자서명을 통한 신고 서류의 상호 승인을 완료할 때 거래 확정을 자동화 한다.

〈그림 69〉 스마트컨트랙트 기반 업무 자동화



출처: 관세청, "18년도 ICT기반 공공서비스 촉진사업 제안요청서", Apr. 2018

(4) 현황 및 결론

전자상거래물품 개인 통관 시스템은 2019년 1월부터 관세청과 업무협약을 체결한 전자상거래업체 (주)코리아센터와 운송업체 CJ대한통운이 참여하여 구매·배송대행 신청시 물품 주문정보가 운송정보와 함께 블록체인에서 취합되어 신속 자동처리 되도록 운영될 예정이다. 향후 관세청은 관련성과를 바탕으로 서비스 운영 전자상거래·운송업체를 확대하고 기타 통관정보 관련기업의 참여를 유도할 계획이다. 이를 통해 발송부터 국내 배송까지 모든 정보를 관리하는 전자상거래 통관 통합 서비스로 확대 개편할 계획이다. 시범사업엔 48개의 기업이 참여한다. 관세청이 블록체인 플랫폼을 제공하면 참여기업은 자사 시스템에 연계하고 실제 데이터를 전송해 통관 등의 업무를 처리한다. 이 과정에서 업무적·기술적 검증을 수행한다. 또한 블록체인 기반 e-C/O 발급,교환 시범사업은 1단계로 베트남과 싱가포르와 추진 할 예정이며, 향후 단계적으로 FTA체결 국가인 인도, 필리핀 등으로 확대할 계획이다. 이와 같은 시범사업을 통한 기대효과는 첫째, 블록체인 기술기반으로 국가 물류 정보망을 재편, 국가 물류의 표준을 선도적으로 정립하여 민간의 물류 혁신 활성화의 토대를 마련할 수 있다. 둘째, 수출입 공급망 관련 기업의 실증 참여를 통해 무역·물류 분야에 특화된 블록체인 기술개발이 활성화 되고 이에 따른 민간의 일자리 창출로 연계를 기대할 수 있다. 셋째, 전략적인 시범사업 추진을 통해 해외 IT업체들의 블록체인 기반 무역기술 연구에 대응한 경쟁력을 확보해 국제 표준으로서 선점 가능하다. 넷째, 수출입

공급망의 정보 공유 시스템 개선 및 통관물류의 디지털화에 따라 물류의 신속한 처리와 물류비용 절감이 가능하다. 마지막으로 수출신고서 작성시의 재입력 작업비용(약 240억원) 절감효과를 얻을 수 있으며 수출 기초서류를 조회하고 다운로드하여 수출신고서 작성을 위한 별도의 신고정보 재입력 작업을 최소화함으로써 편의성을 높이고 수작업 오류 또한 줄일 수 있다.



정책적 제언

2장에서는 세계 각국의 데이터 플랫폼 구축 및 활용 사례를 설명했다. 3장에서는 데이터 플랫폼 정책에 대해서 제언하도록 하겠다. 정부가 데이터 플랫폼 관련하여 추진해야 하는 정책은, ‘데이터 산업 활성화 계획 추진’ 과 ‘블록체인 기반 보안산업 활성화’로 나눌 수 있다. 데이터 산업 활성화 계획 추진은 4차산업 혁명에 대응하는 데이터 산업육성을 위해 제도, 기술, 산업 기반 등 데이터 전반에 걸친 활성화 계획을 추진하는 것에 대한 내용이다. 블록체인 기반 보안산업 활성화는 4차 산업의 주요 요소인 IoT기술 및 개인데이터 활용 기술의 발전을 저해하는 보안 문제를 해결하기 위한 정책 방안에 대한 내용이다.

(1) 데이터 소유권 및 사용 권한 이슈

기업, 공공, 개인의 데이터를 수집하고 제공하는 데이터 플랫폼은 데이터의 소유권과 사용 권한의 통제와 관련하여 법적인 틀에서 보호 받아야 한다. 우리는 일반적으로 데이터를 재산으로 인식하지 못하며 데이터에 대한 소유권을 결정하기 위한 기술적, 정책적 체계가 없다. 법 제도 정비를 통해 개인은 자신의 개인정보가 유출되지 않으면서 편리하게 활용될 수 있다는 신뢰를 쌓을 수 있고, 이를 통해 정부 및 기업은 개인 데이터 수집과 활용 영역을 지속적으로 확장할 수 있다.

(2) 개인정보 보호

전 세계적으로 생성된 데이터의 대부분은 개인 데이터이다. 개인의 프라이버시 권리를 보장하기 위해서는 개인정보의 적절한 보호가 필요하다. 개인 데이터는 기업과 개인 모두에게 경제적, 사회적으로 중요한 가치를 갖는다. 개인데이터를 통해 기업은 개인화된 제품, 개인화 서비스 및 서비스 개선에 대한 피드백을 받을 수 있기 때문에 개인과 기업에 상호

이익을 가져다 준다. 영국의 데이터 보호법 2018 및 EU의 일반 데이터 보호규정(GDPR)은 개인의 정보를 보호하기 위한 새로운 기준을 수립하고 개인에게 데이터 사용에 대한 권한을 부여한다. 기업이 GDPR을 위반 할 경우 최대 70만파운드(약 10억원)에 해당하는 벌금을 부과한다. 우리 정부도 강력한 데이터 보호 법규를 제정, 시행하여 개인 정보의 공유에 대한 대중의 신뢰를 구축하기를 희망한다.

(3) 상호 운용성 및 표준 개발 지원

공공, 기업, 개인이 데이터를 쉽고 편리하게 제공 및 사용할 수 있도록 데이터 표준 개발이 필요하다. 정부는 이해당사자로 하여금 공유 데이터 표준을 개발하고 장려하기 위해 경쟁과 협력의 기준을 정확히 마련하고 적절한 규제를 통해 표준 개발을 주도할 필요가 있다. 이런 방식으로 기업으로 하여금 데이터에 대한 접근성을 높이면 경쟁을 통해 소비자의 요구에 충실한 새롭고 혁신적인 제품을 개발할 수 있어 기업의 경쟁력확보에도 도움이 된다.

(4) 안전하고 합법적인 데이터 공유 지원

중소기업 및 스타트업 기업은 개인의 데이터 수집과 관련하여 경험, 전문성이 떨어지며 데이터 플랫폼등의 기술이 부족한 실정이다. 데이터 공유가 안전하고 윤리적이며 데이터 보호법을 준수한다면 정부는 기업의 데이터 수집, 공유를 통해 성장하고 혁신할 수 있도록 도움을 줄 필요가 있다. 효과적인 데이터 공유가 이뤄지지 않으면 소수의 대기업에 데이터가 집중되는 문제가 발생하여 경쟁의 부재로 인한 제품의 품질 저하를 가져올 수 있다. DECODE의 사례처럼 정부 주도의 데이터 플랫폼 개발을 통해 공공, 기업, 개인이 유용한 데이터를 공유하면서도 성장 할 수 있는 기반을 마련해야 한다.

참 | 고 | 문 | 헌

- [1] 한국정보화진흥원, “Global Open Data, Now”, Issue Paper, 2017
- [2] EU GDPR.ORG. "GDPR key changes". <https://eugdpr.org/the-regulation/>
- [3] Theo Bass and etc., “Reclaiming the Smart City: Personal data, trust and the new commons”, DECODE project, Jul. 2018.
- [4] TADA. "Principles". <https://tada.city/en/>
- [5] Francesca Bria.(2017). DECODE Data sovereignty for citizens. Paper presented at the MyData 2017 conference, Helsinki.
- [6] EU. "What is Horizon 2020?" Accessed Nov, 2018.
<https://ec.europa.eu/programmes/horizon2020/what-horizon-2020>
- [7] DECODE project. "Pilots" Accessed Nov, 2018. <https://decodeproject.eu/pilots>
- [8] George Danezis and etc., “First Version of DECODE Architecture”, DECODE project, Oct. 2017.
- [9] Blum, Manuel Feldman, Paul Micali, Silvio. "Non-interactive Zero-knowledge and Its Applications", 1988
- [10] Oleguer Sagarra and etc., “CityOS connection”, DECODE project, Jul. 2018.
- [11] Data Transfer Project. Available at <https://datatransferproject.dev/>
- [12] BBC, “페이스북: 5000만 개인정보 유출 어떻게 가능했나?”, Available at <https://www.bbc.com/korean/news-43481991>
- [13] 조선일보, “페이스북 이어 구글도...5250만명 개인정보 유출”, Available at http://news.chosun.com/site/data/html_dir/2018/12/11/2018121101304.html
- [14] Developer Guide. Available at <https://github.com/google/data-transfer-project/blob/master/Documentation/Developer.md>
- [15] Google, “Data Transfer Project Overview and Fundamentals”, Jul. 2018
- [16] 한국정보화진흥원, “지능화 시대 스마트 선박 발전 방향”, Nov. 2016
- [17] Hideyuki Ando, “The role of open platform in the coming era of ship digitalization”, Apr. 2018

- [18] Mineyuki Morimoto, "Introduction of ISO/FDIS 19848 Standard data for shipboard machinery and equipment", Apr. 2018
- [19] Trimble connect. Available at <https://connect.trimble.com/>
- [20] Trimble and Komatsu Collaboration, Available at <https://www.trimble.com/news/release.aspx?id=071417a>
- [21] Trimble, "Trimble connect overview", Available at <https://bpmidamerica.com/wp-content/uploads/2016/03/Trimble-Connect-Overview.pdf>
- [22] Equipment world. (2017). "Trimble software platforms will soon share 3D construction data", Available at <https://www.equipmentworld.com/komatsu-trimble-software-platforms-will-soon-share-3d-construction-data/>
- [23] Hitachi Cnstruction Machinery, "Hitachi Construction Machinery Partners with Trimble in the Development of Cloud-based Platform". Available at <https://www.hitachicm.com/global/news/all-press-releases/17-03-07e/>
- [24] I-Construction. Aavailable at <http://www.mlit.go.jp/tec/i-construction/index.html>
- [25] BaasID, "블록체인 공공 네트워크 기반의 인스턴트 개인 인증 플랫폼", Apr. 2017
- [26] 장형준, 심준혁. (2018). "블록체인과 로봇의 미래 동향". 전기의세계, 67(8), 13-18.
- [27] 관세청, "18년도 ICT기반 공공서비스 촉진사업 제안요청서, 블록체인 기반의 수출통관 물류서비스 시범사업", Apr. 2018
- [28] 관세청 보도자료, "블록체인이 해외직구 통관도 혁신한다", Nov. 2018
- [29] 인터넷진흥원, "지능형 개인통관 서비스 플랫폼 구축 시범사업 상세제안 요청서 ", Mar. 2018

데이터로 연결하는 **사람·세상·미래**



Bridging DATA

제3편

데이터 기반 경제에서 알고리즘의
역할과 책임성

NIA 한국정보화진흥원



CONTENTS

I. 서론 : 알고리즘 책임성의 개념과 접근법	169
II. 알고리즘에 대한 이해	170
1. 인공지능(AI)과 알고리즘	170
2. 알고리즘 유형과 특징	172
3. 알고리즘 구현과정	177
III. 알고리즘 책임성	181
1. 개념과 배경	181
2. 다양한 관점과 접근법	185
IV. 알고리즘 책임성에 대한 접근법: 대안과 전망	188
1. 알고리즘 책임성의 필요성	188
2. 새로운 접근법: 대안과 방향	191
3. 전망	199
V. 알고리즘 책임성 적용 사례 및 관련 이슈들	202
1. AI 플랫폼	202
2. 알고리즘 가격설정과 담합(Algorithmic Pricing and Collusion)	209
3. 레그테크(RegTech) 및 리걸테크(Legal Tech) -- regulation by machines	216
VI. 결 론	219
참고문헌	221



서론 : 알고리즘 책임성의 개념과 접근법

18C 산업혁명이 기계를 통해 인간의 노동력을 확장하고 강화시키는 시대를 열었다면, 4차 산업혁명이라고까지 불리는 기술적 환경과 생산방식의 변화는 정보·통신·미디어기술(ICT, Information and Communication Technology)을 통해 인간의 지력(知力, brain power)이 본격적으로 확장·강화되는 시대를 열었다. 그런데 이러한 변화가 사회 전반에 미치는 영향력이 급속히 확대되면서 이에 따른 부작용과 위험요소에 대한 관심과 걱정도 크게 증가하고 있다.

4차 산업혁명을 이끄는 ICT의 핵심에는 오래전부터 인공지능(AI, Artificial Intelligence)이 있었지만, 그 영향력을 세상이 인식하기 시작한 지는 그리 오래되지 않았다. 아직 세간에서 AI에 대한 인지도와 인식이 유동적이기 때문에 낙관적 전망(예: Ray Kurzweil, 커즈와일(2016))과 비관적 견해(예: Frank Pasquale, 파스칼레(2016))가 공존하고 있는 상황이다. 그런데 극단적 견해가 대립하면서 가능한 합의와 절충이 점점 어려워지는 원인 중의 하나로, AI에 대한 이해가 여전히 부족하다는 점을 들 수 있다. 대표적인 예로, AI와 알고리즘(algorithm)을 동일시하거나 AI와 알고리즘을 통해 최적화된 결과를 쉽게 이끌어 낼 수 있다는 오해 등을 들 수 있다.

AI 혹은 AI 시스템에서 엔진과도 같은 역할을 하는 알고리즘에 대한 이해를 높임으로써 4차 산업혁명 시대에 필요한 AI의 발전을 효율적이며 효과적인 방향으로 이끌 수 있을 것이다. 특히, 이 글에서는 기계학습(machine learning) 외에도 다양한 알고리즘의 유형과 특성을 소개하여 알고리즘에 대한 오해를 해소하고 더욱 균형 잡힌 시각에서 AI를 조망할 수 있도록 한다. 또한, 알고리즘이 AI 시스템으로 구현되는 과정을 살펴봄으로써, 법제도적 관점에서 논의가 시작된 알고리즘 책임성(algorithm accountability)이나 AI 거버넌스(governance)와 같은 이슈가 소모적 논쟁으로 흐르지 않고 보다 생산적인 방향으로 나아가는 데 도움을 주고자 한다.

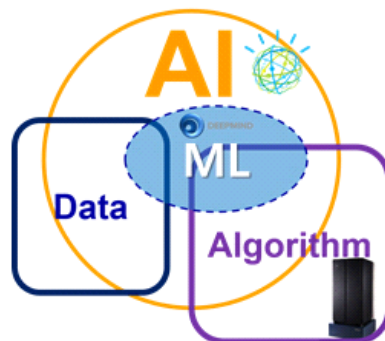


알고리즘에 대한 이해

1 인공지능(AI)과 알고리즘

AI는 사용자 인터페이스와 AI를 구동하는 엔진인 알고리즘 및 데이터와 같은 알고리즘 구동에 필요한 자원 등으로 구성되는 시스템이다(그림 1) 참조). 여기서 알고리즘은 주어진 의사결정 문제(decision problem)를 해결하기 위한 체계적인 절차와 제반 기법을 의미한다. AI가 단순히 사건과 사물을 분류하는 것에서 더 나아가 의사결정에 관여하게 되면 알고리즘의 역할은 더 중요해진다. 원칙적으로 알고리즘은 데이터와 독립적으로 설계·구현되는 요소이지만, 오늘날 알고리즘의 대명사처럼 쓰이는 기계학습 또는 머신러닝(machine learning)에서는 데이터와 알고리즘이 서로 긴밀하게 상호작용하기 때문에 양자를 구별하는 것이 어려울 수 있다(단지, 서로의 역할을 코드 블록에서 나누는 정도만 가능하다. 그러나 이조차 쉽지 않은 경우가 많다).

〈그림 1〉 시스템으로서의 AI와 알고리즘 및 데이터의 역할¹⁾



1) 김도훈(2018)에서 인용

사실, 알고리즘이라는 용어를 낳은 제반 학문 분야(컴퓨터과학(computer science), 경영과학(Management Science 또는 OR(Operations Research)), 응용 최적화(applied optimization) 등)의 관점에서 보면 이 용어의 사용법은 적절하지 못하다. 엄격한 학문적 관점에서 볼 때 세간에 쓰이는 알고리즘이라는 용어는 휴리스틱(heuristic)으로 대체되어야 한다. 알고리즘과는 달리, 휴리스틱은 최적해(optimal solution)를 보장하지 못하지만, 상대적으로 빠른 시간 내에 꽤 괜찮은 답을 주기 때문에 (주로 상업용 목적에서) 알고리즘이라고 불렸다. 컴퓨팅 능력이 비약적으로 발전하면서 머신러닝을 비롯하여 타부서치(tabu search), 유전자 알고리즘(genetic algorithm)과 같은 메타 휴리스틱(mets heuristic)의 현실적 효용성이 (이론적 혹은 좁은 의미의) 알고리즘을 능가하면서 (다소 부정적인 뉘앙스를 주는) 휴리스틱이 아닌 알고리즘으로 불리기 시작했다. 이러한 맥락에서 Parnas(2017)는 AI가 대중화되면서 AI에 쓰이는 알고리즘이 마치 최적의 답을 주는 것처럼 과장되고 왜곡되는 상황을 경계해야 한다고 말한다. 이 글의 알고리즘도 (좁은 의미의) 알고리즘과 휴리스틱을 포괄하는, 일반적으로 쓰이는 알고리즘을 의미한다.²⁾

AI에 활용되는 알고리즘도 사실은 여러 알고리즘이 유기적으로 구성된 프로그램이다. 이러한 프로그램은 마스터 프로그램 내에서 (뒤에서 소개되는) 복수의 최적화와 분류 알고리즘들을 복합적으로 사용한다. 이러한 이유에서도 AI, 알고리즘, 데이터 등을 구분하여 이해하는 것이 중요하다. AI 시스템에서 쓰이는 프로그램을 구성하는 단위 모듈은 결국 하나의 의사결정 문제를 다루기 때문에 이 글에서도 주로 단위 알고리즘에 대해 설명한다. 따라서 의사결정 문제의 종류에 따라 알고리즘의 성격과 유형도 달라지며, 이를 구분하는 것이 알고리즘 책임성을 비롯한 법제도 논의에서도 중요하기 때문에 다음 절에서 이를 간략히 소개한다.

2) (좁은 의미의) 알고리즘과 휴리스틱의 차이와 그 의미에 대해서는 튜링상(Turing Award, 1975년)과 노벨 경제학상(1978년)을 모두 수상한 Herbert A. Simon이 이미 1970년대에 설파하였다.

2 » 알고리즘 유형과 특징

알고리즘이 다루는 의사결정 문제의 종류에 따라 크게 최적화(optimization)를 위한 알고리즘과 분류(classification)를 위한 알고리즘으로 구분할 수 있다.³⁾ 또한, 알고리즘의 적용 대상이 되는 의사결정 문제 자체의 특징이 알고리즘의 성과의 한계를 결정할 수도 있다. 대표적으로, 계산복잡도(computational complexity) 관점에서 문제 해결 과정에서 필요한 정보량에 따라 정확도가 크게 좌우되는 NP-Complete형 의사결정 문제와 그렇지 않은 P형 문제가 서로 분리되어 존재한다(아래 [참고] 참조). NP-Complete형 문제를 위한 알고리즘은 복잡성이 높을 수밖에 없으며 현실에서 마주치는 큰 문제에 대해서는 정확한 답을 제공한다는 것을 보장할 수 없으므로 정확성과 속도 사이에서 타협해야만 한다. 최근의 빅데이터와 머신러닝의 발전으로 정확성과 속도 모두 비약적으로 개선되고 있지만, (양자컴퓨팅(quantum computing)이 보편화되지 않는 한) Alpha-Go와 같은 성능의 AI가 스마트폰에 하드웨어 형태로 ‘탑재’ 되는 것이 사실상 불가능한 이유도 여기에서 찾을 수 있다(물론 클라우드 방식으로 Alpha-Go를 ‘호출’하여 이용할 수는 있다).

(1) 최적화 알고리즘

최적화 알고리즘은 내비게이션에서 최단 경로를 찾거나 최적 재고관리 및 생산 일정을 결정하고 전국적인 에너지 수급계획을 마련하는 것과 같은 최적화 문제를 해결하는 알고리즘이다. 제2차 세계대전 당시 미국에서 군수물자 생산과 병참 물류에 관한 계획을 수립하기 위해 개발된 이후, 경영학 경제학 공학 등 광범위한 분야에서 사용되는 선형계획법(linear programming)에 사용되는 알고리즘들이 대표적인 예이다. 이를 비선형함수와 이산형 변수 등으로 확장한 제반 수리계획법(mathematical programming)에서도 다양한 알고리즘이 개발되어 사용 중이며, 이 경우 NP-Complete형 의사결정 문제가 대부분

3) 참-거짓(혹은 on-off 등)을 판별(decision)하는 것과 최적해(optimal solution)를 찾는 것은 계산학(computation) 관점에서는 차이가 없지만, 현실적으로 의사결정 문제의 목적을 최적화와 분류로 구분하는 것이 편리하므로 이에 따라 알고리즘 유형도 나누어 생각해 볼 수 있다.

이기 때문에 최적해 찾는 것을 포기하는 대신에 꽤 괜찮은 결과를 빠른 속도로 찾는 휴리스틱이 사용된다(Curtis & Scheinberg, 2017; Parnas, 2017). 휴리스틱에 관한 일반화된 이론은 없지만 메타 휴리스틱으로 불리는 여러 방법은 매우 잘 개발되어 있다([표 1] 참조).

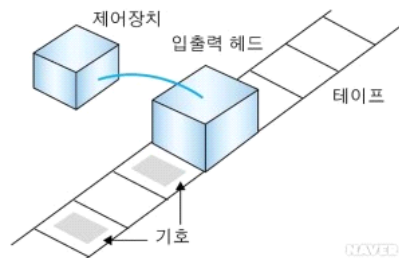
계산복잡도의 의미와 NP-Completeness

앨런 튜링은 1930년대에 훗날 ‘튜링 기계’라고 불리는 추상적인 연산장치 모형을 제시하였는데, 이는 오늘날 컴퓨터 프로세서를 설명하는 기반 이론이 되었다. 일반적인 물리법칙의 지배를 받는 세상에서는 모든 조건이 같을 때 특정 상태는 하나의 움직임(다음 상태)만 가능하므로 이를 결정형(deterministic) 튜링 기계라고 한다. 튜링은 다음 상태가 하나 이상이 될 수 있는 비결정형(non-deterministic) 튜링 기계⁴⁾도 생각해보았다. 마치 다음번 계산 단계에서 여러 개의 CPU가 동시에 생성되는 컴퓨터를 상상하면 된다(양자컴퓨팅이면 어느 정도 가능).

당연히 결정형 튜링 기계보다 비결정형 튜링 기계가 같은 문제를 빠르게 푼다. 비결정형 튜링 기계를 사용할 때, 정확한 답을 찾는 속도가 (최악의 상황에서도) 문제 크기(정보량)의 다항함수(polynomial function)로 표현되는 의사결정 문제를 NP형(Non-deterministic Polynomial type)이라고 부른다. 이에 반해, P형(Polynomial type) 의사결정 문제에서는 최악의 경우에도 결정형 튜링 기계(쉽게 말해 우리가 쓰는 컴퓨터)를 써서 정보량의 다항식으로 표현되는 속도로 정확한 답을 얻을 수 있다. P형 문제에서는 최적해를 보장하는 ‘효율적인’ (좁은 의미의) 알고리즘을 개발할 수 있으나, NP형 문제에서는 효율적인 (좁은 의미의) 알고리즘이 존재하지 않는다는 것이 정설이다. 바둑과 체스 같은 게임에서 필승전략(winning strategy)을 찾거나 금융상품의 최적 포트폴리오를 구성하는 문제가 NP형의 대표적인 예이다. 반면에 내비게이션에서 최단 경로(shortest path)를 찾는 문제가 P형의 예이다 (따라서 내비게이션 개발이 금융상품 설계보다 쉽다).

NP형 문제와 P형 문제가 서로 다른 카테고리로 분리되어 존재한다는(쉽게 말해, 풀기 어려운 문제와 쉬운 문제가 뚜렷이 구분될 것이라는) 명제(P vs. NP conjecture)는 아직 증명되지 못했지만, 거의 모든 수학자, 컴퓨터과학자, 경영과학자가 이 추측을 받아들이고 알고리즘(휴리스틱 포함)을 개발한다. 위 명제에서 NP형 문제들은 서로 연관될 것이므로(이때 P형은 NP형 내의 ‘뚜렷한’ 부분집합이 된다.) 이 클러스터를 NP-Completeness라고 부른다. 이에 대한 이론은 이미 1971년 Cook에 의해 정식화되었다.⁵⁾

〈그림 2〉 튜링 기계의 개념도⁶⁾



(출처 : 김종훈, 김종진(2013)에서 인용)

4) 가상에서만 존재할 수 있다

[표 1] 최적화 알고리즘의 종류

구분	특성	알고리즘
(좁은 의미) 알고리즘	- (이론적) 최적해 보장 - NP-Complete형 의사결정 문제에서는 속도와 저장공간 효율성은 보장 못함(단, P형은 가능 - 예: 선형계획법의 Karmarkar 및 Dijkstra 최단 경로 알고리즘 등)	- 선형계획법을 위한 단체법(Simplex), ellipsoid, Karmarkar법 등 내부점 (Interior Point) 알고리즘 - 비선형계획법(Nonlinear Programming)을 위한 수치해석(Numerical Analysis)기반 알고리즘들 - (혼합) 정수계획법((mixed) Integer Programming)을 위한 분지한계법 (Branch & Bound Method)과 그 변형 등 - 최단 경로, 최대 흐름 등, 네트워크와 이산형 최적화(Discrete Optimization)에서 개발된 다양한 알고리즘들
휴리스틱 알고리즘	- 최적해를 보장 못 함 - 빠른 속도 - 효율적인 저장공간 활용	- 일반 탐색(Search)을 위한 휴리스틱: Local Search 등 - 시뮬레이션 기반 최적화 휴리스틱(Monte Carlo 시뮬레이션 활용 등) - 메타 휴리스틱 * 유전자 알고리즘(Genetic Algorithm) * 타부서치(Tabu Search) * Simulated Annealing * 신경망(neural network)

(출처: 김도훈(2018)에서 인용)

(2) 분류형 알고리즘

분류형 알고리즘은 말 그대로 사물을 적절하게 분류하는 의사결정에 사용되는 알고리즘으로, 의사결정 문제의 기반이 되는 수리 모형, 구현 전략, 사용 목적에 따라 다양하게 유형화할 수 있다. 여러 교재와 문헌(Hastie et al., 2009; James et al., 2013; Mullainathan & Spiess, 2017 등)에서 가장 많이 채택된 유형화에 따르면, 기본 구현방식에 따라 지도 학습(supervised learning)과 비지도 학습(non-supervised learning)으로 크게 나눌 수 있다([표 2] 참조). 분류형 알고리즘을 구현하는 과정에서 최적화 알고리즘이 서브루틴으로 호출되는 경우가 많다.

지도학습은 올바른 분류에 대한 답(예; 회귀분석의 종속변수)과 해당 속성값(예; 회귀분석의 설명변수)이 관계를 맺을 때 가장 좋은 성과를 보이는 분류기(classifier)를 산출한다. 이에 반해 비지도 학습에서는 정답을 사전에 알 수 없거나 알리지 않은 상태에서 속성값만으로 주어진 데이터를 몇 개의 카테고리 나눈다는 분류기가 산출된다. 따라서 지도학습의

5) Garey & Johnson(1979) 등을 참조

분류기 성능은 (검증용) 데이터에 의해 객관적으로 파악될 수 있지만, 비지도 학습에 의한 분할의 적절성에 대해서는 별도의 평가방식이나 기준이 필요하다. 이런 식의 구분은 분류형 알고리즘을 적용해야 하는 목적과 용도에 따른 것이므로, 한 유형의 알고리즘을 다른 유형의 알고리즘과 직접 비교할 수는 없을 때도 많다.

그런데 지도학습과 비지도 학습의 구분은 절대적인 것은 아니다. 예를 들어, [표 2]에서 비지도 학습으로 분류된 자기 강화학습의 경우, 이를 구현하는 방법과 전략에 따라 지도학습으로도 분류될 수 있다. 머신러닝형 자기강화학습의 근간이 되는 뉴럴네트워크 혹은 신경망(neural network)의 기본 운영방식도 지도학습이다. 기본적으로 지도학습형 알고리즘을 구현할 때 그 훈련 단계(다음 절 참조)에서 판별 결과를 검증할 수 있는 수단이 있다고 해도 이를 의도적으로 활용하지 않는다면, 이는 비지도 학습에 가까운 형태가 된다(이 경우, 학습 과정에서 오류의 여지를 인정하면서 사후에 오류를 바로잡는(correction) 방식으로 구현된다). 이는 자기강화학습에서 보상(reward)을 보는 시각과 구현 전략에 따른 해석의 차이로 보는 것이 타당하다. 따라서 지나치게 ‘지도 학습 vs. 비지도 학습’이라는 도식에 얽매일 필요는 없다.

또한, 분류기의 수리적 구조에 따라 함수(적) 형태와 비함수 형태로 나누어 볼 수 있다. (기본형) 신경망을 제외한 지도학습 알고리즘들과 비지도 학습의 분산분석(analysis of variance)형 알고리즘은 함수 형태로 분류기를 산출하므로 (정도의 차이는 있지만) 판별에 대한 근거를 추적하고 검증하는 것이 비교적 쉽다. 클러스터링 분류기도 그 자체가 함수 형태를 취하지는 않지만, 그룹화 기준과 커널 및 거리함수를 통해 판별의 근거를 추론하거나 해석하는 것이 가능하다. 따라서 이들은 (최소한 이론적으로는) 화이트박스(white box)나 그레이박스(gray box) 수준의 투명성(transparency)을 보인다. 그러나 함수 형태를 갖지 않거나 너무 복잡해서 해석이 실질적으로 불가능한 분류기는 블랙박스 같은 불투명성(opaqueness)을 가진다. 그러나 다음 절에서 보듯이 화이트박스 형태의 분류기라고 하더라도 그 구현과정까지 고려하면 실제로는 블랙박스처럼 투명하지 못할 가능성이 크다.

[표 2] 분류형 알고리즘의 종류

구분	특성	알고리즘(분류기)
지도학습	- 판별 결과(맞거나 틀리거나)가 알려진 검증 데이터나 지식이 있고, 이를 활용하여 분류기를 산출 - 분류기 성능 평가가 용이하고 표준화되어 있음 * 예: confusion matrix, ROC - 훈련용 데이터의 품질에 성능이 크게 의존함 * 예: garbage-in garbage-out	- 회귀분석형(Regression Analysis) 분류기 * 일반 선형 및 비선형 회귀분석 * 로지스틱형(Logistic) 회귀분석: Logit, Probit 등 - 반복 분할형(Recursive Partitioning) 분류기 * CART(Classification and Regression Tree), MARS(Multivariate Adaptive Regression Spline) 등 - 커널함수(Kernel Function) 및 거리함수(Distance Function)기반 분류기 * SVM(Support Vector Machine), k-nearest neighbor 등 - (기본) 신경망(neural network) * N-계층 신경망, CNN(Convolution 신경망)에 의한 딥러닝(deep learning) 등
비지도 학습	- 판별 결과를 확인해 줄 수 있는 검증 데이터 및 지식을 사전에 확보하지 않음 - 성능 평가를 위한 또 다른 프레임워크나 전략이 필요함 - 새로운 분류체계를 창출해야 하는 경우에 적합 * 예: 신규 시장에서 고객집단분류, 핵심정보 요약 추출을 위한 차원 축소	- 클러스터링(Clustering) 혹은 군집화 알고리즘 * k-means, p-median, 계층화(hierarchical) 클러스터링 - 성분분석(component analysis, CA) 및 분산분석(Variance Analysis) * 주성분 분석(Principal CA), 독립성분분석(Independent CA) 등 * (공분산)행렬 분해(Covariance Matrix Decomposition): SVD Singular Value Decomposition) 등의 차원 축소(Dimension Reduction) 방식 - 비정상 행위 및 이상 징후 탐지(Anomaly Detection) 알고리즘 - 자기강화학습형(self-reinforcing learning) 딥러닝 신경망^{주)} * generative adversarial 신경망, Hebbian 신경망 등 * 자기조직화지도(Self-Organizing Map) 등

(출처: 김도훈(2018)에서 인용 - Cao(2017), Hastie et al.(2009), James et al.(2013), Mullainathan & Spiess(2017), Ng(2016), Pyle & San Jose(2016) 등의 문헌을 바탕으로 본고의 내용에 맞추어 다시 정리한 것임)

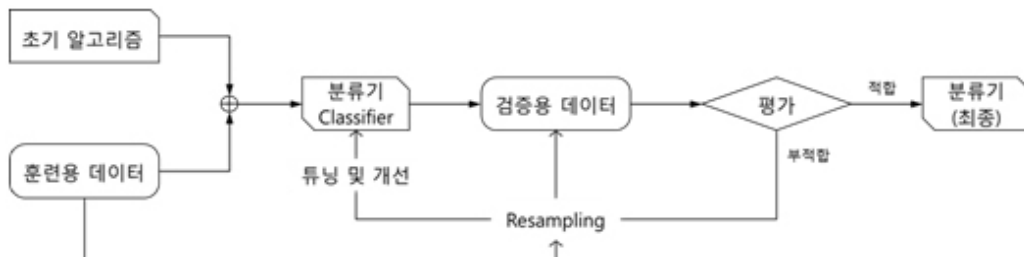
3 《 알고리즘 구현과정

(1) 분류기 산출과정의 개요

여러 유형의 알고리즘 중에서 AI 시스템과 관련하여 현재 가장 많은 관심을 받고 있는 분류형 알고리즘의 구현과정을 간략히 소개한다. 특히, 알고리즘 책임성에 관한 논의에서 필요한 부분을 중심으로 살펴본다. 머신러닝과 같은 분류형 알고리즘은 데이터와 밀접하게 상호연동하면서 분류기를 생성하기 때문에 그 성능은 데이터의 품질에 크게 의존한다. 데이터 자체에 문제가 없고 충분한 규모의 데이터가 확보되었다고 전제할 때, 분류형 알고리즘은 [그림 2]와 같은 절차로 구현된다.

먼저, 데이터 일부는 알고리즘에 의해 산출된 분류기의 성능을 검증하기 위해 남겨진다(검증용 데이터 집합, validation set). 단, 비지도 학습의 경우, [그림 2]의 검증용 데이터의 용도는 (성능 평가보다는) 분류기의 정상 작동을 파악하기 위함이다(이 단계가 생략될 수도 있다). 남겨진 훈련용 데이터 집합(training set)은 알고리즘을 가동하고 파라미터를 조절하는 데 사용된다. 훈련과정에서 지도학습과 비지도 학습 간에 차이가 있지만, 적용되는 알고리즘에 의한 영향력의 차이가 더 크다. 일단 분류기가 생성되면 검증용 데이터에 의한 성능 평가가 진행되고 문제가 발견되면 원데이터의 재추출(resampling)과 분류기 파라미터의 튜닝 및 개선을 거쳐서 새로운 분류기를 산출하는 과정을 반복하여 분류기를 다듬는다.

〈그림 2〉 분류형 알고리즘 구현의 기본 절차



(출처: 김도훈(2018)에서 인용)

(2) 분류기 성능 평가의 개요

분류기의 성능 평가는 지도학습과 비지도 학습 간에 차이를 보이는데, 지도학습의 경우 먼저 정오표(confusion matrix)를 만들고, 정답률(accuracy), 정밀성(precision), 민감도(sensitivity 또는 재현율, recall), 특이도(specificity) 등의 지표(metrics)를 활용한다. 비지도 학습의 경우에도 위와 같은 지표를 산출할 수 있다면 지도학습과 유사한 방식을 따른다.

민감도는 양성(positive)으로 정의된 결과(속성값)를 제대로 양성으로 판별한 비율(TPR, True Positive Rate)을 말한다. '1 - 민감도(= FNR, False Negative Rate)'는 통계학에서의 제2종 오류(type 2 error, β)에 해당한다. 특이도는 음성(negative)으로 정의된 결과를 음성으로 제대로 판별한 비율(TNR, True Negative Rate)이다. 따라서 '1 - 특이도(= FPR, False Positive Rate)'는 제1종 오류(type 1 error, α)에 해당한다. 통계학에서 제1종 오류와 제2종 오류를 동시에 줄이는 것이 어려운 것과 같은 이유로, 분류기의 판별 임계치(threshold)에 따라 특이도와 민감도는 상충관계(trade-off)를 보인다. 이 상충관계를 ROC(Receiver Operating Characteristic) 곡선으로 추적하면서 임계치 파라미터를 조정하고, ROC 곡선의 AUC(Area under the Curve) 면적을 계산하여 분류기의 성능을 평가한다. 일반적으로 AUC가 클수록 우수한 분류기로 평가한다.

성능 평가와 이를 통한 분류기 개선의 기본적인 접근법은 교차검증(cross validation)과 부트스트래핑(bootstrapping)을 활용한 반복과 재활용이다([그림 2]의 resampling 단계를 참조). 교차검증은 훈련용과 검증용 데이터 집합을 체계적으로 구성하여 검증을 반복함으로써 분류기의 성능을 높이고 과적합(over-fitting) 문제(다음 절 참조)를 해결하는 방법이다. 크기 p 의 검증용 데이터를 모든 경우의 수로 생성하는(leave- p -out) 심층적인(exhaustive) 교차검증과 사전에 k 개의 부분집합(k -fold)으로 데이터를 분할하여 k 번만 교차검증 하는 비심층적 방식으로 크게 나뉜다. 부트스트래핑은 원래 표본의 수가 작을 때 통계적 추론의 신뢰성을 높이기 위해(예를 들어, 추정량(estimator)의 분산을 줄이려는 목적에서) 표본을 반복 추출하여 사용하는 방법이다. 머신러닝과 같은 분류형 알고리즘에서 부트스트래핑은 데이터(표본)의 체계적인 재활용을 통해 분류기의 강건성(robustness)⁷⁾을

높이고 과적합 문제를 해결하기 위해 활용된다.

(3) 구현과정에서 알고리즘 투명성 관련 이슈

[그림 2]의 기본 절차만으로는 현실적 문제에 곧바로 적용할 수 있는 분류기를 산출하기 어렵다. 특히 머신러닝과 같이 데이터와 알고리즘이 유기적으로 상호작용하는 분류형 알고리즘의 경우, 훈련 과정에 사용된 데이터에서는 분류기가 우수한 성능을 보이지만, 다른 데이터에서는 성능이 떨어지는 과적합(over-fitting) 문제가 발생할 가능성이 높다. 성능 평가에서 교차검증과 부트스트래핑을 통해 과적합 발생 가능성을 낮출 수는 있지만, 데이터에 내재된 편향에 따른 근본적인 한계는 극복할 수 없다. 이러한 잠재적 편향의 영향을 줄이고 어떤 형태의 데이터에 대해서도 일관된 성능을 보일 수 있는 유연한 분류기를 산출하기 위해서는 보다 정교한 개선이 필요한데, 그 해법에는 정규화와 랜덤화(randomization) 과정이 수반된다.

정규화(regularization)는 훈련과정에 필요한 최적화를 의도적으로 제한해 분류기가 다소 특이한 데이터의 속성까지 고려하는 것을 방지한다. 최적화를 제한하는 방식과 제한 수준에 따라 다양한 정규화 방법이 가능하다. 보통 페널티(penalty) 혹은 손실(loss) 함수를 도입한다. 예를 들어, 회귀분석형 분류기에 적용되는 LASSO(Least Absolute Selection and Shrinkage Operator)는 페널티 함수로 L1-norm을 사용한다. 과적합 문제 해결에서 정규화는 대부분 위에서 소개된 교차검증과 함께 사용된다.

랜덤화는 분류기의 강건성과 유연성을 높이기 위해 분류기 대부분에 적용된다.⁸⁾ 여기에는 크게 다음 두 가지 전략이 있다. 첫째, 배깅(bagging)은 훈련과정에서 데이터를 무작위로 분리해 여러 부분집합을 생성하고 부분 집합별로 서로 독립적인 복수의 분류기를 생성한 후 이를 종합하는 방식이다. 생성된 여러 분류기 집합을 앙상블(ensemble)이라고 부른

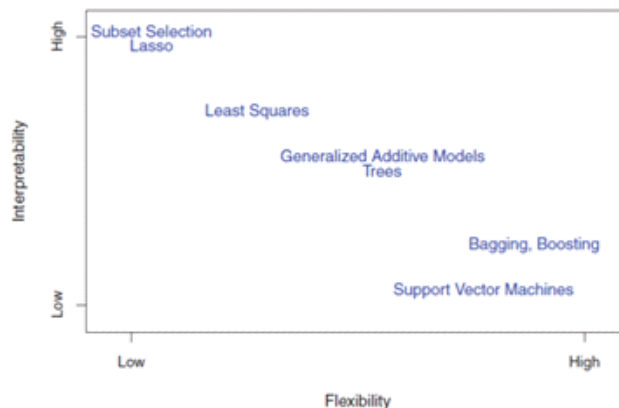
7) 통계학에서의 강건성은 해당 모형이 통계학적 기본 가정과 전제에 완벽하게 부합되지 않거나 표본이 바뀌었을 때에도 그 추론의 성과가 크게 악화하지 않는 것을 말한다. 예를 들어, 통계적 가설검정에 많이 사용되는 F-통계량은 (균형설계일 경우) 등분산(variance homogeneity) 가정에 부합하지 않는 경우에도 일반적으로 꽤 유효한 성과를 보인다. 여기에서 강건성도 위와 비슷한 의미에서 데이터가 바뀌어도 해당 분류기가 어느 정도 일정한 성능을 유지한다는 것을 의미한다.

8) 바로 앞에서 소개한 부트스트래핑도 표본에 대한 랜덤화라고 할 수 있다. 바로 다음에 소개하는 배깅과 표본을 대상으로 한 랜덤화라는 점에서는 동일하지만 구체적인 활용방식은 다르다.

다. 둘째, 데이터(표본)를 대상으로 하는 배깅이나 부트스트래핑과는 달리, 부스팅(boosting)은 설명변수 혹은 속성의 랜덤화라고 볼 수 있다. 부스팅은 사용 가능한 속성 일부를 이용하여 정밀도가 떨어지는 약한 분류기(weak classifier 또는 weak learner)를 여러 개 생성한 뒤에 다시 훈련과정을 거쳐 약한 분류기에 적절한 가중치를 부여하고 종합해 새로운 분류기를 생성한다. 의사결정나무(decision tree)에 기반한 CART 분류기를 예로 들면, 의도적으로 깊이(depth)를 제한하여 작고 간단한 의사결정나무가 약한 분류기에 해당한다. 이런 작은 의사결정나무를 여러 개 생성하여 서로 경쟁시키고 최종적으로 (주로 가중치에 따라) 종합하여 새로운 분류기를 생성한다.

배깅, 부스팅, 교차검증, 부트스트래핑은 모두 랜덤화로 분류기의 과적합 문제를 해결함과 동시에 보다 강건한 분류기로 업그레이드하고자 한다. 실제로 이들은 복합적으로 적용되기 때문에 구현과정에서 명확하게 구별하기는 쉽지 않다. 예를 들어, 기본형인 CART가 발전된 형태인 랜덤 포레스트(random forest)는 교재와 문헌에 따라서 배깅 또는 부스팅의 일종으로 다뤄진다(그러나 이를 어느 범주에 포함할 것인지는 본질적인 문제가 아니다). 그런데 [그림 3]에서 보듯이 기본형의 알고리즘은 해석·구분할 수 있을지 몰라도, 동일한 방식을 적용해 고도화된 분류기는 강건성과 유연성이 높아진 대가로 불투명성이 높아져 그 결과에 대한 해석이 어려워진다. 이는 다음 절에서 소개할 알고리즘 책임성에 관한 법제도적 이슈와도 밀접히 관련된다.

〈그림 3〉 해석 가능성과 유연성 사이의 상충관계



(출처: James et al.(2013)의 p.25에서 인용)



알고리즘 책임성

1 » 개념과 배경

AI가 보편화되면서 알고리즘에 의한 ‘자동화된 불평등’에 대한 우려가 커지고 있다. 알고리즘을 이용해 개인이나 사물을 특정 유형으로 판단하는 과정을 ‘프로파일링(profiling)’이라고 한다. 예를 들어, 개인이나 집단에 대한 정보를 모아서 이들을 분류하기 위해 직장 내 업무, 경제적 상황, 개인적 선호(건강, 관심사 등), 신용도, 행동 패턴, 위치 및 이동 경로 등을 분석하고 이를 이용해 대상의 향후 행동을 예측하거나 성과를 평가하는 것이다. ‘개인화’라는 이름으로 출시되는 영화 및 도서 추천 서비스나 맞춤형 광고는 이런 프로파일링을 과정을 통해 제공된다.

AI가 인간의 편견과 잘못된 데이터의 입력으로 자동화된 차별의 수단이 된다면 일상을 파괴할만한 위험이 될 수 있을 것이다. 이미 AI를 이용해 얼굴 사진만으로 성적 취향이나 잠재적 범죄 가능성, 정치적 성향까지 판단할 수 있다. 2018년 9월 미국 스탠퍼드대학은 얼굴 사진만으로 동성애자를 식별할 수 있는 인공지능을 발표했다. 온라인데이트 사이트에 올라온 사진을 이용해 남자의 경우 91%, 여자의 경우 83%의 정확도로 동성애자를 구별했다. 중국은 범죄자 검거는 물론 잠재적 범죄자까지 예상할 수 있는 안면인식 프로그램을 개발했다.

유럽연합(EU, European Union)은 2018년 5월 25일 ‘일반개인정보보호법(GDPR, General Data Protection Regulation)’⁹⁾을 전격 시행하였다. GDPR은 그 전문에서 EU 시민은 법적 효력을 초래하거나 이와 유사하게 중대한 영향을 미치는 사항에 대하여 프로파일링과 같은 자동화 프로그램의 적용을 받지 않을 권리를 갖는다고 규정하고 있다.

9) <https://www.eugdpr.org>

또한, GDPR 22조는 ‘프로파일링을 포함한 자동화된 의사결정’에 대하여, 알고리즘에 의한 자동화된 결정에 반하여 인간의 개입을 요구할 권리, 알고리즘의 결정에 대한 설명을 요구하고 그에 반대할 권리 등을 명시적으로 규정한다. EU는 AI를 활용하는 기업들이 이용자에게 자동화된 결정이 미치는 범위와 예견된 결과의 심각성 등에 대한 정보 제공을 요구한다. 예를 들어, 실제로 많은 금융기관에서 시행하고 있는 신용평가에서도 데이터 제공의 주체인 개인은 자신의 정보가 처리되는 절차에 대해 알 수 있어야 하며 그 처리를 거부할 수 있는 권리를 가진다.

위의 사례와 같이, AI의 기반이 되는 알고리즘의 신뢰성의 문제는 자동화된 의사결정과 정에서 차별과 오류의 명확한 원인을 파악하거나 이해하는 게 어려울 뿐만 아니라 알고리즘 설계에 주관적 편향이 개입될 가능성에서 기인한다(박종훈, 2016, 2017; Baer & Kamalnath, 2017; Raymond et al., 2017). 치안이나 의료와 같은 공공성이 강한 분야에서 결과의 공정성은 과정의 투명성(process transparency)을 요구하는 경우가 많은데, 알고리즘이 적용되는 경우 투명성을 보장하기 어렵다. 투명성은 경쟁법(competition law)의 관점에서 중요한 이슈이다. 파스칼레(2016)는 평판, 검색, 금융에서 빅데이터와 알고리즘의 불투명한 활용을 지적하면서(파스칼레는 이를 ‘비밀주의’라고까지 부른다), 사회적 개방성과 시장의 공정성 관점에서 알고리즘의 부정적 기능을 경고한다.

미국과 EU에서는 온라인에서의 알고리즘 편향과 가격설정 알고리즘의 담합이나 불공정 거래의 가능성에 대해 많은 논의가 이뤄졌으며(Raymond et al., 2017 및 5.2절 참조), EC와 OECD 같은 국제기구에서도 관심을 기울이고 있다(김건우, 2017; 최난설현, 2017; Blockx, 2017; Capobianco & Gonzaga, 2017 등). 사실 알고리즘 편향성에 관한 법제 도적 논의는 2010년부터 시작되었다. 미국 공정거래위원회(FTC)는 2010년에 Google 검색 알고리즘의 불공정 가능성을 조사하면서 세간의 주목을 받았다. FTC가 2013년 증거 부족을 이유로 심의를 종료했지만, 이 문제는 많은 플랫폼이 채택하고 있는 양면 시장(two-sided market) 비즈니스 모델(김도훈, 2016; Garcés-Tolon, 2017)에서 복합 서비스(bundling 및 tying)가 제공될 때 언제든지 다시 불어질 소지가 있다. 가격설정 알고리즘도 2015년 미국 법무부(DOJ)의 반규제 분과가 Amazon 마켓플레이스에서 거래되는

포스터 상품의 가격설정 알고리즘에 문제가 있다고 판결한 바 있다(Li & Xie, 2017). DOJ는 2016년 Uber의 동적 가격설정(dynamic pricing) 방식에도 문제가 있다고 지적했다.

그러나 이러한 문제가 알고리즘의 투명성을 보장하는 것으로 해결되기는 어렵다. 앞 절(2.3절 참조)에서도 소개한 바와 같이, 기술적 측면에서 머신러닝 같은 분류형 알고리즘의 구현방식을 보면 투명성 확보가 원천적으로 매우 어렵다. 따라서 (Tesla의 주행사고처럼) 알고리즘의 오류가 확실해도 일반적인 소프트웨어 디버깅 방식으로는 문제를 해결할 수 없다. 단지, 수차례 실험을 통해 경험적으로만 신뢰성과 안전성을 점검할 수 있을 뿐이다. 그러나 이 역시 품질보장(quality assurance) 등 상업화 측면에서 볼 때 매우 비효율적인 접근법이어서 실현 가능성이 작다. 만약 알고리즘에도 현재의 제조물 배상책임(product liability)에 따르는 품질수준이나 6 시그마(6 sigma)와 같은 TQM (Total Quality Management)에서의 품질관리를 요구하는 법제도가 시행된다면 그로 인한 비용이 막대할 것이다. 이러한 이유에서 머신러닝에서 데이터 의존도를 낮추려는 연구가 IBM, Nvidia, Qualcomm 등에서 진행되고 있다(one-shot learning이나 정교한 부스팅 및 가지치기(pruning) 등을 바탕으로 한 smaller algorithm 등: Sun, 2018).¹⁰⁾

또한, 투명성이 확보된다고 해서 모든 문제가 해결되는 것도 아니다. 예를 들어, 투명성 보장이 오히려 해킹이나 기업의 지식재산권(IP, Intellectual Property) 침해 등에 악용될 가능성도 있다. OECD에서 잠정적으로 제안한 해결 방향도 투명성을 강조하고 있지만, 그 유용성을 의심하는 경향이 강하다(Capobianco & Gonzaga, 2017). 투명성이 해결책이 될 수 없는 궁극적인 이유는 머신러닝을 비롯한 분류형 알고리즘이 인과관계(causal-effect)를 직접 다루는 것이 아니라 상관관계(correlation or association)를 바탕으로 마치 인과관계를 추론하는 듯이 보이기 때문이다(Diakopoulos, 2016). 따라서 이론적으로도 알고리즘의 투명성은 해결책이 되기 어렵다는 것을 알 수 있다.¹¹⁾ 미국 피츠버그대학병원은 10~11%에 이르는 폐렴 환자의 사망률을 낮추기 위해 머신러닝을 실험에

10) 이와 같은 연구가 성공하면 방대한 데이터 처리에 필요한 CPU나 GPU도 적게 소요되므로 하드웨어 제조업에도 직접적인 영향을 줄 것이다.

11) 투명성이 궁극적인 해결책이 될 수 없다는 것은 의료진단 AI인 Deep Patient를 개발하고 운영하는 조엘 더들리가 한 다음과 같은 말에서도 드러난다. “우리는 모델을 만들 수 있지만, 그 모델이 어떻게 작동하는지는 알 수 없다.”

도입했는데, 머신러닝 알고리즘은 천식 증상이 나타나면 폐렴이 호전된다는 결론을 내놓았다. 이는 폐렴에서 천식이 나타나면 위중해진 상태여서 집중 치료를 한 것을 머신러닝이 상관관계로 추론한 결과로, 인과관계를 설명하지 못한 상태에서 나온 오류였다. 이처럼 투자, 의학적 판단, 군사적 결정에서 AI가 아무리 효율적이어도, 판단의 근거를 설명할 수 없다면 사용이 어렵다. 이러한 한계를 해결하기 위해 미국 국방부는 ‘설명 가능한 AI(XAI, 4.2절 참조)’를 핵심적 연구과제로 추진 중이다. IBM의 최고경영자 지니 로메티 역시 2018년 다보스 세계경제포럼에서 “기업은 AI 알고리즘의 결정이 어떻게 도출되는지를 설명해야 할 의무가 있다. 그렇게 할 수 없는 기업이라면 퇴출당해야 한다.”라고도 말했는데, IBM 역시 AI 편향성을 추적할 수 있는 클라우드 서비스인 오픈스케일(OpenScale)을 출시했다(4.2절 참조).

2 » 다양한 관점과 접근법

법제도 전통에서는 책임성(accountability)을 위해 투명성을 요구하는 경향이 강하다. 그런데 알고리즘의 기술적 특성으로 인해 투명성을 확보하는 것이 매우 어려워, 법제도 연구에서도 이러한 전통적 접근법에 대한 회의적인 시각이 부상하고 있다(Desai & Kroll, 2018; Perel & Elkin-Koren, 2017). 첫째, 알고리즘이 공개되어도 복잡성 때문에 현실적으로 해석할 수 없거나 중의적 해석만이 가능하다. 둘째, 2010년 YouTube와 Viacom 판례처럼 알고리즘 자체가 사적 자산 즉, 지적재산(IP)으로 보호받아야 할 대상인 경우가 많다. 셋째, 알고리즘 (특히 학습기능이 있는 경우) 자체가 동적(dynamic change)으로 변한다. 넷째, 알고리즘 스스로도 주어진 입력에 대해 어떤 산출이 발생할지 모르는 경우 (입력-산출 관계가 명시적이지 않은 경우)가 많다. 이러한 이유에서 알고리즘 감사 (algorithm audit, 김건우, 2017; Doneda & Almeida, 2016 등) 제도를 도입하더라도 실효성 있는 대안 마련이 쉽지 않다.

법학 분야에서 제안하는 방안들은 최대한 투명성 혹은 유사 투명성을 제공하도록 해 책임성을 담보하고자 한다. 예를 들어, Perel & Elkin-Koren는 ‘블랙박스 두드려 보기 (black box tinkering)’를 제안한다(Perel & Elkin-Koren, 2017). 불투명한 알고리즘에 사전 혹은 사용 중에 다양한 입력-산출 실험을 적용하여 그 결과를 공개하고 문제가 발견될 경우 알고리즘을 보완 수정하도록 하는 법적 근거를 마련하자는 것이다. 또한, 두드려 보기 절차를 제공하는 알고리즘 운영자에게는 문제가 발생하면 면책(safe harbor)도 가능하게 한다. 이 접근법은 이미 실행 중인 DMCA(Digital Millennium Copyright Act)와 유사하므로 DMCA가 현실적으로 부딪히는 한계를¹²⁾ 기술적 법제도적으로 극복해야 하는 도전과제가 남는다.

Desai & Kroll(2018) 및 Diakopoulos(2016)도 알고리즘의 태생적 불투명성을 받아

12) 대표적으로, ‘발견 즉시 조치(notice and take-down)’에 대한 규정은 현실적으로 실행에 어려움이 따른다. 일단, 저작권(copyright)이 침해되었는지 그 여부를 발견하기 어려울 뿐만 아니라, 즉각적인 조치를 하는 것이 일반인들에게는 상당한 부담과 비용이 수반된다.

들인 상태에서 관련 법제도를 마련할 것을 제안한다. 더욱 현실적인 대응방안으로, 먼저 알고리즘이 활용되는 목적과 상황에 따라 공공용(public interest)과 사적인 용도(private interest)로 구분해 대응하는 것이 필요하다. 공공성 차원에서는 알고리즘의 '정의로운 구현과 이에 상응하는 절차(justice and due process)'가 가장 중요하며, 사적인 용도(특히 상업용)에서는 '불법적인 차별(illegal discrimination)'을 억제하며 '민주적으로 공평하게 사용되는(fair use)' 환경을 조성하는 데 초점을 맞춘다. 소프트웨어 엔지니어링에서 사용되는 화이트박스 수준의 검증은 불가능하지만, 블랙박스에 대해서도 충분한 검침이 가능하며, 문제가 발생했을 때 적극적으로 사후 개입(ex-post analysis and oversight)을 할 수 있는 법제도적 장치를 마련해야 한다. 그러나 Desai & Kroll(2018)도 궁극적으로 내부고발자나 감시자(whistle-blower)에 의존할 수밖에 없다는 점을 인정하고 있다. 이처럼 알고리즘의 불투명성과 동태성은 법제도로 알고리즘의 설계, 구현, 운영의 책임을 구체화할 때 피할 수 없는 한계 낳는다. 특히, 머신러닝과 같이 데이터와 알고리즘이 긴밀하게 연결되는 경우 (과거의 데이터 등에 대한) 동태적 상황 의존성을 원천적으로 배제할 수 없어 알고리즘 책임성에 관한 원칙과 규정이 있어도 사후 개입의 실효성은 약화될 수밖에 없다.

이러한 배경에서 알고리즘 책임성에 대해 애초부터 보다 기술적인 대안을 마련할 필요가 있다는 인식이 퍼지고 있다. 마치 자동차를 설계·생산하는 단계에서 배출가스나 안전 기준을 강제하는 것과 유사한 맥락을 따르는 것이다. 컴퓨터과학의 대표적 국제학술단체인 ACM(Association for Computing Machinery)의 공공정책분과(Public Policy Council)에서는 알고리즘 책임성을 높이기 위해 본격적인 연구에 착수했다. 먼저 알고리즘 개발과 구현에서 책임성을 강화하기 위해 7가지 원칙을 [표 3]과 같이 제시했다(USACM, 2017).

[표 3] USACM과 EUACM의 알고리즘 책임성을 위한 7대 원칙

원칙	설명
인지 가능성 (Awareness)	알고리즘이 사용되고 있음을 충분히 알리고 가능하면 사용법도 공지함
접속 및 시정 (Access and Redress)	알고리즘에 대한 조사가 원칙적으로 가능해야 하며 오류 및 잘못된 의사결정에 대한 수정지침을 사전에 제공함
책임감 부여 (Accountability)	알고리즘 구현 및 운영을 담당하는 주체를 명확하게 하고 책임감(responsibility)을 부여함
설명력 (Explanation)	인간이 이해할 수 있는 수준으로 작동원리(logic)에 관해 설명할 수 있어야 함
데이터 출처 (Data Provenance)	알고리즘의 올바른 작동을 위한 충분한 데이터를 확보하고 데이터 출처에 대한 기록과 무결성(integrity)을 제공함
감사 가능성 (Auditability)	로그와 작동 기록을 남김으로써 감사와 분쟁 해결이 가능하도록 함
타당성 평가와 검사 (Validation and Testing)	알고리즘 성능에 대한 평가방식을 제공하고 적절한 방식으로 검사할 수 있도록 함

(출처: USACM(2017)에서 인용하면서 이 글의 내용에 맞추어 일부 내용을 재정리하였음)

ACM의 알고리즘 책임성에 대한 노력은 현재 선언적 수준에 머물러 있다. 2~3년 전에 의욕적으로 출발하였던 기술공학적 관점에서의 알고리즘 및 AI 거버넌스에 대한 논의도 편향성, 검열, 차별, 사생활 보호, 지적재산권 보호, 지배력 남용에 대한 원론적인 방향 제시와 업계의 자율규제(self-regulation)에 의존하는 정도에 머물러 있다(Doneda & Almeida, 2016; Saurwein et al., 2015). 그런데 최근 들어 보다 구체적인 방안을 마련하기 위해서 여러 연구가 본격적으로 추진되고 있는데, 다음 절에서는 대안 마련을 위한 몇 가지 방향을 소개한다.



알고리즘 책임성에 대한 접근법: 대안과 전망

1 » 알고리즘 책임성의 필요성

알고리즘 책임성 이슈는 알고리즘과 AI 기술 발전에 따른 부작용을 제도적으로 최소화 하려는 노력 중의 하나이다. 실제로 알고리즘이 부적절한 시점에 개입하여 주식거래에 심각한 문제가 발생했던 사례가 있으며, 음성인식 오류로 인해 주문을 잘못하거나 거짓 신고가 발생하는 상황도 이제는 쉽게 예상해 볼 수 있다. 이제 가짜 뉴스와 같이 여론 조작에 알고리즘이 활용되거나 알고리즘 개발자의 인종과 종교에 대한 편견이 프로그래밍 되는 위험을 사전에 방지하려는 노력이 필요한 시점이다.

여러 나라가 알고리즘과 AI의 공공성, 책무성, 통제성, 투명성 확보를 통해 이러한 부작용에 대비하려 한다. 이러한 시도는 보통 개발자-공급자-이용자가 준수해야 할 가이드라인을 제정하는 것으로 귀결된다. 미국은 2016년 3월에 이미 가이드라인을 마련했으며 EU는 로봇법과 GDPR을 통해 몇 가지 기준과 원칙을 밝혔다. 일본도 2016년에 AI 연구와 개발에 관한 국제 규범을 만들 것을 제안했으며, 우리나라도 2017년부터 관련 가이드라인 제정에 착수했다.

먼저, 공공성은 알고리즘 개발 단계에서부터 사회적 차별 요소를 배제하기 위한 기준이다. 사회적 약자의 AI 접근과 활용을 보장하여 디지털 차별(digital divide)을 최소화하고, 사회문제 해결에 더 적극적으로 AI를 활용하려는 것이 그 목적이다. 공급자는 알고리즘을 활용할 때 공공의 이해관계를 고려해 제품과 서비스를 공급해야 한다. 또한, 이용자는 공공의 이익에 반하여 알고리즘을 남용하거나 악용하지 않으며 본인이 의도하지 않더라도 타인에게 피해가 발생하지 않도록 사전경보(alert) 기능을 도입한다.

책무성은 개발자가 이용 과정과 결과에 책임을 지게 하는 것이다. 개발자는 기술 개발에서 데이터 및 정보 활용 기준을 마련하고 알고리즘 개발 과정에서 윤리적 절차를 준수해야 한다. 공급자는 오작동과 사고를 대비하고 합리적인 사후 보상책임 기준을 수립한다. 이용자에 대해서도 알고리즘 활용·교체·갱신·폐기 과정에서 적절한 책임을 질 수 있는 제반 규정을 마련한다.

통제성의 원칙은 개발자가 발생 가능한 문제를 검토하여 오작동에 대한 수정이 가능하도록 하는 것이다. 가능한 각종 위험 프로파일도 사전에 제작하고 발생한 오류를 차단하기 위한 제어장치도 마련해야 한다. 공급자도 알고리즘 운영에서 발생할 수 있는 여러 위험요인과 문제점을 사전에 검증하고 지속해서 모니터링해야 한다. 이용자 역시 알고리즘을 활용할 때 필요 이상 남용하거나 자의적으로 조작하는 것을 금지한다.

투명성은 알고리즘과 AI 시스템 개발-공급-운용-사용 프로세스가 제3자에 의해 인지되고 이해될 수 있어야 한다는 원칙이다.¹³⁾ 개발자는 알고리즘 개발과정을 은닉하지 못하도록 하며, 운영 시 부작용을 예측해 공급자와 공유할 수 있어야 한다. 공급자는 발생 가능한 위험요인을 이용자에게 알리고 정확한 이용방법을 제공한다. 이용자 역시 알고리즘을 활용한 기록을 남기고 부적절한 경로를 통해 알고리즘을 악용하지 못하도록 한다. 그러나 3장에서 살펴보았듯이 알고리즘 자체에 대한 투명성 확보는 원천적으로 매우 어렵다. 머신러닝의 경우 개발자는 몇 가지 규칙과 데이터만 줄 수 있을 뿐이고 이후 분류기 산출과 갱신은 시스템 스스로 학습을 통해 이루어지기 때문에 실제로는 모두에게 미지의 영역으로 남는다. 실제로 2013년에 쇼기(일본식 장기)에서 일본 최고수를 이긴 AI 프로그램 ‘포난자’를 개발한 야마모토 잇세이도 머신러닝은 본질적으로 불투명하다고 말한 바 있다.

이러한 기술적 난관에도 불구하고 AI와 알고리즘이 판단 근거, 과정, 결과에 대해 일반 사람들도 이해할 수 있는 방식으로 작동해야 한다는 요구가 점차 강해지고 있다. 민간 영역에서도 기업들이 불투명성과 신뢰 문제로 AI 시스템과 자동화된 의사결정 프로그램의 채택을 꺼리고 있다(예를 들어, Facebook은 뉴스와 게시물 필터링에 여전히 많은 인력을

13) 여기에서 투명성은 앞 절에서 소개한 알고리즘 투명성과는 다른 차원에서의 투명성이다. 제도적 차원에서 투명성이므로 알고리즘 투명성보다는 좀 더 추상적인 원칙에 해당한다.

활용한다). 또한, 기술적 측면에서도 블랙박스처럼 취급되는 알고리즘은 악의적 의도로 조작된 투입요소(데이터, 숨겨진 코드 등)에 의해 시스템이 무력화되는 적대적 공격(adversarial attack)에 취약하다는 단점을 가진다. 이러한 사회적 요구사항과 분위기는 AI와 알고리즘의 정착을 위해 넘어야 할 도전과제이다.

따라서 AI와 알고리즘이 사회에서 받아들여지고(사회적 수용성) 신뢰를 얻기 위해서는 단순한 선언의 수준에서 그칠게 아니라 기술 및 운영 차원에서 알고리즘의 책임성을 어느 정도 확보할 수 있는 조치가 필요하다. 알고리즘과 AI 거버넌스의 공정성과 투명성 확보를 위한 노력이 주목을 받는 이유이다. 설명 가능한 AI(XAI)나 AI 위험관리(risk management) 프레임워크 등에 관심이 쏠리는 것도 이러한 배경에서 이해가 된다. 예를 들어, 알고리즘이 편향된 데이터에 의해 한쪽으로 치우친 결과를 낼 때 이러한 상황을 충분히 감지하여 예방하거나 오류를 검증하고 고칠 수 있는 수단은 향후 기업들이 대비해야 할 중요한 위험관리 대상으로 부상하고 있다.

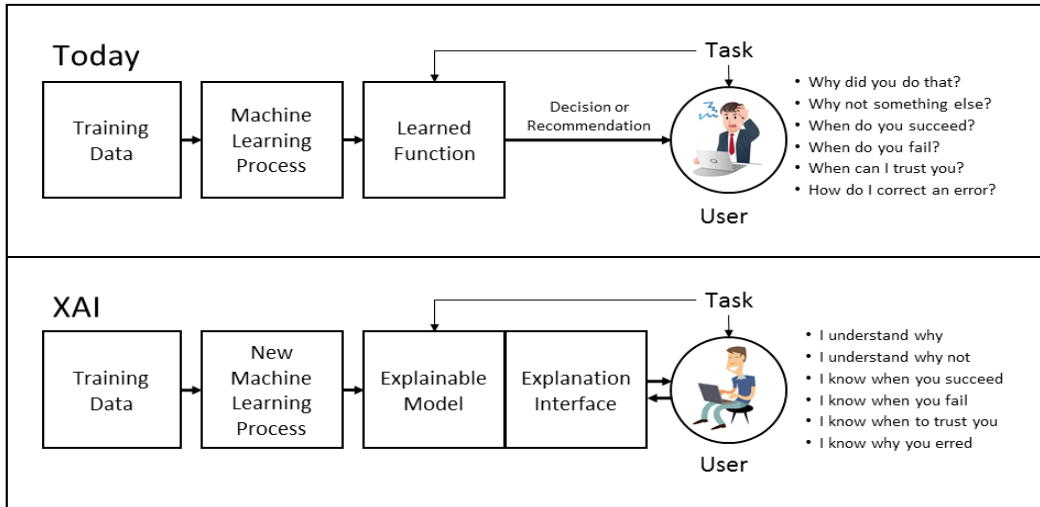
2 » 새로운 접근법: 대안과 방향

(1) 설명 가능한 인공지능(XAI)

위에서 잠시 설명한 바와 같이, (법제도적 관점에서) 책임성 확보를 위한 가장 효율적인 방법은 투명성을 확보하는 것이다. 투명성은 공공성, 책무성, 통제성의 바탕이 되기 때문에 알고리즘의 투명성을 확보할 수 있는 기술적 해결방안을 찾으려는 노력이 주목을 받고 있다. 대표적인 예로, AI 선진국들은 복잡한 신경망 속에서 학습된 정보를 설명하기 위해 설명 가능한 딥러닝 시스템을 개발하고 있다.

지금까지는 알고리즘의 효율성 향상에만 집중했지만, 공정성을 배제한 채 정확성을 높이는 것이 현실적으로는 한계가 있다는 것을 깨닫기 시작했다. (사후라도) 블랙박스와 같은 알고리즘 내부에서 어떤 근거로 결론이 도출되는지 파악할 수 있어야 AI 시스템이 신뢰를 얻을 수 있으며, (언젠가는 발생할 수밖에 없는) 오류를 수정하는데 활용할 수 있기 때문이다. 최근에 본격적으로 연구되고 있는 ‘설명 가능한 AI(XAI, Explainable AI)’는 블랙박스를 최소한 그레이박스로 전환해 머신러닝의 한계를 보완할 수 있는 대안으로 볼 수 있다. 즉, AI가 왜 그런 결과를 도출했는지에 대한 이유를 납득할 수 있는 방식으로 설명해준다는 것이다. XAI는 AI가 의료, 법률, 국방과 같이 공공성이 높은 분야에서 요구되는 투명성과 신뢰를 담보할 수 있도록 의사결정에 대한 이유를 사람이 이해할 수 있는 방식으로 제공하려는 목표에서 시작되었다. 공공분야뿐만 아니라 금융, 유통과 같은 민간 부문에서도 투명성 확보는 법제도적 요구사항(예: GDPR)을 준수하는 데 필요할 뿐만 아니라 사용자의 신뢰를 높이기 위한 기업의 전략 차원에서도 중요하다.

〈그림 4〉 XAI의 개념



(출처: DARPA(2018))

현재 XAI 연구에 가장 적극적인 미국 방위고등연구계획국(DARPA)¹⁴⁾은 2017년부터 약 800억 원의 예산을 투입하여 2021년 4월까지 프로젝트를 진행 중이다. MIT와 Google 등 민간 부문에서도 DARPA와의 협력을 통해 혹은 독립적으로 활발히 연구를 진행하고 있다. 아직까지 XAI는 주로 딥러닝에 의한 시각화 과정에 집중하고 있는 것으로 보인다. 예를 들어, 딥러닝에 의한 분류기가 어떤 사물을 고양이로 인식했을 때 또 다른 AI 도구를 활용해 어느 시점에 고양이로 판별했고 어떤 경우에는 아니라고 답을 했는지 모니터링하고 분석하는 것이다. 이처럼 XAI는 도출한 결과를 스스로 설명함으로써 ‘간접적인’ 투명성을 제공해 사용자 신뢰성을 확보하고자 한다. 이를 통해 AI 사용자나 서비스 대상이 AI의 결정을 이해하고 그 결과를 신뢰할 수 있도록 하는 것이 목적이다. 고양이를 판별하고 설명하는 방식을 다른 분야로 확대·응용하면 의료 현장에서 암이나 질병을 조기에 진단하고 국방에서 아군과 적군을 식별하는 데 신뢰성을 크게 높일 수 있을 것으로 기대된다.

14) 인터넷의 메카이기도 한 DARPA는 이미 1960년대에 제1세대 AI로도 불리는 전문가시스템(ExS, expert system)을 개발을 선도한 바 있다. 전문가시스템은 규칙기반 추론(rule-based reasoning)에서 시작하여 사례기반 추론(CBR, case-based reasoning)으로 발전하면서 정형화된 업무의 자동화에 크게 기여하였다.

XAI는 기업의 자동화된 의사결정의 효과에 대한 검증을 가능하게 하고 책임소재를 보다 명확하게 함으로써, AI 시스템 도입과 운영의 부작용에 대응할 수 있는 도구를 제공한다.¹⁵⁾ 혁신을 추구하는 기업 환경에서 이러한 XAI의 역할은 매우 중요하다. 실제로 IBM 기업가치연구소가 전 세계 5천 명의 기업 임원을 대상으로 조사한 결과에서 응답자의 82%가 AI 도입을 검토하고 있음에도 불구하고, 60%는 AI 시스템에 활용되는 데이터를 신뢰할 수 없고, (AI와 관련된) 기업 내·외부의 여러 규제를 우려한다고 답하였다. XAI를 통해 AI 시스템의 신뢰성이 향상되고 법제도에 보다 유연하게 순응할 수 있는 여건이 마련된다면 AI 시스템 도입과 확산에 크게 이바지할 수 있을 것이다.

XAI는 신경망 기반 머신러닝(딥러닝)을 중심으로 크게 세 가지 방향에서 추진된다. 첫째, ‘왜 이것을 고양이라고 판단했는지’를 설명해주는 것과 같이, 분류기의 판별 근거를 제공한다. 둘째, 금융, 의료, 경제 분석 등에서 모델(의사결정의 프레임)을 산출했을 때 어떠한 기준과 분석을 바탕으로 제시된 것인지를 설명한다(이는 모델 자체를 설명하는 것으로 단순히 판별 이유만을 제시하는 첫 번째 방향과는 다르다). 첫 번째와 두 번째 방향이 기존에 블랙박스로 취급됐던 AI의 내부를 보여주는 ‘설명 모델’에 관한 것이라고 하면, 세 번째는 출력 인터페이스에 관한 것이다. 즉, AI의 산출물(판별 결과 및 모델 등)을 인간이 이해할 수 있는 방식으로(예를 들어, 사람이 작성한 문서처럼 분석적 의미를 담은 자연어로 된 보고서 등) 설명할 수 있는 인터페이스를 설계하는 것이다. XAI에 관한 특허도 2014년 이후 급증하고 있는데, 앞의 첫 번째와 두 번째에 관련되는 신경망의 네트워크 구조와 관련된 특허가 가장 큰 비중을 차지하며, 해석 인터페이스와 시각화에 대한 특허가 그 다음을 잇는다(윤태승, 2018).

첫 번째와 두 번째의 경우, 신경망을 기반으로 한 접근법은 크게 두 가지이다. 하나는 분류기의 네트워크 구조 자체를 되도록 단순하고 투명하게 만드는 것이고, 다른 하나는 분류기의 결과를 해석하는 부가적인 모델을 생성하고 인터페이스를 통해 분류기와 접목해

15) 일반적으로 기업이 혁신적으로 신기술을 도입하면 부작용이 수반되기 마련인데, 지금까지의 기술은 문제가 발생했을 때 그 원인과 책임소재 파악이 AI보다는 상대적으로 용이하였다. 그러나 Tesla 자율주행 사고에서 보듯이, AI 활용에 따른 문제는 그 원인과 책임소재를 파악하는데 상당한 시간과 노력이 필요하다. 이는 AI를 통해 혁신을 추구하는 기업들에게는 큰 부담이며, 혁신의 도입과 확산을 방해하는 요인이다. XAI는 혁신 기업들이 직면한 이러한 문제를 해결하는 데 도움을 줄 것으로 기대된다.

서 통역하는 방법이다. 전자의 접근법은 알고리즘 투명성을 직접 어느 정도 확보할 수 있지만, 성능이 저하될 수 있고 해킹에 취약해진다는 문제가 있다(2.2절 및 [그림 3]도 참조). DARPA의 XAI 프로젝트도 주로 후자의 접근법을 취하는데, 이 프로젝트의 책임자인 데이비드 거닝은 XAI의 목표가 머신러닝의 높은 예측 정확성을 잃지 않으면서 설명력을 갖도록 하는 것이라고 밝힌 바 있다.¹⁶⁾ 거닝은 XAI가 신경망 내부에서 (스스로 창출된) 개념 공간(예: 신경망 노드들의 부분집합)에 라벨을 붙이는 방법을 찾는 것이 관건이라고 본다. DARPA의 연구팀은 신경망의 특정 개념 공간은 인간의 두뇌가 정보를 처리하는 것과 유사한 방식으로 비슷한 일을 반복하는 기능을 수행한다는 것을 발견했다. 예를 들어, 고양이를 식별할 때 고양이의 특징(예: 귀)에 해당하는 부분을 집중적으로 처리하는 일정한 신경망 노드들이 있다는 것이다. 여기에 해당 라벨(예: ‘고양이 귀’를 식별)을 붙이면 머신러닝의 추론 과정을 역추적하여 판단 근거를 규명할 수 있다(예: 이미지의 동물 귀로 볼 때 고양이로 판단하며, 이러한 판단의 가중치는 얼마이다). 이러한 방법은 신경망 내부에서부터 자체적으로 대화할 수 있도록 알고리즘이 무슨 일을 하고 있는지를 스스로 말하게 만드는 것으로, 그는 이를 ‘햄릿전략’으로 부른다([그림 5]의 deep explanation 접근법).

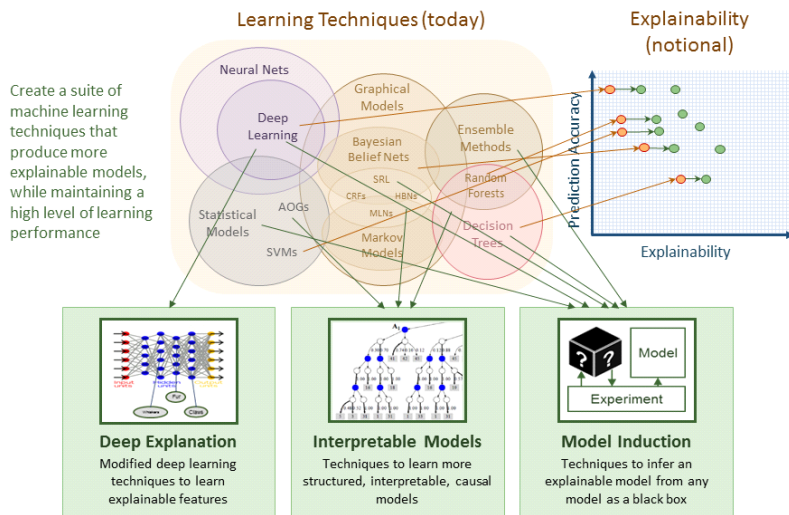
DARPA는 XAI를 먼저 머신러닝형 분류기에 적용하여 기밀 정보에 대한 분석 실험을 하고 있으며(데이터 분석 도메인), 나아가 강화학습을 무인 자율 시스템에 적용하는 방향으로(자율 무인 시스템의 강화학습 도메인) 진행하고 있다. 그 과정에서 AI가 내린 결론의 신뢰성을 높이기 위해서는 알고리즘과 데이터 간의 연관성과 추론 과정의 인과관계에 대한 고려가 반드시 필요하다는 결론을 내린 것으로 보인다(NIA, 2018a).¹⁷⁾ 이러한 목적을 위해서는 신경망 기반 머신러닝보다는 의사결정트리(decision tree, [표 2]를 참조)를 활용하는 것이 더 효과적이다([그림 5]의 interpretable model 접근법). [표 2]와 [그림 3] 등에서 간략히 소개한 CART와 같은 알고리즘은 그 자체가 설명력이 매우 높을 뿐만 아니

16) 거닝이 말한 XAI의 목표는 AI가 일반인의 이해와 신뢰를 얻기 위해서 정확성과 책임성 모두 필요하다는 것을 잘 대변한다. 이 두 가지 목표는 서로 상충하기도 하지만 동시에 (최소한 특정 수준 이상으로는) 모두 달성되어야 한다.

17) 그러나 다시 인과관계를 고려하는 것은 누군가에 의해 만들어진 모델이나 최소한의 설명(internal description)을 투입해야 한다는 점에서 머신러닝의 기본 철학을 (부분적으로라도) 부정해야 한다는 문제에 직면한다. 그러나 이런 이슈가 알고리즘과 AI의 현실적 활용을 제한하지는 않는다.

라 학습 과정과 연계되었을 때 일치성을 추적하는 게 더 쉽다.¹⁸⁾ 신경망 기반 머신러닝이 의사결정트리의 추론방식을 학습하여 신경망 내부의 처리 과정을 의사결정트리에 빗대어 설명할 수 있다면 더욱 설명력을 높일 수 있다. 또 다른 방식은 신경망의 투입-산출 데이터로부터 통계적 기법을 적용하여 의사결정 방식을 모델로 유추(induction)하는 것이다(그림 5)의 model induction 접근법). 베이지안(Bayesian) 방식의 통계적 추론방식을 적용하는 이러한 접근법은 비교적 오랜 전통을 가지지만, 신경망 내부를 정확히 모사하는 데 한계를 가질 수밖에 없다.

〈그림 5〉 DARPA의 XAI 개발을 위한 접근 전략들



(출처: DARPA(2018))

DARPA는 XAI에 대한 다양한 접근을 통해 실험 데이터를 확보하고 실제 사용자를 대상으로 만족도와 에러율 등을 평가하여 XAI를 정교하게 개발할 계획이다(XAI 평가지표에 대해서는 NIA(2018a)의 [표 2]를 참조). 다만, 이 과정에서 상황과 데이터에 대한 의존도가 높다는 것은 실험 결과에 대한 보편성을 확보하는데 여전히 한계로 작용할 것이다. DARPA도 AI가 내린 결과에 사용한 데이터가 미치는 영향을 정확히 평가해야 한다는 점

18) CART 자체는 매우 투명한 알고리즘이나, 2.3절에서 설명한 바와 같이, 실제로 CART를 구현하는 과정에서 배깅이나 랜덤 포레스트와 같은 랜덤화를 적용하면 불투명성과 해석의 어려움은 증가한다.

을 인식하고 있다. 이러한 한계를 어느 정도 극복할 수 있다면 DARPA의 XAI 프로젝트는 향후 AI에 대한 품질관리와 인증 및 감리 등의 분야에서 매우 강력한 표준이 될 수 있다. 또한, EU GDPR의 요구사항과 도전을 극복할 수 있는 중요한 레버리지가 될 것으로 기대된다.

DARPA의 XAI 프로젝트는 그 자체가 목적이자, 2018년 7월에 ‘AI Next’로 명명된 미래 AI를 향한 보다 큰 계획의 일환이다(NIA, 2018c). AI Next는 AI가 인간과 협업(man-AI collaboration)하면서 능동적으로 인간과 업무를 도울 수 있는 시스템을 의미한다. 이를 위해 AI는 변화하는 상황을 제대로 인지하고 이에 적응(contextual adaptation)할 수 있어야 하며, 인간과 원활하게 소통할 수 있어야 한다. 이러한 목적에서 XAI는 반드시 해결해야 할 과제이다. 더 나아가 머신러닝 형태 학습의 근본적인 한계를 극복하기 위한 프로젝트도 진행 중이다. 이 프로젝트는 시행착오(trial-and-error)를 근간으로 하는 머신러닝의 학습방식을 탈피하여 이미 검증된 것으로부터 일반화를 추구하는 방식으로 AI를 설계하는 것을 목표로 한다. 이것이 성공한다면 XAI의 설명 가능성에 더하여 AI의 검증 가능성(verifiability)이 어느 정도 확보된다는 것을 의미한다. 따라서 더 효율적이며 효과적으로 AI를 구현할 수 있고 운영자나 사용자로부터의 신뢰도 획득할 수 있다.

(2) 알고리즘 위험관리(algorithm risk management)

의료, 행정, 치안, 법제도의 집행과 같은 공공분야에서 알고리즘에 의존하는 자동화된 의사결정(automated public decisions)은 공정성(fairness) 확보에 초점을 맞춘다. 공정성 문제와 관련하여 알고리즘의 불투명성 문제를 해결하기 위한 접근법으로 위험관리 프레임워크(risk management framework)에 기반한 도구(toolkit)들이 제안되고 있다. 이러한 접근법은 특히 분류기 산출과정에서 부적절한 데이터에 의한 학습으로 야기되는 편향성을 제거하거나 줄이는 것을 목표로 한다.

이 분야의 대표적인 프로젝트로, 존스홉킨스대학과 하버드대학 및 샌프란시스코와 워싱턴DC 소재 공공단체들의 공동연구(<http://ethicstoolkit.ai/>)가 있다. 이 프로젝트는 2016년 미국 뉴욕시(NYC)가 시험적으로 도입한 AI 기반 범죄예방프로그램이 흑인 시민

들에게 편파적인 결과를 산출하는 문제를 바로잡기 위한 공동연구에서 시작되었는데, 그 결과로 알고리즘에 내포된 위험성을 평가하는 도구(소프트웨어)를 개발했다. 이 도구는 아래에 소개하는 몇 가지 (부정하기 어려운) 상황인식을 전제로 한다.

- 정부 및 행정기관 등 공공 부문에서 알고리즘 사용은 피할 수 없는 대세이다.
- 알고리즘 개발에 투입되는 데이터는 서로 다른 소스로부터 독립적으로 구해서 집계되는 경우가 많으므로 모든 데이터는 편향을 가질 수밖에 없다(All data has bias).
- 따라서 모든 알고리즘은 편파적인 결정에서 자유로울 수 없다(All algorithms have bias).¹⁹⁾

위와 같은 관점에서 알고리즘에 내재된 (그렇지만 피할 수 없는) 잠재적 문제점을 인식하고 사전에 평가할 수 있는 프레임워크를 개발하는 것이 이 프로젝트의 목적이었다. 예를 들어, 알고리즘이 적용되는 공공 서비스 대상자와 이해관계자를 명확히 밝히고 만약 문제가 있다면 어떤 피해가 예상되는지를 공개하는 체계적 기준과 절차를 제공한다. 또한, 알고리즘 개발에 투입된 데이터와 방법론을 설명하고 밝히는 방식(예: 공개할 수 있는 권한이 누구에게 주어지면 어느 수준까지 밝혀야 하는지에 대한 기준 등)도 제공한다. 이를 바탕으로 알고리즘에 내재된 위험성을 평가하고 편향을 줄이는 방식도 제시한다. 예를 들어, 대단위 집단을 대상으로 하는 공공정책에 적용되는 알고리즘 X가 있다고 하자. 이 도구는 먼저 X의 운영을 감시하는 독립적인 위원회(review board) 또는 자문기구(advisor group) 설치를 제안한다. 또한, X에 의해 의사결정의 자동화 수준이 어느 정도인지 평가하고 그 수준에 따라 위험관리와 대응 방식을 달리하도록 한다.

Accenture와 IBM도 각각 2018년 2월 및 9월부터 알고리즘 편향성을 추적할 수 있는 클라우드 서비스를 제공하기 시작했다.²⁰⁾ 특히 IBM은 AI 활용의 부작용에 따른 위험을 예방하고자 하는 기업들의 요구에 부응하여 오픈스케일(OpenScale)이라는 AI 편향성 검증 서비스를 2018년부터 제공하기 시작했다. 이러한 서비스를 IBM 클라우드(프라이빗)를

19) 이는 사람도 마찬가지(All people have bias)라고 토를 달고 있다. 위 주장에 원칙적으로 동조하더라도 편향의 정도에 대해서는 다른 견해를 가질 수 있다. 즉, 원본 데이터에 내재된 편향을 완전히 제거할 수는 없어도 편향을 보정하거나 정제하여 (최소한 활용 목적에 비추어 볼 때) 무시할 수 있을 만한 수준으로 줄일 수도 있을 것이다.

20) <https://github.com/IBM/AIF360>

통해 제공함으로써, IBM은 자사가 개발 중인 AI 설계도구(일종의 AI 플랫폼, 5.1절 참조)와 더불어 AI 플랫폼과 인프라를 공급하는 메타 플랫폼(meta platform)으로 포지셔닝하고자 한다. Google도 자사의 AI 플랫폼인 텐서플로우(TensorFlow) 이용자들이 개발한 AI 프로그램의 편향성을 시각화·분석해주는 서비스를 제공한다.²¹⁾ 이러한 AI 위험관리 서비스는 AI 플랫폼 이용자들이 부지불식간에 코드화한 편향성으로 인해 받을 수 있는 피해를 예방함으로써 AI 플랫폼 확산에 큰 도움을 줄 수 있다. 위험관리가 주는 편익을 보다 구체적으로 정리하면 다음과 같다.

첫째, AI와 알고리즘이 도출되는 배경과 과정을 이해하는 데 도움을 주어 일반인의 알고리즘 및 AI 시스템의 신뢰성을 향상한다. 둘째, 자동화된 편향 탐지·제거·대응 등을 통해 다양하고 막대한 수의 AI 프로그램들을 모니터링하고 잠재적 위험요소를 사전에 탐지함으로써 피해를 최소화한다(특히, 공공 부문에서 효용성이 높다). 셋째, 기업과 민간 조직이 GDPR과 같은 AI 관련 법률과 규정을 준수할 수 있도록 도움으로써, 민간 부문의 비용 부담을 덜고 AI가 법제도와 공진화하는 것을 촉진한다. 넷째, 이러한 플랫폼과 인프라에 축적된 데이터와 모델, 훈련 및 측정지표 관련 기록들은 알고리즘과 AI의 발전에 중요한 자산이 되며 향후에 생각해 볼 수 있는 AI 시스템의 표준화에 도움이 될 것이다.

21) <https://ai.googleblog.com/>

3 » 전망

AI의 엔진으로 활용되는 알고리즘의 특성상 완전한 투명성을 담보하는 것은 사실 거의 불가능하여서 일반 제조물이나 서비스와 비슷한 수준으로(제조물 배상책임이나 서비스 수준 협약 등) 알고리즘과 AI의 책임성을 입법화하는 것은 어려울 것으로 전망된다. AI 개발자인 제프 클론 교수는 지능의 본질이 모든 것을 이성적으로 이해하고 설명할 수 있어야만 하는 것은 아니라고 말한다. 이성으로 설명할 수 없는 경우에는 직관이나 무의식이 작용하여 그 부족한 부분을 메꾸는 것이 우리가 ‘지능적’이라고 부르는 행태라는 것이다(MIT Technology Review, 2018). 이는 지능적인 것이 행하는 모든 것을 다 ‘해석’해야 할 필요는 없다는 것을 의미한다. 실제로 Alpha-Go가 바둑을 두는 방식이 인간의 사고체계와는 다르므로 ‘절묘한 수’로 사후에 판명된 해법에 대해서도 바둑 전문가들이 쉽게 해석을 내릴 수 없었다.

또한, 알고리즘과 AI가 추구하는 정확성과 설명 가능성은 어느 정도 상충관계에 있을 수밖에 없다. 예를 들어, 알고리즘이 복잡해질수록 정확도는 높아지지만, 이해는 더욱 어려워진다. 또한, 신경망 기반 머신러닝에서 데이터에 제한을 둘 경우(앞 절에서 소개한 도구(toolkit)는 데이터 활용에 제한과 검증을 강화할 것을 요구한다), 편향과 공정성 이슈는 줄어들 수 있지만, 그 대가로 예측 정확성(predictive accuracy) 측면에서 손해를 볼 수 있다.²²⁾ 더구나 사람이 이해할 수 있도록 하려면 결과를 그대로 전달하는 대신에 인간의 언어와 논리로 추상화하고 재구성하는 가공단계를 거쳐야 한다. 이는 위에서 소개한 XAI의 햄릿전략에 태생적인 한계가 존재할 수밖에 없다는 것을 의미한다. ‘설명한다.’는 것은 사실 인간적 사고와 소통방식으로, 실제 있는 그대로를 빈틈없이 커버하는 것이 아니다. 오히려 ‘잘 설명하기’ 위해서는 그 대상의 특징(패턴)이나 설명을 듣는 사람이 가진(잠재적) 신념에 부합되는 방식의 설명이 채택되는 경우가 많다. 이러한 특성과 한계로

22) 그러나 투입 데이터에 대한 제한과 알고리즘의 정확성이 반드시 상충관계(trade-off)를 보이는 것은 아니다. MIT Sontag 교수가 지적하듯이 오류가 있는 데이터는 편향된 결과뿐만 아니라 정확성도 떨어뜨린다. 그런데도 공정성을 위해 특정 속성(예: 인종, 지역, 연령대 등)별로 임의표본추출(randomized sampling)을 시행하면(전체 모집단 데이터에 대한) 편향은 줄어도 예측의 정확도는 감소하는 상충관계가 빈번히 발생한다.

인해 XAI는 오히려 결과와 상관없이 사람들의 생각과 신뢰수준을 재단할 수 있다. 이는 AI의 또 다른 위험요인이다.

마지막으로, 사람들의 만족도는 설명의 내용만으로 충분하지 않으며, 설명하는 과정의 공정성이나 진실성 등 다른 가치와 경험에 영향을 받는 경우도 많다. 사실 ‘좋은 설명’은 그 내용보다도 사람들(수신자들)이 동의하는 틀에서 가능하다는, 사회적 성격을 갖는다. 따라서 XAI의 모든 과정이 기계적으로 완벽하고 심지어 투명하다고 인정받을 수 있어도 AI의 설명력이 완전히 해결되는 것은 아니다. 예를 들어, 자율주행차의 ‘트롤리 딜레마’를 해결하기 위해 AI의 작동원리 등을 연구하는 것이 큰 도움을 주지 못할 수 있다(이원태, 2015; 주간기술동향, 2018a; Knowledge@Wharton, 2018).

위와 같은 기술적·사회적 한계(즉, 인간과 같은 방식으로 지능을 구현하고 소통하는 것은 어렵다고 하더라도)에도 불구하고, AI가 궁극적으로는 인간을 위한 기술이어야 한다는 점에서 알고리즘의 책임성 문제에 대해서는 제도적 해결책이나 사회적 합의가 마련되어야 한다. 특히, 공공성을 지닌 응용에 있어서 책임성에 대한 법제도적 기반은 시급한 과제이다. 먼저, XAI와 위험관리와 같은 접근법 역시 만병통치가 될 수 없다는 점을 인식하고 이들을 활용하는 목적을 명확히 하고(예: XAI가 무엇을 어떻게 어느 수준까지 설명할 것인지 등) 일반 사용자와 이해당사자가 과도한 기대나 불안감을 느끼지 않도록 가이드라인을 제정할 필요가 있다. XAI는 인간이 알고리즘과 AI가 내린 결과와 절차를 어느 정도 이해할 수 있도록 도와주는 도구일 뿐이며, AI 위험관리를 통해서 사전적으로 문제가 완전히 제거할 수 있는 것은 아니라는 점(위험관리는 오히려 사후적 조치와 피해 최소화를 위한 것이다)을 널리 알려서 이들에 대한 명확한 이해를 공유하고 확산시켜야 한다.

또한, 정교한 알고리즘을 사용한다고 해도 AI 및 자동화 시스템이 전문가와 작업자의 검증 없이 전격적으로 사용되어서는 안 될 것이다. 미국 국토안보국(homeland security)도 자신이 주도하는 AI 기반 몸수색장치(body scanner)가 수색을 담당하는 인력을 완전히 대체하지는 않을 것이라고 한다(The New York Times, 2017). 이러한 방향은 의료분야의 IBM 왓슨이나 법률서비스에서 ‘Do Not Pay’와 같은 전문가시스템(expert systems)에서도 마찬가지이다. 당분간 AI 시스템은 전문가와 작업자의 업무량을 낮추는

데 초점을 맞추면서 오랜 기간의 현장 검증을 거쳐야만 자동화된 의사결정 시스템으로 완성할 수 있다. 일군의 미래학자들이 주장하는 ‘적응형 자동화(adaptive automation)’에 관심을 기울이는 것도 이러한 맥락에서 이해될 수 있다. 적응형 자동화는 사용자와 시스템이 협력하여 자동화 수준을 조절하면서 현장의 상황에 맞추어 자동화 구현을 확대해 나가는 방식이다. 이는 작업자의 인지적 부담(cognitive load)은 줄이면서 상황에 따른 의사결정과 대응 역량을 개선하여 효율성을 점진적으로 높인다.

마지막으로, 알고리즘 생태계를 구성하는 여러 이해당사자(개발자-공급자(운영자)-사용자 등) 사이의 지적재산권에 대한 소유(own-right) 및 활용(use-right) 범위와 이로 인한 이익과 비용의 분배를 포함한 혁신적 비즈니스모델에 대한 체계적인 고려도 중요하다. 알고리즘과 AI 등이 다목적 기술(general purpose technology)로 발전하면서 비즈니스모델의 성공 여부가 이러한 법제도적 이슈와 밀접히 연결되었다. 예를 들어, Tesla 자율주행 사고나 SNS에서 AI에 의해 생성되는 가짜 뉴스(fake news)와 같은 사례에서 보듯이, 사고나 결함이 발생했을 때 그 책임이 누구에게 귀속되는지(알고리즘 설계자 혹은 데이터 제공자 혹은 운영자나 사용자 등)와 같은 책임소재의 문제는 해당 알고리즘 생태계 전체의 흥망성쇠에 큰 영향을 미치게 되었다. 그런데 사실 이러한 이슈는 과거 새로운 혁신적 기술이 등장할 때에도 발생했다. 예를 들어, 미국에서는 오랜 논쟁을 거쳐서 항공 사고가 잘못된 항공지도로 판명되면 지도제작자가 책임을 지는 것으로 결정이 된 후에 자동항법장치가 더 확대되었다. 데이터와 알고리즘 시대에 비즈니스모델과 관련된 제도적 뒷받침의 문제도 위와 비슷한 진화과정을 거칠 것으로 예상된다. 데이터 제공자의 품질 보증, 알고리즘의 정상적인 작동에 대한 개발자의 보증(과 책임) 범위 및 관련된 위험을 분산(hedge)하기 위한 보험 등이 제도적 뒷받침의 예이다. 또한, 현재의 알고리즘 및 AI 기술로는 예측할 수 없는 결함이 발생한 경우에 대한 면책 및 방법(safe harbor 조항 등)도 중요하다. 관련 당사자들이 원인 규명과 해결책 마련을 위해 협력하는 것을 의무화하고 손해배상에 관한 보험 등의 제도와 이를 촉진할 수 있는 환경을 마련하는 것이 알고리즘 투명성 및 책임성 논의와 더불어 진행되어야 할 것이다.



알고리즘 책임성 적용 사례 및 관련 이슈들

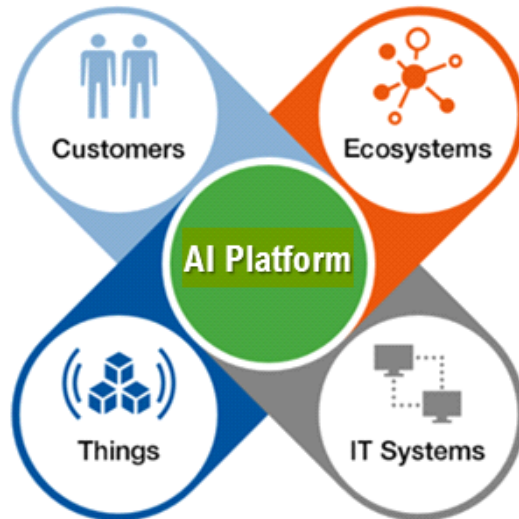
1 » AI 플랫폼

AI는 데이터 및 플랫폼 기반 경제와 디지털 전환(digital transformation) 환경에서 가장 중요한 요소 기술이다. 플랫폼은 AI 기반으로 여러 기능을 서로 연결하면서 빅데이터 환경에서 의사결정을 지원하고 제품·서비스 가치사슬을 조율하는 역할을 한다.²³⁾ B2C 환경에서 AI는 주로 사용자의 편의성을 높이기 위한 수단으로 주목을 받고 있지만, B2B 환경에서는 AI 자체가 직접 서비스되는 경우보다는, 로봇이나 자율주행 자동차와 같은 지능형 제품과 서비스에 투입요소로 활용됐다. 이 경우 AI를 자동화(automation)의 연장선에서 접근하는 것이 더 현실적이고 의미 있는 접근일 수 있다. 기존 플랫폼과 산업에서 전문성과 차별화를 지원하는 보조적인 역할에 머물러 온 AI는, 디지털 전환이나 4차 산업혁명의 시대에서는 플랫폼을 구동하는 인프라가 되어 데이터와 알고리즘 기반의 차별화된 프로세스와 광범위한 O2O(Online to Offline) 및 서비스화(servicification)를 구현하는 촉매제(enabler)의 역할을 할 것이다. 즉, 디지털 전환에서 AI는 기존의 전통적인 생산방식에 비해 서비스의 범위, 경쟁의 속도, 사용 편의성, 생산 효율성 등을 크게 개선함으로써 거래비용(transaction costs) 감소와 품질통제를 위한 핵심요소가 되며, 플랫폼의 경쟁력 유지에 필수적인 요소가 된다. 더 나아가 IoT와 가상화(virtualization) 인프라가 보편화

23) 이는 Gartner(2016)가 제시한 디지털 플랫폼 개념과 일맥상통한다. 즉, 디지털 플랫폼은 기존 플랫폼에서 한 단계 더 발전하여 파트너, 공급자, 고객 커뮤니티 등이 사업적 이해관계를 위하여, 디지털 프로세스 및 역량(digital processes and capabilities)을 공유, 개선, 확장(share, enhance, extend)할 수 있도록 설계된 프레임워크이다(김민식·이가희, 2017). 여기서 디지털 역량이 기존 플랫폼과 디지털 플랫폼을 차별화시키는 핵심요소인데 AI는 기술적 측면에서 디지털 역량의 기반이 된다. Gartner(2016)이 제안하는 디지털 플랫폼 아키텍처를 구성하는 ICT 자원 - 고객 경험 - 생태계 지원시스템 - 데이터 및 분석시스템 모두에서 AI는 필수요소로 기능하게 될 것이기 때문에 AI 플랫폼이 가능하다.

되면 AI 자체가 모든 서비스에 필수요소(essential element)가 되는 시나리오가 가능하다. 이와 같은 상황에서 AI는 대기업은 물론 중소기업을 비롯한 모든 조직뿐만 아니라 사용자도 자유롭게 활용할 수 있는 개방된 공간에서 서비스될 것이다. AI를 독자적으로 개발하려면 알고리즘 개발부터 서버와 클라우드 인프라와 같은 대규모 투자가 필요하지만, AI를 플랫폼과 같이 제공할 수 있다면 스타트업이나 중소기업 등도 AI를 활용한 제품 및 서비스를 적은 비용으로 신속하게 개발할 수 있을 것이다. 즉, AI 기술을 집약하여 특정 생태계 구축에 필요한 기술 공간(technology space)을 제공하면서 확장성(scalability)이 매우 높은 메타 플랫폼(meta platform)을 생각해 볼 수 있다. 모든 기업이 디지털 플랫폼이 될 수는 없지만, 모든 기업이 이런 플랫폼이 필요하여 이들 중 하나의 생태계에 포함되어야 한다는 측면에서 디지털 전환의 필수요소인 AI를 플랫폼으로 생각해 볼 수 있다. 이 글에서는 이러한 개념에서의 AI를 ‘AI 플랫폼’으로 부른다.

〈그림 6〉 AI 플랫폼의 입지



(출처: Gartner(2016)의 그림을 본문의 내용에 맞추어 일부 수정하였음)

여기서는 AI 플랫폼에 대해 아래와 같이 보다 엄밀하게 정의한다.

- AI 플랫폼의 정의: AI 기술과 서비스가 집약된 기술 공간이 복수의 비즈니스 도메인에 필요한 거래와 서비스의 핵심요소로 활용될 때 이를 AI 플랫폼이라고 한다. AI 플랫폼은 기술 공간이기 때문에 대상 비즈니스 도메인의 서비스 개발·운영·관리에 필요한 자원(resource)과 이를 관리할 수 있는 시스템(resource management)을 갖추고 있다.²⁴⁾

AI 플랫폼을 구성하는 자원에는 기본 데이터와 알고리즘 라이브러리를 포함하여 알고리즘 활용에 필요한 모듈화된 도구들(자연어 처리 및 인터페이스 등)과 AI를 서비스로 제공할 수 있는 각종 시스템(응용 애플리케이션 개발을 지원하고 배포를 쉽게 하는 커뮤니케이션 시스템 등)이 해당된다. 구체적으로 후자의 경우, 챗봇(chat bot), 마케팅, 고객관리(CRM, Customer Relationship Management), 결제 지원(payment support), 최적화 탐색, 텍스트 분석과 같은 애플리케이션 SDK(Service Development Kit)와 보안 및 클라우드와 연동을 지원하는 통신 지원 인프라(의 일부)를 포함한다. AI 플랫폼은 플랫폼으로서의 본연의 기능을 수행한다. 즉, 특정 도메인에서 기존 서비스제공방식의 거래비용을 줄이고 이질적인 요구사항을 연결하여 다양한 서비스를 창출하고 해당 도메인에서의 진입장벽을 낮추는 것을 목표로 한다. 예를 들어, 2018년 여름을 기준으로 Amazon 알렉사를 활용하여 180개 국가의 개발자들이 5만 개 이상의 ‘Skill(알렉사를 이용하는 앱이나 응용 프로그램을 부르는 용어)’을 개발하였다(주간기술동향, 2018b). Amazon은 Skill 개발을 지원하기 위해 SDK와 ‘알렉사 보이스(YourVoice)’를 제공한다. 알렉사를 탑재한 가전기기도 본격적으로 출시된 이후 1년 만에 2만 대가 넘었으며 스마트홈이 활성화되면 더욱 확대될 전망이다. 또한, 자율주행 자동차와 금융·의료와 같은 전문 서비스 분야에서 알렉사와 같은 AI 플랫폼을 활용한 제품과 서비스는 급속히 확대될 것으로 예상된다.

AI 플랫폼이 제공하는 서비스를 통칭하여 ‘AI as a service (AAAS)’라고 부르기도 한다. AI 플랫폼이 실현되면 AI(개발자)는 기존 플랫폼으로부터 종속적인 입지에서 벗어나 독자적인 플랫폼 공급자의 지위를 획득할 수 있을 것이다. 디지털 플랫폼의 핵심요소로

24) 박외진(2018)에서도 AI 플랫폼에 대해 이와 비슷한 정의가 제공된다.

AI의 일반 기능(general purpose complement)들로부터 새로운 서비스 창출의 기반을 제공하고 기존 플랫폼의 인터페이스와 도구(toolkit: API, SDK, 앱 마켓 등)를 성공적으로 흡수·통합할 수 있다면 디지털 플랫폼들의 메타 플랫폼으로서의 입지를 구축할 수 있다.

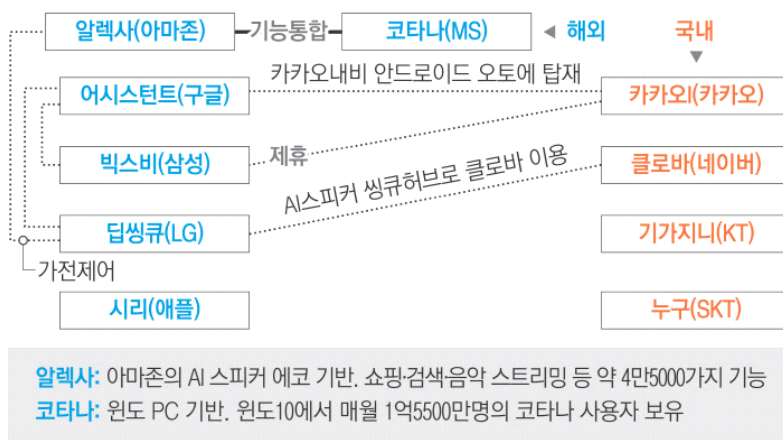
이러한 낙관적인 전망에서 이미 2017년에 ‘AI First’를 주창한 Google을 비롯하여 Amazon, Apple, Facebook 등의 거대 플랫폼들은 AI 플랫폼의 주도권을 확보하기 위한 치열한 경쟁에 돌입했다. 또한, 삼성전자와 LG전자와 같은 국내외 하드웨어 기업들도 독자적인 AI 플랫폼을 제안하고 있으며, 우리나라의 경우 통신사업자들이 AI 서비스 제공에 깊은 관심을 보인다.

이들 사업자는 초기에는 범용성을 지닌 AI(AGI, Artificial General Intelligence) 플랫폼 개발에 몰입하였지만, 데이터 확보와 아직 기술적 제약으로 인해 AGI가 제공할 수 있는 서비스 품질에 한계를 절감하기 시작했다. 즉, AI 플랫폼이 제공할 수 있는 추상적 수준의 서비스(service abstraction level)를 인간이 만족할 수 있을 만한 수준으로 제공하기 위해서는 기술적 한계뿐만 아니라 여러 도전과제가 있다는 것을 깨달았다. 이에 따라 현재 AI 플랫폼은 특정 서비스 도메인이나 산업 부문을 대상으로 하는 특화된 AI 플랫폼 개발에 집중되어 있다. 예를 들어, 의료서비스 분야에 특화된 AI 플랫폼이 제공되면 의료 서비스에 관련되는 여러 기업과 조직들이 이 AI 플랫폼을 사용하는 것이다.

이러한 상황에서 AI 플랫폼의 리더십을 확보하기 위해 사업자 간 합종연횡이 활발하게 진행되고 있다. 최근 Amazon과 MS가 각자의 AI 비서인 알렉사와 코타나를 서로 연동하기로 한 것이 대표적이다. 이로 인해 AI 플랫폼을 둘러싼 경쟁 구도는 신/구 글로벌 자이언트들인 Google, Apple, Amazon/MS, IBM 간 경쟁으로 전개될 전망이다. 이러한 전략적 제휴·연대는 AI 프로그램뿐만 아니라 스마트 디바이스는 물론, 자동차, 가전기기와 같은 영역으로 확대되고 있다. 이와 같은 방향은 콘텐츠-서비스-운영체제-디바이스(H/W)에 걸친 가치사슬 전체의 유기적 통합(seamless integration)이 AI 플랫폼으로 완성된다는 것을 의미하며, 이 시장의 막대한 잠재적 가치를 시사하는 것이기도 하다. 또한, 아무리 글로벌 자이언트라고 하더라도 어느 한 플랫폼만으로 모든 서비스를 제공하는 것은 기술적

으로도 불가능하다는 것을 방증한다. 따라서 향후 서비스, 운영체제, 하드웨어 등의 부문에서 장점을 교환하고 약점을 보완하는 제휴와 연대는 더욱 활발해질 것이다. 또한, 다양한 분야에서 파트너십 확보를 위해 자사의 플랫폼의 일부 기능과 서비스를 지속적으로 개방할 것이다. Google과 Amazon이 예전보다 적극적으로 CES에 참여하는 것도 이런 노력의 일환으로 해석된다.²⁵⁾

〈그림 7〉 AI 플랫폼 사업자 간 협력 관계



(출처: 주영재(2018)에서 인용)

AI 플랫폼에 대한 낙관적인 전망과 잠재력으로 인하여, 2000년대 이후에 등장한 플랫폼들이 주도해 온 AI 플랫폼 시장의 경쟁 구도에 급격한 변화가 발생했다. 20C에 컴퓨터 하드웨어로 전 세계를 석권했던 IBM이나 소프트웨어(특히 OS)의 중요성을 처음으로 전 세계에 부각시킨 MS와 같은 전통적인 IT 기업들이 적극적으로 참여하기 시작한 것이다. IBM은 신경망을 자동으로 구축하는 기술인 뉴넷S(NeuNet S)를 통해 AI 운영을 돕는 AI인 오픈스케일(4.4.2 참조)로 AI 플랫폼에 도전하였다. MS의 프로젝트 AIX와 삼성전자의 베레스(Veles) 등도 오픈소스(open source)로 공개되면서, 기존 Google 텐서플로(Tensorflow)²⁶⁾와 Facebook의 토치(Torch)²⁷⁾ 등과 더불어 다양한 분야에서 본격적인

25) 글로벌 자이언트들의 AI 플랫폼(과 제품)에 대해서는 박외진(2018), 윤장우(2018), 이정원(2017) 등에서 보다 자세한 현황을 볼 수 있다. 우리나라에서 개발 중인 ADAMs이나 Jonathan 등의 AI 플랫폼(및 솔루션)에 대해서도 위 문헌을 참조하라.

경쟁에 나섰다.

또한, 플랫폼의 특성상 빠른 속도로 시장을 장악하기 위해서는 오픈소스를 통해 AI 기술을 공유해 개발 속도를 높임과 동시에 시장에서의 영향력을 확대하는 것이 중요하다. AI 플랫폼은 여러 분야의 정보와 데이터를 확보해야 금융, 의료, 교육 등 다양한 분야에 응용될 수 있는, 일정 수준 이상의 알고리즘 성능을 확보할 수 있다. 어느 한 플랫폼이나 기업이 이 모든 영역을 섭렵하는 것은 불가능하므로 협업이 필수적이다. 즉, AI의 기본 알고리즘을 공유하면서 AI 생태계를 구축해야 규모의 경제를 달성할 수 있으며, 여러 서비스를 쉽게 구현할 수 있도록 관련 기능을 모듈화하고 API 형태의 라이브러리를 구축하는 것이 효과적이다. 개발에 필요한 시간을 단축하며 그에 따른 진입비용을 줄이고 신뢰성을 확보할 수 있다는 장점으로 볼 때, 오픈소스는 이러한 목적에서 최적의 접근법이라 할 수 있다.

현재 AI 플랫폼들은 주로 클라우드 기반의 서비스 인프라를 바탕으로 운영된다. 여러 자원도 클라우드 상에 준비된 자사의 서비스들과 연동을 통해 활용되는 방식이다. 예를 들어, Amazon은 자신이 강점을 가진 AWS(Amazon Web Service)에 여러 알고리즘 라이브러리를 구축한 다음 알렉사와 연동시킬 수 있도록 구현하였다. Google도 텐서플로를 Google 클라우드를 통해 제공하면서 이미지 인식이나 패턴 분석과 같은 서비스를 제공한다. 또한, 클라우드 서비스에 준하는 과금체계를 적용하고 있어서 AI 플랫폼을 자연스럽게 확산하기 위한 전략적 접근을 취한다. 우리나라의 AI 플랫폼 역시 과학기술정보통신부와 산업통상자원부가 지원하는 산업특화형 개방형 클라우드 사업과 연계된다. 이러한 배경에서 우리나라는 초기부터 범용성 AI 플랫폼이 아닌, 산업 및 도메인 중심의 AI 플랫폼 구축에 집중하는 경향이 강하다.²⁶⁾

AI 플랫폼 생태계에 대해 낙관적인 전망만 있는 것은 아니다. 예를 들어, 차별화와 IP 확보 등의 측면에서 협업 당사자 간 잠재적 갈등 요소도 존재한다. 이러한 부작용을 줄이기

26) <https://www.tensorflow.org/>

27) <http://torch.ch/>

28) 삼성전자와 LG전자와 같은 글로벌 기업들은 범용성 AI 플랫폼에 관심을 가지고 있지만, 주로 가전제품의 활용이나 스마트홈 부문과 연결되어 있다는 측면에서 빅스비(Bixby)나 누구(NUGU) 같은 AI 플랫폼도 도메인 중심 성격이 강하다고 볼 수 있다.

위해서는 계약관계와 확장성(scalability)에서 비즈니스적 합의와 법제도적 제약을 사전에 예측하고 위험관리를 제공하는 제도적 장치와 지원이 필요하다(이정원, 2017; 주간기술동향, 2018b).²⁹⁾ 예를 들어, 알고리즘 자체는 오픈소스와 협력 생태계를 통해 발전할 수 있지만, AI 플랫폼 간 경쟁에서의 우위와 차별화는 데이터 축적으로 형성될 수 있다. 차별화의 기반이 되는 요소들은 (생태계 내에서 공유되는 알고리즘이나 API보다는) 고객, 파트너, 사물, 기업 내부의 여러 소스로부터 생성되는 데이터를 수집-저장-가공-축적한 결과와 그 과정에 존재할 가능성이 크다. 딥러닝을 중심으로 한 현재의 AI 시스템에 기반을 둔 AI 플랫폼을, 데이터 구축과 관리와 분리해 생각할 수는 없기 때문이다. 따라서 AI 플랫폼의 발전을 위해서라도 데이터 소유와 활용에 관한 법제도적 준비가 사전에 이뤄질 필요가 있다. 예를 들어, 정부의 ‘마이데이터’ 정책의 내용과 현실에서 도입되고 정착되는 과정에 따라 우리나라 AI 플랫폼 시장과 경쟁 구도도 크게 영향을 받을 것이다.

29) 예를 들어, GPL(General Public License)도 모든 라이선스를 사용할 수 있는 것은 아니므로 계약을 면밀히 검토하여 허용 범위 내에서만 사용해야 하지 않으면 문제가 발생할 수 있다. 그런데 GPL과 같은 오픈소스에서는 사전적으로 모든 가능한 경우에 완벽한 계약조건을 마련할 수 없으므로 갈등의 여지를 완벽하게 제거하는 것은 불가능하다. 따라서 개방형 혁신(open innovation)에서는 위험관리에 대한 역량이 더욱 중요하다.

2 《 알고리즘 가격설정과 담합(Algorithmic Pricing and Collusion)

알고리즘과 AI 시스템이 변화시키고 있는 비즈니스 관행의 대표적인 대상이 마케팅과 가격설정 방식(pricing scheme)이다. 이들 부문은 소비자와의 거래에 직접 관련된다는 측면에서 사회적 파급과 관심도 큰 편이다. 알고리즘에 의한 자동화가 소비자 생활의 편의를 증진하고 있지만, 공급 측면에서 볼 때 알고리즘을 활용하는 사업자 간 경쟁에서 새로운 국면이 형성된다. 예를 들어, Uber와 같은 공유경제 플랫폼은 기본적으로 수요와 공급을 매칭(matching)하는 비즈니스모델이므로 알고리즘에 의한 가격설정 노하우와 시스템이 경쟁우위의 주요 원천이 된다. 이러한 비즈니스모델이 등장한 초기에는 데이터와 알고리즘을 통해 거래비용을 낮추고 공급과 수요 사이의 정보 비대칭성(information asymmetry) 문제를 완화함으로써 보다 투명하고 효율적인 시장을 가능하게 할 것이라는 낙관적인 전망이 지배적이었다. 이는 또한 기업의 혁신을 촉진하고 소비자의 선택권을 확대함으로써 사회 후생(social welfare)을 증가시킬 것으로 기대되었다. 그러나 투명하고 (거래비용 감소로 인해) 빈번한 거래를 가능하게 하는 알고리즘에 의한 자동화가 오히려 가격담합(price fixing)과 같은 반경쟁적 행위도 쉽게 하는 여지를 제공하는 역설적인 상황도 가능하다.

이러한 가능성에 대한 몇 가지 사례가 있지만, 이러한 가능성을 뒷받침하는 엄밀한 이론은 아직 제시되지 않고 있다(따라서 이 절의 설명 역시 개념적인 수준에서의 이론적 접근에 근거한다). Ezrachi & Stucke(2017), Mehra(2015), OECD(2017) 등이 알고리즘에 의한 담합(collusion) 가능성이 크다고 주장하면서 사회·경제적 파장에 대한 분석과 사전적 대비책 마련을 요구한다. 이와는 반대로, (게임이론 등을 활용한) 산업조직론에서의 전형적인 담합 모형에 따를 때, 알고리즘의 성능에 따라 수요예측의 정확도가 높아지는 것이 담합이 지속될 가능성을 낮추어(즉, 담합으로부터의 이탈의 유인을 증가시킴으로써) 소비자 후생에 긍정적인 영향을 미친다는 연구도 있다(Ittoo & Petit, 2018; Miklos-Thal & Tucker, 2018).³⁰⁾ 그러나 이들 연구에서 가정한 수요모형과 알고리즘은 매우 단순한

30) 또한, Ezrachi & Stucke(2017)와 OECD(2017) 등에서 주장하는 관련 사례도 전적으로 알고리즘 담합으로 볼 수

것이어서 향후 보다 정교한 보완이 필요해 보인다. 또한, 알고리즘 담합에 대한 시나리오도 단순해서 다음에 소개할 Hub-&-Spoke 형태와 같은 암묵적 담합 유형을 포괄하지 못한다(알고리즘 담합에서는 플랫폼을 활용하는 이런 형태의 암묵적 담합이 더 보편적인 형태가 될 것으로 전망된다).

담합에 관한 이론과 암묵적 담합(tacit collusion)

일반적으로 담합은 경쟁해야 할 주체들(예: 공급자들)이 소통을 통해 의도적으로 특정 행위(예: 가격)를 일치시키는 것으로, 대부분은 반경쟁적 행위(anti-trust behavior)에 해당하여 경쟁법에서 금지하고 있다. 주로 가격(예: 라면이나 원유 가격)을 고정하는 유형이 많기 때문에 담합이라고 하면 일반적으로 가격담합(price-fixing)을 의미한다.

담합에 참여한 주체들을 카르텔(cartel)이라고 부르는데, 중동의 석유 수출국 기구(OPEC)가 대표적 예이다. 담합이 흔하지 않은 이유는 법규제로 금지되어 있을 뿐만 아니라, 담합 자체가 유지되기 어려운 특성을 가지고 있기 때문이다. 담합은 게임이론에서 말하는 죄수의 딜레마(prisoner's dilemma)에 해당하는 균형(공모에 의한 균형, collusive equilibrium) 즉, 담합에서 이탈함으로써 카르텔에 속할 때보다 더 큰 이익을 보게 될 유인이 존재하는 경우가 많다(실제 OPEC 회원국들도 이탈하는 경우가 꽤 많다). 따라서 이러한 상태에서도 담합을 유지하기 위해서는 카르텔 회원들이 서로 견제와 감시뿐만 아니라 유효한 보복(retaliation)을 할 수 있어야 한다. 하지만 이러한 견제·감시·보복하는데 수반되는 거래비용 또한 상당하다는 점에서 담합 유지가 어렵다는 게 일반적인 견해이다.

그러나 알고리즘에 의한 담합은 그 거래비용이 매우 낮다. 또한, 담합이 형성되기까지 필요한 소통도 묵시적인(tacit 또는 implicit) 경우가 오히려 일반적이다. 이처럼 (비밀)계약 및 합의나 (비)공식적 소통 채널에 의한 제도화된 또는 의도적인 조정(institutional arrangement) 없이도 담합이 발생할 때 이를 '암묵적 담합(tacit collusion)' 혹은 '의식적 병행행위(conscious parallelism)'라고 한다.

알고리즘에 의한 가격설정이 문제가 되는 경우를 크게 다음의 두 가지 방향에서 생각해 볼 수 있다. 첫째, 고객에 대한 데이터가 대량으로 확보되고 빠르게 정확하게 처리·활용될 수 있는 데이터 경제에서 가격 차별(price discrimination)에 관한 문제가 새롭게 대두된다. 둘째, 알고리즘에 의한 여러 형태의 담합은 매우 정교하게 진행되거나 심지어 공급자조차 인지하지 못한 상황에서 발생할 수 있다. 특히, 전자는 동적 가격설정(dynamic pricing)

있는 것은 아니며(즉, 관련 사례에서 인간 및 조직의 개입이 더 중요한 역할을 했음) 전통적 관점에서의 담합에 더 가까워 보인다(Ittoo & Petit, 2018). 실제로 알고리즘에 의한 가격설정이 이미 오래전부터 도입되었던(그러나 세간의 주목을 받지 못했던) 항공이나 전력 산업에서 평균 실질 가격(long-term real price index)의 장기추세는 뚜렷한 하락세를 보인다.

과 관련되어 매우 미묘한 상황을 초래하며, 후자는 반독점법(anti-trust law) 전통에서 부정적으로 취급되었던 이슈에 모호한 상황과 새로운 도전을 제기한다는 점에서 경쟁법 전문가들의 관심을 끌고 있다.

먼저, 동적 가격설정은 기업의 수입 극대화(revenue maximization) 차원에서 이미 오래전부터 광범위한 영역에서 사용됐으며, 정당한 (때로는 우수한) 비즈니스 방식으로 인정받아왔다.³¹⁾ 예를 들어, 호텔과 항공 등 시간에 따른 사용권을 파는 서비스업에서 이러한 가격설정 방식의 효율성은 기업의 성과와 직결되기 때문에 중요한 지적 자산이 된다. 그런데 O2O 및 서비스화 경향이 촉진되면서 주로 서비스업에 국한되었던 동적 가격 설정은 이제 영역을 가리지 않고 적용될 수 있는 상황으로 전개되고 있다. 특히 데이터 기반 경제에서는 단순한 가격 차별을 넘어서 개별 소비자의 특성에 맞춤형 가격설정(personalized pricing)도 가능하다. 이는 2차 가격 차별화(second degree price discrimination)를 넘어서 3차(third degree) 가격 차별화가 쉬워질 뿐만 아니라, 종전에는 현실적 제약으로 인해 이론적으로만 가능했던 1차(first degree) 가격 차별화에 가까운 형태도 구현될 수 있음을 의미한다(김영혁, 2017; 최난설현, 2017).³²⁾ 따라서 개인별 가격설정에서는 소비자 후생(consumer welfare)의 상당 부분이 공급자로 전이될(transfer) 수 있다는 미묘한 문제를 낳는다.

플랫폼 기반 서비스화가 성공하기 위해서는 데이터 경제의 특징을 최대한 활용하는 것이 관건이므로 동적 가격설정은 중요한 자산이자 혁신의 도구이다. 이에 따라 알고리즘을 적극적으로 활용하는 동적 가격설정은 혁신적 비즈니스모델에 대한 지지와 소비자 보호라는 가치가 상충하는 지점이 된다. 아직 이에 대한 법제도적 입장이 정리되지 않은 것으로

31) 많은 마케팅 및 서비스운영에 관한 대학 교과서에서 동적 가격설정을 다루고 있으며, 이를 잘하는 것이 (특히 서비스업에서) 수입 극대화에 중요하다고 강조하고 있다. 동적 가격설정을 중점적으로 다루는 전문 학술지도 매우 많다(예를 들어, Journal of Revenue and Pricing Management 등).

32) 2차 가격 차별화는 구매량에 따라 단위가격을 다르게 하는 것으로, 대용량 구매자에게 단위가격을 할인해주는 수량할인(volume discount)이 대표적이다. 3차 가격 차별화는 구매자 집단의 특성에 따라 다른 가격을 적용하는 방식으로 극장의 조조할인(early-bird price discount)이나 특정 단체에 항공권을 싸게 파는 것 등이 대표적인 예이다. 2차 및 3차 가격 차별화는 일반적으로 별다른 문제 없이 광범위하게 인정되는 비즈니스 관행이다. 이에 반해 1차 가격 차별화는 소비자 개개인의 특성과 지불의사(willingness-to-pay)를 파악해서 개인별로 서로 다른 가격을 부과함을 의미한다. 따라서 소비자 보호를 우선시하는 법제도와 규제의 관점에서 볼 때 (사안에 따라) 문제의 소지가 있을 수 있다.

보이며, 향후 사안별로 이슈가 될 가능성이 크다. 특히 수요와 공급의 매칭을 핵심적 역할로 하는 공유경제 및 O2O 플랫폼에서 과도한 개인별 가격설정이 발생할 때마다 사후 규제(ex post regulation)가 개입될 가능성도 커진다. 실제로 뉴욕시에서는 Uber의 사실상 동적 가격설정 방식인 탄력요금제(surge pricing)가 특정 구간의 승객에게만 지나치게 높은 가격을 설정한 것이 문제가 되었으며, 이와 관련하여 2016년 3월에는 뉴욕 연방지방법원이 가격설정 방식의 문제점을 인정하는 판결을 내리기도 하였다. 그러나 이러한 가능성을 사전에 방지하기 위한 사전 규제(ex ante regulation)가 필요한지는 여전히 논란이 계속되고 있다. 플랫폼 간 경쟁이 심해지면서 데이터 기반 알고리즘을 활용하여 동적 가격 설정을 정교하게 다듬을 가능성이 커지고 경계를 넘나드는 가격 차별화도 심화될 가능성이 크기 때문에 플랫폼들은 알고리즘 위험관리를 위한 최소한의 대책을 마련해야 한다.

알고리즘에 의한 담합은 이미 2010년 이후부터 소프트웨어에 의한 불공정한 가격설정 방식이 문제가 되면서 관심을 끌기 시작했다(김건우, 2017; 최난설현, 2017; Blockx, 2017; Okuliar & Kamenir, 2017).³³⁾ 이후 빅데이터와 AI가 본격적으로 활용되고 주목을 받으면서 알고리즘에 의한 자동화된 담합과 순간적인 담합의 가능성이 제기되었다. OECD 경쟁위원회에서도 이러한 가능성에 주목하고 있는데, 이를 ‘디지털 카르텔(digital cartel)’이나 ‘암묵적 담합(tacit collusion)’이라고 부른다. 아직 이러한 가능성이 구체적이며 명백하게 실현된 사례는 발견되지 않기 때문에 가설 수준에 머무르고 있지만(Ittoo & Petit, 2017), EU를 중심으로 활발히 조사·연구가 진행 중이다. 예를 들어, 머신러닝의 자기 강화형 알고리즘이 EU 로마조약(TFEU, The Treaty on the Functioning of the European Union 또는 The Treaty of Rome) 101조를 위반할 가능성이 크다는 주장도 있다(Blockx, 2017). 아래에서는 OECD의 연구보고서(OECD, 2017)를 중심으로 지금까지의 조사 및 연구 결과를 간단히 정리한다. 먼저, 알고리즘에 의한 담합이 경쟁을 제한하는 경우를 정리하면 다음과 같다.

- 모니터링(monitoring) 알고리즘: 경쟁 기업의 가격설정 방식을 자동으로 모니터링 하면서 자신의 가격을 유리하게 책정하거나, 담합 기업의 가격을 모니터링 하면서 이

33) 대표적으로 Amazon 오픈마켓에서 포스터를 판매하던 Poster Revolution 사건이 미국의 반독점 금지법인 셔먼법 제1조를 위반했다는 판결을 받은 바 있다(2015년 4월).

탈할 경우 제재를 가함으로써 담합을 지속시키도록 위협(threat)하는 방식

- 병행(parallel) 알고리즘: 동일한 가격설정 알고리즘을 공동으로 사용하여 직접 정보나 의사를 교환하지 않으면서 담합과 동일한 효과를 창출하는 방식
- 신호(signaling) 알고리즘: 가격 조정에 대한 신호를 사업자들끼리 주고받으면서 담합을 유지하는 방식
- 자기학습(self-learning) 알고리즘: 알고리즘 스스로 사업자의 입장을 대변하여 최적화된 가격설정을 수행하는 과정에서 담합으로 수렴되는 방식

모니터링 알고리즘의 경우, 실제로 이미 가격 비교 웹사이트 등에서 시행되고 있기 때문에 반경쟁적이라고 보기 어렵다는 점에서 미묘한 문제를 낳는다. 오히려 소비자의 선택권을 확대하고 전반적인 가격수준을 낮춘다는 점에서 긍정적인 효과도 있기 때문이다. 반면에 병행 알고리즘은 위에서 소개한 Amazon 오픈마켓 사건에 적용된 적이 있으므로 반경쟁적 요소로 인정된다. 신호 알고리즘도 반독점법의 전통적인 관점에서 반경쟁적 행위로 인정되지만, 알고리즘을 사용하기 때문에 이를 검증하기가 쉽지 않다는 문제가 발생한다. 동적 가격설정과 결부되면 신호를 선도하는 기업이 짧은 시간에 신호를 방출하며 동조하는 기업들도 교묘하게 그 행위를 위장할 수 있기 때문이다.³⁴⁾

신호 알고리즘을 활용하지는 않지만 의도하지 않았던 행위가 가격담합의 신호로 오인되어 결과적으로 담합이 초래될 수도 있다. 이는 실제로는 가격 경쟁 하는 두 기업 사이에서 유연하지 못한 모방(imitation) 알고리즘을 사용할 때 가능한 시나리오이다. 예를 들어, 한 기업이 어떤 이유에서 (혹은 실수로³⁵⁾) 가격을 올렸는데 다른 기업이 이를 해석하는 과정에서 따라 하게 되고, 이들의 모방 알고리즘들이 유연하지 못해서 잘못된 신호와 해석을 빠르게 수정하지 못하면서 가격 인상을 진정한 의도(credible commitment)라고 잘못된 확신을 하게 되면 발생한다. 물론 모방 알고리즘은 너무 간단한 방식이어서 현실적이지 못하기 때문에 이러한 가능성이 현실화 되기가 쉽지 않을 것이다. 그런데도 위와 같은 시나리오는 기술적으로 충분한 가능성이며 모방 알고리즘보다 정교한 Q-learning 알고리

34) 이는 주식시장에서 작전에 의한 주가조작 신호를 탐지하는 방식과 유사하다.

35) 게임이론에서는 이를 '떨리는 손(trembling hands)'에 의하여 의도했던 바와 다른 행위를 하는 것으로 묘사한다.

즘을 도입한 쿠르노 경쟁형(Cournot competition) 실험에서는 암묵적 담합이 실현되는 경우가 많았다(Waltman & Kaymak, 2008).³⁶⁾

반독점법 관점에서 가장 다루기 까다로운 부분은 자기학습 알고리즘에 의해 초래되는 담합이다. 자기학습 알고리즘은 기업의 이윤 극대화 맥락에서 잘 들어맞을 뿐만 아니라, 시행착오에 의한 조정 과정을 모사한다는 측면에서 매우 유연한 알고리즘으로 인정된다. 이 경우 가격설정 방식 개발자 혹은 운영자가 의도하지 않아도 결과적으로 담합 및 이에 유사한 결과가 쉽게 발생할 수 있다는 점에서³⁷⁾ 규제와 처벌의 근거를 확보하기가 쉽지 않다. 또한, 앞 절에서 살펴본 알고리즘의 태생적 불투명성에 따라 그 결과를 사전에 예측하는 것이 어려울 뿐만 아니라 사후적 분석도 쉽지 않다. 즉, 알고리즘의 어떤 요인에 의해 결과적으로 담합이 발생하는지를 파악하는 것이 어려울 뿐만 아니라 알고리즘의 동태적 변화와 적응에 따라 다른 환경과 조건에서는 다른 요인에 의해 담합으로 수렴할 수도 있다. 그러나 아직은 여러 기술적(예: 알고리즘 복잡도와 컴퓨팅 능력, 2.2절 참조) 및 현실적(GDRP 등 개인정보 활용에 대한 법제도적 제한) 제약으로 인해 자기학습 알고리즘이 담합으로 수렴할 가능성이 크지 않다는 반론도 있다. 예를 들어, Hu & Wellman(2003)이 제시한 간단한 자기학습 알고리즘(Nash Q-learning 알고리즘³⁸⁾)을 가정할 때 암묵적 담합이 발생할 가능성이 매우 낮다.³⁹⁾

암묵적 담합을 가능하게 하는 알고리즘 유형을 바탕으로 Ezrachi & Stucke(2017)는 디지털 카르텔이 가능한 몇 가지 방식을 정리하여 제시한다. 먼저, 명시적 합의에 따른 반경쟁적 담합은 당연 위법(per se illegal)으로 병행 알고리즘과 같이 동일한 알고리즘을 사용하는 경우에 해당한다(Ezrachi & Stucke(2017)는 이를 메신저(messenger) 유형이

36) 이 경우에도 담합이 유지되기는 쉽지 않았다. 즉, 이런 담합을 유지하는 것은 비용과 같은 파라미터 변화에 취약하였다.

37) 게임이론 용어로 말하면, 암묵적 담합이 Nash 균형이 되어 매우 안정적이다. 즉, (최소한 이론적으로는) 담합에서 이탈할 유인이 없다.

38) Q-learning 알고리즘은 간단하면서도 강력하기 때문에 오래전부터 광범위한 영역에서 연구 및 활용되었다. 예를 들어, 로봇 통제나 비디오게임에서 에이전트들의 행동을 시뮬레이션할 때 적용됐으며, 항공과 전력의 동적 가격설정에서도 사용된다. 실제로 Atari2600 게임에서 Q-learning을 사용한 AI가 인간 게이머를 이긴 바 있다.

39) Hu & Wellman(2003)에서 사용한 간단한 알고리즘과는 달리, 실제로 많이 사용되는 머신러닝형 알고리즘은 보다 정교한 예측과 학습을 통해 담합에 가까운 결과(Nash 균형)를 만들 수도 있을 것이다. 그러나 아직은 기술적·현실적 제약들로 인해 반복게임(repeated game)에서의 Nash 균형으로 정의되는(김건우, 2017) 암묵적 담합으로 수렴될 가능성은 크지 않아 보인다.

라고 부른다). 디지털 카르텔로 주목을 받는 경우는 묵시적 합의나 합의가 없는 경우에 발생하는 암묵적 담합으로, Hub-&-Spoke 형태나 예측에 따라 움직이거나(predictable agent) 자동화된 의사결정에 따른 결과로 발생한다.

Hub-&-Spoke형 담합은 허브 역할을 하는 플랫폼과 수직적 관계에 있는 기업들(spokes)로 구성된 생태계에서 발생할 수 있다. 플랫폼이 제안하는 가격설정 방식을 동일하게 사용하면서 수평적 담합을 암묵적으로 유도하거나 의도하지 않은 상황에서 결과적으로 초래되는 경우가 이에 해당한다. 이는 과거에도 있었던 형태의 암묵적 담합으로, 예를 들어 하나의 유통업자와 다수의 거래처(공급업체) 사이에서 벌어지는 3자 간 행위에서 발생할 수 있다. 전자상거래와 O2O가 활성화되면서 이에 해당하는 담합 사례는 여러 분야에서 보이고 있으며, 알고리즘에 의한 가격설정이 도입되면서 좀 더 정교하고 탐지하기 어려워졌다(Okuliar & Kamenir, 2017).

예측 알고리즘과 자동화된 의사결정에 의한 담합은 가격과 같은 정보가 공개되고 탄력적으로 결정되는 환경에서 가능한 형태의 암묵적 담합이다. 예를 들어, 모니터링 알고리즘에서도 한 기업이 가격을 인하하는 순간 경쟁 기업들이 같은 수준으로 가격을 조정하는 것이 가능하므로 담합에서 이탈을 통해 이득을 취하는 것이 어려워진다. 이는 투명성이 오히려 담합을 유지하는 데 도움을 주는 역설적인 상황이다. 즉, 알고리즘에 의한 가격설정에서는 의식적 병행행위(conscious parallelism)가 더 쉬워지며, 의사결정과정 자체가 자동화될 경우 인간의 개입 없이도 빅데이터를 기반으로 한 경쟁 제한적 상황이 자연스럽게 초래될 수 있다. 반면에 알고리즘이 정교하게 진화할수록 담합을 위해 조율되어야 할 파라미터들(예: Q-learning에서 exploration vs. exploitation 파라미터)과 계산복잡도도 증가하기 때문에 자동화된 의사결정이 수렴하지 못할 가능성도 증가한다. 또한, 시장세분화와 동적 가격설정까지 고려한다면 계산복잡도가 지수적으로 증가하기 때문에 가격담합이 형성·유지될 가능성은 더 낮아진다.⁴⁰⁾ 그런데도 알고리즘 담합은 현행 반독점법 체제에서는 제지할 수 없는 문제를 일으키기 때문에 앞으로 더 많은 연구가 필요하다.

40) 이외에도, 게임이론 관점에서 볼 때 담합 자체가 형성되기 어려운 조건들이 많다. 대표적으로 담합 상대방의 진정한 의도, 전략적 대안, 실질적 이득 등이 '모두가 아는 정보(common knowledge)'가 아닌 경우가 일반적이다. 이론적으로는 이러한 내용이 모두가 다 아는 정보로 존재한다는 전제는 담합 균형을 예측하고 그 방향으로 수렴하기 위해 반드시 필요한 조건이다. B2C보다 B2B의 경우에 이런 정보가 더 제한적이다.

3 >> 레그테크(RegTech) 및 리걸테크(Legal Tech) --- regulation by machines

보스턴컨설팅그룹(BCG)은 2008년 글로벌 금융위기 이후 자금 세탁, 시장 조작, 테러 자금 등에 대한 규제가 강화되면서 전 세계은행이 약 3,210억 달러(약 369조 원)를 넘는 과징금을 부담한 것으로 분석했다. 실제로 미국 재무부 산하 금융범죄단속반은 플로리다주 Gibraltar Private Bank와 Trust Company의 자금세탁방지(AML, Anti-Money Laundering) 프로그램에서 오탐(false positives)⁴¹⁾을 포함한 결함이 발견되었다는 것을 근거로 400만 달러의 벌금을 부과한 바 있다. 4.2.2절에서도 언급한 바와 같이 통제나 관리가 불가능한 알고리즘 및 AI 시스템의 위험에 대해서도 예방적 컴플라이언스(preventive compliance 또는 법규제 준수)를 요구하는 상황에서 기업은 한정된 자원과 예산을 효율적으로 활용하면서 규제에 효과적으로 대응하기 위해 레그테크(Reg Tech)에 관심을 기울이고 있다.

레그테크 솔루션은 현재 정형 및 비정형 데이터의 처리 과정에서 법규제 준수를 위한 업무를 자동화하고 효율화하는 것을 목적으로 한다. 특히 금융기관과 다국적 기업을 비롯한 다양한 기업들이 법규제 준수와 대응에 필요한 운영비용을 절감하고 오류를 최소화함으로써 위험을 관리하는 것에 관심을 많이 기울인다. 세계경제포럼(WEF)은 2025년 전 세계 금융기관의 30%가 AI를 활용한 준법 감시시스템을 도입할 것으로 전망했다. 우리나라 금융감독원 역시 빅데이터 플랫폼으로 레그테크 인프라를 구축하여, 리포팅, 소비자 보호 및 보안에 관한 관리·감독을 고도화하고자 한다. 미국, 영국, 호주, 싱가포르 등 금융 분야 선진국에서는 이미 법규제 준수 서비스에 AI를 활용한 레그테크의 적극 도입을 장려하고 있다. 금융범죄단속반, 뉴욕연방준비은행, 통화감독청과 같은 미국의 대표적인 규제기관과 더불어, 글로벌 대형 은행들은 이미 SAS를 비롯한 여러 SI 업체들이 제공하는 레그테크 솔루션을 활용한다.

금융에서 레그테크 솔루션이 담당하는 대표적인 분야는 통상적인 사이버 보안과 AML

41) '긍정오류'라고도 하며, 오류나 위협 등 해당되는 조건의 이벤트가 아님에도 해당 이벤트로 인지하는 것

탐지를 포함하여, 직원 컴플라이언스, 위험관리·보고, 규제 컴플라이언스, 트랜잭션 모니터링 등의 통합 운영을 목표로 한다. 특히 의심되는 거래를 오탐할 때의 손실이 커지면서 기존의 보안시스템과는 본질적으로 다른 차원에서 이러한 시스템을 구현하고자 한다. 최근의 머신러닝 기반 AI는 AML 분석·모델링 및 고객 세그먼트 분류와 시나리오의 정교화·세분화를 통해 트랜잭션 모니터링의 효율성을 획기적으로 향상시킬 수 있는 여건을 마련하였다. 이는 레그테크가 기존의 단순한 보안시스템과 다른 차원에서 구현되어야 그 장점을 극대화할 수 있다. AML 규제가 무역 거래기반 자금세탁(Trade-Based Money Laundering, TBML)으로 확대되고 있는 상황에서 각종 거래에 따른 컴플라이언스 비용은 금융기관뿐만 아니라 무역에 관련되는 여러 이해당사자에게 더 큰 부담이 되고 있다. 글로벌 거래 환경에서 수출입 계약은 끊임없이 변경·수정되면서 매우 유동적인 상태에 있기 마련이므로 프로세스 전반에 걸친 정형·비정형 데이터를 실시간으로 분석해 송장 왜곡(mis-invoicing)과 같은 불법 행위를 사전에 파악하는 레그테크(예: Robotics Process Automation, RPA 등)는 실질적으로 큰 의미가 있다.

레그테크가 AI를 통한 효율성 향상과 위험관리를 통해 기존 법제도를 지원하는 데 중점을 둔다면, AI가 법제도 자체의 직접적인 주체가 될 수도 있다. 현행 변호사법은 변호사만 법률 사무를 볼 수 있지만, 일반 법인(예: IT 기업)과 변호사 간 동업을 제한적으로 허용하도록 하는 변호사법 일부 개정 법률안도 국회에서 논의 중이다(김종훈, 2018). 이미 영미권 국가를 중심으로 해외에서는 소위 리걸테크(legal tech)로 불리는 AI를 개발·운영하는 스타트업이 활발하다. 2016년 기준으로 미국에서는 약 1,100여 개의 리걸테크 기업이 활동하고 있고, 투자 액수는 2억9,200만 달러(약 3,300억 원)에 이르는 것으로 추정된다. 리걸테크 소프트웨어 등 직접적으로 관련된 IT서비스 시장도 앞으로 2년 뒤에는 약 57억 달러(약 6조 원) 규모로 성장할 것이란 전망도 있다(김종훈, 2018). 실제로 영국에서는 PPI(지급보증보험) 불완전판매⁴²⁾ 사건들을 놓고 ‘케이스 크런처 알파’로 불리는 AI 챗봇과 변호사들이 대결을 벌인 적이 있는데, AI의 예측율이 월등히 높은 결과를 보인 바 있다. 미국에서도 ‘Do Not Pay’로 불리는 챗봇이 16만 건에 달하는 주차위반 취소를 끌어낸

42) PPI는 보험가입자가 예기치 못한 일로 대출을 상환하지 못하게 될 때 보험사가 대신 갚아주는 보험상품으로, 1990년대부터 2000년대 중반까지 다른 금융상품에 끼워 팔기 형태로 수백만 명에게 판매되었다.

바 있다. 이러한 성공을 바탕으로 이 챗봇은 현재 난민신청자를 대상으로 무료 법률서비스를 제공하고 있다.

리걸테크는 법률 자문과 서비스에 드는 비용을 획기적으로 낮추면서 기득권을 지닌 대형 로펌 등의 진입장벽과 거래비용을 낮추는 역할을 한다. 예를 들어, 스타트업의 경우 사업 초기 단계에 필요한 어렵고 복잡한 법률 및 규제와 계약 관련 문제들을 낮은 비용으로 빠르게 해결하는 데 도움을 준다. AI 변호사와 판사 및 가상 법원 등을 활용하여 자동화된 법제도 운영도 가능하다는 장밋빛 전망도 등장한다. AI가 가장 빠르고 쉽고 광범위하게 활용될 수 있는 분야가 바로 법률시장이라는 것이다(마쓰오 유타카, 2015). 실제로 미국에서 활동 중인 리걸테크 스타트업만 1,000여 개가 넘으며, 이들은 판례 분석에서 송장 작성과 변호사 업무대행 등 전 방위로 법률시장에 진입하고 있다. 이들은 클라이언트 분석과 정보 및 문서를 정리하고 관련 법률을 검토·분석하는 업무에서 AI가 매우 우수하다고 주장한다. 대표적으로 미국의 파산 전문 로펌인 베이커앤호스테틀리는 파산 관련 사례와 판례를 수집·분석·정리하는 업무에 AI 변호사 Ross를 활용한다. 스탠퍼드대학교 로스쿨은 ‘CodeX 프로젝트’를 통해 법률 솔루션 스타트업을 지원하며, 저소득층을 대상으로 법률 자문서비스를 제공한다. 이외에도 계약서 작성·관리를 전문으로 하는 LawGeex, 스타트업들에게 맞춤형 패키지를 제공하는 Startup Documents, 낮은 비용으로 변호사 업무를 대행하는 Docassemble 등의 프로그램들이 온라인으로 법률서비스를 제공한다. 우리나라도 ‘i-LIS’와 같은 지능형 법률정보시스템이 상품화되어 조만간 법률시장에 등장할 예정이다.

판결이라는 매우 까다로운 의사결정을 해야 하는 판사의 업무 영역에서도 AI를 도입하는 방안을 실험 중이다. 민사재판과 같이 정서적인 면에 영향을 받을 수밖에 없고 아직 이해관계자들의 주장을 섬세하게 조율하는 것은 어렵겠지만, 다소 기계적으로 판단할 수 있는 사안에 대해서는 외압에 휘둘리지 않고 오히려 공정하게 결론을 내릴 수 있을 것이라는 기대를 받는다. 우리나라 대법원도 2018년 4월 ‘사법정보화 발전위원회’를 발족시키고 이러한 변화에 대응하고자 노력하고 있다.



결론

알고리즘 책임성에 대한 지금까지의 논의를 살펴보면 아직도 우리 사회가 알고리즘(및 AI)에 대한 이해가 부족하다는 것을 알 수 있다. 알고리즘의 책임성에 대한 논의 초점이 완전히 다른 방향으로 향해야 할 수도 있다. 알고리즘의 투명성과 같은 알고리즘 자체의 특성이 아니라 오히려 그 활용방식에서 더 큰 위험이 존재할 수 있다. 실제로 2018년 WEF의 'Global Risks Report'에서는 알고리즘의 남용으로 인해 발생할 수 있는 혼란과 자원 낭비 및 부작용의 가능성을 언급하면서 '잡초와 같은 퍼지는 AI(AI weeds)'를 경고하였다(Jenny, 2018). 저질 알고리즘과 AI는 인터넷을 통해 확산되면서 시간, 공간, 네트워크 대역폭, 에너지를 질식시키며 복잡하게 얽히고설킨 골칫덩어리('tangled web')로 남을 수 있다는 것이다. 부지불식간에 컴퓨터에 설치된 뒤로 메모리와 에너지를 소모하는 여러 작은 소프트웨어를 생각해보면 이러한 가능성을 충분히 이해할 수 있을 것이다. 그런데 Jenny(2018)은 알고리즘 스스로 진화하는 단계가 되면 그 자체가 적응시스템(adaptive system)이 되어 스스로를 유지·확대하려는 경향을 보이면서 잡초처럼 강한 생명력으로 완전히 제거되지 않고 퍼져나갈 것이라고 경고한다. 이런 어두운 시나리오가 시끌벅적하게 실현되는 것이 아니라 서서히 그리고 조용히 퍼지면서 그 폐해를 제대로 인식하지 못하게(심지어 당연한 것으로 받아들일 수도 있게) 한다는 점이 문제를 심각하게 만든다(잠시라도 관리하지 않은 잔디밭에 잡초가 퍼지는 것과 같다). 이미 이러한 단계에 이르면 어떤 플랫폼이나 조직(심지어 국가)의 통제권에서도 벗어나 있을 가능성이 크기 때문이다.

그러나 Jenny(2018)의 음울한 시나리오는 긍정적 현실을 애써 외면하려는 측면이 있다. 잡초처럼 남겨진 알고리즘과 AI는 혁신을 위한 부단한 실험의 어쩔 수 없는 부산물일 경우가 더 많을 것이다. 따라서 지금까지 논의한 여러 부작용(심지어 위기까지도)이 충분히 있을 수 있는 것으로 받아들이면서, 그 위험성을 평가하고 대응하는 보호 시스템

(safeguard system)을 구축하는 것이 알고리즘 책임성을 높이는 실효성 있는 접근방식일 수 있다. 이를 통해 우리의 디지털 인프라와 사회-기술시스템이 덜 취약하면서 높은 복원력(resilience)을 가지도록 만들어야 한다.

결국, 데이터기반 경제에서 알고리즘의 역할과 책임성에 대한 혼란은 과거 증기기관, 전기, 컴퓨터, 인터넷과 같은 새로운 혁신 기술이 등장할 때 우리가 겪었던 그것과 크게 다르지 않을 것이다. 혁신 기술이라 해도 사회에서 원하는 형태로 수용되고 정착되기 이전의 초기 단계에서는 불완전하고 부작용을 일으킬 수밖에 없다. 또한, 혁신 기술 자체가 악마를 품고 있는 것이 아니라, 그것을 활용하는 사람과 조직의 실수, 오용, 남용 때문에 바람직하지 못한 결과가 초래되는 경우가 더 크다. 머신러닝의 경우 많은 부정적 사례들은 잘못되고 편향된 원 데이터에 기인한다(garbage-in, garbage-out). 따라서 알고리즘의 불완전성을 받아들이고 그 전제하에 시스템을 설계하고 운영하며 사후 대비책을 마련하는 것이 현재 취할 수 있는 최선의 노력일 것이다. Amazon이 2014년부터 자사의 채용시스템에 도입한 AI의 편향 오류를 끊임없이 개선했지만, 결국에는 (일부 단계에서) 잠정적으로 중단하기로 한 결정이 그러한 선례일 수 있다. Google이나 Facebook 등의 AI를 적극적으로 활용하려는 플랫폼과 EU 집행위원회나 민간단체 등에서 발표하는 알고리즘 및 AI에 관한 여러 윤리 규정이나 가이드라인도 이러한 상황을 반영하면서 탄력적으로 진화해야 한다(예: EU 집행위원회의 ‘신뢰할 수 있는 AI를 위한 윤리 가이드라인 (초안); AI Ethics Guidelines for Trustworthy AI’, AI4People의 ‘좋은 AI 사회 구현을 위한 윤리적 프레임워크; Ethical Framework for a Good AI Society’, Google의 ‘AI 개발과 활용을 위한 7대 원칙’ 등). 현 상황에서는 알고리즘(과 AI)을 생산성 혁신을 위한 도구로 인식하고 사회에서 단계적으로 수용하는 로드맵을 마련하는 데 보다 집중할 필요가 있어 보인다.

참 | 고 | 문 | 헌

- 김진우 (2017) 알고리즘으로 움직이는 경제 디지털 카르텔 가능성 커진다, LG경제연구원 리포트, 2017.8.2.
- 김도훈 (2018) 알고리즘 책임성 논의와 알고리즘에 대한 이해, 주간기술동향, 2018.5.29., 14-28.
- 김도훈 (2016) 플랫폼서비스 생태계의 개념적 유형화, IT서비스학회지, 15(1), 299-319.
- 김민식, 이가희 (2017) 디지털 플랫폼과 인공지능(AI)의 이해, KISDI 정보통신정책연구, 29(18), 1-19.
- 김영혁 (2017) Dynamic Pricing이 확산되고 있다, LG경제연구원 리포트, 2017.11.3.
- 김종훈, 김종진 (2013), 컴퓨터 개론, 한빛아카데미, 2013.3.10
- 김종훈 (2018) 인공지능의 습격, AI와 전쟁 앞둔 변호사들, 머니투데이, 2018.8.3.
- 마쓰오 유타카, 박기원(역) (2015) 인공지능과 딥러닝(원제: 人工知能は人間), 동아엠앤비.
- 박외진 (2018) 인공지능 플랫폼의 개념과 도입 전략, 주간기술동향, 2017.10.3., 2-13.
- 박종훈 (2017) 딥러닝에도 보안 문제, 인공지능(AI)을 속이는 수법에 주의할 필요, 주간기술동향, 2017.11.29., 40-42.
- 박종훈 (2016) 답은 맞는데 풀이과정은 알 수 없는 인공지능을 믿어야 할까?, 주간기술동향, 2016.12.28., 31-38.
- 윤태승 (2018) 설명 가능한 AI 기술을 포함한 인공지능의 IP-R&D 전략, 주간기술동향, 2018.9.26., 2-13.
- 이원태 (2015) 인공지능의 규범이슈와 정책적 시사점, KISDI Premium Report, No.15-07.
- 이정원 (2017) 인공지능 플랫폼 동향과 정책적 시사점, 주간기술동향, 2017.5.10. 14-23.
- 주간기술동향 (2018a) 무인 자율운전차량의 핵심 쟁점 트롤리 문제: 동서양의 답은 다르다, 주간기술동향, 2018.11.21., 27-35.
- 주간기술동향 (2018b) 아마존의 인공지능 개발 플랫폼 개방을 통한 AI 생태계 확대 전략, 주간기술동향, 2018.7.4., 37-40.
- 주영재 (2018) 한몫 된 알렉사·코타나... AI 비서 생태계 흔들까?, 경향신문, 2018.8.17.
- 최난설현 (2017) 알고리즘을 통한 가격정보의 교환과 경쟁법적 평가, 경쟁법연구, 35, 215-241.

- 커즈와일(Kurzweil, R.), 장시형·김명남(역) (2016) 특이점이 온다(원제: The Singularity Is Near: When Humans Transcend Biology), 김영사.
- 파스칼레(Pasquale, F.), 이시은(역) (2016) 블랙박스 사회(원제: Black Box Society), 안티고네.
- NIA (2018a) 미 국방연구원 설명가능 인공지능(XAI), NIA Special Report 2018-2.
- NIA (2018b) EU의 인공지능 新 규제메카니즘: 설명가능 인공지능(XAI), NIA Special Report 2018-3.
- NIA (2018c) 미 국방연구원 AI 캠페인 차세대 인공지능(AI Next), NIA Special Report 2018-18.
- Baer, T. & Kamalnath, V. (2017) Controlling Machine-Learning Algorithms and Their Biases, McKinsey Global Institute, November Issue.
- Blockx, J. (2017) Antitrust in Digital Markets in the EU: Policing Price Bots, Proceedings of the Radboud Economic Law Conference, 2017.6.9.
- Cao, L. (2017) Data Science: A Comprehensive Overview, ACM Computing Survey, 50(3), 43:1-42.
- Capobianco, A. & Gonzaga, P. (2017) Algorithms and Competition: Friends or Foes?, CPI Antitrust Chronicle, August Issue.
- Curtis, F.E. & Scheinberg, K. (2017) Optimization Methods for Supervised Machine Learning: From Linear Models to Deep Learning, Tutorials in OR, 89-113.
- DARPA (2018) Explainable Artificial Intelligence (XAI), DARPA(Defense Advanced Research Projects Agency) Program Information (<https://www.darpa.mil/program/explainable-artificial-intelligence>).
- Deng, A. (2018) What Do We Know about Algorithmic Tacit Collusion?, SSRN Working Paper Series (<http://papers.ssrn.com>).
- Desai, D.R. & Kroll, J.A. (2018) Trust But Verify: A Guide to Algorithms and the Law, Harvard Journal of Law & Technology, (to appear).
- Diakopoulos, N. (2016) Accountability in Algorithmic Decision Making, Communications of the ACM, 59(2), 56-62.
- Doneda, D. & Almeida, V.A. (2016) What Is Algorithm Governance?, IEEE Internet Computing, 20(4), 60-63.
- Ezrachi, A. & Stucke, M.E. (2017) Artificial Intelligence and Collusion: When

- Computers Inhibit Competition, University of Illinois Law Review, 1775.
- Garces-Tolon, E. (2017) The Dynamics of Platform Business Value Creation, CPI Antitrust Chronicle, August Issue.
- Gartner (2016), Building a Digital Business Technology Platform, 2016.6.8.
- Garey, M.R. & Johnson, D.S. (1979) Computers and Intractability: A Guide to the Theory of NP-Completeness, Freeman.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009) The Elements of Statistical Learning: Prediction, Inference and Data Mining (2nd ed), Springer Verlag.
- Himel, S. & Seamans, R. (2017) Artificial Intelligence, Incentives to Innovate, and Competition Policy, CPI Antitrust Chronicle, December Issue.
- Hu, J. & Wellman, M.P. (2003) Nash Q-learning for General-Sum Stochastic Games, Journal of Machine Learning Research, 4, 1039-1069.
- Ittoo, A. & Petit, N. (2017) Algorithmic Pricing Agents and Tacit Collusion: A Technological Perspective, SSRN Working Paper Series (<http://papers.ssrn.com>).
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013) An Introduction Statistical Learning with Applications in R, Springer.
- Jenny, J. (2018) AI Weeds: What They Are, How They Could Choke off the Internet, World Economic Forum (<https://www.weforum.org/>).
- Knowledge@Wharton (2018) Autonomous Car Crashes: Who—or What—Is to Blame?, (2018.4.6). Retrieved from <http://knowledge.wharton.upenn.edu/article/automated-car-accidents/>.
- Li, S. & Xie, C. (2017) Rise of the Machines: Emerging Antitrust Issues Relating to Algorithm Bias and Automation, SSRN Working Paper Series (<http://papers.ssrn.com>).
- Mehra, S.K. (2015) Antitrust and the Robo-Seller: Competition in the Time of Algorithms, Minnesota Law Review, 100, 1323.
- Mullainathan, S. & Spiess, J. (2017) Machine Learning: An Applied Econometric Approach, Journal of Economic Perspectives, 31(2), 87-106.
- Ng, A. (2016) What Artificial Intelligence Can and Can't Do Right Now, Harvard Business Review, November Issue.

- Ni, X., Wu, C. & Zhao, H. (2018) Biased Learning Creates Overconfidence, *Journal of Systems Science & Complexity*, 31: 1603.
- OECD (2017) Algorithms and Collusion, Background Note by the Secretariat, OECD Roundtable on Algorithms and Collusion.
- Okuliar, A. & Kamenir, E. (2017) Pricing Algorithms: Conscious Parallelism or Conscious Commitment?, *CPI North America Column*, July Issue.
- Parnas, D.L. (2017) The Real Risks of Artificial Intelligence, *Communications of the ACM*, 60(10), 27-31.
- Perel, M. & Elkin-Koren, N. (2017) Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement, *Florida Law Review*, 69, 181-234.
- Pyle, D. & San Jose, C. (2016) An Executive's Guide to Machine Learning, *McKinsey Quarterly*, June Issue.
- Raymond, A., Young, E.A.S. & Shackelford, S. (2017) Building a Better HAL 9000: Algorithms, the Market, and the Need to Prevent the Engraining of Bias, *Northwestern Journal of Technology and Intellectual Property*.
- Saurwein, F., Just, N., & Latzer, M. (2015) Governance of Algorithms: Options and Limitations, *Info*, 17(6), 35-49.
- Sun, Y. (2018) More Efficient Machine Learning Could Upend the AI Paradigm, *MIT Technology Review*, 2018.2.2. Retrieved from <https://www.technologyreview.com/>.
- The New York Times (2017) Uncle Sam Wants Your Deep Neural Networks, 2017.6.22.
- USACM (2017) Statement on Algorithmic Transparency and Accountability, ACM U.S. Public Policy Council & ACM Europe Policy Committee, (2017.5.25.) <https://www.acm.org/public-policy/>.
- Waltman, L. & Kaymak, U. (2008) Q-learning Agents in a Cournot Oligopoly Model, *Journal of Economic Dynamics and Control*, 32(10), 3275-3293.

데이터로 연결하는 **사람·세상·미래**



Bridging DATA

제4편

데이터 혁신과 고용 창출

NIA 한국정보화진흥원



CONTENTS

요약	229
I. 데이터 혁신을 통한 고용 창출 전략	231
II. 데이터 혁신과 신기술의 도래	236
1. 데이터 생산과 주요 이슈	236
2. 새롭게 생성되는 공공데이터의 종류와 활용	240
3. 4차 산업혁명에 따른 미래 일자리의 변화	242
4. 데이터 혁신 기술이 기업 일자리 창출에 미치는 영향	246
III. 노동시장의 변화와 사회적 타협	251
1. 신기술이 미치는 노동시장의 변화	251
2. 신기술이 미치는 복지 사회적 변화	256
3. 신기술의 변화가 초래할 기술 불평등	259
4. 공공 고용서비스의 역할	263
IV. 데이터혁신에 따른 도시개선 및 일자리 창출 해외 사례	267
V. 데이터 혁신 평가 사례	275
1. 데이터의 활용 정도	275
2. 데이터 기반의 핵심 기술 플랫폼 활성화 방향	281
3. 인력 및 기업 자산 개발	285
VI. 변화에 발맞추기 위한 정책적 방향	291
참고문헌	299

요 약

최근 우리 사회는 고용문제가 어느 때보다 큰 사회적 이슈로 떠오르고 있습니다. 디지털 산업 혁명으로 불리는 4차 산업의 변화에 대해 많은 사람들은 새로운 일자리 창출을 기대하지만, 다른 한 편으로 이러한 변화가 대량실업의 위기를 가져올지도 모른다는 불안감을 동시에 느끼는 것 같습니다. 4차 산업 혁명이 미칠 노동시장의 커다란 변화에 이제는 정부 차원에서 효과적으로 대응 전략을 개발할 필요가 있습니다. 본고에서는 대응 전략의 일환으로서 데이터 혁신에 발맞춘 일자리 창출을 주제로 정부가 추진해야 할 방안을 모색해 보고자 합니다.

첫 번째 제안 전략은 데이터 생산 혁신을 위해, 정부나 공공기관이 소유한 공공 데이터를 더욱 폭넓게 개방하고, 접근이 용이하도록 시스템을 구축하며, 품질을 유지 및 관리하도록 제도화하는 것입니다.

두 번째 제안 전략은 정책 수행의 효과성 제고를 위해 개방 데이터의 활성화 정도와 일자리 창출 관련 4차 산업 추진과정을 측정할 수 있도록 평가 지표를 수립하는 것입니다.

세 번째 제안 전략은 데이터 혁신을 위해서 재교육 및 연구개발 확대, 공공 고용서비스 역할 확대, 기업가 정신을 촉진하는 것입니다. 이를 통해 사회가 더욱 안정적으로 발전해 나갈 수 있을 것입니다.

이러한 제안 전략이 타당한 이유는 공공뿐만 아니라, 민간분야에서도 사라지는 일자리를 대체할 수 있는 새로운 일자리를 창출할 수 있기 때문입니다. 공공 분야에서는 데이터 생산 일자리 및 4차 산업에 필요한 데이터 과학 분야의 일자리를 창출할 수 있습니다. 또, 민간 분야에서는 기업의 기존 제품 혁신과 데이터를 이용하여 새로운 서비스를 창출하는 과정에서 새로운 일자리를 창출할 수 있습니다. 근로자에게 재교육 기회를 확대함으로써

서 기술격차로 인한 불평등을 완화하여 사회 안정성을 높일 수 있으며, 기업이 정신을 촉진해 4차 산업에 적합한 새로운 제품 및 서비스 개발이 촉진될 수 있습니다.

데이터 혁신을 통한 일자리 창출 사례는 경제, 정부, 인프라, 환경, 시민생활의 여러 분야에서 해외 여러 도시를 통해 찾을 수 있습니다. 그리고 데이터의 활용정도, 데이터 기반의 핵심 기술 플랫폼 활성화, 인력 및 기업 자산개발의 측면과 같은 세부적인 4차 산업의 진행 정도는 미국의 사례를 통해 매우 체계적으로 평가지표를 찾을 수 있습니다.

정부차원에서 데이터 혁신에 따른 시대 변화의 중요성을 이해하고, 앞으로 나아가야 할 방향을 모색하는 것은 매우 중요합니다. 정부가 적절한 제도적 장치와 대응 전략을 마련한다면 많은 사람들은 이러한 혜택들을 다양한 방식에서 골고루 누릴 수 있을 것이고, 4차 산업의 변화는 우리에게 더욱 희망적인 미래로 다가올 것입니다.

최근 청년실업, 중장년 및 고령자 실업 문제가 더욱더 큰 사회적 문제로 떠오르고 있습니다. 청년실업 문제는 대졸청년의 높은 실업률과 양질의 일자리 부족에 기인한다고 볼 수 있고, 중장년의 실업은 조선업, 자동차 산업 등 다양한 분야에서 야기되는 경기 침체로 인한 일자리 부족으로 원인을 찾습니다. 최근 택시직종의 변화와 같이 4차 산업으로 인한 노동시장의 변화에도 커다란 영향을 받습니다. 고령자 실업은 인구구조의 변화로 인한 고령자 인구 증가와 노후 준비 부족으로 노후에도 계속적으로 일을 해야 하는 고령자들이 점차 많아지는 것에 기인합니다. 고령자 인구의 경제적 문제는 중장년층의 노후 준비에 대한 경제적 불안감을 야기하고, 이는 소비감소와 개인 이기주의 등 또 다른 사회적 안정을 저해하는 원인이 될 수 있습니다.

고용정책은 고용만을 따로 떼어 해결책을 강구하는 것은 불가능합니다. 고용은 사회 경제상황에 영향을 받기 때문에 사회경제 정책에 맞추어 고용정책을 수립할 수밖에 없습니다. 또 경제정책 역시 사회정책과 분리하여 생각하기 어렵습니다. 정치적 안정 상태와 정책적 방향이 한 나라의 경제 상황에 중요한 영향을 미치기 때문입니다. 정부가 정치적 안정을 통해 국민의 신뢰를 얻고 그에 기반한 정책을 실행한다면, 국민 개개인은 자신의 노력에 대한 노력을 기대할 수 있어 결과적으로 국민 경제의 선순환으로 작용할 것입니다 (Acemoglu and Robinson, 2018). 4차 산업이라는 시대변화를 기회로 우리나라가 현재 직면한 문제를 해결하기 위해서는 데이터 혁신이 사회 및 경제에 미치는 효과를 분석하고 관련 정책과 제도적 장치를 마련해야 합니다. 본 보고서에서는 4차 산업이 사회에 미치는 전반적인 이슈사항을 논함으로서 새로운 정책 도입에 따른 부작용을 예방하고자 합니다.

취업률을 높이고 실업률을 낮추기 위해서는 디지털 시대에 부응한 새로운 일자리를 창출해야 합니다. 일자리 매칭은 부분적인 해결책에 머무를 수밖에 없기 때문에 구조적 실업

문제를 해결하기 위해서는 공론화를 통해 일자리 확대에 대한 사회적 타협이 필요합니다. 양질의 일자리를 확대하기 위해서는 단기적으로 4차 산업과 관련된 데이터 생산을 위한 공공 일자리를 확대해 나갈 필요가 있습니다. 중장기적으로는 기술 산업 시장에 자본의 유입을 확대하고 4차 산업과 관련한 노동시장의 형성을 지원하여야 합니다.

더 이상 실업의 원인이 개인에게서부터 기인한다고 보는 시대는 지나갔습니다. 오히려 오늘날의 실업 문제는 사회구조적인 문제로 보는 시각이 점차 설득력을 얻고 있습니다. 과거 대학을 졸업한 청년이 대부분 취업하던 시대에는 개인의 노력 부족으로 실업 원인을 찾기도 하였습니다. 하지만 오늘날과 같은 시대에는 높은 청년실업률이 양질의 일자리가 부족하기 때문이라고 사회구조적 문제로 보는 공감대가 형성되어 있습니다. 과거 IMF 시절과 같은 경제위기에는 개인의 실업을 사회구조적 실업에서 그 원인을 찾았습니다. 고령자 인구의 재취업은 경비, 청소 등의 일부 업종에 국한되는 경향이 있습니다. 고령자 인구는 급격히 증가하지만 일자리 수는 한정되어 있습니다. 이 역시 고용의 문제를 개인의 문제로 보기 보다는 사회구조적 문제로 보아야 할 것입니다.

이러한 현실 속에서 4차 산업이 새로운 일자리를 창출할 것이라는 기대감도 있지만, 반대로 기존의 많은 일자리가 사라고 양질의 일자리는 줄어든 것이라는 우려도 높습니다. 때문에 일자리의 양을 늘리려는 시도로서 4차 산업을 통해 신제품 혁신의 경쟁력을 높이고, 신제품 개발을 통해 새로운 양질의 일자리를 창출하려는 시도를 게을리해서는 안 됩니다. 3차 산업 시대는 정보를 다루던 시대인데 반해, 4차 산업 시대는 정보 기반으로 제품 제조를 확대해 나가는 시대입니다. 3차 산업 시대는 사람과 사람 그리고 사람과 정보를 서로 연결하였다면, 4차 산업 시대는 사람과 정보에서 물체(제품, 설비 등), 장소로 연결을 확대합니다.

4차 산업 기술 육성은 장기적으로는 중고생이나 대학생에게 해당하는 정책이라고 할 수 있으며, 단기적으로는 대학교 관련 공학 전공 학생이 정책의 직접적 대상이 될 수 있습니다. 반면, 중장년들은 4차 산업에 필요한 새로운 기술을 습득하여 첨단 기술 일자리에 재취업하기에는 한계가 있습니다. 오히려 이들은 4차 산업으로 일해 줄어드는 일자리에 대한 직접적인 피해를 걱정하고, 새롭게 생산되는 질 낮은 일자리(예, 플랫폼 일자리)로

내몰릴 위험을 염려하고 있습니다. 가령 학원의 1등 강사에게는 디지털 산업이 새로운 기회가 될 것입니다. 하지만 모든 학생들을 1등 강사의 강의만 들으려 할 것이기 때문에, 2등 강사는 오히려 소득이 줄어든 위험이 될 수 있습니다. 그러므로 4차 산업은 양질의 일자리를 줄일 것이라는 우려의 목소리에 대비하여 이를 극복할만한 양질의 일자리를 늘리려는 사회 제도적 개선책을 함께 찾아야만 합니다.

4차 산업의 일자리 혜택을 누리려면 데이터 혁신 인프라의 구축이 필요합니다. 4차 산업으로의 변화 과정에서 기존의 많은 일자리가 사라지는 것은 사실이지만, 반면에 또 다른 일자리가 새로 만들어질 것으로 예상됩니다. 디지털 혁신의 혜택을 누리기 위해서는 데이터 혁신과 4차 산업 기술의 발전이라는 두 가지 측면이 함께 맞물려 이루어져야 합니다. 4차 산업이 발전하기 위해서는 단순히 기술만 개발하는 것이 아니라, 사회 공동체가 어떻게 생각하고 그 철학이 무엇인지가 중요합니다. 우리가 얼마나 포용적이고 공정한 생각을 가지느냐가 우리의 삶에 직접적인 영향을 미치게 될 것입니다.

4차 산업을 통해 제품 혁신 및 신규 일자리 창출을 하려면 경쟁력 있는 기술, 기업가 정신, 자본의 삼박자가 맞아야 합니다. 데이터 혁신이 원활하게 이루어지려면 사회적 안정이 필요하다는 점을 인식할 필요가 있습니다. 사회 안정을 통한 데이터 혁신은 더 나은 미래사회로 우리를 이끌어 갈 것입니다. 데이터 혁신이 이루어지려면 데이터 관리 IT 기술 뿐만 아니라, 데이터 생산을 위한 사회적 합의가 필요합니다. 사회적 합의를 이끌어 내려면 사회적 투명성을 높이는 것이 사회 전체적으로 이익이라는 인식기반이 요구됩니다. 사회적 투명성을 높이려면 기득권층이 자신의 독점적 위치를 내려놓아야 하는 부분도 있을 것입니다. 하지만 크게 우려하지 않아도 될 점은 안정적인 사회로 나아갈수록 기득권층이 자신의 독점적 위치를 내려놓기 쉬워질 것이라는 점입니다. 가령, 독재자가 자신의 현재 권력을 내려놓기는 어렵겠지만, 반대로 어느 누구라도 안정적인 노후의 삶이 보장된다면, 기존의 권력을 내려놓기는 상대적으로 쉬울 것입니다.

정부는 공공 및 민간 부문에서 우리사회에 도움이 되는 데이터 사용을 가속화하고 구체화하는데 중요한 역할을 할 권한을 가지고 있습니다. 정부는 공공데이터를 대중에게 공개하고 새로운 유형의 데이터를 수집하며 민간 부문의 데이터 사용 방식에 대한 규제를 개선

하는 정책을 통해 이를 지원할 수 있습니다. 요즘과 같은 빅데이터 시대에 데이터는 경제 성장과 사회 문제 해결에 필수 도구라고 할 수 있습니다. 경제 및 사회 전반에 걸쳐 양질의 데이터 사용을 확대해 나가면 이를 통해 신제품 및 서비스를 활성화하여 삶의 질을 향상시킬 것입니다. 그리고 새로운 통찰력을 제공하여 정부와 기업이 정확하게 의사 결정을 하도록 돕게 될 것입니다.

데이터 혁신으로의 안정적 발전을 위해서는 정부는 재교육 및 연구개발 확대, 고용유연성과 사회보장 확대, 기업이 정신 축진을 포함한 노력을 수행하여야 합니다. 데이터 혁신을 통해 정부 투명성을 제고하거나, 행정적인 비효율을 줄이고, 소비자에게 권한을 부여하고, 민간 부문에 새로운 기회를 창출하여 경제 사회적인 발전을 가져올 것입니다. 그러므로 데이터 혁신을 주도적으로 지원하려면 정책 입안자는 적절한 시기에 적절한 정책 지원을 제공해야 합니다.

사회 경제적인 이익을 위해 정책 입안자는 데이터 개방 및 활용 중심의 기술 혁신이 번성하기 위해 필요한 데이터, 기술 및 사람들에 대한 투자를 게을리 해서는 안 됩니다. 정부는 데이터 중심의 기술 혁신에 긍정적인 영향을 미칠 수 있는 직접적인 행동을 취해야 합니다. 정부는 사회적 변화의 진전을 위해 관련 정책을 법안으로 제정해야 합니다. 예를 들어, 소비자가 스마트미터 프로그램을 선택하고 업그레이드하는 작업을 막는 시대에 역행하는 제도가 만들어지지 않도록 주의해야 합니다. 각 지자체에서도 스스로의 자치법으로 인한 조치로 인해 큰 영향을 받습니다. 예를 들어 지방 정부는 스스로 대중교통, 건물 허가 또는 식당 청결 검열 등에 관한 데이터를 게시할지 여부를 결정할 수 있습니다. 이러한 노력은 중요한 결과를 가져올 수 있습니다. 예를 들어, 뉴욕시는 BigApps 대회를 통해 임대료 가격, 일몰을 보기 좋은 곳의 지도 등을 제공하는 앱의 개발을 장려한 결과 생활 편리성이 향상되었습니다. 그러므로 각 지방자치 단체 스스로 관련 정책을 수립해 나가야 합니다.

본 보고서의 구성을 살펴보면 먼저 2장에서는 데이터 혁신의 의미를 살펴보고, 이러한 변화를 데이터 생산, 새롭게 생성되는 데이터의 소스 활용, 4차 산업 시대의 신기술 종류, 신기술이 가져오는 미래 사회 일자리의 변화 형태, 데이터 혁신 기술이 경제 산업 및 고용

창출로 이어지는 효과를 살펴보고자 합니다. 3장은 이러한 노동시장의 시대적 변화에 대응하기 위한 공공 고용서비스 정책을 마련하는 데 초점을 두고, 정책 방향이 취업알선에 머무르기보다 사회 안정성을 높이는 사회 통합적 관점으로 바라보아야 할 것을 강조합니다. 이러한 관점에서 노동시장의 변화와 사회적 타협, 신기술이 초래하는 노동시장의 변화, 사회 복지적 변화, 기술 불평등을 살펴보고자 합니다. 또한 한국 노동시장의 과제와 공공 고용서비스의 역할을 모색해 보고, 사회 안정성을 높이기 위한 노동시장 정책의 필요성을 강조하고자 합니다. 4장은 새로운 일자리 창출을 위해 사회 투명성 제고 및 데이터 생산 정책의 필요 사례를 여러 해외 주요 도시에서 찾아보고자 합니다. 5장은 데이터 혁신에 따른 사회 안정성과 일자리 창출을 효과적으로 수행하기 위해서 정부의 데이터 혁신의 정책을 보다 효과적으로 평가할 필요성을 논의합니다. 이와 관련하여 미국의 사례를 살펴보고자 합니다. 미국의 데이터 혁신 평가는 데이터의 활용 정도, 핵심 기술 플랫폼의 활성화 측면, 인력 및 기업 자산 개발이라는 측면에서 수행되고 있는데, 우리나라 실정에 적용할 수 있도록 구체적인 항목까지 자세히 살펴보고자 합니다. 마지막으로 6장은 4차 산업의 시대에 효과적으로 일자리를 창출하기 위해 사회의 투명성 제고, 데이터 생산, 신기술 산업 일자리 육성 정책을 시기적절하게 수행할 것을 다시 한 번 강조합니다. 그리고 정부가 수행해야할 과제를 다시 한 번 정리하여 제시하고자 합니다. 4차 산업의 시대에 효과적으로 일자리를 창출하기 위해 정부 스스로 데이터 생성 분야에 일자리를 늘리고 데이터 혁신을 이끌 지도자 육성에 적극적인 것을 제안합니다. 데이터 혁신과 신기술이 경제 산업에 미치는 실질적 효과가 매우 크기 때문에 이러한 변화에 대응하는 것은 매우 중요합니다.



데이터 혁신과 신기술의 도래

데이터 혁신이란 디지털 혁명 시대에 국가 차원에서 양질의 데이터를 생산하고 이러한 데이터를 최대한 이용할 수 있도록 시민들에게 정보제공을 확산하는 것을 의미합니다. 이를 위해 데이터의 생산과 활용 및 이것이 산업에 미치는 효과를 살펴보고자 합니다.

1 » 데이터 생산과 주요 이슈

데이터는 종종 새로운 오일에 비교됩니다. 석유는 매장량이 많은 일부 지역만 경제적 혜택을 누릴 수 있습니다. 그러나 데이터는 어느 지역이나 지능적인 정보 수집, 공유 및 사용을 장려하기 위한 공공 정책을 제정함으로써 데이터 경제의 혜택을 누릴 수 있습니다 (Morabito, 2015; Vivarelli, 2012). 중앙 정부는 데이터 혁신을 촉진하기 위해 지방 정부 및 민간 단체를 지원하는 조치를 취할 수 있습니다. 이러한 단계를 수행함으로써 정부는 데이터 기반 혁신의 이점을 얻을 수 있는 역량을 향상시킬 것으로 기대할 수 있습니다. 데이터 품질을 높일 때 이를 이용한 기업의 수가 늘어나고, 이는 일자리 수를 증가시킬 것입니다. 데이터를 이용하는 기업의 수와 데이터 과학 일자리 수는 높은 상관관계를 가집니다. 양질의 데이터를 생산해 내기 위해서는 국가의 각 정부기관(재무, 보건, 복지, 증권 거래 등 관련)과 지방 정부가 협력하여 데이터를 생산해야 합니다. 정부는 공공 데이터를 생산하는 일자리를 늘려야 합니다.

● 공공 및 민간 데이터의 소유권 명확화

정부와 공공 기관은 공공 서비스를 제공하고 공공의 이익을 증진하는데 데이터를 사용하고 관리하는 감독자입니다(Gudipati et al., 2013). 관련 특허정보의 경우 특허 소유자가 소유권을 가지지만 정부가 그 존재와 완전성을 보호하지 않는 한 그 특허정보는 모두에게 쓸모가 없습니다. 개인, 기계 등 다양한 출처에서 데이터가 생산된다고 하더라도 기계는 오류의 가능성이 있고, 사람은 데이터를 속일 수 있습니다. 정부와 공공 기관은 데이터의 불일치가 발생하였을 때 이를 관리하는 방법과 규칙을 만들어야 할 것입니다.

● 자유와 평등의 확대와 통제의 딜레마

사생활보호와 개인의 자유는 동전의 양면과 같습니다. 국민 개개인에게 보다 맞춤형 정보를 제공하고자 한다면, 개인의 연령, 지역, 학력 등 더욱 많은 정보를 필요로 할 것입니다. 개인별 맞춤 정보에 대해 부정적인 시각도 있습니다. Kerr와 Earle (2013)은 개인의 건강, 위치, 핸드폰 사용 및 온라인 활동에 근거하여 개인을 프로파일링 하는 것이 차별과 배제를 부추기고 통제 상실의 위험을 높인다고 하였습니다. 공공 서비스를 맞춤형으로 제공할 때 이러한 위험은 보다 심각해질 수 있습니다. 가령, 워크넷의 일자리 정보를 그 사람의 학력과 이전 직장의 고용형태를 고려하여 낮은 임금의 일자리 위주로 추천을 하였다면, 취업 기회는 높아지지만 구직자는 보다 높은 임금의 일자리 기회를 상실할 수도 있습니다. 또 만일 인공지능에 의해 어떤 지역(예, 산악 지역)에 위험이 선제적으로 예측되어 출입을 통제한다면 이는 개인의 이동 자유를 제한하는 것이 될 수도 있습니다. 선제적 예측은 보수적으로 혹은 반대로 낙관적으로 해석될 소지가 있습니다. 정부는 가능한 보수적으로 해석하여 대처하려 할 소지가 높습니다. 가령, 청소년들이 밤거리를 배회한다고 할 때 이들이 절도를 하거나 배행을 저지를 확률은 높지만 범죄를 저지른 것은 아닌데 외출 자체를 억압한다면 반대급부로서 부작용(예, 사회적 격리와 우울증)이 양상 될 소지가 있습니다. 디지털 시대에 또 다른 문제점은 디지털 기기 사용에서 자의 혹은 타의에 의해 배제된 이들의 의견이 반영되지 않을 수 있다는 점입니다(Lerman 2013). 정부가 빅데이터에 너무 많이 의존하게 되면 의사결정과정에서 다양한 사람의 소수 의견을 고려하지 못할 수

있기 때문입니다.

● 개인정보 보호 및 활용

컴퓨터 기술의 진보는 데이터를 수집하고 분석할 수 있는 기회를 열어주고 있습니다. 데이터 분석을 통해 개별 소비자 뿐만 아니라 사회 전반에 도움이 되는 중요한 통찰력과 혁신을 이끌어 낼 수 있습니다. 데이터는 일반적으로 단일 목적으로 수집되지만 경제적 및 사회적 가치가 있는 여러 가지 용도로 사용됩니다. 개인정보라고 하더라도 익명화를 통해 통계적 활용은 허용되어야 합니다. 예를 들어, 대규모의 휴대 전화 데이터 소스는 도시 계획자나 엔지니어가 교통 패턴을 더 잘 이해하고 혼잡을 최소화할 수 있는 도로 네트워크를 설계하는 데 도움이 될 수 있습니다. 사회 현상을 분석하기 위한 대부분의 데이터 분석에는 개인 식별 정보를 반드시 사용해야 하는 것은 아닙니다.

데이터 집합에 개인 식별 정보가 포함되어 있다면, 조직에서 개인 정보를 보호하기 위한 준비가 필요합니다. 익명화된 데이터 집합을 만드는 데에도 통계적으로 엄격함을 요구하며 형식적으로 수행해서는 안됩니다. 개인정보를 수집한 기관은 비식별 기술 등에 최첨단 기술을 도입하여 위험에 대비해야 합니다. 정책 입안자는 강력한 개인정보 실별 도구, 교육 및 모범 사례를 개발하여 이러한 기술이 더 널리 채택되도록 지원해야 합니다. 조직은 개인정보 보호를 위해 위험성을 낮출 수 있는 관리 구조가 있어야 합니다.

반면, 데이터에서 개인정보 식별을 제거하면 데이터의 유용성이 감소할 수 있습니다. 데이터의 유용성과 개인정보 노출에는 균형이 맞춰져야 합니다(Cavoukian & Castro, 2014). 데이터가 더 구체적이고 식별 가능할수록 연구자에게 더 유용할 수 있지만, 개인정보 노출의 위험이 커지기 때문입니다. 그러므로 용도에 맞추어 개인정보의 활용 수위를 조절하는 시스템을 체계화할 필요가 있습니다. 법적 합의나 물리적 보안 통제로 보호되는 데이터의 사용은 일정 자격을 갖춘 담당자가 민감도가 높은 데이터에 접근 할 수 있도록 허용함으로써 개인 정보 보호의 균형을 유지하여야 합니다.

● 낮은 품질의 데이터에 잠재된 위험

낮은 품질의 데이터는 상당히 큰 문제점을 일으킬 수 있는데, 이러한 문제점은 정부와 시민들에게 모두 중요한 문제입니다. 기록된 데이터는 쉽게 오류가 발생할 수도 있고 조작될 수 있습니다. 데이터의 완전성, 적합성, 일관성, 정확성, 복제 및 무결성을 검사해야 하며 데이터 품질을 관리하여야 합니다. 민간기관은 전체 데이터에서 낮은 품질의 데이터를 제외하고, 정확한 데이터만을 선별하여 사용할 수 있을 것입니다. 하지만, 보건 서비스와 같은 분야에서 데이터 품질이 낮아지게 되면 그만큼 문제가 커질 수 있습니다. 데이터의 품질은 데이터의 병합 과정에서 발생할 수도 있으며, 데이터를 검증하는 작업 또한 매우 어렵고 힘든 일이지만 중요한 일이기도 합니다.

독립적으로 수집된 데이터를 가지고 분석 결과를 연관지어 해석하는 과정에서도 문제는 발생합니다. 예를 들어, 워크넷과 사람인 취업사이트의 직종 분류가 서로 상이할 때 직종별 종사자 수를 서로 연결지어 분석하는 것은 매우 조심하여야 할 것입니다. 데이터 분석 문제를 해결하기 위한 규제 및 기타 지원에 관한 기구도 필요하며 이러한 진단 분야의 전문가 또한 필요합니다.

2 《 새롭게 생성되는 공공데이터의 종류와 활용

4차 산업 시대에 경쟁력을 갖추기 위해서는 데이터 활용이 중요합니다. 어떠한 데이터를 생산해야하고, 어떻게 데이터가 활용되는지를 이해하기 위해 데이터의 종류를 살펴보고자 합니다.

● 민간-공공 파트너십을 통한 데이터 생산과 활용

민간-공공 파트너십은 공공 서비스의 거의 모든 측면에서 나타나고 있습니다. 대부분의 공공시설의 민영화와 아웃소싱 추세에 따라 선진 국가의 많은 공공 부문에서의 서비스를 민간단체가 운영합니다. 파트너십의 기본이라고 생각되는 한 가지 측면은 공공 부문의 전문성 부족을 민간이 대체하여 지원하는 것입니다. 그러므로 효과적인 협력과 관리를 위한 체계적인 데이터 생산이 필요합니다. 민간 부문과의 협력을 통해 발생하는 사업 데이터, 의사소통 데이터 등 다양한 데이터가 발생하고, 이를 모은 빅데이터는 협업을 관리하는 핵심 도구로 사용될 수 있습니다.

● 전자 정부 데이터와 공공 서비스 전달

종이서류를 전자화일로 바꾸는 목적은 비용 절감입니다. 전자 정부는 업무의 전반적인 효율성을 높이고, 사기 및 오류를 줄일 수도 있습니다(Laney 2014). 이러한 대민 데이터를 수집하여 빅데이터 분석을 통해 공공서비스에 대한 시민의 반응을 모니터링한 맞춤형 서비스를 제공할 수 있습니다. 전자화일로 생성되는 서류는 정부 투명성을 높이는 유용한 데이터이기도 합니다.

● 대중의 참여(크라우드 소싱)을 통한 데이터 생산과 활용

도시 및 지역 설계 시 기획 단계에서 대중을 참여시키는 방법은 선진국에서는 제도화되어 있습니다. 가령, 예전 철길 자리의 활용방법에 대한 의견을 구하기 위해 대중의 의견을 이용하였고, 현재 시민이 즐겨 방문하는 문화공간으로 활용되고 있는 사례도 있습니다.

크라우드 소싱(Crowdsourcing)은 어려운 문제를 해결하기 위한 아이디어와 솔루션을

구하는데 도움이 되는 웹 기반 모델로 부상했습니다. 클라우드 소싱이란 대중(crowd)와 아웃소싱(outsourcing)의 합성어로 대중들의 참여로 해결책을 얻는 방법입니다. 시민은 지역 사회에 공헌 할 기회 및 문제 해결을 위한 대화의 일부가 되고 싶어 하는 이타적인 이유로 참여하기로 결정한다고 합니다(Evans-Cowley 2011). Wikipedia (2018)는 데이터와 정보가 온라인으로 협업 방식으로 어떻게 구성되고 해석되는지를 보여주는 가장 좋은 예입니다. 끊임없이 진화하는 논쟁을 통해 각 주제에 대한 이해를 제공합니다. 이처럼 클라우드 소싱은 시민 참여 및 투명성을 창출하는 새로운 길을 열어줄 것입니다.

사회적 문제가 발견되면 시민들은 온라인으로 실시간으로 전송하고, 반영 여부에 대한 추적이 가능합니다. 과거에는 이와 같은 문제를 보고하기 위해 국회에 전화하거나 편지를 쓰는 것과 같은 다른 방법을 사용하여 할 수 있었습니다. 차이점은 제기한 문제에 대한 대응의 직접성과 투명성, 그리고 비회피성일 것입니다. 시민은 문제를 발견하면 앱 혹은 페이스 북에 로그인하여 보고할 수 있어 시간이 많이 걸리지 않습니다. 이 과정은 우리의 일상생활에 녹아들게 됩니다. 이제 스마트폰을 가지고 있는 모든 사람들은 돌아다니면서 시민 문제를 신고 할 수 있습니다. 직접적인 피드백을 통해 공통의 이익에 기여한 만족감과 성취감을 느낄 수 있습니다. 시민들이 축적한 정보를 이용하는 것은 공공서비스에만 국한하지 않습니다. 시민 기사, 시간 기상관측 등의 정보를 활용할 수 있습니다. 그리고 지능형 사물인터넷과의 결합도 확산될 것입니다.

● 사물인터넷(Internet of Things, IoTs)을 통한 데이터 생산과 활용

IoT는 주변 환경 정보를 측정하는 지능형 시스템으로 일반적으로 인식됩니다(Zhang and Mitton 2011). IoT는 사물을 유무선 통신망으로 연결하고 센서에서 발생하는 실시간 데이터를 자동으로 주고받는 환경으로 유틸리티, 스마트 홈, 의료, 웰빙, 운송 등 다양한 영역으로 확산되고 있습니다. IoT가 우리의 삶을 어떻게 변화시킬지에 대한 사회 경제적 함의는 거대할 수도 있습니다. IoT를 분산시켜 공동으로 작동하게 하는 조직이 필요합니다(Bentley et al., 2014). IoT를 통해 고도로 네트워크화 된 도시 인프라는 스마트한 도시를 만들 수 있습니다. 향후 스마트 시티의 센서는 학습하고, 암기하고, 관련 경험을 기억하고, 가르치고, 최종 가치 판단을 내릴 수 있도록 훈련받을 수 있을 것입니다.

3 >> 4차 산업혁명에 따른 미래 일자리의 변화

4차 산업의 변화는 현명하게 관리된다면 새로운 일, 좋은 직업 및 모든 삶의 질 향상으로 이어질 수 있지만 관리가 제대로 이루어지지 않으면 기술 격차가 커지고 불평등이 커지고 양극화가 확대될 위험이 있습니다. 4차 산업이 진행됨에 따라 기업들은 새로운 수준의 생산 및 소비 효율을 달성하고 새로운 시장으로 진출하며 경쟁하기 위해 새롭고 신기한 기술을 활용하고자 합니다. 새로운 서비스와 제품은 보다 더 데이터에 기반하고 있습니다. 4차 산업에 대응하기 위해 기업의 리더는 포괄적인 인력 구성 전략을 수립할 필요에 직면하고 있습니다. 인력 대체가 가능해지고, 그 속도가 가속화됨에 따라 이러한 변화를 사전에 관리할 수 있도록 정부와 기업은 글로벌 노동 시장에 대한 새로운 비전을 사전에 계획하고 구현해야 합니다. 그렇지 않으면 변화의 과정에서 발생하는 기회는 빠르게 끝날 것입니다.

세계 경제 포럼 조사는 향후 5년의 노동 전망을 조사하였습니다(World Economic Forum, 2018). 이는 향후 5년에 걸쳐 2018-2022 기간을 다루었습니다. 조사는 정책입안자, 기업 리더, 학자, 싱크 탱크, 전략 컨설턴트 등으로 구성하여 미래 일자리의 변화에 대한 의견을 포함합니다.⁴³⁾ WEF는 이 보고서를 통해 변화의 원동력을 초고속 모바일 인터넷, 인공지능, 빅데이터, 클라우드 등이라고 밝혔습니다. 데이터는 이 모든 기술의 기반에 있고, 이처럼 데이터를 기반으로 한 신기술들은 혁신을 창출하고 있습니다.

결론적으로 WEF는 기존 직무를 보완하거나 강화하기 위해 자동화 기술을 접목할 것을 제안합니다. 그리고 궁극적으로는 근로자의 역량을 최대한 강화하고 도출할 것을 제안합니다. 자동화를 통한 임금 비용 감소라는 협소한 측면만 바라보기 보다는 가치창출 활용이라는 새로운 영역확장의 측면으로 바라볼 것을 제안합니다. 그 가치창출은 반복 업무에서 벗어난 시간적 여유가 있을 때 개인 고유의 능력과 신기술의 보완을 통한 인간 근로자가 만들어내는 것입니다. 그러나 이러한 긍정적인 비전을 달성하기 위해서는 근로자는 미래의 직장에서 성공할 수 있는 적절한 기술과 자신의 삶 전체에서 재훈련할 수 있는 능력이

43) 설문 조사 지역은 아르헨티나, 호주, 브라질, 중국, 프랑스, 독일, 인도, 인도네시아, 일본, 멕시코 등 20개 선진국과 신흥 경제국에 대한 기업 차원의 대응 방안에 초점을 맞춘 다양한 표본을 기반으로 합니다.

필요합니다. 건전한 회사 내에 평생 학습 시스템을 만들고, 인적 자원에 투자하고 인력 전략에 대한 다른 이해 관계자와 협력하는 것은 기업의 중장기 성장에 중요한 핵심 비즈니스 과제이며 사회에 대한 중요한 공헌입니다. 근로자들이 이전에 상상하지 못한 미래로 옮겨 갈 때 민첩한 학습의 사고방식도 필요합니다. 마지막으로, 정책 입안자, 감독자 및 교육자는 실직 상태의 사람들이 새로운 기술을 습득하고 적응하도록 교육 개선에 집중적으로 투자하도록 돕는 역할을 수행해야 합니다.

- 기술 채택 가속화 : 응답자의 85 %는 2022년까지 사용자 및 단체가 빅 데이터 분석을 채택할 가능성이 높다고 예견하였습니다. 많은 회사들이 사물 인터넷, 앱과 같은 기술 채택을 확대할 가능성이 높으며, 클라우드 컴퓨팅을 광범위하게 사용합니다. 기계 학습 및 가상현실은 상당한 비즈니스 투자를 유치할 것입니다.
- 로봇화 추세 : 휴머노이드 로봇의 사용은 2018-2022 년 동안 다소 제한적으로 남아 있겠지만, 고정 로봇, 비상용 로봇, 육상 로봇, 완전 자동 드론, 기계 학습 알고리즘, 인공 지능은 중요하게 채택될 것으로 예상되고 있습니다. 로봇 채택률은 업종별로 크게 다르며, 대부분의 제조 기업이 고정식 로봇을 채택할 가능성이 가장 높지만, 석유 및 가스 산업계는 고정식 및 공중과 수중 로봇에 대한 수요를 똑같이 예상하고 있습니다. 금융 서비스 산업은 2022 년까지 휴머노이드 로봇을 채택할 가능성이 가장 큼니다.
- 생산, 유통의 지형 변화 : 기업은 생산 및 배포를 위해 지리적 위치를 중요하게 생각했습니다. 위치와 관련하여 기업들은 현지에서의 숙련된 인재의 가용성을 압도적으로 우선시 하였으며, 노동 비용을 주요 관심사로 꼽았습니다. 하지만, 현지 노동법의 유연성, 산업 응집 효과, 원재료 접근성과 같은 요인들의 중요성은 낮게 나타났습니다. 인건비 민감도와 관련하여 산업별로는 차이가 있습니다. 예를 들어, 자동차, 우주 항공, 운송, 화학, 신소재, 생명 공학, 건강 및 보진, 정보 통신 기술 산업은 숙련된 현지 인재의 가용성에 우선순위를 두는 경향이 있습니다. 대조적으로 일반 소비재, 에너지 유틸리티 및 기술, 금융 서비스 및 투자자, 채광 및 금속 분야는 노동 비용을 강조하는 경향이 높은 산업에 속합니다.

- 고용 유형 변경 : 절반 이상의 고용주는 자동화가 전임 인력을 일부 감축할 것으로 예상합니다. 그러나 조사 대상 기업 중 38%는 인력을 생산성 향상의 역할로 확대할 것을 계획합니다. 1/4 기업의 고용주는 자동화 시 내부에서 새로운 역할이 창출될 것으로 기대합니다. 또한 기업들은 과업기반의 계약직의 사용을 확대할 계획이며, 물리적인 사무실 이외의 원격 근무 및 운영의 분산화를 통해 보다 유연한 방식으로 근로자를 고용 할 의사가 있는 것으로 나타났습니다.
- 기존 작업 내에서 기계의 확대 : 2018년에는 전체 작업 시간의 평균 71%는 사람이 수행하며 기계는 29%를 수행하지만, 2022년에는 인간에 의한 작업은 58%이고 기계에 의한 작업은 42%로 변할 것으로 예상됩니다. 예를 들어, 2022 년까지 조직의 정보 및 데이터 처리, 정보 검색 및 전송 작업은 현재 42%에서 앞으로 62%가 기계에 의해 수행될 것으로 예상됩니다. 심지어 지금까지 압도적으로 사람에 의해 수행되던 인간과 의사소통이 필요한 작업 (23 % -> 30%), 관리 및 조언 (20 %-> 29%), 추론 및 의사 결정 (18 %->27%)은 자동화되기 시작합니다. 추론, 의사 결정, 관리, 정보 검색 및 수신 분야에서 자동화의 증가가 특히 두드러질 것으로 예상됩니다.
- 총 일자리 수에 대한 긍정적인 전망 : 새로운 일자리가 감소하는 일자리를 상쇄할 것으로 예상됩니다. 2022년까지 모든 산업에서 일자리 공유는 16%에서 27% 로 증가할 것으로 예상됩니다. 산업 전반에 걸쳐 오늘날의 핵심 일자리의 약 절반은 2022 년까지 안정적으로 유지 될 것입니다. 98만 일자리는 쇠퇴하지만, 오히려 174 만 건의 일자리는 더 확대될 것으로 예상되었습니다.
- 요구되는 역할: 2022년까지 수요가 증가하는 직업의 범위에는 데이터 분석가 및 과학자, 소프트웨어 및 응용 프로그램 개발자, 전자 상거래 및 소셜 미디어 전문가 등입니다. 또한 고객 서비스 근로자, 영업 및 마케팅 전문가, 교육 및 개발, 인력 및 문화, 조직 개발 전문가 및 혁신 관리자와 같이 인적기술을 차별적으로 활용하는 역할이 커질 것으로 예상됩니다. 최신 신기술을 이해하고 활용한 새로운 전문 역할에 대한 다양한 요구가 증가하고 있습니다. 예를 들면, AI 및 기계 학습 전문가, 빅 데이터 전문가, 프로세스 자동화 전문가, 정보 보안 분석가, 사용자 경험 분석, 기계 상호 작용 디자인

너, 로봇틱스 엔지니어 및 블록체인 전문가 등의 직업입니다.

- 기술 변화의 증가: 근로자와 기계간의 분업의 변화를 감안할 때 대다수의 고용주는 2022년까지 현재 기술을 이용한 일자리의 대부분이 바뀔 것으로 예상합니다. 동일하게 유지 될 수 있는 기술의 비율은 약 58% 일 것으로 예상됩니다. 이는 2018-2022년 기간 동안 인력 기술에서 평균 42%의 변화를 의미합니다. 새로운 기술은 비즈니스 성장, 고용 창출 및 전문 기술에 대한 수요를 이끌 수 있지만, 특정 업무가 쓸모없게 되거나 자동화되면 전체 기존기술과 직업을 대신 할 수도 있습니다.
- 직원의 재교육 필요성: 2022년까지 모든 직원의 54% 이상이 상당한 재교육을 필요로 할 것으로 나타났습니다. 이 중 약 35%는 최대 6개월의 추가 교육이 필요하며, 9%는 6개월에서 12개월 동안 지속되는 재개발이 필요하고, 10%는 1년 이상 추가 기술 교육이 필요합니다. 2022년까지 요구되는 기술은 기술 설계 및 프로그래밍과 같은 기술의 중요성이 급격히 증가합니다. 창의력, 독창성, 주도권, 비판적 사고, 설득 및 협상과 같은 인간적 기술 또한 마찬가지로 중요하게 여겨지며 일자리가 증가할 것입니다. 서비스 측면 뿐만 아니라, 유연성 및 복잡한 문제 해결 능력, 정서적 지능, 리더십 및 사회적 영향력에 대한 능력도 요구되고 있습니다.
- 기술 격차 해소를 위한 현재 전략 : 기업은 신기술과의 기술 격차를 줄이기 위해 세 가지 전략을 강조합니다. 이는 신기술과 관련된 기술을 소지하고 있는 새로운 직원을 고용하고, 자동화가 가능한 업무를 완전히 자동화하고, 기존 직원을 재교육하는 것입니다. 관련 기술을 갖춘 신규 직원을 고용 할 가능성은 기존 직원을 재교육할 가능성보다 거의 두 배입니다. 약 1/4 기업이 기존 직원의 재교육이 쉽지 않다고 보고 있으며, 2/3 기업은 근로자가 변화하는 직무에 적응하고 신기술을 습득 할 것으로 기대합니다. 기술격차 해소를 위해 절반에서 70%가량은 외부 계약자, 임시 직원, 프리랜서를 채용할 가능성이 있습니다. 근로자의 기술 격차(지도자의 기술 격차를 포함)는 일부 경우 자동화에 대한 경향을 가속화 할 수 있지만 새로운 기술 채택에 장벽이 될 수 있으며 비즈니스 성장을 저해 할 수 있습니다. 신기술, 직업, 기존기술 사이에는 복잡한 피드백 루프가 있기 때문입니다.

- 불충분한 재교육 및 기술향상 : 대략 절반의 기업은 기업의 전략적 역량을 강화하는 방법으로 현재 고부가가치 역할을 수행하고 있는 직원에 대해 재교육 및 역량 강화 노력을 우선적으로 집중하는 것으로 나타났습니다. 고용주의 41%는 높은 성과를 내는 직원에게 재교육을 집중시키는 반면, 33%는 저성과 직원에게 재교육을 지원하겠다고 했습니다. 즉 재 숙달과 숙련이 필요한 현재 저성과의 사람들은 재훈련을 받을 가능성은 상대적으로 낮습니다.

4 » 데이터 혁신 기술이 기업 일자리 창출에 미치는 영향

데이터 혁신이 경제 산업에 미치는 효과를 살펴보고자 합니다. 가장 큰 영향을 미치는 분야는 지식 집약형 서비스, 수상 운송, 항공 운송, 법률 등의 부문입니다. 경영 컨설팅, 건축 및 엔지니어링, 성능 점검, 광고 및 시장 조사, 과학적 및 기술적 분야도 포함됩니다. 첨단 지식 집약형 서비스에는 텔레비전, 음악 및 영화 산업, 통신, 상업용 R&D 서비스, 정보 통신 기술 (ICT) 서비스가 포함됩니다. 데이터 관련 산업 혁신은 다음과 같은 특징을 가집니다(Dachs, 2018).

- (1) 정보, 상품 및 서비스의 디지털화는 복제를 쉽게 함.
- (2) 통신 및 운송의 향상과 세계 경제 통합으로 인해 기업이 잠재적으로 도달 할 수 있는 소비자의 수가 증가함.
- (3) 표준화 및 네트워크 효과가 점차 중요해짐.

4차 산업이 산업별로 미치는 영향을 살펴봅니다. 4차 산업 시대에 크게 확대될 것으로 예상되는 분야는 농업/어업/산림업, 예술/디자인/엔터테인먼트 입니다. 이는 인간밖에 할 수 없는 분야로 인식되기 때문입니다. 또한 여가활용이 확대되기 때문에 이러한 분야의 일자리가 크게 증가할 것으로 예상됩니다. 확대될 것으로 예상되는 분야는 경영, 건설,

교육/훈련, 기술 설계입니다. 자율자동차를 위한 도로교통표지판 변경, 데이터 생산을 위한 센서/전자태그 확대 등 다양한 분야에서 건설 및 철거가 확대될 것입니다. 다소 감소할 것으로 예상되는 분야는 마케팅/판매, 운송/물류, 금융/법률/지식서비스, 제조/생산입니다. 인공지능은 지식을 보다 효과적으로 관리하고 제공하여 금융/법률/지식서비스 전문가를 대체할 것입니다. 크게 감소할 것으로 예상되는 분야는 사무직, 컴퓨터 프로그래머, 설비/유지관리 관련입니다. 컴퓨터 프로그래머 분야는 클라우드, 알고리즘을 통해 모듈화 등의 신기술이 확대되어 개발 및 시스템 유지관리 인력은 감소할 것입니다.

● 신제품 출시를 통한 수요 영향

기업의 혁신에 대한 중요한 영향은 신제품 출시를 통한 수요 효과입니다. 기업이 출시한 신제품이 소비자에게 유용하다고 판단되면 회사에 새로운 수요가 창출됩니다. 수요 증가는 기술 혁신에서 고용으로 이어지는 가장 중요한 긍정적 결과라고 볼 수 있습니다. 신제품으로 인한 고용 창출에도 불구하고 4차 산업의 논의에서 이러한 측면의 장점에 대한 논의는 미미합니다. 이것은 미래를 상상할 수 있는 능력의 제약으로, 미래의 주요 제품 혁신과 일자리 창출에 미치는 영향을 예측하기에 한계가 있기 때문일 것입니다. 하지만, 프로세스 혁신으로 인한 고용 손실은 우리 모두가 잘 알고 있는 제품과 직업의 변화에 적용되기 때문에 오늘날의 관점에서 훨씬 더 분명하게 보입니다. 기술혁신을 통한 생산 단가의 하락 또한 기업의 입장에서는 일자리의 창출을 가져옵니다. 기업이 프로세스 혁신을 통해 제품 가격을 낮춤으로써 생산 비용의 감소를 소비자에게 되돌려 준다면 전체 시장은 확장되어 결과적으로 고용 성장으로 이어질 수 있습니다. 제품 혁신에 있어 회사 내의 엔지니어링, 디자인, 멀티미디어, 그래픽아트, 소프트웨어 개발, 마케팅 연구 등과 같은 다양한 기술이 복합되어 혁신 활동에 긍정적인 영향을 미친다는 것을 보여 줍니다(Peters 2016). 이러한 근무환경을 제공하는 회사가 혁신을 이끌어낼 확률이 보다 높습니다. 고용과 관련된 기존 많은 연구조사에서는 기술, 경험, 창의력 등의 측면에서 사람 간에 차이가 없고 노동의 질은 동일하다고 가정하곤 하지만 이것은 매우 단순한 가정입니다.

● 프로세스혁신이 미치는 영향

신제품 및 프로세스 혁신에 대한 요구는 업체의 유형을 변화시킵니다. 4차 산업의 기술을 활용하여 혁신을 이끌 기업은 기존의 제조기술을 기반으로 하는 기업과는 다른 기술을 요구합니다. 또한 도래하는 신기술은 사람들이 일하는 방식과 미래의 일하는 조직을 바꿀 수 있습니다. 4차 산업을 이용한 새로운 생산 설비의 설치는 프로세스 혁신으로 이어지며, 기업이 노동력을 줄임으로써 동일한 생산량을 산출할 수 있게 합니다. 결과적으로 프로세스 혁신은 막대한 자본이 투입되기 때문에 생산성도 낮추고 고용에 부정적인 영향으로 이어 집니다(Edquist et al., 2001). 프로세스 혁신이 미래 노동시장에 부정적인 영향을 미칠 것이라고 예측되지만 이에 대한 증거는 모호합니다. 일부 연구(Van Reenen, 1997; Entorf and Pohlmeier, 1990)에서는 중요하지 않는 것으로 나타났지만, 또 다른 연구(Blechinger and Pfeiffer, 1999)에서는 오히려 더 많은 고용을 창출하는 것으로 연구되고 있습니다.

생산성 향상을 위한 프로세스 혁신은 서비스보다 제조분야에서 높게 나타납니다(Baumol, 2012). 서비스 회사의 혁신과 제조 회사의 혁신에 동일한 측정 기준을 사용할 수는 없습니다. 서비스의 경우 비 기술적인 경우가 많고 과학적이고 공식적인 지식을 적용하는 정도가 더 낮습니다. 반면에 고객과의 상호 작용이 더 필요하고 상호 작용에 필요한 사회적 기술이 더 많이 필요합니다. 이 때문에 제조업분야에서 보다는 공식적인 지적 재산권에 대한 보호가 쉽지 않습니다. 많은 경우 서비스분야는 사용자와 생산자 상호 작용의 결과이기 때문에, 서비스 제품은 제조 제품보다 훨씬 덜 표준화되어 있습니다(Tether 외 2001, Miles 2005). 따라서 서비스 분야는 제조업의 혁신에 비해 생산성을 높이기 위한 프로세스 혁신 기회가 적습니다.

● 다연결 맞춤 생산으로의 변화에 따른 영향

제품개발 기업의 측면에서는 기존에 제공하던 제품 및 서비스를 파악하여 제품 및 업무의 혁신을 통해 경쟁력을 확보하고 이에 따른 일자리 창출이 필요합니다. 기업은 다품종 소량생산에서 기회 선점을 위해서는 다연결 맞춤 생산으로 변해야 한다. 이를 위해서는

제품의 연결을 개발해야 하며, 기존의 제품을 인공지능과 신기술로 재설계해야 하고, 이에 따른 고객의 경험 또한 재창조해야 합니다. 이를 원활히 수행하기 위해서는 기존의 제품개발자의 경험과 아이디어가 필요하다. 이를 위해 근로자의 재교육이 수행되어야 합니다.

● 데이터를 이용한 신산업 기업 도래에 따른 영향

데이터를 이용하여 새로운 산업에 참여하는 기업의 측면에서 살펴봅니다. 새롭게 만들어지는 공공 및 민간의 데이터를 이용하여 새로운 서비스를 개발하는 기업은 새로운 일 자리를 창출합니다. 이를 지원하기 위해 정부는 양질의 데이터 생산을 정부 차원에서 확대하여야 합니다. 양질의 데이터를 생산하는 과정은 데이터 생산을 위한 공공일자리 뿐만 아니라, 이를 이용한 기업의 서비스 제공 과정에서 민간일자리도 생산됩니다.

● 데이터 과학 육성을 위한 지원 영향

데이터 기반 혁신은 모든 산업에서 발생하므로 국가는 이미 성공한 산업에 데이터 과학을 적용하는 데 집중해야 합니다. 정부는 기업의 충분한 수요가 없는 한 대규모 데이터 과학 인력을 유치할 수 없을 것입니다. 그럼에도 불구하고 이러한 장애물을 극복하기 위해 정부가 노력해야 할 것들이 있습니다. 예를 들어 국가 경제를 다각화하고 연구자를 유치하며 해당 산업 분야에서 기업이 정신을 장려 할 수 있도록 국립 지방 대학과 협력하여 경제 개발 프로그램을 운영할 수 있습니다. 또한 데이터 과학 및 개방형 데이터 네트워킹 그룹, 대회를 개최하고 후원할 수 있습니다. 예를 들어, 2013년에 일리노이 과학 기술 연합은 수천 페이지에 달하는 도시 조례를 수립하고, 공공 데이터를 이용하여 지역별 주택 개발 분석하고 순위를 매기는 앱과 같이 실용적인 앱을 만든 개발자에게 각각 1만 5천 달러를 수여했습니다.

● 환경 중요성에 대한 인식 변화의 영향

또 다른 관련 분야는 '그린 잡 (Green Jobs)'과 환경 기술의 고용 효과에 관한 연구입니다. OECD(2012)의 연구는 저탄소 경제로의 변화에 따른 고용 효과에 대한 변화를 추정한

바 있습니다. 녹색 성장으로의 전환은 태양광 및 풍력 전기 장비 제조업체에는 고용 증가로 이어지지만, 반면에 석유 및 석탄 채굴 분야에서의 고용은 감소가 예상됩니다. 제품 혁신과 프로세스 혁신의 과정에서 환경적 요소를 고려하여야 하며, 환경 분야 자체에서 새롭게 창출되는 일자리는 증가할 것입니다. 환경 분야와 관련한 일자리의 상당수는 IoTs 기술을 활용할 가능성이 높습니다.



노동시장의 변화와 사회적 타협

1 » 신기술이 미치는 노동시장의 변화

신기술이 고용에 미치는 영향에 비관적인 견해가 많지만, 관련 자료를 살펴보면 낙관적이라는 것을 알 수 있습니다. 기술의 혁신과 창조적 파괴가 적어도 장기적으로는 경제 성장과 새로운 고용의 주요 동인이라고 볼 수 있습니다. 수많은 과학적 발전은 혁신적 파괴가 수반되지만 새로운 고용을 창출함을 보여줍니다. 새로운 제품을 통한 일자리 창출과 새로운 공정 기술로 인한 일자리 파괴 사이의 경쟁은 그동안 새로운 일자리 창출 효과에 의해 상쇄되어 왔습니다. 물론 4차 산업 혁명이 것처럼 새로운 일자리를 가져다 줄 것이라는 긍정적 기대를 보장 할 수는 없습니다. 긍정적 기대를 하기 어려운 이유는 새로운 기술 패러다임에 따라 창출 될 일자리 유형에 대한 충분한 지식과 상상력이 부족하기 때문에 대부분의 사람들은 근본적인 기술 변화의 일자리 창출 가능성을 과소평가하는 경향이 강한 탓이기도 합니다. 기술 혁신, 세계화 및 인구 통계학적 변화는 국내 노동 시장에 새로운 도전을 가져 왔습니다. 한국 경제가 장기적으로 성장할 것인지 여부는 노동 시장이 변화하는 환경에 성공적으로 적응할 수 있는지 여부에 달려 있습니다(Kuroda, 2016). 실제로 노동 시장 개혁은 정부 정책의 중요한 부분입니다. 고령화와 저 출산율 추세를 감안할 때, 장기적인 경제 성장을 위해서는 노동시장에서의 다양한 일자리 증가와 노동 생산성의 증가가 필수적인 요소라고 할 수 있습니다.

일반적으로 생명 공학, 첨단 소재, 건강 응용 분야에서는 4차 산업의 경제적 중요성을 인식하지만, 고용효과를 직접적으로 다루는 연구는 많이 진행되고 있지 않습니다. 4차 산업이 확대되어 감에 따라 일부 직업은 축소될 수 있습니다. 예를 들어, 회계사의 업무는 대체될 위험에 처해질 수 있습니다. 현재 기술로도 회계 표준에 따라 송장을 정확하게

인식하고 기록한다면 회계사의 역할이 감소할 수 있기 때문입니다. 자동화로 인해 대체될 위험이 높은 직업은 데이터 입력자, 도서관 기술자, 신규 회계 담당자, 사진 공정 작업자, 기계 운영자, 세금 설계사, 화물, 시계 수선공, 보험업자, 수학 기술자, 텔레마케터 등입니다(Dachs, 2018). 한국고용정보원(2016년)에 따르면 국내 주요 400개 직업 가운데 직업 대체 확률을 분석한 결과 플라스틱 제품 조립, 청원경찰, 조세 행정 사무원이 자동화로 대체될 가능성이 높은 것으로 나타났습니다.

● 일상적 직무에 대한 수요 감소

Acemoglu와 Autor(2011)는 일상적인 직무에 대한 수요가 상당히 감소한다고 밝혔습니다. 결과적으로 중숙련 인력에 대한 수요가 감소한 반면, 고숙련 인력과 저숙련 인력에 대한 수요가 증가했습니다. 이러한 경향은 '직업 양극화'로 볼 수 있습니다. 기술 성장의 결과로 숙련도간 불평등이 증가하는 것에 대처할 방안을 찾는 것은 미래의 새로운 해결해야 할 문제입니다. 기술 편향과 일상적인 직무에 관련된 기술 변화는 불평등을 증가시킬 수 있는 두 가지 원인입니다. 그들은 특정 그룹의 노동력을 발전시키고 다른 그룹의 고용 가능성을 줄이기 때문입니다. 특히 이미 조정 비용의 상당 부분을 차지하고 있는 저 숙련 노동자들의 일자리가 사라질 위험성이 더 높고 자격이 낮은 일자리 수가 줄어들 것이기 때문입니다.

● 비일상적 직무의 일자리 증가

사회는 필요한 기술이 부족하거나 일자리를 찾지 못한 많은 사람들에게 경제의 다른 서비스 분야에서 일자리를 제공합니다. 낮은 기술 직업의 고용 시장에서 구직자의 증가는 사회 경제에 부담을 줄 것입니다. 그러나 건강, 사회사업, 교육 및 ICT로 대체하기 어려운 다른 분야에서 많은 비일상적 업무의 일자리가 증가할 것입니다. 또한, 새로운 4차 산업의 신기술은 사회 경제 정책보다 효율적인 방식으로 사회적 노동력과 개인의 능력을 조직할 수 있는 새로운 도구를 제공 할 수도 있습니다.

● 업무 방식의 변화

디지털 기술은 일의 양과 난이도가 다른 기술 수준에 대한 수요뿐만 아니라 업무 방식 또한 바꿀 것으로 예상됩니다. 사람들이 수행하는 개별 작업은 점점 인터넷을 통해 거래가 가능하게 될 것입니다. 전문가들은 회사 외부에서 수행하는 작업의 비율이 점차 증가하고 다양한 고객을 대상으로 프로젝트별로 작업하는 자영업자의 비율은 증가할 것입니다. 이러한 고용형태를 '플랫폼 근로자'라고 합니다. 결과적으로 회사는 점차적으로 고정된 근로자를 통해서가 아닌, 프로젝트 지향적 조직 구조로 이동하게 될 것입니다. 이러한 '플랫폼' 경제는 더 많은 자기 결정과 더 나은 일과 삶의 균형을 직원들에게 가져올 수 있지만, 다른 한편으로는 더 많은 불안감과 반복되는 비자발적 실업을 초래할 수도 있습니다. 더욱이 자영업 분야의 플랫폼 근로자는 법적 보호와 직원들이 받을 수 있는 다양한 사회적 혜택이 부족한 경우가 많습니다.

● 생산성 증가

생산성은 고용 및 생산량 증가에 관한 모든 논의의 핵심입니다. 첨단 기술 부문에서도 고용 성장의 상당 부분을 생산성 증가를 통해 설명해야 합니다(Peneder 2009, p.14). 생산성 증가는 혁신과 고용 성장 간의 관계입니다. 그러나 혁신, 생산성 및 고용 간의 연계성은 직접적이지 않으며 다양한 혁신 형태는 고용 성장에 다른 영향을 미칠 수 있습니다(Edquist et al., 2001). 고용 효과는 자원 효율성과 연관 지어 생각하는 경향이 있습니다. 왜냐하면 자원 수요의 감소는 경제적 활동이 적고 고용이 감소하는 것을 의미할 수 있기 때문입니다. 자원 수요의 감소가 고용의 감소를 반드시 가져오지는 않습니다. 가령 건물의 지속적인 개조는 고용 창출 기회를 증가시킬 것이기 때문입니다(Keivani et al., 2010).

● 제품, 프로세스, 마케팅 혁신에 따른 변화

기술 혁신과 고용 관계를 살펴보기 위해서는 기술 혁신을 제품 혁신과 프로세스 혁신으로 구분할 필요가 있습니다. OECD는 제품 혁신과 프로세스 혁신을 통합하여 기술 혁신으로 간주합니다(OECD 2005). 또한 기업은 조직 혁신이나 마케팅 혁신과 같은 비기술적

혁신에 투자 할 수도 있습니다. 프로세스 혁신은 기술적 변화를 통해 개선된 생산 공정의 변화를 의미하지만, 조직 혁신은 새로운 비즈니스 프로세스의 도입과 새로운 근로 환경을 포함하여 새로운 조직의 구성 방법입니다(OECD 2005). 마케팅 혁신은 제품 디자인이나 포장, 제품 배치, 제품 판촉, 가격 책정을 포함합니다(OECD 2005). 마케팅 혁신과 제품 혁신의 살펴보면 기능이나 사람들이 제품을 사용하는 방식이 바뀌면 제품 혁신이라고 할 수 있습니다. 기술 혁신과 비기술적 혁신은 모두 기업 혁신의 경제적 가치와 일자리 창출의 가능성에 영향을 주지만 이에 대한 구분은 별로 주의를 기울이지 않습니다. 조직 및 마케팅 혁신은 기술 혁신을 보완하고 잠재력을 최대한 발휘하는 데 도움이 됩니다(Brynjolfsson and Hitt 2000).

● 근로시간 단축

미래의 작업과 관련된 또 다른 주제는 근로시간 단축입니다. 새로운 기술과 그에 따른 생산성 증가로 과거의 노동 시간이 크게 단축되어 왔습니다. 20 세기 동안 미국과 유럽에서 근로시간이 거의 절반으로 단축되었습니다. 이는 실업률을 낮게 유지하고 여성의 노동 시장 참여를 확대시켰습니다. 역사적으로 새로운 기술로 인한 생산성 증가는 노동 시간의 단축으로 이어졌습니다. 디지털화를 통해 자동화가 더욱 확대될 때 노동시간의 축소를 기대할 수 있습니다. 근무 시간의 단축은 쉽게 구현 될 수 있으며, 단계적으로도 수행될 수 있습니다.

● 로봇의 소유가 개인별 생산성에 미치는 차이

노동시장에 로봇의 도입이 확대될 경우 로봇의 소유 여부가 개인이 생산성에 영향을 미칠 것입니다. 가령 과거 농경사회에서 소의 소유 여부는 농사에 큰 영향을 미쳤습니다. 로봇 소유에 대해 적절한 세금을 부과하기 위해서는 세금 징수 제도가 노동 과세에서 자본 과세로 옮겨갈 필요가 있습니다. 자본이 경제적 가치의 주된 기여자로서 가치가 부가 될 필요가 있습니다. 로봇 이용을 자본으로 간주하여 로봇 활용에 세금을 확대하면 로봇은 사회 보장 재정에 기여할 수 있습니다. 일부 사람들은 막대한 자본을 소유한 수퍼 부자에게

과세를 증가하는 것은 불평등을 줄이고 사회보장 기금을 모으는 방법이라고 생각합니다. 하지만 이는 매우 어려운 정치적 조정이 필요하기 때문에 현실성이 있다고 보기는 어렵습니다.

● 신규일자리 노동시장 활성화 노력

우리나라의 노동 시장이 기술 혁신의 발전에 어떻게 대응할 것인가 생각해 보아야 합니다. 최근의 로봇, 인공지능, 빅데이터와 같은 기술혁신이 근로자와 기계간의 대체가 우리가 지금까지 경험했던 것보다 넓은 범위의 영역으로 까지 확장 될 수 있다는 우려의 목소리가 있습니다. 이전에는 인간을 위해 독점적으로 여겨졌던 직업 분야에서도 마찬가지입니다. 역사는 기술 혁신이 장기적으로 경제적 번영을 가져온다는 것을 우리에게 가르쳐 줍니다. 기술의 진보는 새로운 상품과 서비스에 대한 수요뿐만 아니라, 수입과 구매력을 증가시킬 수 있습니다. 이러한 상황에서 우리는 기술 혁신의 혜택을 누리기 위해 새로운 일자리를 창출해 내는 노동시장을 활성화 시키는 노력을 게을리 해서는 안 됩니다. 우리나라 노동 시장은 재취업을 위한 프로그램과 실업자를 위한 사회 안전망으로 한 직업에서 다른 직업으로의 전환을 모색하는 근로자를 수용하는 새로운 관행을 받아들이도록 더욱 노력해야 합니다.

● 세계화에 따른 노동시장의 변화

세계화가 가속화됨에 따라 노동 시장이 어떻게 변할 것인가를 생각해 보아야 합니다. 세계화는 종종 혁신과 함께 진행되기 때문에 노동 생산성의 향상을 가속화할 것입니다. 뿐만 아니라 해외에 있는 한국 근로자의 증가와 한국에 유입된 외국인 노동자를 증가시킬 것입니다. 기업은 국경을 넘는 협력을 통해 새로운 성장 기회를 찾고 기술 혁신을 통해 노동 생산성을 높이려 할 것입니다. 이 과정 속에서 경쟁은 서로 간 더욱 치열해질 것이며 국내 근로자들은 훨씬 더 광범위한 산업과 전문직 분야에서조차도 해외 경쟁자들을 대면하게 될 것입니다. 재화와 서비스뿐만 아니라 근로자도 세계화의 영향에 노출되는 것입니다. 노동 시장의 관습과 제도는 필연적으로 세계 경쟁에서의 압력에 더욱 노출될 것입니다.

● 인구변화에 따른 노동시장의 변화

인구 변화에 대응하여 노동 시장이 어떻게 변화 할 것인가는 이미 오랫동안 고민해 왔습니다. 노동력 참여를 장려하는 정책은 잠재 성장률을 높이기 위해 필수적입니다. 고령자와 여성의 노동시장 참여 증진은 중요합니다. 이러한 정책들이 효과를 발휘하기 위해서는 근로 형태와 고용 및 임금 구조의 규범에 관한 수정이 필요합니다. 평생 고용이나 연공서열 임금 체계의 유지가 어렵다는 것을 우리 사회는 어느 정도 인식하고 있습니다. 노동시장 참여를 확대하기 위해서는 육아 및 간병과 관련된 책임에서도 변화가 필요합니다.

2 » 신기술이 미치는 복지 사회적 변화

데이터 혁신과 함께 살펴보아야 할 것은 변화하는 사회입니다(Morabito, 2015). 사회의 변화를 함께 파악해야 적절한 대응 방안을 찾을 수 있습니다.

● 정부의 역할에 대한 기대의 변화

현대 사회에서 시민과 정부의 관계는 서비스 개념을 기반으로 합니다. 시민은 세금을 내고 이에 대한 교환으로 건강, 교육, 도로 유지 보수 등 다양한 사회서비스를 정부로부터 제공받는 것이었습니다. 하지만 앞으로는 정부와 시민 간의 파트너십이 확대될 것입니다. 즉 공공 서비스 결과물의 품질은 정부에 달려 있다기 보다는 정부와 시민 모두에게 달려있다는 개념입니다. 시민은 정부를 그리고 정부는 시민은 서로 돌볼 필요가 있다는 의미입니다(McDermott 2010). 소셜 미디어와 스마트 폰은 시민과 정부 간의 상호 작용을 촉진하고 있습니다. 또한 필요에 따라 지역 사회의 의사소통과 참여를 증폭시킬 수 있습니다.

예를 들어, 범죄와 싸우기 위해 시민들은 의심스러운 활동을 감시하고 경찰과 협력해야 합니다. Citysourced.com과 같은 응용 프로그램은 시민들이 낙서, 부서진 포장 도로 또는 가로등과 같은 모든 종류의 시민 문제를 지방 정부에 보고할 수 있도록 합니다. 사람들

은 문제점을 사진을 찍어 쉽게 업로드 하고 지도에서 위치가 파악되도록 할 수 있습니다. 시민이 제출한 보고서는 협의회로 보내지고 온라인으로 이슈에 대한 진행 상황을 추적할 수 있습니다(CitySourced Inc. 2018). 이것은 시민들이 지역의 감찰관 역할을 하도록 촉진시키는 방법의 전형적인 예입니다. 자선 단체와 이해 집단은 메시지를 증폭시키기 위해 협력합니다. 예를 들어, 자전거 타는 공동체는 움푹 들어간 곳에 큰 관심을 가지고 있기 때문에 이러한 신고를 촉진할 것입니다.

그리고 사회 변화와 기술 발전으로 인해 더욱 근본적인 변화가 일어날 수 있습니다. 고대 그리스의 직접 민주주의 유형과 마찬가지로 시민들이 사회 문제에 대한 결정권을 가질 것입니다(Nelson et al. 2015). 온라인 기술을 통해 정치적 의사 결정에 모든 이해 관계자를 참여시킬 수 있습니다. 이런 종류의 기술이 향후에는 글로벌 이슈를 논하는데 사용될 수도 있을 것입니다.

● 메가시티 간의 국제 경쟁

2011년부터 시골 지역보다 도시에 더 많은 사람들이 살고 있습니다. 전 세계적으로 인구 1,000 만 명이 넘는 메가시티가 1975년 5개에서 2006년에는 26개로 증가하고 2017년 46개로 증가하고 있습니다(United Nations Department of Economic and Social Affairs, 2006; Wikipedia, 2018). 대도시화는 지역 또는 국가의 문제가 아닙니다. 그들은 전 세계에서 국가 경제의 힘의 균형을 형성하고, 인구 이동성 및 재능의 구성에 영향을 미치고, 세계의 사회적 정치적 역동성에 영향을 미칠 것으로 전 세계의 미래 번영 및 안정성에 영향을 미칠 것입니다(United Nations Department of Economic and Social Affairs, 2006). 메가시티는 기능적으로 상징적인 역할을 합니다. 런던이 없는 영국, 서울이 없는 한국을 상상할 수 있을까요? 메가시티는 모든 면에서 사회 경제적 개발의 중요한 도구일 뿐만 아니라 한 지역의 사적 및 공공 부문의 사회 혁신을 도모합니다. 메가시티는 높은 수준의 생활, 더 나은 직업, 더 나은 교육 측면에서 삶의 질을 향상시키는 사람들에게 매력적인 곳으로 다가옵니다. 전 세계적으로 경쟁이 치열한 환경에서 메가시티는 글로벌 인재와 자본을 포함한 다양한 종류의 자원을 놓고 경쟁할 것입니다. 메가시티

는 도시 인프라 및 공공 정책에 압력을 가하는 생활 수준(예: 보안, 생활비, 이동성, 고용, 환경)의 향상에 대한 도전에 직면하게 됩니다. 메가시티는 높은 인구 증가율에도 불구하고 웰빙 지수를 통해 생활 수준을 높이는 방법에 대해 나름대로의 고민에 직면할 것입니다. 지능형 도시 인프라는 이러한 문제에 대한 방향을 제시할 것입니다. 생산되는 데이터의 양과 인프라 관리의 중요성은 높은 데이터 처리 기술을 요구할 것입니다.

● 기본 소득 제도

4차 산업의 확대에 의한 막대한 일자리 손실에 대한 비관적 시나리오와 디지털화에 기인하는 경제적 불평등에 대한 해결 방안으로 모든 시민에게 기본 소득을 제공하는 제도가 제안되고 있습니다. 하지만 이보다 앞서 심각한 경제적 불평등을 사전에 예방하고 사회 참여 수단을 제공하는 것에 우선적으로 초점을 맞추어야 합니다. 기본소득 제도는 많은 사람들이 고용과 기업가 정신에 불리한 인센티브를 제공한다고 생각하고 사회 복지 혜택에 따라 생활하는 문화를 창출하고 육성 할 수 있다고 생각하기 때문에 이 아이디어는 논쟁의 여지가 많습니다.

3 >> 신기술의 변화가 초래할 기술 불평등

불평등은 사회 불평등, 법 앞에서의 불평등, 건강과 사망률의 불평등과 같이 여러 분야에서 발생합니다. 현재의 논의는 주로 통화 불평등을 의미합니다. 통화 불평등은 소득과 부의 불평등을 의미합니다. 소득 불평등은 세전 및 세후의 개인 또는 가구의 소득의 차이를 의미합니다. 유의할 점은 고용 형태로서 파트타임 또는 풀타임, 다양한 기술 수준 등 추가 검토가 필요합니다.

기술의 불평등 또한 4차 산업의 미래사회에 매우 중요한 이슈가 되고 있습니다(Dachs, 2018). 기술의 편향은 불평등을 증가시킬 수 있습니다. 고급 기술로의 높은 장벽이 진입 가능성을 줄이기 때문입니다. MIT의 Erik Brynjolfsson은 기술의 변화가 최근의 불평등 증가의 주요 동인이라고 하였습니다. 많은 사람들이 불평등을 사회에서 큰 관심사로 생각하며 불평등 수준이 높아질 것을 염려합니다. 부의 불평등은 측정하기가 어렵습니다. 부의 불평등에 대한 자료 수집에서 주요 과제는 부동산, 부채, 부유 계층의 무응답률 차이, 사회 보장 청구의 변수 등을 고려해야 합니다. 대부분의 경우 부의 분포를 비교할 수 있는 데이터를 구하기 쉽지 않습니다. 불평등 척도를 직종이나 기술수준과 연계시켜본다면, 불평등의 정도가 현격히 나타나는 구간을 파악할 수 있을 것입니다.

혁신기술이 불평등을 증가시키는 방법들은 여러 가지가 있습니다. 첫째는 기술이나 작업의 차이입니다. 새로운 기술이 숙련 편향적이고 일상적인 수준의 기술을 중심으로 대체가 시작될 때 불평등은 더욱 커질 수 있습니다. 둘째로 혁신기술이 불평등을 증가시킬 수 있는 방법은 신기술의 독점이 발생하는 것입니다. 신기술이 전체 시장을 점유하는 한 개의 회사로 연결되는 Winner-take-all 시장은 일종의 자연 독점으로 간주 될 수 있습니다. 이럴 경우 불평등은 확산된다고 볼 수 있습니다. 혁신기술이 불평등을 초래하는 세 번째 방법은 비즈니스 도둑 효과입니다. 혁신은 오래된 제품과 서비스를 쓸모없게 만듭니다. 비즈니스 도둑 효과는 경제의 구조적 변화로 이어집니다. 자동차는 20세기 초반에 마력 수송을 불필요하게 만들었으며 오늘날 다른 분야에서도 이와 유사한 사례를 볼 수

있습니다. 이러한 감소 추세에 해당하는 분야의 실업자가 고용이 증가하는 부문으로 재할당 되지 못한다면 실업은 확대되어 사회의 불평등의 주요 원인이 될 수 있습니다. 이는 산업 간 기술 요구가 다르기 때문에 해결이 쉽지 않은 경우가 많습니다. 일부 경제학자들은 기술혁신으로부터의 불평등은 기술 진보에 대해 지불해야하는 가격이라고까지 주장합니다(Mankiw, 2013). 기업가들이 신제품을 도입하는 주된 동기는 시장 지배력과 독점이익을 얻기 위해서이기 때문입니다.

다른 한편으로 기술은 사람들의 시간과 재정을 자유롭게 함으로써 평등에 기여합니다. 기술혁신이 불평등을 초래한다는 주장이 있지만, 신기술이 불평등을 감소시킨 사례도 많이 있습니다. 기술의 역사에서 흔히 볼 수 있는 패턴은 과거에는 소수의 사람들만 사용이 가능했던 제품 및 서비스가 기술의 변화로 인해 보다 많은 사람들이 이용할 수 있게 되는 부분입니다. 가령, 세탁기는 여성으로부터 가사 일을 상당부분 해방 시켰기 때문에, 시간 절약이 평등에 기여한다는 사실을 보여주는 가장 좋은 예일 수 있습니다(Cardia, 2008). 제1차 산업 혁명 기간 동안 직물, 도자기 및 기타 가재도구의 대량 생산으로 인해 일반인도 사용이 가능해 졌습니다. 제2차 산업 혁명 기간에는 대중교통 및 전기의 사용이 가능해 졌습니다. 20세기에는 의학이 보다 넓은 사람들에게 제공되고 있습니다(Freeman and Soete 1997). 신기술이 평등을 창출하는 최근 사례로는 개발도상국의 휴대 전화입니다. 휴대 전화 서비스의 도입으로 통신비용이 절감되고 농산물 가격과 같은 중요한 정보에 대한 접근이 용이해 졌습니다. 휴대 전화는 농부와 상인 사이의 정보 비대칭을 줄여 농업 시장의 효율성을 개선하고 생산자와 소비자의 이익을 향상시킵니다(Aker and Mbiti, 2010).

기술적 변화가 창출하는 양질의 일자리에 대한 관심도 주요 이슈입니다. 고용의 질은 안정적이고 고임금인 양질의 일자리와 저임금의 불안정한 고용 조건을 질 낮은 일자리로 구별합니다. 고급 기술 일자리의 관점에서 보면 기술이 새로운 기술의 생산성을 향상시키기 때문에 미래에 결국 양질의 일자리를 창출 할 것이라는 점은 분명합니다. 하지만 일반적으로 전체적인 일자리의 관점에서 보면 차이가 있습니다. 기계화로 대체가 가능한 일상생활의 작업은 기술 혁신에 따라 얼마든지 대체가 가능합니다. 플랫폼 근로자, 운송 및 배달

서비스의 자영업자, 청소, 보건 서비스 등의 경우가 그러합니다. 데이터 혁신으로 인한 미래 사회는 오늘날의 업무 수행 방식과 취업 방식이 과거에 바뀌어 왔듯이 앞으로 크게 변화할 것으로 예상됩니다. 미래 사회에는 일이 무엇을 의미하는지에 대한 의문을 다시 생각해 보아야 할 수 있습니다. 디지털화, 플랫폼 작업 등과 관련된 문제에 대해 노동조합이나 기업에서도 통일 된 견해가 없습니다. 기업의 디지털 책임에 대해서도 생각할 때입니다.

불평등은 지역 간의 차이에서도 나타납니다. 수도권 대도시와 농촌 지역 간의 격차가 커짐에 따라 인구의 이동도 증가합니다. 지역 개발에 대한 정책의 딜레마도 커지고 있습니다. 지역균형의 측면에서 삶의 수준과 사회 인프라 확충을 균형 있게 하기 위해서는 정채된 지역에 투자를 확대해야 하는지, 수도권 지역의 미래 경쟁력을 더욱 강화하기 위해 대도시 에 투자를 확대해야 하는지 딜레마입니다. 빈곤한 지역과 부유한 지역 간의 격차가 커짐에 따라 지역 경제력이 바뀌고, 결국 정치력에 따라 대도시와 농촌 지역 간에 불균형이 확대될 수 있습니다.

불평등을 해소하기 위해 수퍼 리치에게 세금을 더욱 과세하자는 의견도 있습니다. 하지만 이는 정치 조율이 필요하기 때문에 당장 이를 실행하기란 불가능합니다. 자본이 노동시장보다 이동성이 크다는 것을 유념해야 합니다. 이러한 과세는 자본의 움직임으로 이어질 수 있습니다. 자본은 국가 간 장벽 없이 쉽게 이동하여 모국 과세에서 벗어나는 곳에서 은닉처를 찾을 것입니다. 이러한 딜레마에 대한 해결책은 국제 협력으로만 가능합니다. 수퍼 리치에 대한 과세 확대와 더불어 노동 세제에서 자본 세제로의 전환 문제 또한 어느 정도는 국가 간의 협력과 조정이 필요합니다.

최근 일부 사람들은 모든 시민에게 무조건 기본 소득을 도입 할 것을 제안합니다. 기본 소득은 모든 거주자 또는 시민이 취업 여부에 상관없이 정부로부터 일정 금액의 돈을 받는다는 것을 의미합니다. 이러한 정책을 도입한다고 하여도 시민들에게 사회 참여 수단을 빼앗아서는 안 됩니다. 기본 소득 지지자는 기본 소득 도입이 많은 사람들이 자동화로부터 얻는 이익을 재분배하는 건전한 접근법으로서 경제적 불평등을 예방하는 효과가 있다고 봅니다. 디지털화로 하여 승자가 모든 것을 차지하는 불균형으로 경제적 이익을 얻는 대기업과 그 소유주의 손에 소득이 집중되는 것을 반대합니다. 기본 소득 지지자들은 이것이

많은 사람들이 빈곤에 빠지지 않도록 하는 효과적인 조치가 되기를 희망합니다. 이들은 기본 소득이 사람들에게 교육 기회를 제공하여 노동력을 질을 향상시키는 효과도 있다고 주장합니다. 기본 소득에 대한 반대자들은 기본 소득에 대한 도덕적 주장이 있습니다. 사회 복지 혜택에 따라 생활하는 문화를 육성 할 수 있으며, 사람들은 노동의 대가로 급여를 받아야 한다고 주장합니다. 다음으로 반대자들은 기본 소득의 도입이 고용 및 기업이 정신에 불리한 인센티브를 제공한다고 생각합니다. 이는 고용 감소, 경제 활동 감소로 이어질 수 있으며 성장을 안정시키기보다는 경기 침체를 야기할 수도 있다고 주장합니다. 더욱이 그들은 모든 시민들에게 기본 소득을 지원하는 것은 국가재정난을 초래하여 엄청난 세금인상으로 이어질 수 있다고 염려합니다. 기본 소득은 노동 대신에 자본 이득과 자동화 및 디지털화의 이점에 초점을 맞추는 새로운 조세 정책을 수반해야 합니다.

이렇듯 기본소득에 대한 지지자와 반대자의 주장을 평가하는 것은 쉽지 않습니다. 왜냐하면 경험적 증거가 제한적이기 때문입니다. 실험이 어려운 이유는 기본소득이 어떤 국가에서 실행되면, 이 결정을 취소하는 것이 쉽지 않을 것으로 여겨지기 때문입니다. 따라서 당장은 기본 소득 도입보다 근로 시간의 단축과 같은 덜 혁신적인 조치가 더 현실적일 수 있습니다. 기본 소득에 대한 정치적 결정은 투표 인구의 요구에 의해 영향을 받을 것입니다.

디지털화를 막을 수는 없지만 교육, 인프라, R&D에 투자를 확대하는 정책을 펼쳐야 하고 노동법, 조세 및 사회 보장 정책을 디지털화에 맞도록 수정하여 다양한 변화의 기회를 받아들이는 필요가 있습니다. 무엇보다 가장 중요한 것은 디지털시대를 위협이 아니라 모든 시민들의 복지와 성장의 기회로 삼아야 한다는 것입니다.

4 《 공공 고용서비스의 역할

공공 고용 서비스(PES: public employment services)가 노동시장의 변화를 모두 효과적으로 조율할 것이라고 기대하는 것은 현실적이지 않습니다. 공공고용서비스의 역할을 현실적으로 바라보고 적절한 고용서비스 정책을 만들어 나가야 할 것입니다. 이를 위해 공공고용서비스의 역할과 최근 이슈에 대해 살펴보려고 합니다.

미시적 관점에서 볼 때 PES의 상담 및 모니터링 활동은 사회 경제에 효과가 있습니다. Kluve (2006)는 직업 탐색, 직업 클럽, 직업지도, 상담 및 모니터링, 제재는 PES가 일반적으로 수행하는 활동이며, PES가 사회에도 도움이 된다고 주장 합니다. 반면에 거시적 수준의 연구는 매칭 효율을 높이기 위한 공공 부문 개입의 유용성을 의심합니다. 복지의 관점에서 실업자의 구직 활동에 보조금을 지급하면 고용주의 검색 프로세스가 촉진되지만(긍정적 외부 효과), 다른 실업자의 구직 기회는 줄어듭니다(부정적 외부 효과). Petrongolo & Pissarides (2001)는 이러한 매칭 활동이 결과적으로 사회에 미치는 영향이 동일하기 때문에 공공이 민간 기업의 구직 활동에 개입하는 것은 사회적 복지 향상에 효과가 없다고 주장합니다. 결국 중개인, 고용주 및 구직자의 검색 활동만 최소화 할 뿐이라고 주장합니다.

이러한 시각 차이의 격차를 해소하기 위해서는 PES 활동의 관점이 기존의 직업 매칭 프로세스보다 더 커야합니다(Koning, 2006). PES 사무국은 모니터링 및 제재 활동을 포함할 수도 있습니다. PES는 실업자가 실업 보험이나 사회 지원을 받기 전에 취업 기회를 찾는지 모니터링을 해야 합니다. 이러한 활동은 흔히 '일 우선'으로 분류됩니다. 그러나 고용 서비스의 이러한 감독활동(거버넌스)에 대해서는 업무 처리로만 인식되고 그의 효과에 대해서는 별로 언급되지 않고 있습니다.

공공 고용 서비스의 운영에 있어 크게 두 가지 논의 사항이 있습니다(Koning, 2006; Mergele and Weber, 2017). 첫 번째는 고용서비스가 단독으로 수행되어야 하는지 아니면 사회부조 서비스와 함께 수행되어야 하는지 입니다(Martin, 2000). 둘째는, 고용 서비스의 민영화 문제와 관련됩니다. 민영화가 실제로 실행되는 국가의 예로는 호주가 포함됨

니다(OECD, 2006). 호주는 고용서비스를 민간에 위탁하여 수행하며, 공공은 민간위탁업체의 실적을 관리합니다. 업체마다의 다양한 특성을 고려하여 표준통계기법을 도입하여 민간위탁업체들 간의 공정한 성과 비교를 중요시하고 있습니다(Department of Employment, 2017). 공공 고용 서비스는 구직자와 고용주가 원하는 정보에 맞게 매칭 서비스를 제공하여 취업(재취업)률을 높이려는 노력을 합니다. 서비스의 효율성을 제고하기 위해 점차 중앙정부 차원에서 제공하던 서비스를 지방정부 차원으로 분산시키는 추세가 강해지고 있습니다(Mergele and Weber, 2017). 고용서비스를 지방으로 분산시키려는 이유는 지방경제상황과 선호에 대해 각 지역들이 더 잘 이해하고 있어 일자리 매칭을 보다 효과적으로 수행할 수 있다는 주장에 따른 것입니다.

지방 분권화가 중앙 집권화보다 더 나은 고용 서비스를 창출하는지 여부에 대해서도 고민할 필요가 있습니다. 정치가는 지방정부로의 분산 목적이 실업을 줄이는 것보다 다른 목적을 가질 수 있습니다. 첫째, 그들은 자신의 지역 내의 취업 알선에 엄격히 초점을 맞추으로써 지역 내 과세를 극대화하는 것을 목표로 삼을 수 있습니다. 이러한 편협된 사고는 구직자들을 지리적으로 구속하여 지역 간 노동 자원의 이동성을 낮출 것입니다 (Wildasin, 1991; Lundin and Skedinger, 2006). 둘째, 지방 정치가들은 재정비용을 다른 정부 수준으로 높이기 위한 수단으로 고용서비스를 활용할 수 있습니다 (Weingast et al., 1981; Besley and Coate, 2003). 예를 들어 비용이 국가 예산으로 충당되기 때문에 지역 정책에 대한 재정적 이득을 가져 오는 한, 어떤 정책이 재취업 촉진에 효과적이지 않더라도 특정 적극적 노동 시장 정책 (ALMP)이나 모니터링 전략을 선호할 수 있습니다. 셋째, 지역 정책 입안자들은 재선을 모색하면서 복지 수급자의 구직 의무를 완화하기 위해 지방 고용 서비스를 압박 할 수 있습니다 (Brollo et al., 2015). 또한 지방 수준의 공공 고용 서비스가 국가 수준보다 정치적 영향을 받기 쉽기 때문에 재선을 위한 인기의 목적으로 취업률을 낮추는 일자리를 방만하게 확대할 경향이 발생할 수 있습니다. 이 문제는 경험적 제약 또는 지방 분권의 정도가 대개 국가마다 다르기 때문에 비교를 통한 답을 찾기 어렵습니다. 이러한 이유로 아직까지 이에 대한 실험적 연구 결과는 부족한 상태입니다(Lundin and Skedinger, 2006, Boockmann et al., 2015).

독일의 경우 2012년 개혁 이전에는 개별 고용 센터 정책이 지방 당국과 협력하여 연방 고용국 (Federal Employment Agency, FEA)의 지침, 감독 및 감독 하에 결정되었습니다. 2012년 이후에는 41개 처리 지구의 당국이 자유롭게 이러한 정책을 관리하고 규정했습니다. 독일 고용 센터는 독일 복지 시스템에서 중심적인 역할을 하는 원 스톱 지역 고용 센터입니다. 고용 센터의 자금 조달은 개혁으로 인해 영향을 받지 않았습니다. 모든 고용 센터에 대해 연방 정부는 복지 혜택과 적극적인 노동 시장 프로그램 비용을 지원하고 지방 당국은 비용을 지원했습니다.

독일의 지방분권화 사례를 살펴보면 지방 분권화된 고용 센터는 일시적으로 구직자의 의무 이행을 줄였으며, 지역별로 고용서비스 전략을 변경하였습니다. 그러나 지방 분권화가 직무 안정성 측면에서 취업알선의 질을 향상 시켰다는 증거 또한 발견되지 않았습니다. 우려하였던 지리적으로 이동성을 낮추는 효과에 대한 증거는 발견되지 않았습니다. 결과적으로 구직자는 지방 분권화의 혜택을 받지 못했다고 결론 내려 졌습니다. 지방 분권화된 고용 센터는 노동 시장 정책을 지역에 맞추어 조정하긴 하지만, 구직자들의 재취업을 향상시키지 못하였습니다. 재취업은 일자리 창출 프로그램의 증가에 따라 영향을 받았습니다.

지방분권화의 유형을 몇 가지 살펴 보겠습니다. 첫째 유형은 공공 고용 서비스 제공이 중앙 집권적 고용 센터 간에 비교적 표준화 되도록 하고, 지시 및 기술 감독을 통해 지역 고용센터를 감독합니다. 특히 취업알선, 적극적노동시장정책 및 제재 정책은 통일된 전국적인 지침을 따릅니다. 지방 정부는 주로 마약 중독이나 심리적인 문제와 같은 복지적 정책을 포함한 서비스를 제공합니다. 두 번째 유형의 고용 센터는 모든 고용 서비스에 대한 책임이 있는 지역 행정부와 분리되어 존재합니다. 중앙 집중형 기관과는 달리, 이들 취업 센터는 구직자/구인자 등록 데이터 교환을 제외하고 중앙고용정책기구와 완전히 독립적으로 운영됩니다. 중앙고용정책기구는 일반적인 기술 감독을 제공하지 않습니다. 다만 지방 정부는 각자의 주정부 또는 중앙고용정책기구와 목표 계약서에 서명합니다.

이러한 극단적인 두 가지 유형 사이에 다양한 유형을 만들어 갈 수 있을 것입니다. 독일의 지방분권화 사례의 시사점은 분권화가 적극적노동시장정책의 추진을 변화시키지 않으면서 일자리 창출 계획을 증가 시키는 것으로 나타났습니다. 지방 분권 개혁은 공공 예산에

상당한 영향을 미쳤습니다. 독일의 경우 2012년 지방 분권화로 인해 연간 최소 6,600만 유로의 추가 비용이 발생했다고 보고하였습니다(Mergele and Weber, 2017). 장기간의 직업 상담, 추가 적극적 노동시작정책 참여 및 세금 환급으로 인해 더 많은 재정적 부담이 발생했습니다. 이러한 독일의 결과는 분권화된 고용 센터가 총 공공 예산과 개별 재취업률에 대한 효과를 보지 못할 수도 있음을 암시합니다. 지방마다의 잘못 설계된 제도적 구조가 고용 센터의 효율을 현저히 감소시킬 수 있으므로 분권화로 인한 인센티브 효과를 신중하게 검토하는 것이 중요함을 시사합니다.

IV

데이터혁신에 따른 도시개선 및 일자리 창출 해외 사례

디지털 산업 혁명은 도시지역에서 보다 활발하게 발달할 것입니다(World Economic Forum, 2017). 도시에 사는 도시민에게 어떤 영향을 미칠지, 도시 정책은 어떻게 바뀌어야 하는지, 도시 내에 발생할 수 있는 분열과 기회를 어떻게 준비해야 하는지를 살펴볼 필요가 있습니다. 시간이 지날수록 디지털 및 물리적 물품의 결합으로 방대한 양의 정보가 더 많이 생성될 것입니다. 시민들의 행동은 자발적이든 비자발적이든 간에 디지털 정보로 기록됩니다. 전화통화, 문자 메시지, 이메일, 소셜 미디어 게시물, 온라인 검색, 신용카드 구매 등의 정보가 클라우드에 저장됩니다. 이러한 정보는 지리정보와 결합되어 도시의 움직임을 실시간으로 보여주는 자료로 사용될 수 있습니다. 데이터가 어떻게 물리적 공간을 변화시킬지를 파악해 보고자 합니다. 이를 통해 디지털 혁명을 보다 잘 이해할 수 있을 것으로 생각됩니다.

디지털 산업 혁명으로 인해 도시 인구는 더욱 증가할 것으로 예상됩니다. 도시 인구 증가는 새로운 현상이 아닙니다. 이러한 도시의 성장에 대처하기 위해 새로운 도시 인프라를 만들 필요가 있을 뿐만 아니라 점차 커지고 있는 대도시가 지속될 수 있도록 기존 인프라 자산을 더 잘 활용해야 합니다. 이러한 도시 인프라 구축에 스마트 기술이 도움이 될 수 있습니다. 오래된 도시와 건물 및 인프라를 20세기 기술에 적응하려고 노력할 경우 오래된 도시 중심지는 디지털 혁명의 새로운 기술에 보다 쉽게 적응할 수 있습니다. 시민, 경제, 정부, 인프라, 환경의 다양한 영역에서 4차 산업 기술의 요구와 일자리가 창출되는 사례를 살펴보겠습니다.

● 경제 영역

두바이시는 2016년에 도시 전역의 블록체인 전략을 시작했습니다. 2017년 비행기 조종

사들로부터 시작하여 정부에서 블록체인 기술을 구현하여 효율성을 높일 수 있는 범위를 탐구하고 있습니다. 이 전략은 블록체인 신생 기업에게 사업 환경을 조성하는 것을 목표로 합니다.

호주의 멜버른시는 시각화 및 분석 도구를 제공할 뿐만 아니라 기업 및 시민이 사용할 수 있는 데이터를 공개적으로 제공하는 데 앞장서는 도시 중 하나입니다. 멜버른은 토지 이용 및 보행자 이동과 같은 분야에서 유용한 데이터를 만들고 있습니다. 멜버른시는 개방형 데이터 문화를 창출함으로써 더 많은 시민과 신생 기업이 도시 서비스를 개선하고 경제적 가치를 창출하기 위해 데이터를 활용할 수 있는 방법을 만들 기반을 제공하고 있습니다.

아일랜드의 Dublin시는 소프트웨어 프로그램 신생기업 창출을 지원합니다. 이를 위해 토지이용, 레크리에이션, 편의시설에 이르기까지 250개 이상의 데이터세트를 구축하고 있습니다. 공공 및 민간 회사는 지방 당국이 제공한 데이터와 함께 웹 사이트에서 데이터를 받을 수 있습니다. Dublin시는 시민들이 도시 계획 정보에 쉽게 접근하도록 하였습니다. 도시계획과 관련하여 지역 개발 계획, 유적지, 홍수 위험, 건축 허가, 기획 신청서 등 계획 문제와 토지 이용에 대해 쉽게 알 수 있도록 온라인 플랫폼을 개발합니다.

Dublin시는 31개의 개별 기관에서 수행되는 400여개의 계획을 모두 한곳에 통합하여 시민들과 각 지방 기관이 손쉽게 열람할 수 있도록 하였습니다. 단순 통합이 아니라 손쉽게 내용을 검토할 수 있도록 개요 작성, 형식 통일 등을 수행하였습니다. 이 플랫폼은 거리지도, 항공사진, 역사지도, 건축 유산, 특별 보호 구역, 자연계 보호 구역, 홍수 범람원, 인구 조사 및 교육 서비스와 같은 데이터를 결합합니다. 각 당국과의 협의, 정보의 매핑, 정보의 통합 과정은 수년이 걸릴 수 있습니다. 이러한 정보는 사업 계획 수립을 용이하게 하여 창업을 간접적으로 지원할 뿐만 아니라, 새로운 서비스를 창출하는 사업의 기회를 제공할 수 있습니다. Dublin시는 개방형 데이터 문화를 확대함으로써 경제적 가치를 창출하기 위한 기반을 만들어 가고 있습니다.

인도의 Ahmedabad시에서는 통합 운송 관리 시스템을 도입하여 택시 및 인력거와 같은 비공식 대중교통을 포함하여 카드를 통해 원활한 지불이 가능하도록 하였습니다. 이는 여행 계획을 손쉽게 짤 수 있고 지불이 편리하도록 도와줍니다. 도시의 대중 교통망의

다른 부분을 통합하는 것이 고객에게 더 편리하여 여행 계획을 수립 할 수 있도록 해주어 전반적으로 관광산업을 발전시킵니다. 특히 비공식 대중교통 수단을 포함시킴으로써 "최후의 수단"연결성을 향상 시켜서 전반적으로 대중교통을 더 많이 이용할 수 있게 합니다.

또한 실시간으로 생성되는 데이터를 분석하여 교통 등 운영 효율성을 향상시킬 수 있습니다. Ahmedabad의 자동 지불 시스템은 스마트카드를 사용하여 사용자의 대중교통 데이터를 수집하여 자동화된 현금 없는 거래와 대중교통 차량의 실시간 모니터링을 가능하게 하였습니다. 이동성 시스템 전반에 걸쳐 원활한 경험을 제공하고 고객과 네트워크 운영자에게 계획 및 일정을 개선하는 데 필요한 정보를 제공함으로써 도시가 새로운 건설에 투자하는 것보다 기존 인프라를 효율적으로 활용하여 교통량을 극대화 할 수 있도록 합니다. 자동 지불 시스템과 스마트카드의 운영은 이를 분석하여 효율적인 교통 정책을 찾아낼 데이터 관리 및 분석 전문가를 필요로 합니다.

프랑스 파리의 루브르 박물관은 방문자의 행동 방식을 추적 분석하여 관광 명소를 개선 하였습니다. 블루투스 센서에 의해 생성된 익명의 중단 데이터를 분석한 결과, 박물관 관리자는 방문객들이 박물관에서 어떤 경로를 택하는지, 어느 곳에서 얼마나 오래 머무는 지 이해할 수 있었습니다. 방문자 행동 및 경험에 관한 데이터는 박물관과 같은 관광 명소를 관리하는 방법을 알려주는 데 중요합니다. 과거엔 이러한 연구를 하려면 시간이 많이 소요 되는 관측 및 조사에 의해 수행되었습니다. 군중 밀도 및 현지 혼잡과 같은 귀중한 정보를 시뮬레이션하고 예측하기 위해 박물관도 데이터 분석과 관련된 일자리를 필요로 하고 있습니다.

중국의 인촨(Yinchuan)시는 안면 인식 및 카드를 통해 지방 도시 버스요금 지급이 가능 하도록 했습니다. 카드 없는 거래는 더욱 빠르고 간단합니다. Yinchuan의 안면 인식 소프트웨어는 승객의 얼굴을 통해 은행 계좌에 연결하고 요금을 직접 수납하게 합니다. 안면 인식의 도입은 개인정보의 남용 및 빅브라더와 같은 부정적 시각으로 인해 아직은 추가적인 검토가 필요할 것입니다. 특히 중국의 경우엔 인터넷 검열 등으로 인해 이러한 과도한 개인정보의 수집 및 사용에 대한 불안한 시각이 많습니다. 하지만 충분한 제도적 보완이 이루어진 실험적 시도 측면이라는 가정에서, 안면인식 소프트웨어 개발 및 데이터 분석

관련한 새로운 일자리가 창출되고 있는 것을 볼 수 있습니다.

● 정부 영역

호주의 멜버른시는 시민들이 예산 작성에 참여하도록 하였습니다. 시민들이 기획 및 예산 책정에 참여하도록 하는 것은 선구적인 방법입니다. 예를 들어, 데이터 투명성 및 경제 발전 방향 마련 시 세금 사용에 대한 의사 결정을 안내하는데 "시민 패널"을 사용할 수 있습니다. 멜버른 시는 시 예산의 일부를 시민들의 결정에 따라 배분합니다. 참여형 예산 정책은 지역사회가 도시운영을 위해 더 많은 관심을 가지도록 할 수 있습니다. 멜버른시가 직접적으로 정보를 공개하기 위해서는 데이터 관리 인력을 필요로 할 것입니다.

미국의 보스턴(Boston)시는 택시와 네비게이션 앱과 제휴하여 도로 정체에 대처하였습니다. 모든 대도시는 교통 정체를 완화하기 위해 데이터를 수집하고 사용하는 다양한 방법을 모색하고 있습니다. 교통 혼잡은 많은 악영향을 끼칩니다. 대기 오염과 온실 가스 배출을 증가시키며, 교통 체증으로 개인이 낭비하는 시간은 도시의 생산성과 경제 성장을 저해합니다. 정체를 효율적으로 처리하기 위해서는 교통 속도와 흐름에 대한 정확한 실시간 데이터가 필요합니다. 많은 시당국의 자원이 제한적인 운송 기관은 자체 인프라 구축을 통해 경제적으로나 기술적으로 정보 수집에 어려움을 겪습니다. 양측은 이러한 조치를 통해 이익을 얻을 수 있습니다. 보스턴(Boston)은 네비게이션 앱인 Waze와 데이터 공유 파트너 관계를 맺었습니다(waze.com, 2018). 보스턴은 Waze의 데이터를 사용하여 불법 주차가 빈번한 곳과 교통 정체의 핫스팟과 같은 문제를 확인합니다. 그 대가로 Waze는 교통상황을 방해 할 수 있는 건설 현장, 행사의 행렬 및 기타 이벤트에 대해 사전 정보를 입수하여 사용자에게 제공하는 서비스를 향상시킵니다. 현재 사용되고 있는 택시 운전사 전화기의 GPS 데이터는 교통 계획 등에 사용되고 있습니다. 차세대 혼잡 관리 솔루션은 교통 신호 타이밍, 대중교통, 도로 인프라, 긴급 교통 관리 및 여행 수요 관리에 관한 보다 나은 데이터 기반의 결정을 내릴 수 있을 뿐만 아니라 여행 공유를 개발하고 최적화하는 것을 도울 것입니다. 이것은 최적화된 경로를 더 적은 비용으로 이동할 수 있는 정보를 제공하고, 더 깨끗하고 오염이 적은 환경을 만드는 데에도 도움이 될 것입니다. 수집된

정보는 여행자의 패턴을 이해하는 데에 도움을 주기 때문에 보다 나은 관광 상품을 개발할 자료로 사용되며, 결과적으로 관광 산업의 활성화를 간접적으로 도울 것입니다.

덴마크 코펜하겐은 폐수를 줄이고 처리를 개선하기 위해 데이터를 이용하고 있습니다. 도시가 하수 및 폐수를 관리하는 방법에 대한 데이터가 개선되고 있습니다. 덴마크 코펜하겐은 폐수 시스템이 어디에 그리고 언제 가장 압력을 받게 될지 예상하기 위해 강우량 분포를 예측하고 이에 따라 인프라를 관리하기 위한 조치를 취합니다. 이러한 개입은 다양한 문제를 해결합니다. 코펜하겐에서 극심한 강우 사건이 발생하여 처리되지 않은 하수도 가 배출될 위험이 있습니다. 코펜하겐은 지역 기상 레이더의 실시간 데이터 뿐만 아니라 덴마크 계량실의 데이터를 활용하고 있습니다. 레이더 기반 일기(강수량) 예보 및 하수도 시스템의 동작 모델은 수 시간 전에 유량 또는 수위를 성공적으로 예측했습니다. 이러한 노력은 하수도 시스템과 폐수 처리 시설을 최적화하며, 공공의 데이터 분석 일자리를 창출합니다.

● 인프라 영역

스페인 은 유비쿼터스 बैंकिंग 인프라를 보유한 BBVA는 수백만 건의 금융 거래에 대한 익명의 데이터를 제공합니다. 이는 도시경제를 이해하도록 돕습니다. MIT Senseable City Lab의 연구원은 은행 데이터를 사용하여 스페인 도시 및 지역의 지출 패턴을 모델링, 시각화하여 이를 예측했습니다. 지출 금액, 지출 유형, 개인 유형 및 개별 이동성에 대한 지표가 포함됩니다. 경제 모델은 거시 경제 동향을 적절히 예측하고 설명하지 못하고 있습니다. 이 프로젝트는 개별적인 거래내역에 관한 데이터를 활용하여 더 큰 경제적 패턴을 이해할 수 있는 방법으로 미시적 행동 모델과 거시적 현상 모델 사이의 차이를 줄입니다. 도시 차원의 미시 경제 패턴에 대한 이해를 향상시키고 도시 간 시간에 따라 비교할 수 있게 함으로써 도시 계획과 정책 결정 및 사업 의사 결정에 도움을 줄 것입니다. 수백만 건의 금융 거래 데이터가 사회 경제 현상을 이해하고 정책 마련을 지원하는 사례로서, 데이터 생성의 중요성을 보여주는 예라고 할 수 있습니다.

미국 휴스턴의 소방서는 태블릿 원격 진료를 통해 불필요한 병원 방문을 줄이는 효과를

보았습니다. 휴스턴 소방서는 시스코 및 파나소닉과 협력하여 최초 대응자인 소방관 및 응급 의료 기술자를 위한 태블릿 기반 앱을 개발했습니다. 휴스턴의 소방서는 연간 약 600,000건의 응급 의료 호출을 받습니다. 응급실의 대기 시간은 최대 14시간입니다. 태블릿 원격 진료를 통한 초기 진단은 응급 처치 요원이 우선순위를 정하는데 도움을 주고 응급 상황이 아닌 문제에 대한 대체 서비스와 교통 옵션을 찾는 데 도움을 줍니다. 태블릿 원격 의료 운영을 위한 데이터관리 및 데이터 분석 인력을 필요로 하는 사례입니다.

● 환경 영역

도시에 나무를 더 많이 심도록 유도하는 방법으로서 도시 거주자가 지역 사회 내에서 나무의 위치와 크기를 보고, 해당 도시에 더 많은 나무를 심도록 의견을 제출할 수 있게 합니다. 녹색 그늘막(캐노피)은 도시 생활의 중요한 부분입니다. 나무는 극심한 기온을 완화하고 교통, 소음 및 혼잡으로 자연스럽게 휴식을 취하고 도시 환경에 거주하는 사람들의 삶의 질을 향상시킵니다. 그러나 일반 시민들은 종종 자신이 사는 지역의 환경을 인식하지 못합니다. 시민들이 자신이 살고 있는 지역의 환경을 잘 이해한다면 선거 운동에 더 적극적인 환경적 역할을 요구할 수 있습니다. Treepedia는 Google 거리 뷰 파노라마를 활용하여 과학자와 도시 거주자가 도시의 수목 분포를 모니터링하고 녹색 뷰 인덱스를 만듭니다 (Treepedia, 2018). 환경단체는 이러한 녹색 정책을 수행하기 위해 구글로부터 데이터를 필요로 할 뿐만 아니라, 이를 분석하여 시민과 소통할 시스템을 운영하기 위한 데이터 관리 전문가를 필요로 하고 있습니다.

자연재해를 대비하고 대응하기 위해서도 데이터 분석 전문가를 필요로 합니다. 모든 도시는 화산, 지진, 홍수, 화재 및 진흙 슬라이드와 같은 여러 종류의 자연 재해에 노출되어 있습니다. 지진이 발생했을 때 경보를 발령하고, 구조적 손상 가능성이 있는 지역은 구분된 색으로 지도를 표시하고, 노인 인구의 비율이 높은 구역, 어린이 수와 같은 정보를 제공함으로써 재해로부터 대응 계획을 세울 수 있도록 돕습니다. 자연 재해가 평일에 발생하면 학교와 사무실을 최우선으로 표시 할 수 있고, 주말에는 자원을 다른 곳으로 우선 배치 할 수 있도록 돕습니다. 이러한 정보는 자연 재해에 대비할 수 있는 보다 탄력적인 도시를

건설하는 데 도움이 됩니다. 이러한 안전한 도시 건설을 위해 데이터 분석 전문가는 필수적인 핵심 기술로 요구됩니다.

● 시민생활

코펜하겐은 도시를 순환시키기 쉽게 하기 위해 제4차 산업 혁명 기술을 사용하고 있습니다. 다이나믹한 자전거 간판과 모바일 앱은 자전거 운전자에게 가장 효율적인 경로에 대한 실시간 교통 정보를 제공합니다. 화면은 혼잡이 적거나 눈이나 바람에 덜 노출되는 대체 경로에 대한 정보를 표시합니다. 이는 자전거 운전자의 현재 속도와 속도를 줄이거나 가속할 예정인지를 파악하고, 동기화된 다음 교통 신호와 맞추어 어느 정도 속도를 내야하는지 표시합니다. 야간 사이클링을 위해 코펜하겐은 가로등이 거리가 비었을 때는 스스로를 끄고 다가오는 자전거가 감지될 때 밝게 빛나도록 지능을 도입하였습니다. 스마트 가로등은 밤새 빛을 내지 않아 전력이 절감되었습니다. 코펜하겐은 자전거의 안전성을 높여 사용자의 수를 늘리고자 하는 정책을 수행하고자 합니다. 이러한 정책을 수행하기 위해서는 가로 등 센터 및 데이터 관리와 분석 기술자를 필요로 합니다.

러시아 모스크바는 “Our City” 포털 및 모바일 앱을 이용하여 도시 서비스에 대한 요청이나 불만 사항을 쉽게 제출하도록 하였습니다. 가령, 거리의 쓰레기, 도로의 패인 부분 등을 쉽게 제출할 수 있도록 하였습니다. 시민들은 종종 이러한 요청 방법을 모릅니다. 시민들이 문제를 제기하지 않는 한 공무원도 문제를 인식하지 못할 수 있습니다. 시민은 문제에 대한 간단한 설명과 사진을 첨부할 수 있도록 하였습니다. 각 사례는 해결 기간과 관련하여 엄격한 규칙이 있어 해당 부서나 업체에 할당되어 전달되도록 하였습니다. 마지막으로 문제가 해결되면 사진과 함께 개별적으로 통지가 되도록 하였습니다. 이러한 플랫폼에서 수집된 데이터는 빠른 유지보스 및 관리에 대한 정보를 제공함으로써 문제를 사전에 예방하는 데에도 도움이 됩니다. 이러한 포털 시스템과 모바일 앱의 운영을 위해서는 데이터 관리 및 분석 인력을 필요로 합니다.

데이터를 활용한 서비스 기회를 확대하려면 점점 더 많은 사람들이 데이터에 액세스하고 그 사용에 관한 집단 토론에 참여하는 것이 중요합니다. 대용량 데이터를 가능한 한

많이 공개할 경우 이는 도시에 지대란 영향을 미칠 것입니다. 데이터를 이용한 서비스가 확대될수록 데이터 분석 일자리는 더욱 많이 증가할 것입니다. 이러한 데이터를 이용한 서비스의 창출은 선진국과 개발도상국 모두에게서 공통적으로 나타나며, 데이터를 이용한 서비스가 확대될수록 데이터 분석 인력은 더욱 요구될 것입니다.



데이터 혁신 평가 사례

정부 차원에서 데이터 혁신에 따른 시대 변화의 중요성을 이해하고, 앞으로 나아가야 할 방향을 모색하는 것은 매우 중요합니다. 만일, 정부가 데이터 혁신을 주저한다면 그 추진 과정에서 효과가 미미할 것이라는 선입견과 실패에 대한 두려움 때문일 것입니다. 데이터 혁신을 효과적으로 관리하여 성공 가능성을 높이기 위해서는 체계적인 평가방법이 수반되어야 합니다(Castro et al., 2017). 본 연구에서는 미국의 지방정부 데이터 혁신 평가 지표를 사례로 살펴보면서 한국의 실정에 이를 보완하여 제시하고자 합니다.

1 » 데이터의 활용 정도

더 많은 데이터를 이용할 수 있도록 노력하는 정부는 데이터 중심 혁신으로부터 가장 많은 보상을 얻을 수 있습니다. 정부는 두 가지 주요 방법으로 데이터 가용성에 영향을 미칠 수 있습니다. 첫째, 정부 차원에서 데이터를 수집하고 게시하도록 정책을 결정할 수 있습니다. 예를 들어, 정부는 국공립 대학 또는 보험 감독 당국이 수집하고 발행하는 공개 데이터의 양에 큰 영향을 줄 수 있습니다. 둘째, 정부는 민간 부문의 디지털화를 촉진하여 민간 부문에서 수집하는 데이터의 양을 늘리고 이의 사용을 촉진할 수 있습니다. 개인과 조직은 데이터를 통해 새로운 통찰력을 이끌어 내고 더 나은 결정을 내릴 수 있습니다(Castro et al., 2017). 정부는 공공 서비스를 개선하고 투명성을 높이기 위해 데이터를 사용할 수 있습니다. 기업은 데이터를 사용하여 제품 및 서비스를 개선하고, 회사 전략을 모색하고, 투자 여부를 결정합니다.

정부 행정 자료(입법 및 재정), 교육 데이터, 건강 데이터 및 에너지 데이터 (건물 효율성 및 에너지 사용)를 포함한 다양한 유형의 데이터 가용성에 대한 7가지 측정 지표가 포함됩니다. 또한 정부 정책이 공공과 민간 부문의 데이터 생산 및 공개를 권장하는지 여부를 반영하는 두 가지 지표가 포함됩니다.

[표 1] 데이터 가용성에 대한 측정 지표

종류	설명
입법자료(Legislative Data)	정부가 개방적이고 기계가 읽을 수 있는 형식으로 입법 데이터를 공개하는 정도
정부재정데이터 (Government Financial Data)	세금이 사용된 내역을 공개하는 정도
정부 교육 데이터(Education Data)	교육 데이터 시스템을 개발하기 위한 조치를 취한 정도
e-처방전(E-Prescribing)	통제 약물에 대해 전자처방을 사용하는 정도
의료서비스 가격의 투명성 (Health-Care Price Transparency)	의료 보험 가격 투명성 법률 및 규정의 존재
에너지 사용량 데이터 (Energy Usage Data)	그린 버튼 이니셔티브에 참여하는 유틸리티가 제공하는 고객의 비율
빌딩 에너지 효율 데이터 (Building Energy Efficiency Data)	만 명 거주자 당 에너지 효율 빌딩 성능 데이터베이스에 포함 된 건물 수
정부 정보 공공 접근 (Public Access to Government Information)	정부 정보에 대한 합법적 권리와 효과적인 접근에 관한 질적 조사 결과를 정량적으로 변형한 측정결과
ANTI-SLAPP 법률 도입 (Anti-SLAPP Laws)	공공 참여 (SLAPPs)에 대한 전략적 소송을 금지하는 법률을 채택했는지 여부

출처: Daniel Castro and Jordan Misra (2013) The Internet of Things, Center for Data Innovation.

● 입법 자료

정부가 개방적이고 기계가 읽을 수 있는 형식으로 입법 데이터를 공개하는 정도를 살펴볼 필요가 있습니다. 입법 자료는 시민들에게 선출직 공무원의 결정에 대한 통찰력을 제공하며 정부의 투명성과 책임을 증진시키는 데 도움이 됩니다. 지폐, 득표 및 위원회에 관한 적시 정보는 시민들이 정부 프로세스를 이해하고, 참여하고, 감독 할 수 있도록 도와줍니다. 개발자는 추세 분석, 투표 패턴 예측 또는 대화식 시각화 작성과 같은 선출 된 공무원과 그 구성원 간의 대화를 촉진하는 등 시민에게 공개적이고 입법적인 데이터를 더 유용하고

이해할 수 있도록 응용 프로그램을 만들 수 있습니다. 더욱이 주정부의 입법 자료의 개방성은 데이터를 보다 일반적으로 공개하려는 주정부의 접근 방식을 나타내는 좋은 지표입니다. 조지아, 뉴햄프셔, 텍사스 및 워싱턴은 모두 쉽표로 구분된 값(CSV) 파일과 같이 기계가 읽을 수 있는 형식으로 데이터를 제공하거나 개발자가 사용할 수 있는 API (Application Programming Interface)를 개발했습니다. 중요한 차이점은 데이터 내용의 영속성입니다. 현행법 뿐만 아니라 과거의 법안도 사용할 수 있어야 합니다. 세션별로 구성된 링크 가능한 특정 URL만을 제공해서는 안되고, 체계적이고 쉽게 탐색 할 수 있는지와 관련된 접근 용이성도 중요합니다.

● 정부 재정 데이터 공개 정도

세금이 사용된 내역을 공개하여 시민은 인터넷을 통해 확인할 수 있습니다. 이는 시민이 공공 정책에 참여하고, 투명성을 제고하도록 돕습니다. 재정 데이터의 공개는 정부에도 도움이 됩니다. 웹사이트를 통해 재정 데이터를 공개함으로써 인쇄비용, 행정 비용 등을 줄일 수 있습니다. 가령 미국의 사우스다코타주는 예산 중복 등을 효과적으로 파악이 가능하여 19,000,000 달러의 예산을 절감하였다고 합니다.

● 정부 교육 데이터

정부가 교육 데이터 시스템을 개발하기 위해 취한 조치의 정도를 측정할 필요가 있습니다. 데이터는 교사가 보다 개인화된 교육을 제공하고 관리자가 학교를 보다 효과적으로 관리할 수 있도록 도움으로써 교육 결과를 향상시킬 수 있습니다. 이상적으로, 교육자는 각 학생에 대한 개인화된 데이터를 수집하여 학습 스타일, 강점 및 약점, 학습 장벽 및 개인 진도를 파악하여 교육 프로그램 및 교육 기관을 더 잘 평가하고 학생들에게 자신의 필요에 맞는 교육을 제공할 수 있습니다. 이러한 시스템은 학습자가 학습하지 않은 환경에 있을 때 교육자가 학생을 식별하고 개입할 수 있게 하며, 학생의 능력에 따라 개별화된 교육을 가능하게하고, 너무 늦기 전에 고생하는 학생을 식별함으로써 중도 이탈률을 감소 시킵니다. 예를 들어, 오레곤 데이터 프로젝트는 교사가보다 효과적인 교육자가 되기 위해

학생 데이터를 수집, 분석 및 사용하는 데 필요한 기술을 교사에게 제공합니다.

오레곤 데이터 프로젝트에 참여한 학교의 학생들은 다른 학교의 독서 및 수학 능력 향상에서 다른 학생들보다 우위에 있는 것으로 나타났습니다. 많은 주에서는 학생, 교사 및 학부모에게 데이터 기반 통찰력을 제공하기 위해 학생 정보를 위한 종단 데이터베이스를 개발하고 있습니다. 이러한 데이터베이스는 종종 사회 서비스와 같은 다른 주 시스템과 연결됩니다. 이러한 연계를 통해 저소득층 가정 아동의 성공적인 학교생활을 위한 정책 결정 효과를 더 잘 이해할 수 있습니다. 가령, 미국 펜실베이니아주와 알칸사스주는 조기 아동 프로그램의 영향을 평가하고 자금을 할당합니다.

● 전자 처방전 도입

데이터 기반 건강관리의 핵심 구성 요소인 전자 처방(E-prescribing)을 통해 의사는 전화, 팩스 또는 수기 처방전을 통하지 않고 약국에 처방전을 전자로 보낼 수 있습니다. 아날로그 프로세스에서 디지털 프로세스로 이동하는 것은 여러 가지 이점이 있습니다. 먼저 전자 처방은 환자의 안전을 향상시킬 수 있습니다. 손으로 쓰는 처방전은 판독하기 어려울 수 있어 약물 남용으로 이어질 수 있습니다. 또한 많은 전자 처방 도구는 다른 약물, 건강 상태 및 알레르기와 가능한 상호 작용을 위해 처방을 확인하는 의사 결정 지원 시스템과 통합됩니다. 둘째, 전자처방은 업무 흐름을 간소화하고 설명을 위해 처방 의사에게 전화를 줄이기 때문에 제공자와 약사 모두에게 서류 작업을 줄임으로써 효율성을 높입니다. 또한 처방전을 직접 제공하지 않아도 돼서 환자의 대기 시간을 줄여줍니다. 마지막으로, 전자 처방은 환자의 건강 결과를 개선하고 상당한 비용 절감을 가져옵니다. 약물 순응도를 향상시킴으로써 환자는 부작용이 적어 응급실 및 1차 진료 의사를 방문하는 횟수가 줄어듭니다. 그 결과 전자 처방 시스템은 의사가 환자에게 저렴한 약제를 선택하도록 도울 수 있게 됨에 따라 환자에게 비용 절감의 효과를 가져다 줄 수 있습니다. 또한 약제의 위조 기회를 줄이는 장점이 있습니다. 미국 뉴욕은 모든 의사가 규제 약물 및 기타 처방약을 전자로 처방해야한다는 규칙을 준수하고 있기 때문에 전자 처방전을 주도하고 있습니다.

● 의료서비스 가격의 투명성

의료 서비스는 국가 경제의 큰 부분을 차지합니다. 의료 시스템의 효율성과 투명성을 적극적으로 증진시키는 국가는 거주자를 위한 이러한 비용에 대한 데이터의 질을 향상시킴으로써 의료비용의 장기적인 감소를 볼 수 있습니다. 특히 보험 회사, 메디 케이드 및 메디 케어와 같은 모든 주요 지불인으로부터 건강 보험 보상 청구 및 가격 정보를 수집하는 전액 청구 데이터베이스(all-payer claims database)가 중요합니다. 이러한 가격 정보를 의사에게 전자 방식으로 제공하며 건강 보험 회사에 개인 맞춤형 서비스 제공을 요구함으로써 건강 관리 비용 투명성이 향상됩니다. 예를 들어 HealthCost 웹 사이트에서 2007년 APCD에 액세스 할 수 있는 데이터를 제공한 뉴햄프셔는 소비자에게 더 많은 영향력을 부여했습니다. 이러한 정보는 건강보험 회사가 보다 합리적인 보험 금액을 청구하도록 만들었습니다.

● 공공 에너지 사용량 데이터

에너지(전기, 물, 가스 등) 제공 회사의 데이터는 고객이 에너지를 언제, 어디에서, 어떤 방법으로 사용하는지 제공할 수 있습니다. 유틸리티 데이터에 액세스 할 수 있는 수백만 명의 주택 소유자 및 기업을 위해 이것은 돈과 에너지를 절약 할 수 있는 새로운 방법을 제공할 수 있습니다. 공개 표준(표준화된 데이터 구조 형식)이 있으면 소비자는 에너지 데이터를 사용하는 응용 프로그램을 보다 쉽게 만들 수 있습니다. 미국 메사추세츠주의 경우 시민의 98%가 자신의 에너지 정보에 접근할 수 있습니다.

● 빌딩 에너지 효율 데이터 구축(building energy-efficiency data)

빌딩 성능 데이터베이스는 상업 및 주거용 건물의 실제 에너지 성능 및 물리적 특성과 운영 특성에 관한 데이터 세트입니다. 가능한 많은 빌딩이 데이터베이스 정보를 공개하도록 하여야 합니다. 데이터베이스에는 건물의 에너지 사용에 대한 최소 1년간의 데이터와 건물 사용에 대한 세부 정보가 있습니다. 위치, 기후, 난방, 냉방, 조명 등의 정보는 데이터 베이스에 입력되기 전에 정리되고 익명화합니다. 이 데이터베이스는 많은 용도가 있습니

다. 정부 및 상업 부동산 관리자는 이 데이터를 사용하여 유사한 건물과 비교하여 건물의 효율성을 평가하고 절감 효과를 평가할 수 있게 합니다. 또한 에너지 효율 솔루션을 판매하는 기업은 데이터를 분석하여 새로운 기회 또는 불충분한 시장을 파악할 수 있습니다. 정책 입안자는 이 데이터를 사용하여보다 효과적인 에너지 절약 프로그램을 설계 할 수 있습니다. 금융 기관은 이 데이터를 사용하여 위험 분석을 수행 할 수 있습니다.

● 정부 정보 공공 접근

정부 정보 공개를 측정 할 때에는 시민은 정보에 액세스 할 수 있는 법적 권리가 있는지, 효과적으로 접근할 수 있는지, 온라인을 통해 접근 할 수 있는 지를 측정할 필요가 있습니다. 시민이 정보에 접근할 수 있는 기능은 개인 및 시민 단체가 주 의원, 심사 위원 및 공무원의 행동을 모니터하고 자신들의 행동에 대해보다 효과적으로 책임을 지도록 하는 역할을 합니다.

● ANTI-SLAPP 법률 도입

주정부가 공공 참여에 대한 전략적 소송을 금지하는 법률(ANTI-SLAPP LAWS)을 채택했는지 여부를 점검합니다. 온라인상에서 비방글이 난무하는 것과 같은 부정적 행위를 예방하기 위해 법적 조치가 필요합니다. 이러한 불필요한 비방글을 제거함으로써 기업과 개인은 불필요한 경쟁 요소를 제거하고, 품질 향상 측면에서의 경쟁에 집중할 수 있게 됩니다. 이는 결과적으로 아담스미스의 '보이지 않는 손'이 보다 효과적으로 작동하도록 돕습니다.

2 데이터 기반의 핵심 기술 플랫폼 활성화 방향

오늘날 데이터 혁신을 위한 플랫폼을 제공하는 다양한 기술들이 존재합니다. 데이터를 쉽게 수집, 공유 및 사용할 수 있게 해주는 기술을 통해 기업, 연구원, 정부 기관 및 시민 모두가 새로운 통찰력을 찾아내고 더 나은 의사 결정을 내리고 가치 있는 신제품 및 서비스를 개발하며 상당한 사회적 혜택을 얻을 수 있습니다(Castro et al., 2017). 예를 들어 건강관리에서 전자 건강 기록 시스템을 채택한 병원은 환자 데이터에 더 잘 액세스하고 질병을 개선하여 환자 치료를 개선함으로써 의료 오류를 줄이는 것 외에도 서류 파일 관리 비용을 줄임으로써 비용을 절감할 수 있습니다. 지능형 온도조절기는 가정의 난방과 냉각을 보다 효율적으로 하기 위해 센서와 기계 학습을 사용하여 가정의 에너지 비용을 절감합니다. 지능형 온도조절기는 지역 사회의 에너지 효율을 향상시켜 오염을 줄이고 전력망에 부담을 줄 수 있습니다.

데이터 기반 혁신의 잠재력은 데이터 수집, 공유 및 사용에 달려 있습니다. 따라서 기술 인프라가 이러한 기능을 지원할 수 있게 하는 국가는 데이터 혁신을 선도하는 매우 유리한 위치에 있습니다. 관련 지표는 광대역 인터넷 속도, 스마트 미터, 대중교통 정보 시스템, 전자 건강 기록, 사물인터넷(IoTs: Internet of Things) 소비자 장치, 개방형 데이터 포털 및 전자 정부 기술과 같은 기술 인프라의 7가지 구성 요소의 존재와 품질을 포함합니다.

[표 2] 핵심 기술 플랫폼의 활용 가능성 관련 측정 지표

종류	설명
브로드밴드(Broadband)	광대역 서비스를 사용하는 인터넷 사용자 및 평균 연결 속도
스마트 미터(Smart Meters)	스마트 미터 인 전기 계량기의 비율
대중교통 정보 시스템	대중 교통 시스템에서 기계가 읽을 수 있는 데이터의 가용성
전자 건강 기록	의사와 병원이 전자 건강 기록을 사용하는 정도
IoT 소비자 장치	1,000 명의 거주자 당 스마트 TV와 웨어러블의 사용 정도
개방형 데이터 포털	개방형 데이터 포털과 정책을 가지고 있는지 여부
전자정부(E-Government)	정부가 디지털 기술을 사용하는 정도

출처: Daniel Castro and Jordan Misra (2013) The Internet of Things, Center for Data Innovation.

● 광대역 인터넷 속도

우리나라의 경우 빠른 인터넷 속도를 제공하는 인프라를 전국적으로 확보하고 있습니다. 다만 인터넷을 사용하는 인구의 비율, 초고속 광대역인터넷에 가입한 가구의 비율, 평균 연결 속도를 함께 측정하여야 할 것입니다. 유선뿐만 아니라 무선 인터넷의 속도 또한 포함되어야 할 것입니다.

● 스마트 미터

스마트 미터는 에너지 공급자와 소비자에게 가격, 사용 패턴 및 비효율에 대한 정보를 제공하는 강화된 양방향 통신 기술을 갖춘 서비스입니다. 미국의 경우 전력망의 비효율성으로 인해 2009년 3,330억 달러와 추가적인 탄소 배출로 인한 1조 2,200억 달러의 손실을 초래했다고 합니다(Surka and Ridlington, 2016). 스마트 미터를 가정에 설치하면 이러한 비효율을 해결할 수 있습니다. 에너지 공급자는 에너지 가격을 동적으로 조정하고 소비자는 이에 따라 소비량을 조정할 수 있습니다. 유틸리티 회사는 실시간 정보를 통해 정전을 방지하고 운영비용을 줄이며 궁극적으로 소비자의 평균 전기 가격을 낮출 수 있습니다(SGCC, 2013).

● 대중교통 정보 시스템의 데이터 가용성

대중교통 정보 시스템을 사용하여 대중교통 정보를 게시함으로써 대중교통 정보 도구에 대한 기계 판독 가능한 데이터 표준을 통해 대중교통 수단 개발 및 사용을 촉진할 수 있습니다. 이는 노선 및 운임 정보와 같은 대중교통 정보를 제공합니다. 이 정보를 제공하면 통근자가 정보에 입각한 교통이용 결정을 내릴 수 있을 뿐만 아니라 개발자 및 정부가 중요한 앱과 기타 서비스를 구축할 수 있는 플랫폼 역할을 돕습니다. 특히, 실시간으로 GTFS 데이터를 게시하면 교통 정보를 더 많이 사용하도록 유도할 수 있습니다. 이 정보를 통해 운전자는 출퇴근 시간 및 지연에 대한 정확한 정보를 볼 수 있습니다. 대중교통 정보 시스템을 이용하여 실시간 데이터를 제공하는 경우 새로운 사업이 발달하고 고용창출로 이어지게 됩니다.

● 전자 건강 기록

의사와 병원이 전자 건강 기록을 사용하는 정도를 측정합니다. 건강 데이터는 의료비용을 줄이고, 환자의 안전을 높이며, 진료의 질을 향상시키기 위한 기초가 됩니다. 의료 연구자가 건강 데이터를 사용하여 특정 집단에 대한 치료의 효과를 판단하거나 유해한 약물 부작용을 발견 할 수 있게 하려면 데이터 공유가 필요합니다. 또한 공중 보건 개선 및 처방 약물 남용 방지와 같은 특정 문제를 해결할 필요가 있습니다. 건강 데이터 인프라의 주요 구성 요소에는 데이터를 저장하는 전자 건강 기록과 환자 및 제공 업체가 데이터에 액세스하고 공유하는 전자 건강 정보 교환이 포함됩니다. 이러한 기술은 치료 시점에서 환자의 의료 정보를 이용할 수 있게 하고, 의사가 의사 지원 시스템을 사용하여 실수를 줄이고 치료의 질을 향상시키는 데 도움을 줍니다. 건강 데이터 인프라는 웨어러블 및 원격 의료와 같은 추가 의료 혁신을 위한 플랫폼 역할도 합니다.

● 사물 인터넷(IoTs)

글로벌 네트워크는 더 이상 사람들이 컴퓨터를 사용하여 서로 통신하는 것에 국한되는 것이 아니라, 다른 장치 및 주변 세계와 전자적으로 통신하는 장치를 위한 플랫폼이기도 합니다(Castro and Misra, 2013). 그 결과 데이터가 한 장치에서 다른 장치로 흐르고 수많은 목적으로 공유되고 재사용 될 사물인터넷의 성장은 센서, 저전력 프로세서, 확장 가능한 클라우드 컴퓨팅 및 유비쿼터스 무선 연결성의 발전에 의해 주도되고 있습니다. 이 기술은 도로 및 교량 모니터링, 가전제품 자동화, 건강 및 피트니스 모니터링, 농업 효율성 향상 등의 많은 분야에 사용될 수 있습니다.

● 개방형 데이터 포털

개방형 데이터 포털은 정부 데이터를 공유하는 홈페이지입니다. 개방형 데이터 포털을 통해 사용자는 농업, 교육 및 경제와 같은 다양한 주제에 대해 수천 개의 정부 데이터 세트를 탐색하고 다운로드 할 수 있습니다. 인기 있는 데이터는 지역마다 다를 것입니다. 가령, 어떤 지역은 어류 정보에 관심이 많을 것이고, 어떤 지역은 기상 정보를 많이 사용할

수 있습니다. 개방형 데이터 포털은 개발자가 애플리케이션 프로그래밍 인터페이스(API) 및 개인이 특정 데이터 세트를 요청할 수 있는 방식으로 정보를 제공하는 것이 좋습니다. 공개적이고 기계가 읽을 수 있는 형식으로 광범위한 공개 정보에 대한 액세스를 보장하는 것은 정부의 투명성과 책임성, 대중 참여 및 혁신을 장려합니다. 개방형 데이터에 대한 기계 가독성을 높이려면 실무 차원에서 수행되기 보다는 법률로서 기계 가독성을 높이도록 요구할 필요가 있습니다. 가령 민간 개발자가 API를 통해 모든 데이터를 읽을 수 있는 형식으로 다운로드 할 수 있도록 법적 제도 마련이 필요합니다.

● 전자정부

기업과 개인이 정기적으로 인터넷을 사용하여 세금을 내고 면허를 갱신하고 허가를 신청할 때, 정부는 이전의 종이 기반 프로세스를 디지털 작성 방식으로 변경하여 다른 용도로 사용할 수 있는 새로운 트랜잭션 데이터를 생성하는 작업을 할 필요가 있습니다. 예를 들어 건물 검사 데이터가 완전히 디지털화 되면 이 정보를 자동화 도구에 통합하여 화재 또는 기타 안전 검사의 우선순위를 매길 수 있습니다. 가령, 뉴욕시 소방국에서는 건물의 나이, 재료, 배선 상태 및 기타 요소에 대한 데이터를 사용하여 위험 점수를 작성하여 검사를 계획합니다(Dwoskin, 2014). 전자 정부 프로젝트는 보다 고급 데이터를 만들기 위한 선구적 역할을 합니다. 현재 데이터 아키텍처를 현대화하고 다른 사람들이 보유한 정보와 상호 작용할 수 있게 하는 공개 애플리케이션 프로그래밍 인터페이스를 구축하는 데 투자를 할 경우 장래에 기계 학습과 같은 신기술을 최대한 활용할 수 있는 기회를 가지게 됩니다(Castro, 2016). 디지털 정부를 만들기 위해서는 각 기관 및 담당 부서의 노력에 달려 있다고 보이지만, 실상은 대통령과 국회의 노력이 중요합니다(Castro et al., 2017).

3 » 인력 및 기업 자산 개발

데이터 혁신 기술을 개발한 일자리 관련 지표를 측정할 필요가 있습니다. 정부는 데이터 관련 일자리를 만들어야만 이러한 기술자를 유지할 수 있습니다. 데이터 혁신은 단순히 발생하는 것이 아닙니다. 데이터 혁신 기술을 이용한 경제에서 사람들이 살아가려면, 필요한 기술, 직업, 조직적 지원이 필요합니다. 첫째, 데이터 과학, 컴퓨터 과학 및 통계 분야에서 숙련된 인력이 데이터 가치를 해독하는 것이 필수적입니다. 정부는 컴퓨터 과학 및 통계 활용 능력 테스트 및 과학 및 공학 대학생을 육성해야 합니다. 전문가들은 데이터 경제시대에 성공하는 데 필요한 분석 기술을 갖춘 근로자와 관리자의 필요성이 증가 할 것으로 예상하고 있으며, 정부는 이 이들을 육성하기 위해 경쟁해야 할 것입니다.

인력 및 기업 자산 관련 지표는 다음과 같습니다. 첫째는 소프트웨어 근로자수와 통계 분석 근로자수를 파악할 필요가 있습니다. 이들이 결국 데이터 혁신과 성공적인 데이터 활용 기업을 만들 것입니다. 기존 기업이 데이터 기반 혁신을 이용하여 생산성을 향상시키는 길을 찾아낼 것이고, 다음 세대의 데이터 기반 제품을 만들 것이기 때문입니다. 데이터 근로자는 비록 비기술 산업 분야의 기업에 근무를 하게 되더라도 새로운 산업 기회와 기존 문제를 해결함으로 매우 가치 있는 역할을 할 것입니다. 두 번째는 기업에서 요구하는 데이터 관련 일자리 개수입니다. 관련 일자리수가 많을수록 데이터 관련 산업도 더욱 활성화 될 것입니다. 세 번째는 데이터 과학 커뮤니티의 규모를 측정할 필요가 있습니다. 데이터 혁신은 팀의 노력으로 이루어지며 동료로부터 배웁니다. 데이터 기술을 지원하는 문화가 있을 때 비즈니스 및 전문 조직이 발전할 수 있습니다. 따라서 오픈소스(open-source)를 사용하는 회사의 수, 데이터 관련 연구 금액, 데이터 과학 커뮤니티의 규모 등을 측정할 필요가 있습니다.

미국의 메사추세츠와 캘리포니아와 같이 역사적으로 강력한 기술 분야를 보유한 국가가 이 범주에서 최고의 경쟁력을 가질 것입니다. 데이터 과학 일자리 수와 같은 지표는 주 정부가 직접 영향을 미치기가 어렵고, 장기간의 투자를 필요로 하기도 합니다. 데이터 관련

일자리를 육성하기 위해 정부는 미국의 미주리주와 같이 세금 감면 혜택을 사용할 수도 있을 것입니다(Verge, 2015). 그리고 데이터 과학 커뮤니티를 육성하기 위해 정부 기관이 참여하여 이러한 조직을 구성할 수 있습니다.

[표 3] 인력 및 기업 자산 관련 측정 지표

종류	설명
컴퓨터 과학 및 통계(Computer Science and Statistics) 능력 테스트	데이터 과학 및 통계 전공 학생 테스트 평균 점수
과학기술 학위자 수(STEM)	과학, 기술, 공학 및 수학 (STEM) 고등 교육 학위자(18세~34세) 인구 비율
소프트웨어 서비스 종사자 수 (Software Service Jobs)	컴퓨터 프로그래머, 소프트웨어 개발자 및 컴퓨터 및 정보 시스템 관리자로 일하는 총 근로자수
통계전문가 일자리 수의 비율 (Statistics Jobs)	통계 전문가, 보험 계리사, 데이터베이스 관리자 및 운영 조사 분석가 (operation research analysts)로 일하는 근로자의 비율(전체근로자수 대비)
데이터 과학자 일자리 수의 비율 (Data Science Job Listings)	데이터 과학자 구인 게시 비율(총 구인 게시건수 대비)
공개 데이터 활용 회사 (Open Data Companies)	Open Data 활용 기업의 수 및 비율
정보 및 데이터 처리 부문 (Information and Data-Processing Sector)	정보 및 데이터 처리 산업의 경제적 산출량(총 경제적 산출량 대비)
데이터 과학 연구 개발을 위한 국가 기금 (National Funding for Data Science R&D)	데이터 과학 기금
데이터 과학 커뮤니티 (Data Science Community)	데이터 관련 커뮤니티 그룹의 평균 구성원 수

출처: Daniel Castro and Jordan Misra (2013) The Internet of Things, Center for Data Innovation.

● 데이터 과학 및 통계 능력 시험 평균 점수

데이터 과학 및 통계 전공 학생들의 역량을 전공 시험 또는 자격증 시험을 통해 측정한 평균점수를 의미한다. 데이터 과학 및 통계 전공 학생들의 역량을 측정한다.

● 과학기술 학위자 수(STEM)

과학, 기술, 공학 및 수학 고등 교육 학위의 가중치 측정은 18세에서 34세 사이의 인구

점유율로 측정할 수 있습니다. 과학, 기술, 공학 및 수학 분야의 능력은 경제에서 점점 더 중요한 역할을 하게 됩니다. 따라서 대학은 이러한 전문가를 확대할 필요가 있습니다. 미국은 데이터 과학 분야에서 훈련받은 근로자의 공급이 이미 수요에 미치지 못하고 있는 것으로 나타났습니다.

● 소프트웨어 서비스 종사자 수

컴퓨터 프로그래머, 소프트웨어 개발자 및 컴퓨터 및 정보 시스템 관리자로 일하는 총 직원 수를 측정할 필요가 있습니다. 컴퓨터 및 정보 시스템과 관련된 서비스 직업은 데이터를 수집, 공유 및 분석하며, 보다 나은 결정을 하도록 통합하고 시각화 합니다.

● 통계전문가 일자리 수의 비율

통계 전문가, 보험 계리사, 데이터베이스 관리자 및 운영 조사 분석가로 일하는 총 인원의 수를 측정할 필요가 있습니다. 통계 및 데이터베이스 관리 기술을 많이 사용하는 직업은 기술 및 비기술적 산업 모두에서 혁신의 열쇠입니다. 통계 전문가, 보험 계리사, 운영 분석가 및 데이터베이스 관리자는 데이터를 보다 광범위하게 액세스하고 액세스 할 수있게 만들고, 큰 데이터를 분석하여 추세를 파악하고, 데이터 기반 생산성을 향상시키는 데 기여합니다.

● 데이터 과학자 일자리 수의 비율

전체 일자리 수 중에서 데이터 과학자 일자리 수의 비율을 측정할 필요가 있습니다. 이 지표는 데이터 과학 분야에서 훈련된 근로자에 대한 수요를 측정합니다. Indeed.com은 미국 최대의 구직 사이트로, 1,600만 건의 일자리를 홍보하고 있습니다. 구직 정보를 추적하여 데이터 중심 경제의 여러 구성 요소를 분석하고 있습니다. 첫째는 우선 데이터 근로자에 대한 주 수요(일자리수)를 측정합니다. 높은 수요는 높은 수준의 성장을 의미하며 이는 국가의 다른 지역 및 세계의 다른 지역에서 고도로 숙련된 근로자를 유치하는 데에도 도움이 됩니다. 둘째, 지표는 지식 기반 경제가 얼마나 빨리 움직이는지를 측정합니

다. 일자리에서 높은 이직률은 역동성, 높은 수준의 혁신, 위험 부담 및 기업가 정신을 의미합니다. 세 번째는 비 정보통신기술 기업이 데이터 집약적인 기술 도구를 채택하는 정도를 측정합니다. 비 정보통신기술 기업이 데이터 과학자를 고용하는 비율이 높을수록 이들이 기업이 생산하는 제품과 서비스를 향상시킬 확률이 높습니다.

● 공개 데이터 활용 회사

정부의 공개 데이터를 사업의 주요 부품으로 이용하는 기업의 수를 측정할 필요가 있습니다. 보험, 금융, 지도 작성, 교육, 운송 등 다양한 분야의 회사가 정부의 공개 데이터를 이용할 것입니다. 회사는 공정한 가격을 설정하고, 소비자 습관을 더 잘 이해하며, 새로운 제품과 서비스를 개발하기 위해, 그리고 더 큰 시장 추세를 확인하기 위해 개방된 데이터를 사용할 수 있습니다. 상무부, 보건 복지부 및 증권 거래위원회와 지방 정부는 이들 회사에 공개 데이터를 제공하는 주요 공급기관입니다. 어떤 회사라도 자유롭게 공개 데이터를 사용할 수 있지만, 모든 회사가 그렇게 하지는 않습니다. 따라서 개방 정부 데이터를 사용하는 회사의 수는 공개 데이터를 얼마나 잘 사용하는지를 나타내는 좋은 지표입니다. 미국의 메사추세츠, 뉴욕, 캘리포니아, 워싱턴 같은 기술 산업이 발달한 지역은 이러한 기업의 수가 많습니다. 이러한 기업의 수는 데이터 과학 기술 인력의 가용성 및 데이터 과학 기술 노동 시장과 높은 상관관계가 있는 것으로 나타났습니다.

● 정보 및 데이터 처리 분야

총 경제적 산출량 중에서 정보 및 데이터 처리 산업의 경제적 산출량의 비율을 측정할 필요가 있습니다. 정보 및 데이터 처리 비즈니스는 다른 회사가 데이터에서 가치를 추출할 수 있게 하는 데 중요한 역할을 합니다. 기업은 점점 더 많은 양의 데이터를 수집하면서 데이터 분석 전문 회사에 더욱 의존하게 됩니다. 이들 데이터 전문 기업은 더 많은 데이터를 저장, 관리, 분석 및 공유 할 수 있는 기술 인프라를 관리합니다. 정보 및 데이터 처리 부문에서는 이러한 서비스를 제공함으로써 종종 다른 기업이 자체적으로 개발할 때보다 낮은 비용으로 이러한 서비스를 보다 높은 품질로 구입할 수 있도록 합니다. 예를 들어

Amazon Web Services와 같은 클라우드 스토리지 공급 업체는 회사가 자체 데이터 센터를 구축하고 운영하는 데 소요되는 비용의 일부만으로 확장 가능한 스토리지를 제공할 수 있습니다. 클라우드 컴퓨팅은 공유 컴퓨팅 리소스를 사용하면 자체 데이터 센터를 유지 관리하는 것보다 비용 효율성이 훨씬 뛰어나므로 다양한 컴퓨팅 요구 사항이 있는 소기업이나 비즈니스에 특히 유용합니다. 실제로 클라우드 컴퓨팅은 비즈니스 성장을 촉진합니다. 미국과 유럽의 신생 기업에 대한 Deloitte의 2014년 조사에 따르면 83%는 클라우드 기술로 인해 다른 방식으로는 제공할 수 없었던 서비스에 액세스 할 수 있다고 믿었으며 이러한 서비스는 쉽게 확장 가능하기 때문에 중소기업은 클라우드 기술로 인해 이러한 인프라를 직접 개발하고 유지 관리해야 하는 경우보다 빠르게 비즈니스를 성장시킬 수 있다고 생각합니다(Deloitte, 2014).

● 데이터 과학 연구 개발을 위한 자금

지식 경제에서 공공 및 민간 부문 연구 개발(R&D)에 의한 혁신은 경제 성장의 주요 동인 중 하나입니다. 민간 부문의 막대한 연구 개발 투자에도 불구하고 민간 부문은 기업 단기 이익을 위한 최적의 연구 개발 투자 수준을 유지할 수 있습니다. 왜냐하면 장기적 연구 결과 성과의 이익을 기업이 독점하긴 어렵기 때문입니다. 공공 부문 투자를 보완하지 않으면 경제의 발전 속도는 느리게 됩니다. 대학은 공공 부문의 연구 개발에서 점점 더 중요한 역할을 하고 있으며, 대학의 연구 개발은 민간 부문에 실질적인 경제적 이익을 가져다줍니다. 가령 뉴 멕시코에는 뉴 멕시코 대학이 개발하는 제품을 상용화하여 R&D 노력을 기울이도록 장려하기 위해 기술 연구 협력(Technology Research Collaborative)을 시작했습니다(New Mexico Economic Development Department, 2016).

● 데이터 과학 공동체(community)

데이터 관련 그룹의 구성원수를 측정할 필요가 있습니다. 번성하는 데이터 과학 커뮤니티는 지식 공유를 장려하고 협업을 촉진하며 데이터 과학 기술과 관심사가 있는 사람들 간의 네트워크를 구축할 수 있습니다. 전국적으로 산업 전문가, 학생, 교육자, 공무원, 시민

기술자 및 기업가를 비롯한 많은 개인이 정기적으로 모여서 기계 학습의 발전에 대한 업계 전문가의 강연에서부터 워크숍에 이르기까지 광범위한 데이터 중심 주제에 대해 논의합니다. 예를 들어, 뉴욕시의 Meetup.com에 있는 NYC Open Data 그룹은 데이터 과학 기술을 가르치는 워크숍을 정기적으로 주최하고 열람하는 4,500명 이상의 회원을 보유하고 있으며 개방형 데이터를 사용하여 프로젝트 및 도구를 개발하기 위해 공동 작업을 합니다(Meetup, 2017). 유사하게 샌프란시스코와 보스턴에는 각각 수천 명의 회원을 보유한 meet up 그룹이 있으며 데이터 시각화에서 인공 지능에 이르는 데이터 주제에 관심이 있는 전문가를 위한 실습 커뮤니티로 활동하고 있습니다. 이러한 데이터 과학 공동체는 과학, 기술, 공학, 수학(STEM) 인력의 기술 향상에 중요 할뿐만 아니라 데이터 과학에 대한 관심을 불러일으키고 현장에서 이러한 기술을 연마하려는 사람들을 유치할 수 있을 뿐 아니라 연구, 비즈니스 아이디어 및 기타 기회에 대한 협력으로 이어질 수 있습니다. Meetup.com 은 인구가 5만 명 이상인 모든 도시의 경우, API를 사용하여 반경 25 마일 이내에 있는 모든 데이터 관련 그룹을 식별할 수 있도록 합니다. 데이터 관련 그룹은 빅데이터 분석, 데이터 시각화, 데이터 마이닝, 자동화, 개방형 데이터, 데이터 센터, 네트워킹 등과 같은 각 주제 영역의 관심이 일치하는 그룹입니다.



변화에 발맞추기 위한 정책적 방향

데이터 혁신은 아직도 시작 단계입니다. 정부 차원에서 데이터 혁신에 따른 시대 변화의 중요성을 이해하고, 앞으로 나아가야 할 방향을 모색한다면 국가와 각 지방자치단체는 긍정적인 변화를 이루어 낼 수 있습니다. 기술적 시대 변화로부터 격리되어 경쟁력을 유지한 국가의 사례는 찾아보기 힘듭니다(Ferguson, 2011).

대용량 데이터를 가능한 한 많이 공개할 경우 이는 국가 및 도시 경쟁력에 지대한 영향을 미칩니다. 이러한 데이터를 활용한 서비스의 창출은 선진국과 개발도상국 모두에게서 공통적으로 나타나며, 어떠한 국가나 도시든 데이터 혁신을 통해 도시를 발전시키고 일자리를 창출할 수 있습니다. 데이터를 이용한 서비스가 확대될수록 데이터와 관련하여 파생되는 일자리는 더욱 많이 증가하게 됩니다.

정부가 공공 및 민간 데이터를 수집하고, 데이터 공유를 활성화할 수 있는 방법을 살펴보기 위해서 아래에 제시한 방법들은 이미 산업계, 비영리 단체 등 이해관계 하에 얹혀 있어서 각계각층의 지지와 합의를 필요로 합니다(Castro et al., 2017). 데이터 혁신이 원활하게 이루어지려면 사회적 안정이 필요하며, 사회 안정이 데이터 혁신을 더욱 원활하게 합니다. 데이터 혁신에 대한 부정적인 결과를 피하기 위해 기술적 혁신을 늦추거나 방해하는 것은 적절하지 않으며 불가능해 보입니다. 데이터를 수집, 공유 및 사용하는 방법을 가속화하기 위해 정부가 취할 수 있는 방법을 다시 한 번 요약해 보고자 합니다.

(1) 개방형 데이터 개발을 확대하고 영구적인 데이터 공개 정책을 수립해야 합니다.

정부는 방대한 양의 유용한 데이터를 지속적으로 수집해야 합니다. 정부는 공공데이터의 공개 및 새로운 데이터를 만들기 위한 지속적으로 노력해야 합니다. 미국은 데이터를 공개적으로 게시하여 누구나 자유롭게 사용하도록 허용함으로써 연간 1조 1천억 달러의

경제적 가치를 창출할 수 있었습니다. 정부는 가장 중요한 데이터 기여자 중 하나입니다. 정부의 투명성, 책임성 및 대응성을 향상시킬 수 있습니다. 공공기관은 공개된 데이터를 사용하여 내부적으로 또는 다른 공공기관과의 데이터를 더 잘 평가하고 공유할 수 있습니다. 대중은 엄청난 양의 공공데이터에 쉽고 빠르게 액세스할 수 있어야 합니다. 이러한 데이터를 통해 민간 기업에서는 경쟁력을 보강하여 새로운 제품과 서비스를 개선하고, 구축할 수 있어야 합니다. 개방된 데이터를 사용하여 누구나 서비스 제품과 도구를 개발할 수 있도록 해야 합니다. 그리고 민간기관 이해 관계자들과 협력하여 공개 데이터의 가용성과 이용을 확대해야 합니다. 국회는 정부가 공식적인 책임자로서 데이터를 공개하는 것을 명시한 미국의 OPEN 정부 데이터 법(OPEN Government Data Act)과 같은 법안을 만들어 통과시켜야 합니다. 공공기관 및 기업을 위한 개방형 데이터의 이점을 완전히 확보하기 위해서는 이 법안은 관련 집행 조치의 수행 및 공개 사항을 체계화한 내용을 포함해야 합니다. 데이터의 품질을 높이기 위해서는 정부 데이터의 정확성과 적시성에 대한 높은 기준을 수립해야 합니다. 데이터의 접근 용이성을 높여 독점을 예방해야 합니다.

공공은 민간과 협력하여 지속적으로 양질의 데이터 개발을 확대하여야 합니다. 가령, 실시간으로 대중교통 데이터를 국가 담당 부서 및 지방 자치단체가 제공할 필요가 있습니다. 전국적인 3차원 입체 데이터베이스를 개발할 경우 관련 산업을 활성화 할 수 있습니다. 가령, 건설에서부터 재난 방지 계획을 위해서는 많은 분야에서 정확하고 고해상도의 지형 데이터가 필요합니다. 이런 데이터는 지질 자원 평가, 정밀 농업, 기반 시설 관리를 포함한 다양한 응용 분야에서 중요한 역할을 할 수 있습니다. 이렇게 되면 경제, 공공 안전 및 혁신에 큰 도움이 될 수 있습니다. 물론 다양한 민간 조직에서 이러한 데이터를 수집하고 제공할 수 있으나, 신뢰할 수 있고 개방된 공공부문의 데이터 개발이 중요합니다. 또한 안보와 관련한 부분도 충분히 검토하여 데이터를 개발할 필요가 있습니다. 모든 정부기관 및 준정부기관 등 공공기관에 이 규칙을 적용해야 합니다. 정부는 데이터 혁신을 촉진하기 위해 다음과 같이 구체적이고 직접적인 조치를 취할 수 있습니다.

(2) 기업 데이터의 투명성을 높여야 합니다.

투명성은 정부에만 해당되는 것이 아니라 기업에도 요구됩니다. 기업의 법적 소유자 및 이익의 실질적 소유자를 공개하여 기업의 투명성을 높여야 합니다. 데이터 투명성은 세금 회피, 돈세탁, 제재 회피, 범죄 행위 자금 조달과 같은 문제점을 예방할 것입니다. 정부가 기업 데이터의 투명성을 요구할 때, 중요한 소유권 데이터가 교섭용 거래 조건으로 사용될 수 있는 한 정부는 이를 수집 할 방법이 없을 것입니다. 정부는 모든 법인 기업의 소유권 데이터를 수집하고 개방적이고 기계가 읽을 수 있는 형식으로 온라인으로 게시함으로써 규제 기관과 대중이 쉽게 접근할 수 있는 법안을 통과시켜야 합니다. 이러한 데이터 수집은 모든 분야에 적용되어야 하며, 국가가 데이터를 검증 할 수 있는 방법이 있어야 합니다. 데이터를 수집만 하고 데이터에 대한 접근을 제한하여서도 안 됩니다. 수상한 사업 거래 및 범죄 행위를 규제기관, 법 집행기관, 언론기관, 시민 사회 단체 및 경쟁 회사가 함께 모니터할 수 있어야 합니다.

(3) 공공 데이터 접근을 용이하게 하고 기계 가독성(예, API)을 제고시켜 나가야 합니다.

정부, 국회, 입법부는 많은 데이터 세트를 가지고 있습니다. 이를 외부에 사용하도록 할 때 대량 다운로드 방식뿐만 아니라 API(Application Programming Interface) 방식으로 접근하도록 허용하여야 합니다. 외부 컴퓨터 개발자가 데이터 세트의 개별 항목과 직접 원활하게 상호 작용하는 소프트웨어를 작성할 수 있게 해주는 일련의 컴퓨터 기능이 API 서비스입니다. 전체 데이터를 다운받아 사용하여야 하는 경우 앱 및 온라인 서비스에 액세스하여 데이터를 사용하기에 매우 불편합니다. 더욱이 자주 업데이트되는 데이터 세트의 경우 전체 데이터를 다운받은 방식은 특히 비효율적입니다. 국회가 API를 통해 정보를 게시하면 많은 기업들이 국회 데이터를 자신의 앱(App)과 서비스에 통합할 수 있습니다. 이러한 도구는 입법 정보를 추적하는 데 소요되는 시간을 줄임으로써 국회 직원의 생산성을 높일 수 있게 해줍니다.

국회 데이터 세트에는 입법, 투표, 청문회, 회원 약력 등 수십 년의 정보가 포함되고 누적될 것입니다. 국회 도서관 데이터도 API를 통해 공개 제공할 필요가 있습니다. 이러한

정보를 제공하였을 때 정보를 이용하는 서비스가 활성화됩니다. 만일 외부의 영리 및 비영리 기관에서 API 서비스를 하게 된다면 비상업용으로만 제한하거나 상업용 사용자에게 추가 요금 부과 방식 등과 같이 사용자 제한을 설정할 수 있기 때문에 이러한 정보제공은 정부가 주도적으로 역할을 수행해 나가야 할 것입니다.

(4) 개인의 공공재(전기, 가스, 수도 등)에 대한 사용 정보 접근을 적극 허용해야 합니다.

전력 회사가 사용량 수준의 스마트 미터를 도입하도록 촉진하고 스마트 온도조절기능을 포함하도록 해야 합니다. 공공기업은 소비자가 자신이 사용한 전기, 가스 및 수도 사용량에 대한 자세한 정보를 모니터링하는 스마트미터를 제공해야 합니다. 그러나 아직 소비자가 이 정보에 액세스 할 수 있는 것은 아닙니다. 소비자가 원하는 시점에 언제든지 정확하게 유틸리티 사용량에 액세스 할 수 있게 하면 비용을 절감하고 효율성을 높일 수 있습니다. 에너지 요금의 약 10%를 절감 할 수 있다고 연구되고 있습니다(Castro et al., 2017). 소비자는 데이터를 통해 개인의 에너지 사용습관을 보다 잘 이해하고 다양한 에너지의 영향을 측정할 수 있습니다. 개인의 에너지 사용습관을 이해하면 태양 에너지와 같은 친환경 에너지 대안에 대한 의사 결정을 보다 쉽게 내릴 수 있습니다. 마찬가지로 소비자가 자신의 물 사용 데이터에 액세스 할 수 있게 되면 더 빨리 누출을 파악하고 수도사용을 더 효율적으로 계획할 수 있게 됩니다. 정부는 전기, 가스, 수도 등 제공 기관이 금액, 시기 등의 정보가 담긴 개방형 표준을 마련하여 소비자에게 정보 접근을 제공하도록 요구해야 합니다. 국회는 소비자들이 자신의 유틸리티 데이터에 접근 할 수 있도록 하는 법안을 만들어야 합니다. 이것은 소비자와 사회 전체에 이익을 제공할 것입니다. 개인뿐만 아니라 건물의 공공재 사용 정보에 대한 내용도 공개해야 합니다. 에너지 절약 빌딩(DOE)의 데이터베이스(Building Performance Database)에 데이터를 제출하는 것과 같은 프로그램에 공공 기관이 참여하도록 함으로써 모범을 보이도록 해야 합니다.

(5) 교육 및 연구 개발 대한 투자를 확대해야 합니다.

교육에 대한 투자는 모든 수준의 교육 시스템을 포괄해야 합니다. 보다 구체적으로, 최근

의 연구에 따르면 유아기는 인지 및 비인지 기술을 형성하고 학습을 위한 기초와 후기에 기술 습득에 중요한 시기입니다(Heckman 2008, Heckman et al 2013). 5세가 되는 때까지 뇌가 빠르게 발달하고, 나중에 학교, 건강, 직업 및 생활에서 성공하기 위해 필요한 인지 및 성격 기술의 토대를 마련합니다. 유아 교육은 주의력, 동기 부여, 자기 통제 및 사교성과 함께인지 능력을 촉진시킵니다(Heckman 외. 2013). 또한 불우한 환경의 아이들은 사회적 경제적 측면에서 부족한 지원을 받을 가능성이 적기 때문에 초기 사회 정책의 한 형태이기도합니다. 이 때문에 유치원, 유치원 및 초등학교의 아동을 조기에 지원하는 것이 필수적입니다. 교육과 관련된 두 번째 질문은 기술을 가르쳐야 한다는 것입니다. 기본 기술이 전문가가 가정하는 것보다 훨씬 빠르게 진행되기 때문입니다. 오늘날 컴퓨터가 점점 더 많은 것을 할 수 있어 음악, 그림, 디자인, 과학적 가설 개발 등 창조성 분야에서도 성과를 내고 있습니다. 그러나 컴퓨터는 인간이 하는 것처럼 세상을 경험하고 인간의 행동을 이해하고 예측하는 데에는 여전히 멀리 있습니다. 이것은 공감, 팀워크, 코칭, 인간 상호 작용을 필요로 하는 작업이 인간에 의해 계속 수행 될 것이라는 것을 의미합니다.

연구 개발은 디지털화와 관련된 정책의 핵심 영역입니다. 그것은 교육, 특히 고등 교육과 밀접한 관련이 있습니다. R&D에 대한 투자는 기업이 역량을 키울 수 있도록 도와주므로 적어도 장기적으로는 경쟁력과 고용 성장의 중요한 원동력이 됩니다(Peneder 2009; Peters et al. 2014; Steinmueller 2010). 전략적 산업 선택 측면에서 국가는 이미 성공한 산업에 데이터 과학을 적용하는 데 집중해야 합니다.

(6) 정부와 기업이 주도적으로 근로자 재교육을 확대해야 합니다.

기술 부족, 대량 실업 및 불평등 증가라는 4차 산업에 대한 부정적 시나리오를 극복하기 위해서는 기업은 기존 인력의 재학습을 적극적으로 지원해야 하며, 개인 스스로도 적극적으로 학습과 교육을 받아야 합니다. 정부는 이러한 개인들의 노력을 지원하기 평생학습 기회를 제공하고 창의적인 교육 환경을 조성할 수 있어야 합니다. 많은 고용주들이 재교육 및 기술 향상에 대해, 고도로 숙련된 일부 소수 직원들에게 초점을 맞추고 있는 경향이 있습니다. 그러나 4차 산업 혁명이 기업에 잘 정착하려면 기업은 인적 자본 투자를 책임이

아닌 자산으로 인식해야 합니다. 이것은 신기술과 기존 숙련 기술 사이에 선순환 구조를 만들기 위해 특히 필수적입니다. 신기술 채택은 지속적인 재교육과 숙련을 통해 미래의 기술을 배우게 하는 동기를 부여합니다. 기업 내부의 창의적 능력을 최대한 활용하여 새로운 기회를 활용할 수 있다면, 사업 성장 및 새로운 일자리 창출을 촉진할 수 있습니다. 반대로 조직 내 고위 지도자의 잘못된 혁신 기술 인식으로 인해 직원 간 기술 격차가 커지고 이로 인한 불협은 새로운 기술 도입 및 비즈니스 성장을 크게 저해 할 수 있습니다.

우리는 인공지능을 포함한 새로운 기술의 잠재력을 보다 잘 이해하고 새로운 고품질의 일자리를 창출하며 기존의 작업방식에 비해 품질과 생산성을 크게 향상시키는 데 중점을 둘 필요가 있습니다. 4차 산업 혁명의 기술 진보의 물결이 특정 작업에 필요한 근로자의 수를 줄이기 위해 발전하고 있음은 분명합니다. 하지만 새로운 역할에 대한 수요 증가는 줄어드는 수요 감소를 상쇄 할 수 있을 것입니다. 물론 이러한 긍정적인 변화는 자동으로 얻어지는 것은 아닙니다. 수백만 명의 근로자의 전환과 학습을 촉진하기 위한 적극적인 투자가 요구됩니다.

(7) 기업가 정신을 갖춘 리더를 양성해 나가야 합니다.

새로운 기술을 성장과 고용으로 전환시키는 기업가가 필요합니다. OECD(2016)의 최근 보고서에 따르면 소기업에 의해 상당한 양의 일자리가 창출되었습니다. 정보 통신 기술은 신생 기업에 의해 형성되어 왔으며 이러한 신생 기업 중 일부는 매우 짧은 기간에 상당한 규모에 이르렀습니다. 이것이 정보통신 및 생명 공학 분야의 새로운 회사에 대한 더 많은 지원을 필요로 하는 이유입니다. 비즈니스를 시작하려는 사람들을 위한 교육 및 컨설팅과 같은 유연한 조치 외에 벤처 캐피탈의 사용 가능한 양을 늘리기 위한 조치가 요구됩니다. 벤처 캐피탈을 제공하는 것이 정부의 임무는 아니지만, 국가는 개인이 신생 기업 및 중소기업에 더 많이 투자하고 자본 시장을 강화하기 위한 인센티브를 창출해야 합니다.

현재의 사회 경제적 변화를 이해하고 경제 성장과 사회적 진보를 긍정적으로 이끌어 가기 위해서는 리더를 양성해 나가는 것이 중요합니다. 4차 산업으로 인한 노동시장의 변화가 전 세계 수백만 명의 근로자와 기업에서 급속히 현실화되고 있습니다(World

Economic Forum, 2018). 이 새로운 세계에서 경제적 번영, 사회적 진보 및 개인 번영을 위한 새로운 기회는 매우 크다고 볼 수 있습니다. 이를 위해서는 교육훈련 시스템, 노동시장 정책, 기술개발에 대한 사업 접근법, 고용 계약 및 기존사회 계약, 기업과 정부의 리더십과 기업가 정신, 직원들의 평생 학습에 대한 열정 등이 필요합니다.

(8) 합리적 세금 제도를 만들고 공공 고용서비스 영역을 확대해 나가야 합니다.

디지털 산업이 확대될 때 긍정적인 시나리오에서 경제 활동은 활발하게 유지되고 성장은 더욱 빨라질 것입니다. 세금 징수를 통해 정부가 복지국가를 유지하고 교육, 연구 개발 및 인프라에 투자 할 수 있도록 증가할 것입니다. 그러나 디지털화가 되면 주요 세입원이 노동에서 자본과 부가 가치세로의 전환이 필요할 수도 있습니다. 이렇게 되면 기제도 사회 보장 재정에 기여할 수 있습니다. 디지털화로 인해 경제 활동이 감소한다는 부정적인 시나리오에서 노동당 부가가치와 자본에 대한 과세가 더 중요한 것으로 보일 수 있습니다. 이러한 변화는 전반적인 세금 부담이 증가하지 않을 경우에만 가능하며, 노동에 대한 세금 부담이 현저하게 감소한 경우에만 가능합니다. 그렇지 않으면 이러한 변화는 전반적인 세금 부담을 증가시킬 것입니다. 투자 상품에 대한 세금 인상은 기업들이 디지털 기술을 활용하려는 의지를 감소시킬 수 있으며 미래 경쟁력을 가로막는 장애물이 될 수 있습니다.

고용의 유연성 측면은 디지털 산업으로 인해 자영업 및 유연한 근무가 미래에 증가 할 것으로 예측됩니다. 사람들이 안정적으로 이직 및 직종을 전환하도록 지원할 필요가 있습니다. 고용은 미래에 구조적 변화가 있을 것입니다. 기술 변화에 대처하는 미래 노동 시장 정책을 수립하는 것이 매우 중요합니다. 가령, 플랫폼 근로자의 경우 이러한 유형에 종사하는 근로자의 상태를 검토하고, 필요하다면 사회 안전망을 확대해야 하며, 보건 및 안전 규정을 확대해야 합니다. 이것은 또한 플랫폼 소유자가 사회 보장에 기여해야 한다는 것을 의미합니다. 그러나 정책은 현재의 비규제에서 과다 규제로 이동하지는 않아야 합니다. 사회 안전망의 확대에 따라 공공 고용서비스의 역할을 기존의 매칭 역할 뿐만 아니라 감독 역할(거버넌스)을 확대할 수도 있을 것입니다.

(9) 데이터 혁신 평가 지표를 개발하고 주기적으로 확인하는 법안을 마련해야 합니다.

데이터 혁신을 효과적으로 관리하여 성공 가능성을 높이기 위해서는 체계적인 평가방법이 수반되어야 하므로, 본 연구에서 제시하는 데이터 혁신 평가 지표를 참고할 경우 혁신 전략을 마련하는 데에 도움이 될 것입니다. 미국은 정부 기관이 가치 있는 데이터를 정기적으로 온라인으로 게시하고 있습니다. 이를 위해서는 2년마다 개방된 데이터를 평가하고 업데이트 하도록 요구할 수 있어야 합니다. 국가는 2년마다 공개 데이터 계획을 평가하고 데이터를 임의적으로 또는 이유 없이 삭제하지 않아야 합니다. 정부가 데이터 혁신을 주저하는 이유는 추진 과정에서 효과가 미미할 것이라는 선입견과 실패에 대한 두려움 때문인 경우가 많습니다. 이를 극복하여, 데이터 활성화와 일자리량의 증가에 대한 적절한 평가 지표를 개발하여 점진해 나간다면 우리사회에 4차 산업을 안정적으로 발전시킬 수 있습니다.

참 | 고 | 문 | 헌

- Acemoglu, D., and D.H. Autor. (2011) Skills, Tasks and Technologies: Implications for Employment and Earnings. In Handbook of Labor Economics, ed. D. Card and O. Ashenfelter, Vol. 4, Part B, 1043-1171. Amsterdam: Elsevier.
- Acemoglu, D., and A.A. Robinson (2012) Why Nations Fail, Crown Business.
- Aker, J.C., and I.M. Mbiti. (2010) Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives* 24, no. 3: 207-232.
- Baumol, W.J. (2012) The cost disease. Why computers get cheaper and health care doesn't. New Haven: Yale University Press.
- Bentley, R.A., M.J. O'Brien, W. Brock (2014) A: mapping collective behavior in the big-data era. *Behav. Brain Sci.* 37: 63-76.
- Besley, T. and S. Coate. (2003) Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach, *Journal of Public Economics*, 87, 2611-2637.
- Boockmann, B., S. L. Thomsen, T. Walter, C. Göbel, and M. Huber (2015) Should Welfare Administration be Centralized or Decentralized? Evidence from a Policy Experiment, *German Economic Review*, 16: 13-42.
- Brollo, F., K. Kaufmann, and E. La Ferrara. (2015) The Political Economy of Enforcing Conditional Welfare Programs: Evidence from Brazil, Tech. rep., mimeo.
- Brynjolfsson, E., and L.M. Hitt. (2000) Beyond Computation: Information Technology, Organizational Transformation and Business Performance. *Journal of Economic Perspectives* 14, no. 4.
- Cardia, E. (2008) Household Technology: Was it the Engine of Liberation? Montreal: Universite de Montreal
- Castro, D., and J. Misra. (2013) The Internet of Things, Center for Data Innovation. <http://www2.datainnovation.org/2013-internet-of-things.pdf>.
- Castro, D. (2016) How Artificial Intelligence Will Usher in the Next Stage of e-Government, Government Technology,

- <http://www.govtech.com/opinion/How-Artificial-Intelligence-Will-Usher-in-the-Next-Stage-of-EGovernment.html>.
- Castro, D., J. New, M. Beckwith. (2017) 10 Steps Congress Can Take to Accelerate Data Innovation, Center for Data Innovation.
- Cavoukian, A., and D. Castro, (2014) Big Data and Innovation, Setting the Record Straight: De-identification Does Work, The Information Technology & Innovation Foundation, Ontario, Canada.
- CitySourced Inc. (2018) Citysourced, <http://www.citysourced.com>. Accessed 15 Nov 2018
- Dachs, B. (2018) The impact of new technologies on the labour market and the social economy, Austrian Institute of Technology(AIT).
- Deloitte. (2014) "Small Business, Big Technology: How the Cloud Enables Rapid Growth in SMBs, Deloitte, September, 2014, <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Technology-MediaTelecommunications/gx-tmt-small-business-big-technology.pdf>.
- Department of Employment. (2017) Performance Framework, Jobactive Guideline(2.0)
- Doris, B., Pfeiffer, F. (1999) Qualifikation, Beschäftigung und technischer Fortschritt/Skill Structure, Employment and Technological Progress: Empirische Evidenz mit den Daten des Mannheimer Innovationspanels / Empirical Evidence Using Data of the Mannheim Innovation Panel, Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik), vol. 218, issue 1-2, pp. 128-146
- Edquist, C., L. Hommen, and M.D. McKelvey. (2001) Innovation and Employment: Process Versus Product Innovation. Cheltenham: Edward Elgar.
- Dwoskin, E. (2014) How New York's Fire Department Uses Data Mining, The Wall Street Journal. <http://blogs.wsj.com/digits/2014/01/24/how-new-yorks-fire-department-uses-data-mining>.
- Entorf, H., Pohlmeier, W. (1990) Employment, Innovation and Export Activities. In J.P. Florens, M. Ivaldi, J.J. Laffont and F. Laisney (eds), Microeconometrics: Surveys and Applications. London: Basil Blackwell.

- Evangelista, R., and M. Savona. (2001) Innovation, employment and skills in services – Firm and sectoral evidence, ECIS Conference, Netherlands, 2001, 20-23.
- Evans-Cowley, J. (2011) Crowdsourcing the curriculum: public participation in redesigning a planning program. Soc. Sci. Res. Netw. Work. Pap. Ser.
- Ferguson, N. (2011) Civilization: The West and the Rest. London: Allen Lane.
- Freeman, C., and L. Soete. (1997) The Economics of Industrial Innovation (3rd ed.). Cambridge [Mass.] and London: MIT Press.
- Gudipati, B.M., S. Rao, N.D. Mohan, N.K. Gajja. (2013) Big data: testing approach to overcome quality challenges structured testing technique. Infosys Labs Briefings 11, 65-73.
- HealthCost. (2018) Shop for healthcare now and start saving, <https://www.healthcost.com/>. Accessed 15 Nov 2018
- Heckman, J. (2008) Schools, Skills and Synapses. Economic Inquiry 46, no. 3: 289-324.
- Heckman, J., R. Pinto, and P. Savelyev. (2013) Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. The American Economic Review 103, no. 6: 2052-2086.
- Keivani, R., J.H.M. Tah, E. Kurul, and H. Abanda. (2010) Green Jobs Creation Through Sustainable Refurbishment in the Developing Countries. Geneva: International Labour Organization.
- Kerr, I., and J. Earle. (2013) Prediction, preemption, presumption how big data threatens big picture privacy. Stanf. Law Rev. 66, pp. 65.
- Kluve, J. (2006). The Effectiveness of European Active Labor Market Policy, IZA Discussion Paper.
- Koning, P. (2006) Measuring the Effectiveness of Public Employment Service (PES) Workers, CPB Discussion Paper, CPB Netherlands Bureau for Economic Policy Analysis.
- Kuroda, H. (2016) New Challenges for Japan's Labor Market – Speech at the Japan Summit 2016 in Tokyo, Governor of the Bank of Japan.
- Laney, D. (2014) Who owns (really owns) “Big Data”.

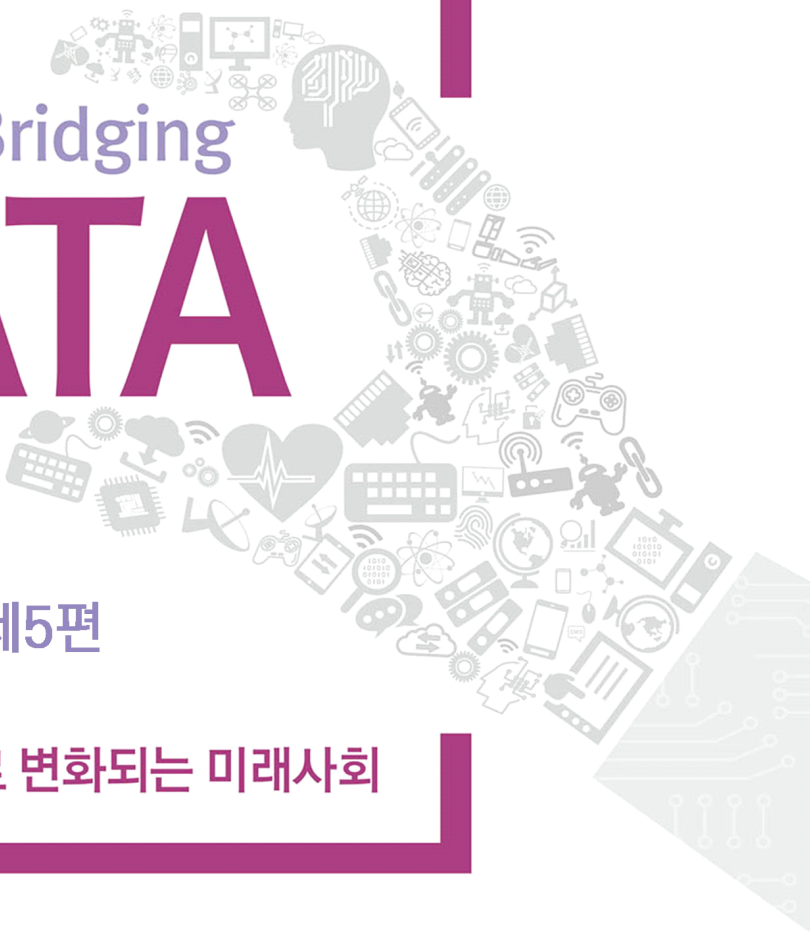
- <http://blogs.gartner.com/doug-laney/who-owns-really-owns-big-data/>. Accessed 15 Nov 2014
- Lerman, J. (2013) Big data and its exclusions. *Stanford Law Rev. Online*. 66, pp. 55-63.
- Lundin, M. and P. Skedinger. (2006) Decentralisation of Active Labour Market Policy: The Case of Swedish Local Employment Service Committees, *Journal of Public Economics*, 90, 775-798.
- Mankiw, N.G. (2013) Defending the One Percent. *Journal of Economic Perspectives* 27, no. 3: 21-34.
- Martin, J.P. (2000). What works among Active Labour Market Policies: Evidence from OECD countries' experiences, *OECD Economic Studies*, No.30, 2000/1.
- McDermott, P. (2010) Building open government. *Gov. Inf. Q.* 27, 401-413.
- Martin, J.P. (2000). What works among Active Labour Market Policies: Evidence from OECD countries' experiences, *OECD Economic Studies*, No.30, 2000/1.
- Meetup. (2017) "NYC Open Data," Meetup.com, accessed May 1, 2017, <https://www.meetup.com/NYC-Open-Data>.
- Mergele, L., and M. Weber. (2017) Public Employment Services Under Decentrallization: Evidence from a Natural Experiment, EEA.
- Miles, I. (2005) Innovation in Services. In *The Oxford Handbook of Innovation*, ed. J. Fagerberg, D.C.
- Morabito, V. (2015) *Big Data and Analytics for Government Innovation*, Springer, Switzerland.
- Nelson, S., Milberry, K., Richardson, J., Martin, M., Thomson, K. (2015) PartyX. <http://partyx.ca/>, Accessed 8 Jan 2015
- New Mexico Economic Development Department. (2016) Technology Research Collaborative Helps Develop and Commercialize High-Tech Products Born in NM, news release, April 4, 2016, <https://gonm.biz/news-andevents/news/technology-research-collaborative-helps-develop-and-commercialize-high-tech>.
- OECD. (2005) *OECD Employment Outlook*, Chapter 5: Public Employment Services: Managing Performance, Paris.

- OECD. (2006) Activation Strategies and Performance of Employment Services in the Netherlands, Germany and the United Kingdom, Paris.
- OECD. (2011) Divided We Stand. Why Inequality Keeps Rising. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2012) The jobs potential of a shift towards a low-carbon economy. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2013) Job Creation through the Social Economy and Social Entrepreneurship. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2015) The Future of Productivity. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2016) OECD Science, Technology and Industry Outlook 2016. Paris: Organisation for Economic Co-operation and Development.
- Peneder, M. (ed.). (2009). Sectoral Growth Drivers and Competitiveness in the European Union. Brussels: Background Studies to the Competitiveness Report 2008, European Commission, DG Enterprise and Industry.
- Peters, B. (2016) Innovation, skills and job creation. In Science, Research and Innovation performance of the EU 2016. A contribution to the Open Innovation Open Science, Open to the World agenda, ed. European Commission, 276-319. Brussels: European Commission, DG Research and Innovation.
- Peters, B., B. Dachs, M. Dünser, M. Hud, C. Köhler, and C. Rammer. (2014) Firm Growth, Innovation and the Business Cycle. Background Report for the 2014 Competitiveness Report. Mannheim:
- Petrongolo, B. and C.A. Pissarides. (2001). Looking Into the Black-Box: A Survey of the Matching Function, *Journal of Economic Literature*, 39, 390-431.
- SmartGrid Consumer Collaborative. (2013), Consumer Information Kit for the Smart Grid, SGCC 2013.
<http://smartgridcc.org/wp-content/uploads/2014/01/SGCC-Consumer-Information-Kit-for-the-Smart-Grid.pdf>.
- Steinmueller, W.E. (2010) Economics of Technology Policy. In *Handbook of*

- Economics of Innovation, ed. B.A. Hall and N. Rosenberg, 1182-1218. Amsterdam: Elsevier.
- Surka, M., and E. Ridlington. (2016) Following the Money 2016: How the 50 States Rate in Providing Online Access to Government Data Spending, U.S. Public Interest Research Group Education Fund and Frontier Group, April 2016, <http://www.uspirg.org/sites/pirg/files/reports/USP%20FollowMoney16%20Report%20Apr16.pdf>.
- Tether, B.S., C. Hipp, and I. Miles. (2001) Standardisation and particularisation in services.
- TreePedia. (2018) Exploring the Green Canopy in cities around the world, <http://senseable.mit.edu/treepedia>
- United Nations Department of Economic and Social Affairs. (2006) World Urbanization Prospects: The 2005 Revision.
- Van Reenen, J. (1997) Employment and Technological Innovation: Evidence from U.K. Manufacturing Firms. *Journal of Labor Economics*, 15, pp. 255-84.
- Van Woensel, L., G. Archer, L. Panades-Estruch, and D. Vrascaj. (2015) Ten technologies which could change our lives. Potential impacts and policy conclusions. Brussels: Study by the European Parliament, Science and Technology Options Assessment (STOA).
- Verge, J. (2015) Missouri (Finally) Passes Data Center Tax Incentives, Data Center Knowledge, <http://www.datacenterknowledge.com/archives/2015/04/13/missouri-passes-data-center-tax-incentives/>.
- Vivarelli, M. (2012) Innovation, Employment and Skills in Advanced and Developing Countries: A Survey of the Literature, *Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor*.
- waze.com. (2018) Get the best route, every day, with real-time help from other drivers, <https://www.waze.com/ko>
- Weingast, B. R., K. A. Shepsle, and C. Johnsen (1981) The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics, *Journal of Political Economy*, 89, 642-664.
- Wikipedia. (2018) "Megacyi," Wikipedia.com, accessed Dec. 10, 2018.

- Wildasin, D. E. (1991) Income Redistribution in a Common Labor Market, *The American Economic Review*, 81, 757-774.
- World Economic Forum. (2017) *Data Driven Cities - 20 Stories of Innovation*, Global Future Council on Cities and Urbanization.
- World Economic Forum. (2018) *The Future of Jobs Report 2018*, Centre for the New Economy and Society, Switzerland.
- Zhang, L. and N. Mitton. (2011) *Advanced Internet of things*. 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, 1-8.

데이터로 연결하는 **사람·세상·미래**



Bridging DATA

제5편

엣지 컴퓨팅으로 변화되는 미래사회

NIA 한국정보화진흥원



CONTENTS

I. 개 요	311
II. 컴퓨팅 환경과 엣지 컴퓨팅의 배경	315
1. 클라우드 컴퓨팅의 현재와 미래	315
2. 사물인터넷의 확대와 데이터 분석의 중요성	320
3. 인공지능의 실시간 처리의 중요성	322
III. 엣지 컴퓨팅의 개념과 특징	323
1. 엣지 컴퓨팅의 정의	323
2. 엣지 컴퓨팅의 탄생	326
3. 엣지 컴퓨팅의 구조	330
4. 엣지 컴퓨팅의 특징	337
5. 엣지 컴퓨팅의 사업적 의미	341
6. 엣지 컴퓨팅 관련 개념과 비교	346
IV. 엣지 컴퓨팅 관련 기술	351
1. 엣지 컴퓨팅의 요소 기술	351
2. 엣지 컴퓨팅의 실시간 처리	354
3. 엣지 컴퓨팅의 보안	356
4. 엣지 컴퓨팅의 통신과 네트워크	359
5. 엣지 컴퓨팅의 표준	363



CONTENTS

V. 엣지 컴퓨팅의 응용 분야	365
1. Smart Factory	367
2. Smart City	369
3. Smart Grid	371
4. Smart Home	372
5. Connected Car	373
6. 의료 서비스	375
7. 증강현실	377
VI. 엣지 컴퓨팅에 대한 기업의 대응	378
1. Amazon	379
2. Microsoft	380
3. Dell	381
4. HPE	383
5. CISCO	385
6. GE	386
VII. 결 론	387
참고문헌	389



개 요

인공지능(AI: Artificial Intelligent)과 사물인터넷(IoT: Internet of Things)으로 모든 사람과 사물이 인터넷에 연결되면서 이로 인해 생성되는 데이터가 기하급수적으로 증가하고 있다. 17년 4월 IT 시장조사기관 IDC가 발간한 ‘데이터 에이지 2025(Data Age 2025)’ 백서 자료에 의하면 2025년에는 전 세계 데이터 총량이 현재보다 10배 늘어난 163 제타 바이트(ZB)에 달할 것으로 예측하고 있다.

이런 데이터 폭증의 원인은 거의 모든 종류의 전기기기가 IoT 디바이스가 되었기 때문이다. 이제는 일반화된 스마트폰, 스마트TV와 같이 인터넷에 연결된 다양한 어플라이언스 제품, 거리에 설치된 CCTV, 대기 센서도 모두 IoT 디바이스가 되었다. 특히, 산업적으로 제조 공정의 단계마다 제품의 생산 과정을 모니터하기 위해 설치된 다양한 센서들이 모두 IoT 디바이스이다. 네트워크에 연결되는 사물의 수는 몇 년 안에 수십억 개 이상으로 증가할 것으로 예상된다. 이들 수많은 디바이스에서 계속해서 데이터가 생성되고 있다.

사실 과거에도 데이터는 계속 생성되었지만, 데이터 수집에 대한 필요성을 크게 느끼지 않았고, 또 데이터를 수집하고 저장을 하려면 많은 기술적 문제와 비용이 발생하기 때문에 꼭 필요한 데이터가 아니면 수집하지 않았다. 하지만 이제 클라우드를 통해 대규모 데이터의 수집과 저장이 가능해졌고, 빅데이터 분석 기술이 발전하게 됨에 따라 데이터에서 다양하고 유의미한 통찰(Insight)을 도출해 내면서 데이터의 중요성이 인식되었다.

결과적으로 많은 기업이 다양한 데이터를 수집하고 분석하게 되었고, 데이터 기반의 경영을 하게 되었다. 이제는 클라우드의 컴퓨팅이 일반화가 되어 다양한 클라우드 서비스가 제공되고, 기업들의 클라우드 전환이 빠르게 진행되면서 클라우드에 대한 인식이 개선되고, 진보된 IT 환경으로 자리매김하고 있다.

그러나, 이러한 변화보다 더 빠르게 폭증하는 데이터의 양으로 인해 여전히 데이터 처리 속도는 느리고, 네트워크의 대역폭은 부족한 상황이다. 즉, 수많은 디바이스에서 생성되는 데이터의 양이 증가함에 따라 데이터를 중앙 집중화된 데이터센터로 전송하는 클라우드 컴퓨팅에 병목 현상이 발생하고 있다. 예를 들어 매초마다 여객기 보잉 787에서는 약 5기가바이트의 데이터가 생성되지만, 비행기와 지상의 위성 또는 기지국 사이의 대역폭으로는 이 데이터를 전송하는데 부족하다. 또한, 멀지 않은 시기에 자율주행차의 경우, 매초마다 1기가바이트의 데이터가 생성될 것이며, 차량을 운전하기 위해서는 매 순간 실시간으로 정확한 결정을 내려야 한다. 그런데 차량에서 수집된 데이터를 처리하기 위해 모두 클라우드로 데이터를 보내야 한다면 응답 시간이 너무 길어지게 된다. 그리고 현재 무선 네트워크 구조는 한 지역에서 서비스할 수 있는 차량의 수가 한정되었기 때문에 동시에 여러 대의 차량을 제어하는 것이 문제가 될 수 있다.

즉, 데이터가 다양한 디바이스에서 발생하고, 실시간으로 수집되면서 수집된 데이터들을 분석 처리하기 위한 데이터센터의 부담이 가중되고 있다. 대규모 데이터로 인해 데이터센터에서 저장과 처리를 위한 비용이 급격히 증가하는 것은 물론이고, 데이터센터에서 처리할 수 있는 용량 이상으로 데이터가 몰리면 정체 현상이 발생하게 된다. 정체 현상이 발생하면 사용자 요청에 따라 빠르게 데이터를 분석하고, 보내는 것이 불가능해진다. 특히, 빠른 응답을 요구하는 서비스 시나리오에서는 기존의 클라우드 컴퓨팅 환경은 적당하지 않다.

그리고 강력해지는 모바일 디바이스와 모바일사용량의 증가는 기존 클라우드 컴퓨팅에 새로운 과제를 제시한다. 클라우드 컴퓨팅 패러다임에서 최종 기기는 대개 데이터 소비 역할을 한다. 예를 들어 스마트폰으로 YouTube 비디오를 시청하는 경우이다. 하지만, 요즘 모바일 사용자는 데이터의 소비뿐만 아니라, 데이터의 생산을 하고 있다. 예를 들어, 스마트폰을 사용하여 사진을 찍고, 비디오 녹화를 한 다음 유튜브, 페이스북, 트위터, 인스타그램과 같은 클라우드 서비스로 공유하는 것이 매우 일상적인 일이 되었다. 게다가 이러한 사용자의 패턴이 일반화되고 있다. 매분마다 유튜브 사용자들은 72시간 분량의 새로운 비디오 콘텐츠를 업로드하고, 페이스북 사용자는 거의 250만 개의 콘텐츠를 공유하며,

트위터 사용자는 거의 30만 번의 트윗을 하며, 인스타그램 사용자는 약 22만 장의 새로운 사진을 게시한다. 특히, 이미지나 비디오 클립은 크기가 상당히 크기 때문에 업로드에 많은 네트워크 대역폭을 사용한다.

결과적으로 IoT 디바이스에서 생성된 엄청난 데이터는 기존의 클라우드 환경에서 처리하는데 한계가 있고, 더 짧은 응답 시간, 더 효율적인 네트워크 처리를 위한 방안이 요구된다. 이러한 결론에 따라 엣지 컴퓨팅(Edge Computing)이 부상하고 있다. 엣지 컴퓨팅은 데이터가 수집되는 끝단(엣지)에서 바로 데이터를 분석하고 이를 다시 현장에 적용한다. 예를 들어 센서가 부착된 하드웨어에 데이터 처리를 위한 컴퓨팅 자원을 따로 두어 데이터 분석, 현장으로 결과를 보내는 것이다.

산업 사물인터넷(IIoT: Industrial Internet of Things)처럼 실시간으로 데이터를 처리해야 하는 사용자 요구가 늘어나고 있는데 원격지에 있는 클라우드 서비스로는 이러한 요구를 만족시키기에 무리가 있기 때문이다. 따라서 엣지 컴퓨팅이 현실적인 해결책이 될 수 있다. 물리적인 거리가 가까우면 클라우드로 데이터를 전송하고 결과를 기다리는 것보다 훨씬 빠르게 대응할 수 있다. 즉, 엣지 컴퓨팅에서는 클라우드에서 불가능했던 실시간 대응이 가능하다.

엣지 컴퓨팅이 적용되는 대표적인 사례로 자율주행차가 있다. 자율주행차는 차량에 부착된 각 센서에서 데이터를 수집해 주변 지형, 도로 상황, 차량 흐름 현황 등을 파악하고, 주행 중 일어날 수 있는 다양한 상황에 신속하게 대처할 수 있어야 한다. 앞에서 달리던 차의 속도가 느려지면 이에 맞게 속도를 낮춰 충돌을 방지하며, 빨간 신호가 켜졌을 경우 신속하게 차량을 정지시켜 교통 법규를 준수해야 한다. 만약 자율주행차가 엣지(차량)에서의 데이터 처리가 아닌 중앙에 있는 클라우드 데이터센터에서 처리를 기다려야 한다면, 돌발 상황이 발생하더라도 즉시 대응하지 못해 사고로 이어질 수 있다. 물론 현재의 컴퓨팅 환경에서도 자율자동차의 구현이 가능하지만, 자율주행차량의 속도를 높여 고속 주행을 하려면 실시간 엣지 컴퓨팅이 필수적이다.

자율주행차뿐만 아니라 공장에서 제조 과정의 모니터링과 이에 따른 대응, 가상현실을 활용한 사용자 인터페이스 등 실시간 대응이 필요한 시나리오는 계속 늘어날 전망이다.

이러한 배경에서 글로벌 정보통신 시장 조사업체 가트너(Gartner)는 매년 기술 트렌드를 선정 발표하는데, 2018년, 2019년 연속적으로 10대 기술로 엣지 컴퓨팅이 선정되었다.

본 보고서에서는 지능형 사물인터넷으로 급격한 데이터의 부상에 따라 주목을 받는 엣지 컴퓨팅을 소개하고, 엣지 컴퓨팅을 적용한 산업적 변화를 다루고자 한다. 또한 엣지 컴퓨팅이 탄생하게 된 배경을 살펴보고, 엣지 컴퓨팅의 개념과 특징을 설명한다. 이후 엣지 컴퓨팅에 사용되는 기술과 이를 활용한 응용 분야를 설명하고, 마지막으로 엣지 컴퓨팅에 대한 기업들의 대응을 살펴보도록 한다.

1 >> 클라우드 컴퓨팅의 현재와 미래

엣지 컴퓨팅은 클라우드 컴퓨팅(Cloud Computing)의 발전된 형태이므로 클라우드 컴퓨팅의 개요와 현황에 대해 먼저 간략하게 살펴보고자 한다.

인터넷에 연결 가능한 IoT 기기들이 다양해지고 규모가 대폭 증가하면서, 서버로 전송하고, 서버에서 처리해야 하는 데이터의 양이 급증하게 되었다. 이런 상황에서 인터넷을 통해 시간과 장소에 구애받지 않고 데이터를 수집, 분석, 처리할 수 있는 방식이 필요하게 되었다. 이에 따라 기존 IT 자원을 기업의 서버나 개인 PC처럼 직접 소유하던 방식에서 벗어나 필요한 시점에 원하는 양만큼 빌려 쓰는 클라우드 컴퓨팅이 등장했다.

클라우드 컴퓨팅은 '인터넷을 통해 필요에 따라(on-demand) IT 자원을 제공받는 서비스'이다. 기업 입장에서 클라우드는 다양한 혜택을 제공한다. “인터넷을 통해” 기업 내부이든지 외부이든지 대규모의 IT 자원을 확보해 놓고, 인터넷을 통해 접근할 수 있게 하여 규모의 경제를 통해 비용 효율화가 가능하다. 또한, 기업이 사전에 충분한 시간을 갖고 IT 자원을 확보하는 것이 아니라, “필요에 따라” IT 자원의 수요가 발생한 시점에 바로 또는 수 분 내에 IT 자원을 확보할 수 있게 되어 기업은 신속한 사업 대응이 가능하게 된다. 서버, 스토리지, 네트워크 같은 물리적인 자원뿐만 아니라 다양한 애플리케이션을 포함하여 기업이 사업에 필요한 다양한 “IT 자원”을 클라우드 형태로 활용할 수 있게 한다. 추가로 클라우드 사업자의 전문적인 운영을 통해 안정적인 서비스 제공이 가능하다. 결과적으로 기업은 클라우드로 인해 그동안 점차 늘어난 IT에 대한 부담을 덜게 되어 기업이 핵심 사업에 집중할 수 있게 되었다. 이러한 이유로 클라우드는 현대 기업들의 핵심 IT

전략으로 자리매김하고 있다.

기업이 클라우드를 활용하는 방식은 크게 퍼블릭 클라우드(Public Cloud)와 프라이빗 클라우드(Private Cloud), 커뮤니티 클라우드(Community Cloud) 그리고 하이브리드 클라우드(Hybrid Cloud)로 구분된다.

먼저, 퍼블릭 클라우드는 특별한 자격 조건 없이 일반 사용자 누구나 사용할 수 있는 클라우드 서비스이다. 클라우드 사업자는 일반인과 기업들이 자유롭게 사용할 수 있는 클라우드 IT 환경을 구축하여 제공한다. 기본적으로 전 세계를 대상으로 서비스하며, 서버와 스토리지는 물론 콘텐츠 전송 네트워크(CDN: Contents Delivery Networks) 같은 대규모 투자가 필요한 자원을 서비스한다. 클라우드 사업자는 대량의 자원을 많은 고객에게 나누어 제공하여 규모의 경제를 이룬다. 클라우드 사업자는 IT 자원을 직접 소유하고, 자신이 관리하는 데이터센터에 IT 자원을 직접 설치하고, 관리하고, 운영한다.

프라이빗 클라우드는 IT 자원이 하나의 조직, 즉 특정 기업이나, 정부 단체를 위해 만들어지고 제공되는 클라우드 서비스이다. 프라이빗 클라우드는 기업 내부나 기업이 위탁한 업체가 소유한 데이터센터에 인프라를 두고 각기 분리된 공간에 하나의 조직을 위한 독립적인 클라우드 인프라를 제공한다. 프라이빗 클라우드는 기업이 통제하지 못하는 외부에 IT 자원이 위치해서 생기는 보안 이슈를 해결하기 위해서 만들어진 클라우드 서비스이다.

커뮤니티 클라우드는 동일한 목적을 가지는 단체를 위해 제공되는 클라우드 서비스이다. 이 서비스를 이용하는 단체는 동일한 보안, 관리 요구사항이 있고, 하나의 클라우드 서비스 사업자에 의해서 제공된다. 퍼블릭 클라우드와 프라이빗 클라우드 사이의 중간적인 서비스라 볼 수 있다. 정부통합전산센터에 있는 G-클라우드(Government Cloud)가 대표적인 커뮤니티 클라우드이다. 정부통합전산센터는 중앙행정 기관이 필요한 IT 자원을 효율적으로 운영하기 위해 정부 소유의 데이터센터 내에 클라우드 서비스를 구축하고 정부 기관을 대상으로 서비스하고 있다.

하이브리드 클라우드는 기업이 자체 소유한 IT 자원을 스스로 통제가 가능한 공간에 설치하고, 운영하고, 사용하는 온프레미스(on-premise) 방식뿐만 아니라 클라우드 서비

스를 혼합하여 IT 자원을 사용하는 방식이다. 예를 들어 내부 IT 자산이 온프레미스 방식으로 존재하는 상황에서 효율적인 운영을 위해서 자체 프라이빗 클라우드를 구축하거나, AWS나 Azure를 활용하여 퍼블릭 클라우드 서비스를 도입하여 동시에 사용하는 방식이다. 이렇게 되면 온프레미스, 프라이빗 클라우드, 퍼블릭 클라우드를 같이 사용하는 하이브리드 클라우드 환경이 만들어진다.

다른 용어로 멀티 클라우드(Multi-Cloud)가 있다. 멀티 클라우드는 하이브리드 클라우드의 일종으로 2개 이상의 클라우드 사업자가 제공하는 2개 이상의 퍼블릭 또는 프라이빗 클라우드로 구성된 클라우드 환경이다.

지금까지 설명한 각 클라우드 서비스의 장단점을 비교하면 퍼블릭 클라우드는 규모의 경제로 인해 비용 절감 효과는 크지만, 자원을 공유하기 때문에 이로 인해 발생하는 문제점을 예방하려는 방안이 필요하다. 즉, 클라우드 사업자는 다른 사용자로 인해 또 다른 사용자의 서비스가 영향받지 않도록 기술적인 방안을 마련해야 하며, 클라우드 사용자에게 미리 정해진 방식으로만 서비스를 사용하도록 제약해야 한다.

반면에 프라이빗 클라우드는 독립적인 IT 자원의 사용을 보장하기 때문에 개별적인 맞춤형 운영이 가능한 장점이 있다. 그러나 프라이빗 클라우드는 시설 확장이나 글로벌 서비스를 위해서는 대규모 추가 투자가 이루어져야 하고, 무엇보다 특정 기업이나 단체만 지원하기 때문에 규모의 경제를 이루어 비용 절감을 하는 데 한계가 있다. 따라서 퍼블릭 클라우드와 비교하면 동일한 IT 자원을 사용할 때에 프라이빗 클라우드가 비용이 많이 든다.

각 클라우드의 용도를 살펴보면 퍼블릭 클라우드는 대고객 서비스 등 성능과 보안의 영향이 크지 않은 서비스에 주로 사용하고, 프라이빗 클라우드는 기업 내부 시스템같이 높은 성능과 높은 보안을 요구하는 시스템에 적합하다. 그러나 최근 퍼블릭 클라우드 사업자가 성능과 보안을 크게 강화되면서 기업 내부 시스템같이 중요 시스템도 퍼블릭 클라우드를 활용하는 사례가 점차 증가하고 있다. 이러한 흐름은 앞으로도 지속할 것으로 예측된다.

이러한 클라우드의 장단점과 활용 방안에 따른 기업의 클라우드 도입 전략을 생각해 보면 IT 예산이 많지 않고, IT 전문 인력도 거의 보유하고 있지 않는 중소기업의 경우

퍼블릭 클라우드 중심의 클라우드 전략을 추진하는 것이 효과적이다. 반면에 대기업은 고유의 요구사항이 있으므로 프라이빗 클라우드가 적합하다. 그러나 프라이빗 클라우드만으로는 실시간으로 변화하는 IT 수요에 대응하거나 비용 절감에 한계가 있어 퍼블릭 클라우드도 함께 사용한다. 따라서 대기업의 IT 환경은 하이브리드 클라우드 환경으로 발전하고 있다.

현재 클라우드 컴퓨팅의 수요는 꾸준히 증가하고 있고, 정보통신 분야뿐만 아니라 금융, 의료, 공공 등 다양한 산업 분야에 확대 적용되고 있다. 이로 인해 클라우드 컴퓨팅 시장이 급성장하고 있고, 글로벌 IT 기업들이 세계 중요 지역에 데이터센터 구축과 같은 인프라에 많은 투자를 하면서 클라우드 서비스 사업을 계속 확장해 나가고 있다. 클라우드 시장을 살펴보면 아마존웹서비스(AWS), 마이크로소프트 등 글로벌 대표 IT 기업들이 주도하고 있다.

시장조사기관 시너지리서치그룹이 발표한 2018년 1분기(1월~3월) 전 세계 클라우드 인프라 서비스 매출은 150억 달러(한화로 약 16조 원)에 달한다. 이는 전년 같은 기간 대비 51% 늘어난 수치다. 1위는 아마존의 AWS로 30% 이상의 시장 점유율을 유지하고 있고, 2위는 마이크로소프트의 애저(Azure)로 13%의 점유율을 보인다. IBM은 8%, 구글은 6%의 점유율을 확보했다. 특히, 아마존 AWS의 매출이 전년 대비 49% 늘어난 54억 달러를 기록했고, AWS의 2018년 1분기 영업이익은 아마존 전체 영업이익의 73%를 차지했다. 마이크로소프트도 애저 매출이 93%나 증가했다고 밝혔다. 이처럼 글로벌 정보통신 시장은 클라우드를 중심으로 성장하고 있다.

클라우드 세계 시장을 살펴보면 미국이 가장 큰 시장을 차지하고 있다. 그리고, 중국의 잠재적 수요를 바탕으로 아시아 시장이 급성장하고 있다. 2017년 기준 국가별 비중은 미국이 62.2%, 서유럽 19.3%, 아시아태평양 8.6%(한국 0.5%), 일본 3.5%다. 반면 지역별 성장률('16~'21)을 보면 아시아태평양 32.6%, 일본 24.3%, 서유럽 22.7%, 미국 19.4% 순으로, 아시아태평양이 가장 높다.

클라우드 전문 기업인 베스핀글로벌이 수행한 2018년 국내 기업의 클라우드 도입 현황을 조사한 설문 결과에 의하면 기업의 36%가 클라우드를 도입 중이거나 완료했고, 응답자의 57%는 클라우드 도입을 계획하고 있었다. 즉, 전체 응답자의 93%가 클라우드를 도입

하거나 고려하고 있는 것으로 나타났다. 또한, 엔터프라이즈 기업의 46%, 중견·중소기업의 34%가 하이브리드 클라우드를 선호하는 것으로 조사됐다. 또한, 퍼블릭 클라우드를 사용하는 응답자 중 하나의 사업자가 제공하는 단일 클라우드를 사용하는 비중은 57%, 2개 이상의 멀티 클라우드를 사용하는 비중은 43%로 나타나 멀티 클라우드가 점차 확산 중인 것으로 판단된다.

공공 부분을 살펴보면 각 국가가 정책적으로 클라우드 활용을 촉진하고 있다. 우리나라는 2015년 9월 세계 최초로 클라우드 발전법을 제정하여 공공부문의 비용 절감과 업무혁신, 민간산업 육성을 도모하고 있다. 미국은 2016년 연방정부 정보화 예산의 8.5%(8.6조 원)를 클라우드에 배정했다. 특히 2017년 5월에 트럼프 정부의 'Cloud Only 행정명령'으로 전 정보화 시스템의 클라우드 전환을 의무화했다. 그리고 높은 보안 수준을 요구하는 CIA 같은 기관에서도 퍼블릭 클라우드 서비스를 이용하고 있다. 영국도 정부 정보화 예산의 10%(1.1조 원)를 클라우드에 활용하고 있으며, 이를 위해 공공조달 거버넌스를 구축하였다. 일본은 2021년까지 모든 정부 정보시스템의 클라우드화를 통해 운영비용을 30% 절감할 계획이다.

특히, 국가 차원에서 클라우드 산업 육성을 지원하는 중국은 2021년까지 중국 클라우드 시장을 71조 원으로 육성하고, 기업의 경쟁력 강화를 위해 클라우드컴퓨팅발전 3개년 계획('17~'19년)을 발표했다. 이를 위해 공공서비스 플랫폼을 구축하고, 소프트웨어 기업의 클라우드화 전환을 지원하고, 클라우드 핵심기업 육성을 추진하고 있다.

클라우드의 미래를 생각해 보면 인터넷에 연결된 사물, 사람이 증가할수록 또 제공되는 서비스가 고도화될수록 대규모 데이터의 수집, 저장, 처리를 위한 컴퓨팅 인프라로서의 클라우드의 역할은 더욱 증가할 것이다. 이로 인해 진화된 형태의 클라우드 서비스가 나오고 있다. 그중에 대표적인 것이 인터클라우드와 엣지 컴퓨팅이다. 인터클라우드 컴퓨팅(inter-cloud computing)은 둘 이상의 클라우드 서비스 제공자 간의 자원을 연결하고, 서로 연계하여 컴퓨팅 수요에 따라 자원이 동적으로 확장할 수 있게 한다. 클라우드 컴퓨팅의 진화된 형태인 엣지 컴퓨팅은 본 보고서에서 다루는 주제로 지역이나 근거리에서 데이터를 처리해 중앙집중형 클라우드의 부하를 해소하려는 방안이다.

2 >> 사물인터넷의 확대와 데이터 분석의 중요성

사물인터넷은 지난 몇 년간 빠르게 일상생활을 변화시키고 있다. 사물인터넷은 스마트폰, 자동차, 센서, 액추에이터 등 기존의 기기를 데이터센터와 연결, 통신하고, 정보를 교환하여 새로운 부가 기능을 제공한다. 스마트 교통, 스마트 시티, 스마트 그리드, 스마트 헬스케어 같은 최근의 서비스들은 사물인터넷 기능에 기반을 둔 서비스이다. 따라서 사물인터넷은 미래에 보편적인 기술로 사용자의 일상생활에 큰 영향을 미칠 것이다.

많은 영역에서 사물인터넷이 적용되고 있고, 그 대상이 점점 확대되고 있는데 그 이유는 기업들의 경제적, 사업적 목표에 부합하기 때문이다. 그중에서도 인더스트리 4.0 또는 산업인터넷으로 산업현장에서 센서와 사물 간의 연결을 통해 생산성 향상을 위해 사물인터넷의 활용을 활발히 하고 있다. 그 예로 기업 IoT의 대표적인 솔루션 GE의 프리디스(Predix)는 전 세계 생산 현장의 기계, 센서를 통해 생성된 데이터를 수집, 분석하며, 자동차 부품 기업 보쉬(Bosch)도 IoT 기반의 시스템을 전 세계 공장에 확대할 계획을 세우고 있다.

반면에 일반 사용자를 대상으로 하는 소비자 가전의 사물인터넷은 아직 초기 단계에 머무르고 있다. 스마트 온도계, 체중계, 전등, 가구 등을 제외하면 소비재의 IoT 활용 예는 아직 많지 않다. 소비자 대상 사물인터넷 서비스가 활성화되려면, 사물 간에 연결이 원활하게 되는 확장성이 보장되어야 다양한 서비스가 가능한데, 현재까지는 매우 제한된 제품, 일부 브랜드 제품끼리만 연결할 수 있다. 사실 확장성보다 IoT를 통해 소비자가 체감하는 가치나 혜택이 부족하여 기업들의 비즈니스 모델이 아직 뚜렷하지 않은 것이 더 근본적인 문제이다. 따라서, 새로운 IoT 서비스를 발굴하고, 이 서비스의 소비자 요구사항을 지속 개선하여, 사용자의 경험을 통해 가치를 체감하여야만 IoT 시장이 활성화될 수 있다.

자동차, 기차, 의료장비, 가로등 등 수많은 사물이 실시간 데이터를 생성하고, 분석하는 미래 환경에서는 사람보다 기계가 생성하는 데이터의 양이 빅데이터의 대부분을 차지하게 될 것이다. 또, 인공지능과 연동해 데이터를 처리, 해석, 활용하는 일이 일반화될 것이다.

이러한 추세에 따라 사물인터넷 데이터의 기계학습, 음성 인식 같은 편리한 사용자 인터페이스, 연결된 사물 간의 협력을 통한 새로운 서비스들이 선보이게 될 것이다. 또, 교통, 전력 등 사회 인프라의 지능화에 사물인터넷이 적용되면서 편리하고, 안전한 사회로의 발전이 계속될 것이다.

결론적으로 사물인터넷이 산업 전반과 모든 디바이스에 적용되면서 이를 통해 생산되는 대규모 데이터의 분석이 필요하고, 분석된 결과를 통해 다양한 스마트 서비스가 제공된다. 앞으로 데이터의 양은 점점 더 많아지고 다양하고 지능적인 분석에 대한 요구도 늘어날 것이다. 이로 인해 IoT 환경에서 빅데이터 분석을 처리하는 환경으로 엣지 컴퓨팅이 주목 받게 될 것이다.

3 » 인공지능의 실시간 처리의 중요성

사물인터넷은 모든 사람과 사물이 인터넷을 통해 서로 소통하고 새로운 가치를 창출할 수 있는 기술이며, 정보의 확산과 활용을 가능하게 하는 중요한 도구이다. 이러한 사물인터넷에 더해 인공지능 기술을 활용하면 사물이 더욱 똑똑해져 스스로 생각하고 판단하는 지능형 시스템으로 진화한다.

인공지능은 전통적인 데이터 분석의 한계를 벗어나 자동으로 데이터의 가치를 추출하게 된다. 즉, 사람이 해야만 했었던 지능적인 일을 알고리즘이 수행하여 생산성을 높일 수 있게 된다. 따라서 미래 인공지능의 진화 방향은 사물인터넷, 빅데이터, 클라우드가 결합하여 복잡하고, 대규모 데이터의 처리가 필요한 문제도 쉽게 해결할 수 있게 된다. 이때 클라우드는 단순한 컴퓨팅 인프라가 아니라 인공지능이 작동하는 인프라로 그 의미가 변화된다. 클라우드는 의사, 변호사 같은 전문가 집단이 특정 분야의 문제를 해결하는 지능을 보유하게 되어, 서비스 형태로 지능을 활용할 수 있는 환경을 제공한다. 결국, 사물인터넷의 진정한 가치는 인공지능이 제공하는 서비스를 통해 만들어진다.

기존 클라우드 환경에서 사물인터넷을 활용한 인공지능은 모든 데이터를 중앙 집중된 클라우드 컴퓨팅 환경에 수집해 놓고, 분석 처리한다. 이 경우, 모든 디바이스들이 중앙 집중된 데이터센터에 연결되어 데이터를 주고받다보니 정체가 일어나 너무 혼잡하고 속도도 느려진다.

이러한 문제의 극복 방안으로, 사물이 지능을 가지고 스스로 학습과 판단함으로써, 많은 양의 데이터를 굳이 먼 곳에 있는 대용량 데이터 서버에 저장하지 않고 데이터가 발생한 인근 지점에서 더 빠르게 판단하고 반응하는 것이 있다. 예를 들어 얼굴인식을 위해 스마트폰이 매번 클라우드에 연결하여 인식하는 것보다 스마트폰에 지능을 이식하여 스마트폰에서 바로 수행하는 것이 효과적이다. 이것이 인공지능이 필요한 상황에서 엣지 컴퓨팅이 사용되는 중요한 이유이다.



엣지 컴퓨팅의 개념과 특징

1 » 엣지 컴퓨팅의 정의

엣지 컴퓨팅이라는 용어는 2002년경에 만들어졌다. 일부 기업들이 네트워크를 통해 대용량 콘텐츠를 전송하기 위해서 사용하는 CDN(Contents Delivery Network)에서 네트워크 속도를 높이기 위해 엣지 서버(edge server)를 도입하였다. 엣지 서버는 CDN에서 적은 비용으로 대규모의 콘텐츠를 효율적으로 전송하기 위해 지역적으로 분산된 소형 서버들을 의미한다. 따라서, 엣지 컴퓨팅이라는 의미가 도입된 초기의 ‘엣지’는 지역적으로 분산된 CDN 서버라는 의미로 한정되었으나, 최근에는 엣지의 개념이 크게 확장되어 엣지 컴퓨팅으로 일반화되었다. 즉, 엣지가 CDN의 서버로 제한되지 않고, 네트워크에 연결된 수많은 사용자 디바이스와 센서를 포함하게 되었다. 여기에 클라우드, 사물인터넷, 빅데이터, 인공지능 같은 기술이 결합하여 그 개념이 더욱 확대되었다.

엣지 컴퓨팅(Edge Computing)은 중앙 집중화된 클라우드 환경에서 모든 처리가 이루어지는 것과는 다르게 스마트 디바이스(Smart Device) 또는 엣지 디바이스(Edge Device)에서 연산과 처리가 부분적으로 또는 완전히 수행되는 분산 컴퓨팅으로 정의된다. 엣지 디바이스는 데이터를 직접 생산하는 단말기나 단말기와 근접한 위치에 있는 각종 디바이스를 의미한다.

사실 엣지 컴퓨팅의 등장은 클라우드 컴퓨팅이 10여 년 전 등장한 이후 기술 발전 속도가 급격히 빨라지면서, 최근에 분산 컴퓨팅 개념을 도입하면서 만들어진 것이다. 따라서, 엣지 컴퓨팅이 기존의 클라우드 컴퓨팅을 대체하는 개념이 아니라 발전된 개념 또는 서로 보완하는 개념으로 인식된다. 즉, 엣지 컴퓨팅으로 사물인터넷에 있는 디바이스는 생성된

데이터 중 시급성을 요구하는 데이터를 처리하고, 클라우드 컴퓨팅은 중앙집중화된 컴퓨팅과 스토리지로부터 대규모 데이터를 처리한다. 결국, 엣지 디바이스와 기존의 클라우드 데이터센터가 서로 역할을 분담해서 전체적인 시스템의 성능을 높이는 것이다.

엣지 컴퓨팅과 유사한 개념으로 CISCO가 2012년에 제안한 포그 컴퓨팅(Fog Computing)이 있다. 엣지 컴퓨팅이 디바이스 가까이 있는 컴퓨팅 디바이스에 중점을 둔다면, 포그 컴퓨팅은 게이트웨이나 라우터 같은 네트워크 디바이스를 더 강조한 개념이다. 즉, 포그 컴퓨팅은 네트워크 패브릭(Fabric) 개념으로, 데이터가 생성되는 엣지로부터 클라우드의 데이터센터까지 네트워크를 확장한 개념이다. CISCO가 네트워크 분야의 강점이 있으므로 이를 기반으로 새로운 개념을 도입한 것으로 보는 것이 타당하다.

시스코, 인텔, 마이크로소프트, Dell이 주도하는 오픈포그(OpenFog) 컨소시엄은 포그 컴퓨팅을 클라우드에서 사물까지 연결된 컴퓨팅, 스토리지, 제어 및 네트워킹의 자원과 서비스를 제공하는 시스템으로 정의하였고, 포그 컴퓨팅을 위한 참조 아키텍처를 제시하였다.

엄밀한 의미에서 포그 컴퓨팅과 엣지 컴퓨팅은 출발이 다르지만, 최근에는 엣지와 포그의 개념이 서로 발전하고, 컴퓨팅 환경이 거의 유사해지면서 두 개념을 유사한 것으로 보고 있다. 사실 엣지 컴퓨팅이라는 용어가 논리적이거나 수학적으로 정의된 것이 아니라 ICT 분야에서 통용되는 용어로 출발하여 그 개념을 놓고 차이를 설명하는 것은 한계가 있다. 물론 일부 저자에 따라서는 포그 컴퓨팅과 엣지 컴퓨팅을 엄밀히 구분하기도 한다. 하지만 본 보고서에서는 중앙 집중된 기존의 클라우드 컴퓨팅이 분산된 클라우드 컴퓨팅으로 발전한 개념을 엣지 컴퓨팅이라고 지칭하였으며, 포그 컴퓨팅과 의미를 구분하지 않았다.

엣지 컴퓨팅의 정의에 따라 엣지 컴퓨팅을 구성하는 요소들은 계층을 이루고 있다. 즉, 엣지 컴퓨팅의 전체 시스템은 1) 클라우드 (Cloud, 중앙 데이터센터, 빅데이터 플랫폼), 2) 포그(Fog, 지역별 인프라, 게이트웨이), 3) 엣지(Edge, 스마트 기기, 단말)의 3층 구조로 나누어 관리하게 된다. 자세한 엣지 컴퓨팅의 계층별 구조는 다음 절에서 설명한다.

이러한 엣지 컴퓨팅 구조에서 IoT 애플리케이션의 동작 과정을 살펴보면, 먼저 네트워크

크 가장자리에서 위치하는 엣지 디바이스인 IoT 장치에서 데이터를 수집한다. 그다음 아래와 같은 전략으로 데이터를 최적의 분석 장소로 이동시킨다.

- 시간에 민감한 데이터는 데이터를 생성한 엣지 디바이스나 가장 가까운 포그 노드에서 분석한다. 예를 들어 발전소에서 생산된 전력을 여러 변전기를 거쳐 가정까지 전달하는 시스템을 생각해 보자. 특정 지역을 담당하는 변전기는 그 지역에 전력 공급 상황을 모니터링하고 사고를 감지하는 역할을 한다. 만일 특정 지역의 가정에서 화재가 발생하면 가장 말단에 위치한 변전기가 그 지역의 전기를 차단하는 것이 가장 효과적이다. 따라서, 시간에 민감한 데이터는 가장 가까운 포그 노드가 문제의 징후를 찾고 제어한다.
- 시간에 덜 민감한 데이터는 추세 분석, 상관 분석과 장기 저장을 위해 클라우드로 전송된다. 앞서 설명한 예에서 지역별 전기의 소비 변화를 분석하기 위해서는 클라우드로 데이터를 보내 분석한다. 사실 수많은 엣지들이 각각 과거 데이터를 분석하고 저장하는 것은 비효율적이어서 정기적으로 데이터를 요약해서 클라우드로 전송하는 것이 효과적이다.

2 » 엣지 컴퓨팅의 탄생

본 장에서는 사물인터넷, 인공지능 시대에 기존의 클라우드 문제로 엣지 컴퓨팅이 탄생한 배경을 먼저 살펴보도록 하자.

● 대용량 데이터의 처리 한계

인공지능과 사물인터넷으로 모든 사람과 사물이 연결되면서 다양한 경로로 수집된 데이터들은 데이터 분석을 위해 데이터센터로 모이지만 처리할 수 있는 데이터의 양은 한정되어 있어 지속해서 데이터가 몰리면 정체 현상이 발생하게 되며, 이로 인해 빠르게 분석을 처리하는 것이 불가능해진다.

클라우드 컴퓨팅은 중앙집중형 컴퓨팅 처리로 인해 데이터 전송과 처리, 결과 수신까지의 응답 시간이 길어지면서 데이터를 저장하고 불러오는데 병목 현상이 일어난다. 즉, 수많은 스마트 기기 또는 지능형 기기들이 인터넷으로 연결되어 발생하는 모든 데이터를 클라우드로 보내서 처리하거나 분석하는 것이 비효율적이게 된다. 이에 대한 단순한 해결책으로 데이터센터의 규모를 늘려 몰려오는 데이터의 수요를 처리할 수도 있지만, 그 경우 늘어나는 비용은 점점 감당하기 어려워진다.

● 실시간 처리에 대한 요구

새로운 서비스가 등장하면서 실시간 처리에 대한 요구가 생겨났다. 카메라를 이용해 물체를 인식하며 이동하는 자율주행차를 예를 들면 통상적인 차량의 속도에서 물체 인식에 필요로 하는 시간은 25ms~50ms이다. 차량에 물체를 인식하기 위한 충분한 컴퓨팅 자원을 설치하는 것은 비용과 이동에 제약이 있고, 실시간 교통 정보 같은 외부 데이터를 통합해야 하므로 클라우드 서비스를 이용한다. 그러나 현재 클라우드 인프라를 사용하면 디바이스와 클라우드 간에 정보를 보내고 결과를 받는데 걸리는 시간이 100ms를 넘어가게 되어 자율주행차처럼 대기 시간에 민감한 서비스의 요구조건을 만족하지 못한다.

비디오 스트리밍은 큰 트래픽을 발생시켜 비디오를 이용한 실시간 인터페이스를 제공하는 서비스도 클라우드를 활용한 실시간 서비스가 어렵다. 화려한 그래픽이 제공되는 게임이나 가상현실(VR: Virtual Reality), 현실 세계에 가상정보를 더해 보여주는 증강현실(AR: Augmented Reality)이 대표적인 예이다.

이 같은 데이터 폭발 시대에는 중앙 집중식 데이터센터에 기반을 둔 클라우드 컴퓨팅이 모든 연산을 처리할 수 없어, 데이터를 생성하는 맨 끝단 엣지에서의 처리가 중요해진다. 즉, 데이터센터와 엣지 디바이스의 분업을 통해 성능을 높이는 방법이 필요하다. 우리가 가장 흔하게 접하는 엣지 디바이스는 스마트폰이다. 스마트폰은 SNS, 사진, 동영상, 지문 인식 등 무수한 데이터를 생성하지만 상당한 데이터 처리도 가능한 컴퓨팅 기기이다. 또한, 제조공장에 적용된 온도, 습도 측정을 위한 디바이스와 일사량, 풍속 등을 측정하는 태양광과 풍력 발전기의 디바이스들도 모두 엣지이다. 이들 디바이스는 클라우드와 비교하면 충분하지는 않지만, 연산 능력을 갖추고 있어 부분적인 데이터 처리가 가능하다.

엣지 컴퓨팅은 이 같은 기존의 클라우드 컴퓨팅의 한계를 극복하기 위한 대안으로 떠올랐다. 엣지 컴퓨팅에 기반한 시스템은 디바이스와 가까운 네트워크 인프라의 끝단에 엣지 노드들을 배치하여 클라우드 서비스를 제공한다. 데이터가 생성된 위치에서 가까운 곳에 있는 스마트 기기에서 중요한 데이터를 실시간으로 처리하거나 저장하고, 꼭 필요한 데이터만 데이터센터나 클라우드 스토리지에 저장해 효율적으로 관리하는 방식이다. 기존에 클라우드 서버에서 실행되었던 클라우드 애플리케이션 중 일부를 디바이스 또는 디바이스에 가까운 노드로 이동하여 처리하는 것이다.

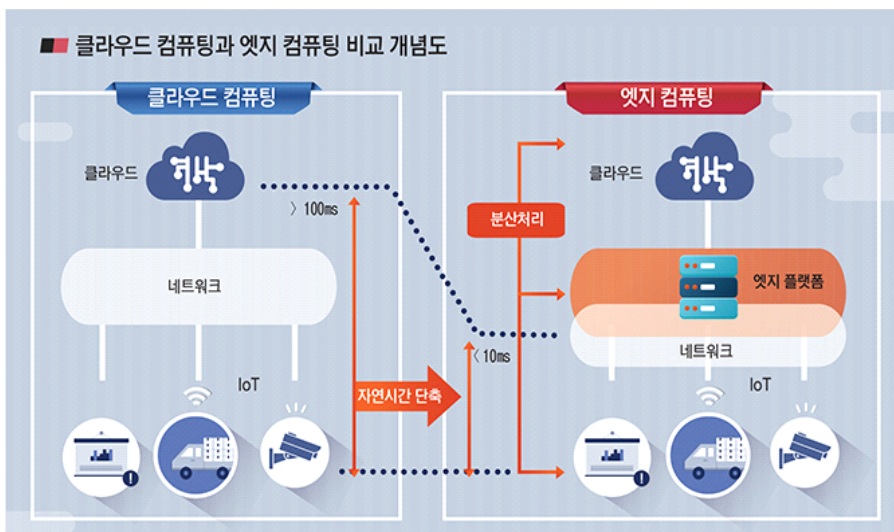
따라서 엣지 컴퓨팅은 데이터가 수집되는 엣지에서 바로 데이터를 처리하고, 연산 결과를 현장에 바로 적용할 수 있어 클라우드 컴퓨팅 환경에서 데이터를 전달하고 분석된 결과를 기다리는 것보다 훨씬 빠르게 요청에 응답할 수 있다. 또한, 클라우드 서버에 집중되었던 처리를 엣지 노드들로 분산시킴으로써 시스템의 확장성을 높이고 잠재적인 네트워크 병목을 해소한다.

엣지 컴퓨팅이 탄생한 또 다른 배경에는 클라우드 컴퓨팅의 중요한 이슈인 보안 문제가 있다. 개인정보 보호에 관한 산업 규제에 의해 특정 유형의 데이터 전송과 저장에 대한

제약조건이 커지고 있다. 클라우드는 기본적으로 중앙집중형이기 때문에 개인정보의 전송과 저장 때문에 보안 문제가 지속해서 제기되어 왔다. 엣지 컴퓨팅에서는 엣지 디바이스에 서만 개인정보의 저장과 처리를 담당하고, 클라우드로 전송을 최소화하면, 클라우드의 보안 문제를 해결할 수 있다.

그림 1은 지금까지 설명한 클라우드 컴퓨팅과 엣지 컴퓨팅을 개념적으로 비교하였다. 클라우드 컴퓨팅은 중앙집중형이라서, 지연시간이 길지만, 엣지 컴퓨팅은 분산처리를 하여 클라우드 컴퓨팅 부담을 감소시키고, 지연시간을 단축하고 보안을 강화한다. 이를 위해 엣지 컴퓨팅은 디바이스 가까이 있는 엣지 플랫폼의 역할이 중요해진다.

〈그림 1〉 클라우드 컴퓨팅과 엣지 컴퓨팅의 비교 개념도



출처: 정보통신신문, 엣지 컴퓨팅 급부상...IoT 네트워크 구조 바뀐다, 2018.8.7.

아래 표는 클라우드 컴퓨팅과 엣지 컴퓨팅을 비교한 표이다. 지연시간(Latency)은 엣지 컴퓨팅이 짧아 실시간 처리가 되고, 하드웨어 측면에서 보면 클라우드 컴퓨팅은 대규모 확장 가능한 저장공간과 컴퓨팅 용량을 보유하고 있고, 엣지 컴퓨팅은 제한적인 디바이스 이지만 이를 분산 시스템으로 연결하여 컴퓨팅 환경을 제공하는 특징이 있다. 클라우드 컴퓨팅은 상용 인터넷 연결을 기반으로 하고 있으나, 엣지 컴퓨팅은 인접한 로컬 네트워크

를 기반으로 하고 있다. 또한, 클라우드 컴퓨팅은 사용자의 요청을 처리하기 위해서 여러 번의 단계를 거쳐야 하지만, 엣지 컴퓨팅은 로컬에서 한 번에 처리할 수 있다.

[표 1] 클라우드 컴퓨팅과 엣지 컴퓨팅의 비교

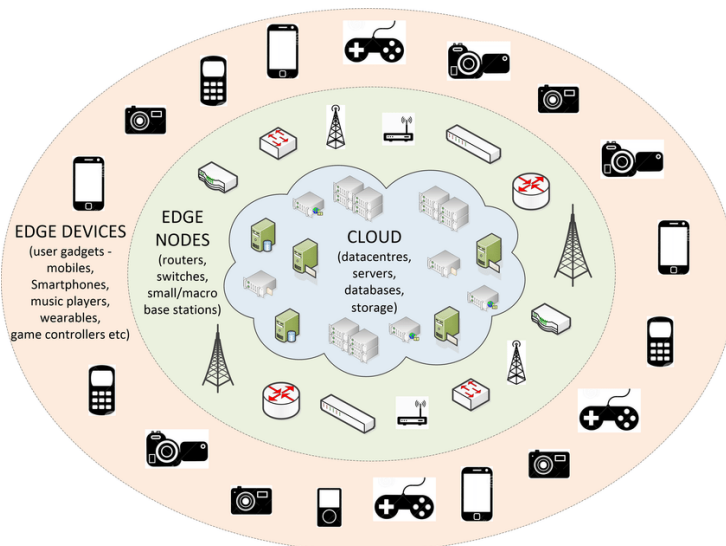
Items	Cloud Computing	Edge Computing
Latency	High	Low
Hardware	Scalable storage and computing power	Limited storage and computing power
Location of server nodes	Within the Internet	At the edge of the local network
Distance between client and server	Multiple hops	One hop

3 » 엣지 컴퓨팅의 구조

이번 절은 엣지 컴퓨팅의 구조를 설명하고자 한다. 앞 절에서 설명한 것처럼 엣지 컴퓨팅에 대한 일치된 정의가 없으므로 표준화된 엣지 컴퓨팅의 구조도 정의되어 있지 않아 저마다 다른 엣지 컴퓨팅 구조를 제시한다. 따라서, 본 보고서에서는 여러 형태의 엣지 컴퓨팅 구조 중에서 대표적인 몇 가지 구조를 설명한다.

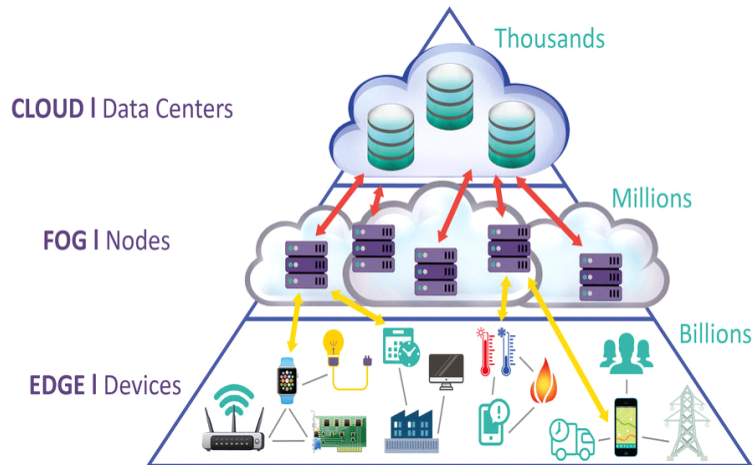
그림 2는 엣지 컴퓨팅의 구조, 그림 3에는 포그 컴퓨팅의 구조를 보인다. 여기서 엣지 컴퓨팅 또는 포그 컴퓨팅은 클라우드 컴퓨팅과 사물인터넷을 위한 분산 컴퓨팅 환경을 의미한다. 두 그림에서 볼 수 있듯이 이 컴퓨팅 환경은 중앙에 있는 클라우드의 데이터센터, 분산된 엣지 노드 또는 포그 노드, 여기에 연결된 엣지에 있는 디바이스로 세 개 계층 구조로 구성된다.

〈그림 2〉 IoT를 위한 엣지 컴퓨팅의 구조



출처 : B. Varghese et al., "Challenges and Opportunities in Edge Computing," 2016 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2016.

〈그림 3〉 포그 컴퓨팅의 구조



출처 : IoT and Predictive Analytics:: Fog and Edge Computing For Industries Versus Cloud
<https://leanbi.ch/en/blog/iot-and-predictive-analytics-fog-and-edge-computing-for-industries-versus-cloud-19-1-2018/>

두 그림 모두 세 개의 계층으로 이루어져 있지만, 각 계층을 부르는 용어에는 차이가 있다. 데이터센터의 서버와 스토리지를 의미하는 클라우드와 말단의 디바이스를 엣지 디바이스라 부르는 것은 동일하다. 그러나, 자료에 따라 중간에 위치하는 계층을 엣지 노드라고 부르기도 하고, 포그 노드라고 부르기도 한다. 본 보고서에서는 일관성을 위해서 중간 계층을 포그 또는 포그 노드라고 부르기로 한다.

포그 노드는 클라우드와 엣지 디바이스들 간의 매개자 역할을 하며, 엣지 디바이스에 가까운 곳에서 데이터의 처리, 저장 및 네트워킹 서비스를 처리한다. 포그 노드는 네트워크가 연결되는 어디에나 배치될 수 있다. 즉, 포그 노드는 컴퓨팅, 스토리지를 가지고 있고 네트워크 연결이 가능한 기기로 산업 컨트롤러, 스위치, 라우터, 내장형 서버 및 비디오 감시 카메라와 같은 것이 될 수 있다. 따라서 엣지 디바이스와 포그 노드는 IoT 디바이스에서 실시간으로 데이터를 입력받아, 수 밀리 초 수준의 응답 시간을 보장하는 실시간 제어 및 분석을 위한 IoT 애플리케이션을 실행한다. 또한, 정기적으로 클라우드로 요약된 데이터를 전송한다.

클라우드는 지역적으로 분산된 포그 노드로부터 데이터를 수신하고 처리하고, 데이터의 소스인 IoT 디바이스들의 상태를 파악하고, 전체적인 비즈니스 통찰력(insight)을 추출하고, 이에 따라 각 디바이스에 적합한 규칙과 명령을 생성한다. 또한, 이러한 규칙과 명령을 엣지 노드에 적용하기 위해서 각 포그 노드로 이를 전파한다.

엣지 컴퓨팅 환경에 주어진 서비스를 수행하기 위해서는 응용 프로그램, 즉, 애플리케이션이 제공되어야 한다. 엣지 컴퓨팅에서 동작하는 애플리케이션은 각 계층에서 동작한다. 즉, 하나의 엣지 컴퓨팅 애플리케이션은 엣지 디바이스에서는 센서 등 디바이스를 동작하기 위한 기능을 수행하고, 포그 노드와 클라우드에서는 데이터의 처리와 분석을 위한 기능을 수행한다.

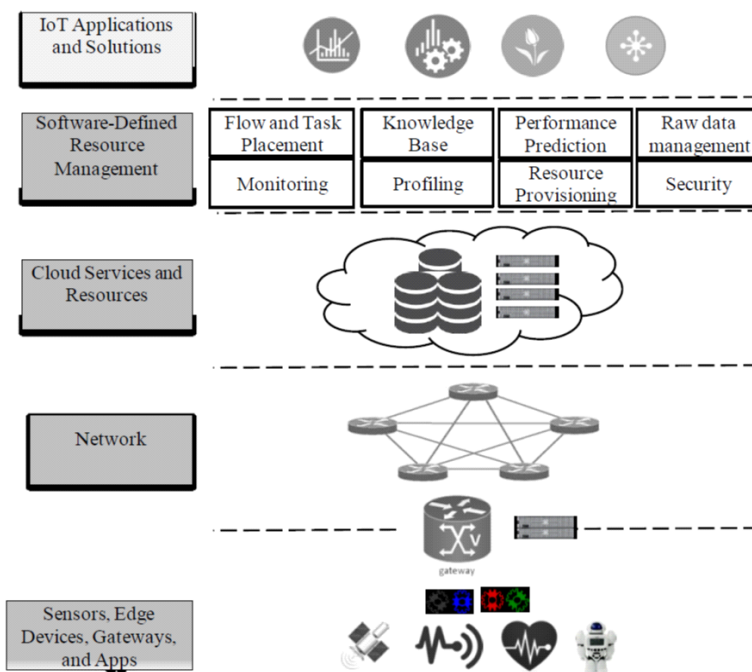
엣지 컴퓨팅 애플리케이션을 좀 더 자세히 살펴보면 엣지에서 수행되는 애플리케이션은 지속해서 발생하는 데이터를 처리하는 작업을 수행하고, 포그에서 수행되는 애플리케이션은 지역적 특성을 가진 데이터의 배치 처리 작업을 주로 수행한다. 여기서 지역적 배치 처리란 전체 지역이 아닌 포그 노드가 위치한 지역에 해당하는 데이터를 모아서 처리하는 것을 의미한다. 클라우드에서 수행되는 애플리케이션은 전역적 데이터 처리를 수행한다. 엣지나 포그는 클라우드 컴퓨팅보다 컴퓨팅 용량과 스토리지가 제한적이기 때문에 클라우드와 협업을 통해 효과적인 스토리지 사용과 빠른 응답을 제공한다.

일반적으로 엣지 컴퓨팅은 이러한 지역적 처리와 전역적 처리의 두 가지 구조적 특성으로 인해 클라우드 컴퓨팅의 이점을 유지하면서 추가적인 장점이 있다. 즉, 네트워크 끝단에 위치한 엣지 노드는 디바이스와 물리적인 거리가 가깝다는 장점으로 인해 실시간 분석과 처리가 필요한 작업에 적합하다. 반면 실시간성을 필요하지 않지만 많은 자원이 필요한 전역적 배치 처리 작업은 중앙의 클라우드 서버가 수행한다.

아래 그림 4는 또 다른 엣지 컴퓨팅의 구조도로 애플리케이션 플랫폼 관점에서 설계되었고, 총 5계층으로 구성된다. 맨 아래 계층에는 엣지 디바이스와 게이트웨이뿐만 아니라 IoT 센서를 포함하고 있다. 이 계층에는 디바이스에 설치되는 애플리케이션도 포함한다. 이 계층의 요소들은 디바이스 간의 통신 또는 디바이스와 클라우드의 통신을 위해 그 윗 계층인 네트워크를 사용한다. 다음 계층에는 클라우드 서비스와 이를 수행하기 위한 연산

및 저장 자원이 위치한다. 클라우드 계층 위에는 전체 인프라를 관리하고 엣지 컴퓨팅 애플리케이션에 대한 서비스 품질을 유지하기 위해 소프트웨어 정의 자원 관리 (Software-Defined Resource Management)층이 있다. 마지막으로, 최상위 계층에는 엣지 컴퓨팅을 활용하여 최종 사용자에게 다양하고 지능적인 애플리케이션을 제공하는 IoT 애플리케이션이 있다.

〈그림 4〉 엣지 컴퓨팅의 참조 구조도



출처 : Dastjerdi, Amir Vahid (Edited), Internet of Things: Principles and Paradigms, Morgan Kaufmann Pub, pp. 61-75, 2016.

이 중에서 엣지 컴퓨팅의 핵심 기능을 수행하는 소프트웨어 정의 자원 관리(Software-Defined Resource Management) 계층을 좀 더 자세히 살펴보도록 하자. 소프트웨어 정의 자원 관리 계층은 상위에 있는 IoT 애플리케이션이 수행하는데 필요한 기능을 제공하는 미들웨어(middleware) 역할을 한다. 즉, 애플리케이션이 실행하는데 필요한 자원을 클라우드와 엣지 디바이스에서 확보하고, 작업을 분배하고, 실행 과정에서 애플리케이션

의 성능을 허용 가능한 대기 시간으로 유지하는 것과 엣지 디바이스를 최대한 활용하여 클라우드의 사용 비용을 줄이는 역할을 한다.

소프트웨어 정의 자원 관리 계층은 다음과 같이 다양한 서비스들이 제공된다.

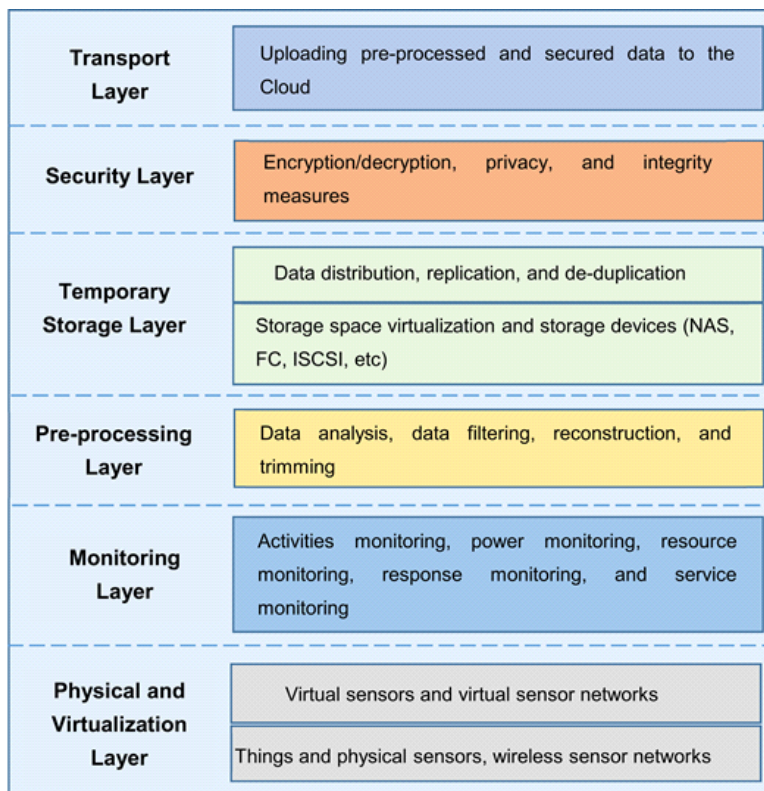
- 흐름 및 작업 배치(Flow and Task Placement): 이 서비스는 사용 가능한 클라우드 자원과 네트워크 자원의 상황을 보고 새로 실행되는 IoT 애플리케이션이 사용할 자원을 선정한다. 이 서비스는 뒤에 설명하는 자원 프로비저닝 서비스와 통신하여 수행되는 작업의 상태를 파악하고, 또 새로운 애플리케이션이 수행되어 자원의 수요가 증가하면 새로운 자원을 할당한다.
- 기술 자료(Knowledge Base): 이 서비스는 과거 애플리케이션 수행된 기록과 그때 사용한 자원의 양을 기록해 두었다가 이 정보를 다른 서비스가 참고하도록 제공한다.
- 성능 예측(Performance Prediction): 이 서비스는 기술 자료 서비스의 정보를 활용하여 사용 가능한 클라우드 자원의 성능을 예측한다. 이 정보는 많은 작업이 동시에 수행되어 기대한 성능에 미치지 못할 때 자원 프로비저닝 서비스가 확보해야 할 자원의 양을 결정하는 데 사용된다.
- 원시 데이터 관리(Raw Data Management): 이 서비스는 분산되어 저장된 데이터를 직접 접근해서 처리하고, 처리된 데이터를 다른 서비스에 제공하는 기능을 한다. 간단한 데이터 처리는 SQL, NOSQL REST API 등을 사용하여 접근할 수 있고, 더 복잡한 처리가 필요하면 MapReduce 같은 패키지를 사용하게 된다.
- 모니터링(Monitoring): 이 서비스는 애플리케이션 및 서비스의 성능과 상태를 추적하고, 필요에 따라 이 정보를 다른 서비스에 제공한다.
- 프로파일링(Profiling): 이 서비스는 기술 자료 및 모니터링 서비스에서 얻은 정보를 기반으로 전체 시스템의 자원 및 애플리케이션의 세부 명세서인 프로파일을 구축한다.
- 자원 프로비저닝(Resource Provisioning): 이 서비스는 애플리케이션이 수행되기 위한 클라우드 자원과 네트워크 자원을 확보하며, 확보된 자원을 애플리케이션에 할

당한다. 전체 시스템으로 보면 시간이 지남에 따라 수행되는 애플리케이션의 수와 애플리케이션에 필요한 자원이 지속해서 변화함에 따라 자원의 할당과 반환이 계속해서 이루어진다. 얼마나 많은 자원을 할당해야 할지에 대한 결정은 앞서 설명한 프로파일링, 성능 예측, 모니터링 서비스가 제공하는 정보를 사용하여 결정한다.

- 보안(Security): 이 서비스는 서비스 및 애플리케이션이 안전하게 동작하는데 필요한 인증과 암호화 등을 제공한다.

그림 5는 애플리케이션의 기능적 관점에서 설계된 엣지 컴퓨팅의 구조이다. 이 구조도는 물리 가상화, 모니터링, 전처리 처리, 임시 스토리지, 보안, 전송 계층으로 6개 계층으로 구성된다.

〈그림 5〉 엣지 컴퓨팅의 계층 구조



출처: Hany F. Atam et al., "Fog Computing and the Internet of Things: A Review." Big Data and Cognitive Computing, vol. 2, no. 2, 2018.

물리 가상화 계층은 사물, 센서와 이들을 네트워크로 연결한 물리적 노드들을 가상화한 센서와 센서 네트워크가 이 계층에 위치하고 있다. 이 계층에 있는 다양한 센서들은 설치된 환경의 값을 측정하고, 수집된 데이터를 추가적인 처리와 필터링을 위해 상위 계층에 전송한다.

모니터링 계층은 디바이스들의 자원 활용 상태, 네트워크 상태를 모니터링한다. 노드가 수행하는 모든 작업은 이 계층에서 모니터링하게 되며, 어떤 노드가 어떤 작업을 언제, 어느 기간 동안 수행하는지, 앞으로 어떤 작업이 필요한지 모니터링한다. 또한, 전체 시스템에서 동작하는 모든 애플리케이션과 서비스의 성능과 상태를 모니터링한다.

전처리 계층은 데이터 관리 작업을 수행한다. 수집된 데이터를 분석하여 의미 있는 정보를 추출하기 위해 필요 없는 데이터를 삭제하고, 데이터를 보정하는 작업을 수행한다. 그런 다음 전처리된 데이터는 임시 저장소 계층에 저장된다. 데이터가 클라우드에 전송되면 더 이상 저장할 필요가 없으므로 임시 저장소의 데이터는 삭제될 수 있다.

보안 계층에서는 데이터의 암호화와 복호화가 이루어진다. 또한, 데이터의 무단 변경을 방지하기 위해 무결성을 보장하는 조치를 데이터에 적용한다. 마지막으로, 전송 계층에서는 전처리된 데이터를 클라우드에 업로드하여 클라우드에서 더 유용한 서비스를 제공하게 된다.

이렇게 엣지 컴퓨팅은 수집된 데이터의 일부만 클라우드에 업로드하여 네트워크 대역폭과 디바이스의 전력을 절약한다. 즉, 센서 네트워크와 IoT 디바이스에서 수집된 데이터는 데이터의 압축이나 처리를 수행한 다음 클라우드로 전송된다. 클라우드는 수신한 데이터를 저장하고 사용자를 위한 서비스를 생성하는 데 사용한다.

지금까지 세 가지의 엣지 컴퓨팅 구조도를 제시하고 설명하였다. 이들 구조에서 보듯이 엣지 컴퓨팅의 구조는 기존의 클라우드 환경과 사물인터넷 환경을 결합한 구조이다. 즉, 중앙 집중된 클라우드 환경에 분산된 사물인터넷 환경을 결합하여 대규모 데이터의 연산과 저장을 가능하게 하고, 다양한 서비스 제공을 지원한다.

4 » 엣지 컴퓨팅의 특징

엣지 컴퓨팅은 기존의 클라우드 컴퓨팅의 개념을 확장하여 엣지 디바이스를 사용하여 컴퓨팅을 효율적으로 활용할 수 있도록 한다. 따라서 엣지 컴퓨팅은 서비스 대기 시간을 단축하고 서비스 품질을 개선하여 우수한 컴퓨팅 환경을 제공한다. 엣지 컴퓨팅과 기존의 클라우드 컴퓨팅을 비교하면 다음과 같은 특징이 있다.

- 짧은 지연시간: 시간에 민감한(mission critical) 애플리케이션은 실시간 데이터 처리가 필요하다. 이러한 애플리케이션의 예로는 로봇 제어, 항공기 제어, 차량 제어 등이 있다. 이들 제어는 센서가 수집한 데이터와 제어 시스템의 프로그램에 따라 수행된다. 이때 클라우드에서 제어 시스템을 동작시키면 통신 장애가 발생해 제어 타이밍이 늦어지거나 제어 자체가 불가능해진다. 따라서 엣지 컴퓨팅에서는 데이터가 외부 데이터센터 또는 클라우드가 아닌 데이터가 생성된 소스에 가까운 곳(엣지)에서 처리되어 클라우드 서버로 데이터를 전송하고, 처리 결과를 받을 때까지 걸린 시간인 지연시간을 단축한다. 예를 들어 자동차에 장착된 여러 센서는 계속해서 데이터를 생성하는데 이 모든 데이터를 클라우드로 전송하는 것은 효율적이지도 않고, 현실적이지도 않다. 따라서, 엣지 컴퓨팅은 자동차 또는 가까운 기기(엣지)에서 생성된 데이터를 필터링하고 분석하는 플랫폼을 제공하며, 이를 통해 실시간으로 데이터를 처리한다. 결과적으로 응답 시간을 크게 줄일 수 있다.
- 대규모 분산처리와 위치 인식: 중앙 집중식 클라우드와 달리 엣지 컴퓨팅은 컴퓨팅 및 스토리지 자원을 지역적으로 분산 배치 가능하고 엣지 컴퓨팅 애플리케이션은 분산 환경에 설치되고 수행된다. 이러한 환경에서 애플리케이션이 서비스를 제공하기 위해서는 분산된 여러 디바이스에서 생성된 정보를 통합하게 되고, 이 정보를 클라우드를 통해 전체 시스템 차원에서 처리하여 제공된다. 또한, 수집된 데이터들은 지역적으로 특정된 노드에서 수집, 처리되기 때문에 각 데이터의 생성 위치를 인식해 위치에 기반한 지역적인 처리가 가능하다.

- 확장성: 엣지 디바이스는 일반적으로 처리 용량과 스토리지 등 컴퓨팅 자원의 제약이 있어 수행할 수 있는 작업이 제한된다. 반면에 비교적 풍부한 자원을 제공하는 클라우드에는 많은 자원이 필요한 작업을 수행할 수 있다. 그러나, 클라우드가 거의 무한한 자원을 사용한다 하더라도 엣지 디바이스에서 생성되는 모든 데이터가 계속 전송될 경우, 병목 현상이 발생할 수 있다. 엣지 컴퓨팅이 엣지 디바이스를 클라우드에 연결한다는 관점에서 보면 클라우드 컴퓨팅과 엣지 디바이스의 성능 범위를 확대하는 방안이 된다. 엣지 컴퓨팅은 데이터가 발생하는 소스의 가까운 곳에 있는 수많은 엣지에서 데이터를 처리하여 클라우드의 처리 부담을 줄여준다. 반대로 엣지 디바이스는 클라우드의 도움을 받아 인공지능 같은 복잡한 서비스를 제공한다.

예를 들어, 스마트폰과 같은 사용자 디바이스는 클라우드 데이터센터의 서버에 비해 하드웨어 자원이 상대적으로 빈약하다. 따라서 사용자 디바이스는 텍스트, 오디오, 비디오, 터치 형태의 정보를 수집하고, 이 정보를 클라우드에 전송하고 클라우드의 자원으로 대부분 처리가 이루어진다. 사용자 디바이스들은 미들웨어 및 하드웨어의 제한으로 인해 복잡한 분석을 수행할 수 없으며, 이를 수행하면 상당한 배터리 소모를 감수해야 한다. 따라서 데이터를 처리하는데 필요한 연산을 수행하기 위해 데이터를 클라우드로 보내야 하고, 클라우드에서 수행한 이후 의미 있는 정보는 다시 사용자 디바이스로 전달된다. 그러나 사용자 디바이스에서 일부 데이터 저장과 처리를 할 수 있게 된다면 모든 데이터를 클라우드에 전송하고, 처리하는 부담이 줄어든다. 따라서 클라우드 관점에서 보면 엣지 컴퓨팅을 통해 추가적인 IT 자원을 확보하게 된다. 결국, 지속해서 데이터가 증가하거나 연결된 디바이스와 서비스의 수가 늘어나도 엣지에서 처리를 강화하여 확장성 문제를 해결할 수 있다.

- 이동성: 엣지 애플리케이션의 특징 중 하나는 모바일 장치와 연동되어 이동성 있는 컴퓨팅 환경을 제공한다. 이로 인해 디바이스들은 사용자와 함께 이동하거나 필요에 따라 이동에 따른 서비스가 가능하다. 대표적인 모바일 장비인 스마트폰을 활용한 위치 기반 서비스뿐만 아니라 차량을 활용한 교통 서비스도 엣지 컴퓨팅의 이동성을 활용한 서비스이다.

- 신뢰성: 신뢰성이 확보되기 위해서는 개인정보와 같이 민감한 정보의 통제와 정보 흐름을 안전하고 신뢰할만한 방식으로 관리해야 한다. 개인과 사회적으로 민감한 데이터는 엣지에서 생성된다. 민감 정보는 가능한 엣지 디바이스나 포그 노드에서 우선 처리하고 클라우드에 전송되는 데이터를 최소로 하여 신뢰성을 높이게 된다.
- 상호 운용성(interoperability) : 새로운 컴퓨팅 환경이 도입되면 과거 시스템들을 재개발하는 것이 일반적이지만 엣지 컴퓨팅에서는 기존 시스템을 엣지 컴퓨팅 환경에서 그대로 사용할 수 있다. 엣지 컴퓨팅은 새롭게 연결되는 디바이스와 여기에 이미 구현된 레가시(Legacy) 애플리케이션의 상호 운용성을 제공한다. 즉, 엣지 컴퓨팅의 분산 특성을 이용하여 다른 도메인에서 사용된 서비스를 표준화된 통신 프로토콜로 엣지 컴퓨팅 인프라에 연결하면 엣지 컴퓨팅 플랫폼에서 기존 서비스를 수정 없이 그대로 사용할 수 있다.
- 지능적 제어: 디바이스가 소형화되고 컴퓨팅 용량이 증가함에 따라 엣지 센서와 디바이스는 더욱 강력해진다. 이로 인해 분산처리가 가능하고, 의사결정에 활용하는 지능적 애플리케이션이 만들어질 수 있다. 이들 애플리케이션은 엣지, 포그, 클라우드의 각 계층 간에서 원활하게 동작하기 위해 적절한 자원의 배치와 애플리케이션의 관리, 데이터의 동기화 등의 기능이 제공된다. 이를 통해 엣지에서 높은 수준의 지능적 제어가 가능한 서비스가 만들어진다.

하지만 이러한 엣지 컴퓨팅의 특징이 항상 보장되는 것은 아니다. 엣지 컴퓨팅의 특징을 최대한 활용하기 위해서는 엣지 컴퓨팅을 구현하고 적용할 때에 신중히 고려해야 할 사항들이 있다.

1. 상호 운용성과 확장성을 보장하기 위해서는 엣지 컴퓨팅 환경을 구축할 때에 표준화된 프로토콜과 엣지, 포그, 클라우드까지 일관된 인프라가 구성되어 있어야 한다.
2. 지능형 제어를 위해서는 관리를 자동화하고, 유지 보수를 위한 장비의 교체가 쉬워야 하고, 예기치 않은 장애 발생 시 복구가 가능한 운영 시스템이 제공되어야 한다.

3. 지리적으로 널리 분포된 환경에서 신뢰성을 보장하기 위해서는 원격으로 제어할 수 있어야 한다. 또, 원격 제어가 가능하다 하더라도 애플리케이션 설계부터 장애를 대비한 내결함 설계가 되어있어야 하고, 우선 인프라를 구성하고 있는 장비의 신뢰성이 확보되어야 한다.
4. 유지 보수는 간단해야 한다. 원격 관리자가 소프트웨어를 다시 설치하거나, 업그레이드할 수 있어야 하고, 현장 기술자가 직접 장비의 수리와 교체를 할 수 있어야 한다.
5. IoT 환경은 공장, 도로 같이 극한 야외 환경이 대부분이다. 따라서 엣지 컴퓨팅 환경은 험하다고 가정하고, 전력, 온도, 습도, 먼지, 진동을 고려하여 설계하고 설치되어야 한다.

5 » 엣지 컴퓨팅의 사업적 의미

지금까지 엣지 컴퓨팅의 정의와 특징을 살펴보았다. 이번 절에서는 엣지 컴퓨팅이 주는 사업적인 의미를 살펴보고자 한다.

● 사업의 민첩성 향상

엣지 컴퓨팅은 수집된 데이터를 밀리초 내에 처리 분석하여, 적절히 조치하므로 산업 현장에서 민첩성이 향상된다. 예를 들어, 풍력발전소에서 풍속도나 방향이 변경되는 경우, 현장에 설치된 엣지 컴퓨팅 애플리케이션은 이 데이터를 실시간으로 분석하고 개별 터빈을 조정하여 전체적인 발전 생산을 최적화할 수 있다.

또한, 클라우드로 전송하는 데이터는 원시 데이터를 압축하여 전송하여 네트워크 대역폭을 절감하고, 클라우드 내에서 심층적인 분석을 한다. 예를 들어 발전소 터빈은 테라바이트급의 데이터를 생성한다. 고급 분석을 실행하기 위해 클라우드에 이 데이터를 보내는 것은 기술적으로는 가능하지만, 매번 수행하기에는 비용이나 시간이 너무 많이 든다. 이때 엣지 컴퓨팅을 통해 터빈에서 스트리밍 데이터를 수집, 실시간으로 처리하여 예기치 않은 고장을 감지하는 데 사용하고, 운영과 관련된 통계 데이터는 클라우드로 전송하여 데이터의 크기를 관리 가능한 크기로 줄이면서 장비의 상태와 유지 보수를 위한 결정에 데이터를 사용한다.

엣지 컴퓨팅을 활용하여 지연시간이 줄어들게 되면 애플리케이션이 빠른 속도로 운영될 수 있을 뿐만 아니라 운영 안정성도 높아진다. 즉, 엣지 컴퓨팅으로 클라우드 서비스의 역할을 축소하는 효과가 있어 클라우드 서비스의 장애로 인한 서비스의 중단 가능성이 줄어든다. 예를 들어, 중앙 집중식 클라우드를 사용하여 데이터를 저장하고 있을 때 클라우드가 다운될 경우, 클라우드 서비스의 문제가 해결될 때까지 데이터에 액세스할 수 없게 되어 심각한 비즈니스 손실을 초래할 수 있다. 따라서 클라우드의 의존성을 줄이는 것이 디바이스가 인터넷에 연결되지 않은 오프라인에서도 안정적으로 작동할 수 있다는 것을

의미한다. 이는 인터넷 연결이 원활하지 않은 지역에서도 효과적인 애플리케이션 동작이 가능하다는 것을 의미한다.

결과적으로 기업이 엣지 컴퓨팅을 도입하게 되면 실시간 대응, 빠른 분석, 안정적인 서비스로 기업의 민첩성이 증가하게 된다.

● 운영비용 절감

엣지 컴퓨팅은 클라우드에 비해 운영비용을 절감할 수 있다. 클라우드 컴퓨팅을 사용하게 되면 데이터를 전송하고, 저장하고, 처리하는 비용이 발생하게 된다. 엣지 컴퓨팅은 클라우드 서비스를 효과적으로 사용하여 운영비용을 절감한다. 예를 들어 오지에서 석유와 가스를 시추하는 회사가 석유 생산량 관리를 위해 컴퓨팅 자원이 필요하다면, 자체 데이터센터를 구축할 수 있으나 현실적으로 시간과 비용이 많이 들기 때문에 비효율적이다. 이에 대한 대안으로 클라우드 공급자를 선정하여 서비스를 제공받을 수 있다. 그러나 네트워크 환경이 좋지 않다면 활용하기 어렵고, 데이터센터가 멀리 떨어져 있어 대기 시간이 길어진다. 대략 석유회사는 매주 500GB의 데이터를 생성하는데, 이런 수많은 엣지 장치에서 발생한 많은 양의 데이터를 클라우드로 전송하는 것은 실용적이지 않다. 이때 현지에 포그 장비로 엣지 컴퓨팅을 구축하여 데이터 분석을 위해 클라우드로 보내지 않고 로컬에서 데이터를 처리하여 비용을 절감한다.

IoT 기기 수가 증가함에 따라 데이터 생성 속도는 지속해서 증가하고 있다. 그 결과, 네트워크 대역폭이 더욱 제한되고, 클라우드 서비스에서 데이터의 병목 현상이 더 커지게 되어 원활한 서비스가 어렵게 된다. 이때 엣지 컴퓨팅을 활용하여 전체 시스템의 네트워크 트래픽을 줄일 수 있다. 예를 들어 비디오와 같은 IoT 카메라에서 발생한 데이터는 클라우드로 전송할 때에 추출된 정보와 메타데이터만 전송한다. 즉, 클라우드 데이터센터보다 포그 노드, 엣지 디바이스에서 데이터를 처리하여 데이터 전송에 들어가는 네트워크 비용, 클라우드 서비스의 컴퓨팅 사용과 데이터 저장에 필요한 비용을 절감하게 된다.

● 실시간 응용 지원

스마트 그리드, 스마트 시티와 같은 실시간 애플리케이션은 몇 밀리초 이상의 지연시간을 허용할 수 없으며 대기 시간에 매우 민감하다. 자율주행 자동차는 낮은 대기 시간과 높은 대역폭을 요구하며 차량이나 도로 장비로 데이터 처리나 콘텐츠 전송을 하므로 엣지 컴퓨팅이 필수적이다. 또한, 엣지 컴퓨팅은 실시간 응답으로 인해 몰입형 애플리케이션의 가능성을 열어 준다. 이 중 일부는 AR/VR, 4K 비디오, 의료 영상에 기반한 서비스도 포함한다. 앞으로 실시간성이 요구되는 서비스가 지속해서 증가하고 있어 엣지 컴퓨팅에 대한 관심이 더 높아질 것으로 예상된다.

● 컴플라이언스, 개인정보 보호 및 데이터 보안

클라우드, 특히 퍼블릭 클라우드 서비스는 여러 사용자가 동시에 동일한 인프라를 사용하기 때문에 보안에 대한 우려가 있는 것이 현실이다. 이에 대해서 클라우드 사업자는 비밀 데이터 또는 개인정보 같은 민감 정보와 관련된 규정을 만들고, 이를 준수하는 방안을 제시하고, 계약 시 이를 보증한다. 하지만, 클라우드 사업자로서는 개인의 접근과 보안 통제를 보증하기 위해 복잡한 보안 체계를 구축하기 때문에 기술적으로 어렵고, 비용도 많이 들어간다. 또한, 지속해서 보안 운영을 위해서 큰 노력과 비용이 들어간다.

엣지 컴퓨팅은 분산 컴퓨팅 기반이어서 기업이 해당 영역, 지역, 도메인별로 필요에 따라 로컬 보안 영역을 분리하여 보안 체계를 구축할 수 있다. 따라서, 이 경계를 기준으로 민감 정보를 로컬 영역에 한정하여 처리하여 독립적인 보안 체계를 구축할 수 있다. 결과적으로 엣지 컴퓨팅은 클라우드 서비스에 비해 보안을 강화하는 효과가 있다.

엣지 컴퓨팅의 분산처리는 컴플라이언스(Compliance) 보안 규정 준수에도 유용하다. 최근 여러 보안사고로 인해 보안 규정 준수에 대한 요구사항이 높아지고 있다. 특히 정부가 기업이 소비자 데이터를 활용하는 방법에 대해 점점 더 강한 규제를 통해 엄밀한 관리를 요구하면서 컴플라이언스에 대한 중요성이 높아졌다. 특히, 개인정보를 데이터 남용으로 부터 보호하고자 하는 EU의 GDPR(General Data Protection Regulations)이 최근 시행되었다. 이 규정에 의하면 개인정보를 다른 국가로 전송과 저장을 금지한다. 엣지 컴퓨팅

에서 엣지 장치는 로컬 수준에서 데이터를 수집하고, 처리하므로 클라우드로 데이터를 전송하는 것을 피할 수 있다. 기존의 클라우드 환경과 비교하면 민감 정보는 네트워크를 통과할 필요가 없고, 설령 클라우드에 대한 사이버 공격이 발생하더라도 위험성이 많이 감소하게 된다.

● 자동화된 운영

IT 기술이 발전된 오늘날에도 아직 인터넷 연결이 원활하지 않고, 데이터의 송수신을 신뢰할 수 없으며, 지연시간을 예측할 수 없는 환경이 존재한다. 여기에는 운송(차량, 선박, 항공), 광업(유정, 파이프라인, 광산), 전력 인프라(풍력, 태양광 발전소) 등이 포함된다. 이런 극한 환경이 아니라 일반 인터넷 환경에서도 정도의 차이는 있으나 같은 문제점을 안고 있다.

엣지 컴퓨팅은 기본적으로 분산 컴퓨팅 환경이고 고장 허용력(Fault Tolerant)을 보장하고 있어, 네트워크 연결을 사용할 수 없을 때 주어진 프로그램에 따라 정상적인 기능이 유지될 수 있도록 환경을 자동으로 변경하여 안정적인 서비스가 제공될 수 있도록 한다. 예를 들어, 네트워크 장애, 클라우드 장애 또는 서비스 거부 공격으로 인해 클라우드 서비스를 사용할 수 없게 될 경우, 근처의 노드로 서비스를 이동하여 일시적인 장애로부터 시스템을 보호할 수 있다. 결과적으로 엣지 컴퓨팅으로 자동화된 운영으로 운영에 대한 부담을 낮출 수 있다.

● 에너지 절약으로 기후 변화 대응

향후 10년간 클라우드 데이터센터는 현재 소비되는 에너지의 3배에 달하는 에너지를 소비할 것으로 예측되어 새로운 에너지 소비 문제를 발생시키고 있다. 점점 더 많은 애플리케이션이 클라우드로 이동함에 따라 증가하는 에너지 수요는 점점 더 늘어나게 된다. 데이터센터를 운영하는데 필요한 IT 장비의 전력뿐만 아니라 이들 장비를 냉각시키기 위해 엄청난 양의 전력을 소비한다. 이에 따라 많은 클라우드 사업자는 에너지 사용을 최소화할 수 있는 에너지 효율화 전략을 채택하고 있다. 그 예로 네이버 데이터센터는 서늘한 온도를

유지하는 산속에 데이터센터를 구축하여 냉방 비용을 절약하고, 마이크로소프트는 좀 더 극단적으로 바다에 데이터센터를 설치하는 것을 시도하고 있다.

엣지 컴퓨팅에서는 에너지 소비 문제를 기지국이나 라우터같이 에너지에 대한 부담이 적은 엣지 노드에 분산하여 분석 작업을 수행하게 하여 해결한다. 즉, 엣지에서 수행할 수 있는 간단한 작업은 데이터센터에 부하를 주는 대신 직접 처리하여 효율적인 전력 관리가 가능한 것이다.

6 » 엣지 컴퓨팅 관련 개념과 비교

엣지 컴퓨팅은 클라우드 컴퓨팅이 분산 구조로 발전된 형태이지만 사실 유사한 개념들이 존재하였다. 본 절에서는 엣지 컴퓨팅과 관련된 개념을 소개하고 엣지 컴퓨팅과 관련성을 설명하고자 한다.

● 모바일 클라우드 컴퓨팅(Mobile Cloud Computing)

모바일 네트워킹 기술의 발전으로 언제 어디서나 인터넷으로 정보에 접근이 가능한 모바일 디바이스의 사용이 점차 늘어나는 추세다. 다양한 모바일 디바이스의 증가는 LTE 등 모바일 네트워킹 기술의 빠른 발전을 가져오고 있다. 이제 새롭게 열린 5G 모바일 네트워크는 스마트폰과 자동차를 포함한 수많은 IoT 디바이스를 고속 인터넷에 연결시킬 것으로 기대된다.

모바일 클라우드는 사용자의 모바일 기기에 클라우드 서비스를 제공하고, 모바일 기기들로 클라우드 컴퓨팅 인프라를 구성하여 기기 간 정보와 자원을 공유하는 컴퓨팅 환경이다. 모바일 클라우드는 모바일 기기의 종류나 운영 체제(OS: Operating System)에 상관없이 기기 간 정보를 공유하고, 모바일 앱, 스토리지 등의 클라우드 서비스를 편리하게 이용할 수 있게 한다. 즉, 모바일 클라우드는 기존 클라우드 컴퓨팅의 장점과 모바일의 이동성을 결합하여 비즈니스적 가치를 높인 것이다.

모바일 클라우드 환경에서 모바일 단말기 내에 내장되어있는 많은 애플리케이션은 모바일 단말에서만 수행되는 것이 아니라, 모바일 클라우드 내에 있는 서버들의 지원을 받아 서비스를 제공하는 형태로 진화하고 있다. 즉, 모바일 단말기 내의 저장공간과 처리 기능을 활용하여 일부 기능은 모바일 단말기에서 수행하고, 모바일 클라우드 내의 서버들로부터 추가적인 기능을 지원받아 서비스가 이루어지는 것이다. 이런 부분은 엣지 컴퓨팅의 분산 컴퓨팅 개념과 동일하여 엣지 컴퓨팅이 모바일 클라우드를 포함하는 개념으로 설명되기도 한다.

현재 다양한 모바일 클라우드 서비스가 나와 있고 많은 사람이 사용하고 있다. Apple 기기 사용자들은 iCloud를 통해 각 디바이스에서 사용하는 앱과 데이터를 인터넷을 통해 동기화하는 기능을 사용하고 있다. 마이크로소프트는 OneDrive라는 클라우드 서비스를 통해 다양한 기기에서 데이터를 공유하는 기능을 제공하고 있다. 그런데 모바일 클라우드에서 모바일 기기 간에 호환이 이루어지지 않아 모바일 클라우드 컴퓨팅 확산에 걸림돌로 작용한다. 즉, 제조사별로 자사 플랫폼에 의존적인 솔루션을 제공하여 제조사가 다르면 서로 서비스나 데이터의 호환이 어려운 실정이다. 이러한 이유로 모바일 클라우드 컴퓨팅 플랫폼의 업체 종속성은 앞으로 시장과 기술이 발전하기 위해서는 해결되어야 할 문제이다.

따라서 이 문제를 해결하기 위해서는 향후 제품 간 상호호환성, 이식성, 보안 등에 대한 표준화가 요구된다. 서비스를 API 형태로 제공하여 어떠한 모바일 단말이라도 접근할 수 있게 만들거나, 데이터 저장 및 전송 포맷을 공유하여 원활한 데이터의 교환이 가능하게 만들어야 할 것이다. 또한, 계속해서 새로운 디바이스들이 발표되고 하드웨어 규격도 다른 상황에서 단말과 독립적인 애플리케이션을 개발할 수 있는 표준이 제시되어야 할 것이다.

모바일 클라우드는 모바일 단말기의 이동 중 연결을 보장해야 하는 상황에서 모바일 기기의 컴퓨팅 처리 능력의 한계와 저장공간 제한 문제 등을 고려하여 클라우드 서비스를 제공하게 된다. 앞으로 모바일 기기가 보편화 되고, 기기의 성능이 높아짐에 따라 모바일을 활용하는 응용 분야가 지금까지의 단순한 협업 중심에서 미디어, 의료 등으로 다양해질 것이기 때문에 이러한 모바일 컴퓨팅의 한계가 더욱 커질 것이다.

이는 응용 분야가 다양해지고, 실시간 지원이 가능한 클라우드 컴퓨팅이라는 새로운 컴퓨팅 패러다임에 대한 요구로 이어지고 있다. 엣지 컴퓨팅은 현재 모바일 클라우드 컴퓨팅에서 모바일 기기의 한계로 제공하지 못하는 낮은 지연시간과 실시간성 보장이라는 요구조건을 만족시킨다. 따라서 엣지 컴퓨팅으로 실시간성이 보장되는 방향으로 모바일 클라우드 컴퓨팅이 발전될 것이다.

● 모바일 엣지 컴퓨팅(Mobile Edge Computing)

모바일 엣지 컴퓨팅은 모바일 클라우드의 발전된 형태로 엣지 서버와 셀룰러 기지국을 함께 운영할 수 있도록 제공한다. 따라서, 모바일 엣지 컴퓨팅은 모바일이 강조된 넓은 의미에서 엣지 컴퓨팅의 일종이다. 모바일 엣지 컴퓨팅은 최종 모바일 기기, 엣지 서버와 클라우드로 구성된 세 개의 계층으로 이루어진 구조이다. 모바일 엣지 컴퓨팅은 모바일 통신망을 사용하여 유연하면서도 빠른 이동 통신 서비스를 제공하고, 네트워크 사용에 따른 비용 효율성을 추구한다.

최근에는 5G 통신 시대가 열리면서 모바일 엣지 컴퓨팅에 대한 기대가 커지고 있다. 즉, 5G를 통해 대규모 모바일 연결이 가능하고, 실시간 응답이 가능해져 새로운 응용 분야에 대한 가능성이 커지고 있다. 반면에 5G가 상용화되기 위해서는 표준화뿐만 아니라 대규모 투자가 수반되는 망을 구축이 이루어져야 하므로 실질적인 5G 시대와 이를 이용한 서비스는 좀 더 시간이 걸릴 것으로 예상된다.

● 포그 컴퓨팅(Fog Computing)

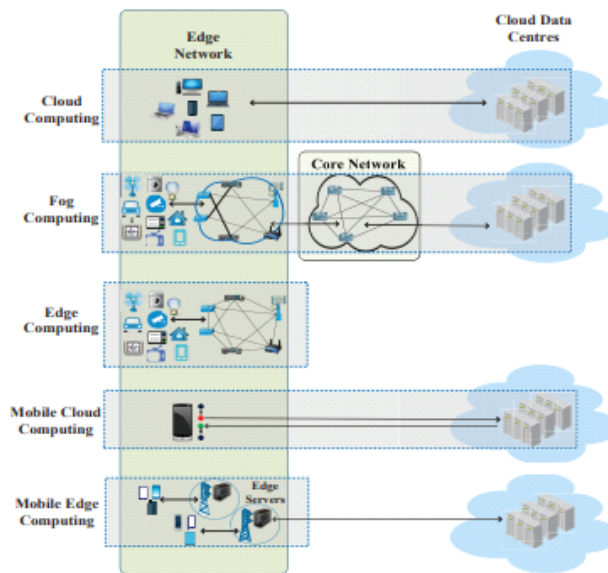
포그 컴퓨팅은 CISCO가 처음 소개한 개념으로, ‘클라우드 컴퓨팅과 서비스를 네트워크 엣지까지 확장하는 패러다임’으로 정의된다. 포그 컴퓨팅은 기존 클라우드 컴퓨팅처럼 네트워크를 통해 제공될 수 있지만, 셋톱박스 또는 무선공유기 같은 사용자에게 근접한 네트워크 장비에서도 컴퓨팅 자원을 제공한다. 포그 컴퓨팅은 애플리케이션을 실시간 또는 예측 가능한 대기 시간으로 서비스를 제공하기 위해 엣지에 지능을 가진 디바이스를 배치한다.

이미 엣지 컴퓨팅의 정의에서 설명한 것처럼 포그 컴퓨팅은 엣지 컴퓨팅과 상당히 개념이 중복된다. 따라서 본 보고서에서는 포그 컴퓨팅과 엣지 컴퓨팅을 같은 개념으로 설명하였으나 엄밀한 의미에서 두 개념을 구분하기도 한다. 특히, 포그 컴퓨팅을 지지하는 진영에서는 포그 컴퓨팅이 엣지 컴퓨팅을 포함하는 개념으로 설명한다. 즉, 엣지 컴퓨팅은 단지 데이터가 생성된 디바이스, 엣지 디바이스만 활용하여 처리하지만, 포그 컴퓨팅은 엣지 디바이스뿐만 아니라 클라우드와 클라우드까지 연결된 네트워크상에 있는 디바이스를 모두 포함하는 것으로 설명하기도 한다. 그러나, 앞서 설명한 것처럼 포그 컴퓨팅과 엣지

컴퓨팅을 세밀한 구분 없이 사용하는 것이 현실이다.

그림 6은 클라우드 컴퓨팅, 모바일 클라우드 컴퓨팅, 모바일 엣지 컴퓨팅과 포그 컴퓨팅, 엣지 컴퓨팅의 개념을 도식적으로 비교한 것이다. 이 그림은 포그 컴퓨팅 관점에서 만든 그림으로 엣지 컴퓨팅을 좁은 의미의 개념으로 표현하였다. 이 그림에서 만일 엣지 컴퓨팅을 네트워크를 통해 클라우드 데이터센터를 사용하는 것으로 표현하면 포그 컴퓨팅과 개념상 차이가 없게 된다.

〈그림 6〉 클라우드, 포그, 엣지, 모바일 클라우드, 모바일 엣지 컴퓨팅의 비교



출처 : Mahmud R., Kotagiri R., Buyya R. "Fog Computing: A Taxonomy, Survey and Future Directions," In: Di Martino B. et al. (eds) Internet of Everything, Springer, 2018.

● Cloudlet

2013년 카네기멜론 대학의 Satyanarayanan에 의해 제안된 Cloudlet 아키텍처는 3개 계층으로 이루어진 클라우드 컴퓨팅 구조를 지칭하는 용어이다. Cloudlet이란 중앙집중의 클라우드 서버와 사용자의 디바이스 사이에 배치된 소규모의 클라우드 서버를 말한다. Cloudlet은 네트워크 가장자리에 위치한 이동성을 강화한 마이크로 데이터센터로서 네트

워크에 연결된 모바일 또는 스마트 디바이스를 지원한다. 즉, IoT 디바이스는 기존의 클라우드 컴퓨팅 구조와 다르게 인터넷을 거치지 않고 바로 Cloudlet에 접속하기 때문에 느린 속도, 비용 증가와 같은 클라우드 컴퓨팅의 한계점들을 극복할 수 있다. 디바이스는 인터넷을 거치지 않고 가까운 Cloudlet으로부터 클라우드 서비스를 제공 받을 수 있으며, 각 지역에 위치한 여러 Cloudlet으로 대역폭이 분산되기 때문에 자유로운 확장이 가능하다.

엣지 컴퓨팅은 Cloudlet 아키텍처의 기본 구조를 따르는 컴퓨팅 패러다임이다. 엣지 컴퓨팅의 시스템 모델은 Cloudlet 아키텍처와 비슷하게 대규모 클라우드 서버와 분산된 소규모의 클라우드 서버들로 이루어져 있다. 즉, 엣지 컴퓨팅은 기존의 클라우드 컴퓨팅과 다르게 엣지 디바이스와 클라우드 사이의 네트워크 중간에 여러 개의 소규모 클라우드 서버인 Cloudlet들이 위치한다. Cloudlet은 담당하는 지역의 디바이스들과 연결을 유지하고 클라우드 서버 대신에 디바이스에 클라우드 서비스를 제공한다. Cloudlet은 다시 네트워크 인프라를 통해 중앙의 클라우드 서버와 연결되며, 클라우드 서비스에 필요한 데이터를 주고받는다.

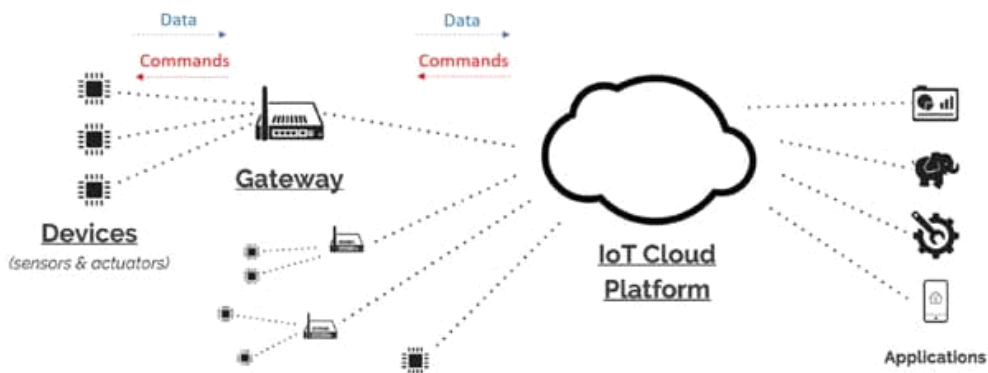
IV

엣지 컴퓨팅 관련 기술

1 » 엣지 컴퓨팅의 요소 기술

본 장에서는 엣지 컴퓨팅과 관련된 기술을 다루고자 한다. 먼저 엣지 컴퓨팅을 이루고 있는 요소와 이에 따르는 기술을 하나씩 살펴보고자 한다. 엣지 컴퓨팅은 일반적으로 센서/디바이스, IoT 게이트웨이, 클라우드 플랫폼으로 이루어져 있고, 이들 요소를 통해 데이터를 생성하고, 데이터를 네트워크로 전송하고, 데이터 처리를 하게 된다.

〈그림 7〉 엣지 컴퓨팅 요소(Source: Ubuntu.com)



(1) 센서/디바이스

사물인터넷에는 넓은 지역에 수만 개의 센서가 배치된다. 사물인터넷의 핵심 구성 요소인 이 센서들은 물리적 값을 측정하여 데이터를 생성하고, 네트워크를 통해 전송한다. 이러한 센서와 디바이스들은 사물인터넷의 상황을 인식할 수 있도록 다양한 데이터를 생성하고, 처리한다.

이 디바이스는 사용자와 상호작용을 위해 사용자 입력을 받아들이고, 그 입력에 관한 결과를 전달하는 휴먼 컴퓨터 인터페이스 역할을 한다. 이 모든 센서와 디바이스들은 서로 연결되어 데이터를 교환하고 서로 협력하여 다양한 서비스를 제공한다. 기기를 연결하는 네트워크를 통해 각 노드는 IoT 애플리케이션을 수행하기 위한 IT 자원을 확보한다. 이를 위해 디바이스의 수행 라이브러리에는 IoT 애플리케이션이 사용할 수 있는 통신을 포함한 다양한 기능을 제공한다.

(2) IoT 게이트웨이(Gateway)

게이트웨이, 라우터, 스위치, 셋톱박스 등과 같은 디바이스는 엣지 컴퓨팅을 위한 중요한 인프라 역할을 한다. 일부 네트워킹 디바이스는 데이터 프로세서, 확장 가능한 메모리, 프로그램 개발환경 등을 포함하여 제공된다. 기존 네트워크 디바이스 외에도 스마트 게이트웨이, IoT 허브 등 여러 전용 네트워킹 디바이스가 엣지 컴퓨팅에서 사용된다. 이들 장비를 통칭하여 엣지 컴퓨팅에서 IoT 게이트웨이라 부른다.

IoT 게이트웨이는 지역적으로 분산되어 있으며 버스 터미널, 쇼핑센터, 도로, 공원 등 매우 일반적인 장소에 배치된다. 기존의 클라우드 서버와 마찬가지로 이러한 IoT 게이트웨이는 서버 기능이 있어 가상화되어 스토리지, 컴퓨팅 및 네트워킹 기능을 갖추고 있다. 이러한 경량 서버를 마이크로 서버(Micro Server), 마이크로 데이터센터(Micro Data Center), 나노 서버(Nano Server) 등으로 부르기도 한다.

IoT 게이트웨이는 센서/디바이스를 네트워크를 통해 클라우드 서버로 연결한다. 엣지에 있는 IoT 애플리케이션이 수행되면 엣지에 있는 데이터를 클라우드 서버에 전송하고, 데이터 처리나 저장 작업을 수행한다. IoT 게이트웨이는 센서/디바이스가 생성한 데이터를 네트워크 통해 클라우드 서버로 전달하기 전에 데이터에 적절한 처리를 수행한다. 예를 들어, IoT 게이트웨이는 센서/디바이스에서 측정하고, 수집한 데이터를 그대로 클라우드에 전달하는 것이 아니라 통계로 처리하여 데이터 용량을 줄여서 클라우드 서버로 전달한다. 일반적으로 IoT 게이트웨이는 중복성과 불필요한 오버헤드를 줄이기 위해 데이터 전 처리를 수행하게 된다. 또한, 클라우드 서버에서 처리가 끝난 후 결과가 만들어지면 최종

결과를 사용자에게 전달한다.

IoT 게이트웨이가 구현되기 위해서는 서버의 기능을 수행하기 위해 앞서 설명한 Cloudlet 기능이 포함되어있다. Cloudlet은 엣지 디바이스와 클라우드 사이에 위치하는 마이크로 클라우드 서버이다. Cloudlet은 게이트웨이가 클라우드 서버의 역할을 하도록 설계되었으며 중앙집중형 클라우드 서비스의 단점을 보완한다. Cloudlet 기반 엣지 컴퓨팅은 게이트웨이를 가상화하여 여러 엣지 디바이스들의 요청을 동시에 처리할 수 있다.

애플리케이션 관점에서 보면 엣지 디바이스에 가깝게 위치하고 있는 게이트웨이는 사용자가 요청하는 대부분의 간단한 명령들을 처리하고, 클라우드 서버는 대용량 데이터 처리, 데이터의 저장과 빅데이터 연산 같이 거대한 자원을 요구하는 연산을 처리한다.

(3) 클라우드 플랫폼

클라우드 서버는 네트워크를 통해 사용자로부터 데이터와 명령을 받는다. IoT 애플리케이션을 지원하기 위해 클라우드 서버는 계산과 스토리지를 위한 상당한 크기의 자원을 가지고 있다. 그러므로, 엣지 컴퓨팅 환경에서 클라우드 서버는 애플리케이션의 자원 요구사항을 충족해야 한다. 데이터 처리가 완료되면 클라우드 서버는 최종 사용자에게 결과를 전송한다.

성능 측면에서 보면 클라우드 서버는 수많은 엣지의 명령을 처리해야 하고, 엣지 디바이스에서 멀리 떨어져 배치됨에 따라, 네트워크에서 전송 지연시간이 중요하다. 이를 위해 클라우드 서버는 더 많은 컴퓨팅 성능과 더 많은 데이터 스토리지를 보유할 뿐만 아니라 대규모 병렬 데이터 처리, 데이터 마이닝, 빅데이터 관리, 기계학습 등을 통해 효과적으로 사용자의 명령을 처리하기 위한 다양한 기술 솔루션을 채택하게 된다.

2 » 엣지 컴퓨팅의 실시간 처리

현재 많은 데이터가 생성되고 있으며, 생성되는 데이터의 양은 갈수록 급증하고 있다. 또한, 데이터를 수집·분석해 의미 있는 결과를 도출하고 이를 빠르게 의사결정에 반영하는 것이 오늘날 기업의 경쟁력이라 할 수 있다.

빅데이터 분석 기술의 발전으로 대용량 데이터의 처리가 가능하게 되었다. 하지만 데이터를 수집하고 분석하는데 오랜 시간이 걸린다면 데이터 분석의 의미는 없게 된다. 특정 사안에 대해서 의사결정이 신속이 필요한데 중요한 시점이 지나고 나서 데이터 분석이 완료된다면, 많은 시간, 비용과 인력만 낭비하는 것이다.

엣지 컴퓨팅은 이처럼 데이터 분석에 오랜 시간이 걸리는 것을 방지하기 위해서 데이터를 데이터센터로 전달되기까지 필요한 시간과 트래픽을 절감하고, 간단한 데이터 처리는 현장(엣지)에서 바로 처리함으로써 효율성과 비용 절감을 동시에 충족시킨다.

그러나 데이터의 종류가 다양하고, 크기도 제각각인 만큼 엣지에서 처리할 수 있는 데이터는 한정적이다. 이에 엣지에서는 복잡한 분석이 필요 없는 간단한 처리만 하고 그 외 다른 데이터들은 데이터센터로 보내 분석을 진행하는 방식이 사용된다. 예를 들어 엣지 컴퓨팅을 활용하고 있는 스마트 팩토리를 보면 고도의 데이터 분석이 필요 없는 공장 내 온도도 조절 등은 엣지에서 처리하고, 사고 위험 예측이나 기계 수명 관리 등 보다 고차원적인 분석이 필요한 데이터들은 데이터센터로 전송해 분석하는 형태로 시스템을 운영한다.

물론 엣지 컴퓨팅이 처음부터 지금과 같은 역할을 했던 것은 아니다. 스마트 팩토리 초기에는 아두이노(Aduino)나 라즈베리 파이(Raspberry Pi)와 같은 초소형 컴퓨터를 이용해 특정 데이터가 수집되었을 때에 알람을 주는 기초적인 기능으로 엣지 컴퓨팅을 활용했다. 그러나 점차 데이터의 종류가 많아지고 크기도 커지면서 초소형 컴퓨터만으로 데이터를 처리하기 어려워졌고, 이에 좀 더 강력한 연산을 위해 CPU가 장착된 미니 PC 형태로 발전해왔다. 현재는 이를 넘어서 데이터센터 수준의 스펙을 갖춘 마이크로 데이터센터까지 발전된 상태이다. 엣지 컴퓨팅의 크기나 규모는 점차 변하고 있지만, 현장에서 데이터를

바로 분석해서 의사결정에 활용할 수 있도록 한다는 엣지 컴퓨팅의 본질은 그대로다.

정보통신 기술로 각 산업영역의 경계가 없어지는 디지털 전환(Digital Transformation)의 시대로 접어들면서 비즈니스 변화 속도는 모든 산업에서 강조되고 있다. 다양한 기기가 연결되고 모든 기업이 이를 통한 새로운 비즈니스를 창출하는 시대에 엣지 컴퓨팅은 중요한 역할을 한다. 엣지 컴퓨팅은 실시간 분석을 통해 비즈니스 인사이트를 얻는 시간을 단축하여, 기업이 적절한 시점에 의사결정을 하도록 한다. 엣지 컴퓨팅이 확산됨에 따라 기업은 비즈니스를 혁신할 수 있는 새로운 방법을 찾고, 사물인터넷에 인공지능을 결합한 인텔리전트 엣지(Intelligent Edge) 전략을 활용해 비즈니스 혁신을 가속하는 중요한 도구로 사용된다.

3 » 엣지 컴퓨팅의 보안

엣지 컴퓨팅이 기본적으로 네트워크 환경에서 동작하기 때문에 네트워크에 연결된 산업용 제어 장치 같은 엣지 디바이스는 해커에 의한 공격받을 수 있게 노출된다는 것을 의미한다. 사실 사물인터넷이 도입되면서 지속해서 보안에 대한 우려가 커지고 있는 것은 현실이다. 자율주행차량, 스마트 가전제품같이 무선 인터넷과 기타 네트워크로 연결되는 제품이 늘어남에 따라 새로운 보안 요구사항이 증가하고 있다. 사실 소비자용 사물인터넷에 대한 공격이 널리 알려졌지만, 피해 규모를 예상하여 보면 제조업과 기간 산업 같은 곳의 보안 위협이 사회 안정과 밀접하여서 그 위협의 정도는 더 크다고 할 수 있다.

2018년 포스포인트 정보 보안 예측 보고서는 사물인터넷에서 네트워크에 있는 중간자 공격(MITM: Man-in-the-Middle Attack) 가능성이 커지고 있음을 보고했다. 특히, 2019년에는 공격자들이 기본적인 클라우드 인프라를 공격하여 산업용 사물인터넷 장치에 침입할 것으로 예측한다. 이는 공격자 관점에서 여러 기업이 공통으로 쓰는 기본적인 시스템에 접속하게 되면, 피해 규모가 크기 때문에 더 많은 보상을 받을 수 있어 공격자가 선호할 것으로 예상된다. 도로에 있는 CCTV 같은 공공 장비도 그 대상이 될 것이다.

엣지 컴퓨팅이 도입되면 보안 측면에서 다음과 같은 과제가 중요해질 것이다. (1) 엣지 컴퓨팅이 확대됨에 따라 네트워크 연결과 의존이 증가한다. (2) 엣지 컴퓨팅에서 많은 연산이 클라우드에서 엣지 디바이나 IoT 게이트웨이로 이동함에 따라 이러한 장치들의 보안을 유지하기 위한 대비가 필요하다. (3) 엣지 디바이스들이 업데이트 및 유지 관리를 클라우드를 통해 모니터링되고, 업데이트되고, 관리되어 안전한 운영을 보장하여야 한다.

서버 보안 관점에서 클라우드 서비스는 보안과 효율성의 균형을 유지하고 있었다. 그동안 클라우드 서비스 사업자들은 효율성과 확장 가능한 서비스를 위해 인프라, 플랫폼, 애플리케이션을 여러 기업이 공유하는 관점에서 시스템에 구축해 놓았다. 예를 들어 클라우드 인프라의 구성을 하나의 인프라를 여러 사용자가 공유하는 멀티-테넌트(Multi-Tenant) 아키텍처로 제공하였다. 그러나, 각각의 사용자 환경이 충분히 격리되어 있지 않아 안전한

운영을 제공하지 못할 수도 있다. 만일 제조업, 에너지 생산업, 기타 중요한 분야의 산업용 사물인터넷의 경우, 서버에 대한 공격으로 불가피하게 중요 서비스가 갑자기 중단되는 사태가 발생할 수 있다.

클라이언트 보안 관점에서 최근 IoT 장치들이 확산되면서 휴대전화, IoT 센서 등 엔드 포인트(end point)를 대상으로 하는 새로운 공격 형태가 등장하고 있다. 이러한 이유로 엣지 컴퓨팅을 도입한다고 하더라도 보안에 대한 강화는 지속해야 한다.

보안 신뢰도를 높이기 위해 보안에 필요한 기능은 말단에 있는 엣지부터 적용하고, 고성능 보안 애플리케이션을 도입하여 침입과 위협으로부터 방어하기 위한 계층별로 보안 정책을 적용해야 한다. 이를 통해 중요한 데이터의 접근을 모니터하고, 제어하고, 이 데이터를 클라우드에 전송하기 전에 정보 보호 정책을 시행하여 보안을 강화해야 한다. 즉, 엣지 컴퓨팅에서 IoT 데이터의 전송 과정과 저장 과정에서 모두 보안을 보장하기 위해, 데이터가 전송되기 전후에 외부 공격에 대해 모니터하고, 공격이 있다면 이에 대해 적절한 조치를 해야 한다.

사실 엣지 컴퓨팅의 동작 환경은 지리적으로 넓고 분산된 지역이고, 도로, 철도, 변전소, 차량과 같이 열악하다. 따라서 이런 환경에 설치된 기기에서 수집된 데이터를 보호하여야 하기에 보안을 보장하는 것이 쉽지 않다. 그러나, 엣지 컴퓨팅 인프라 내의 엣지 노드와 포그 노드에 대해서 동일한 정책, 제어 및 절차를 사용하여 전체 시스템의 보안을 강화해야 한다.

2018년에는 프로세서 하드웨어의 문제점을 이용해 소프트웨어 및 펌웨어 계층을 우회할 수 있는 보안 취약점인 멜트다운(Meltdown) 및 스펙터(Spectre)가 발견되어 그동안 믿고 있었던 하드웨어에 대한 신뢰가 무너지게 되었다. 1995년 이후 나온 거의 모든 중앙처리 장치(CPU)가 취약한 것으로 보이며, 그 이후 스펙터의 새로운 변종이 발견되었다. 공격자는 산업용 사물인터넷 시스템이 사용하는 클라우드 인프라를 파괴하는 변종을 개발할 것으로 예측되어 더 많은 취약점이 발생하게 될 것이다.

이처럼 엣지 컴퓨팅은 클라우드 컴퓨팅과 사물인터넷이 가지고 있는 보안의 문제점을

모두 가지고 있으나, 분산 환경으로 인해 효과적인 보안 체계를 구축할 수 있다. 다시 강조하지만 엣지 컴퓨팅이 된다고 클라우드 컴퓨팅과 사물인터넷의 보안 문제점이 해결되는 것은 아니라, 기존 환경보다 높은 수준의 보안에 대한 대비가 필요하다.

엣지 컴퓨팅의 분산 환경에서 개인정보 보호를 개선하는 한 가지 매력적인 방법은 데이터를 제어하고 처리하는 기능이나 체계를 엣지가 있는 엔드포인트로 옮기는 것이다. 엣지 컴퓨팅은 클라우드와 엔드포인트가 협력하여 보안을 강화는 방식이다. 예를 들어 개인정보의 저장과 관리를 엣지 디바이스나 포그 노드로 한정한다면 클라우드 컴퓨팅과 비교하면 보안이 강화될 수 있다. 이때 개인정보 관련 처리는 엣지나 포그에서 이루어지고 클라우드로는 개인정보 대신에 요약 정보인 메타데이터만 전송되기 때문에 사용자의 개인정보를 보호할 수 있다.

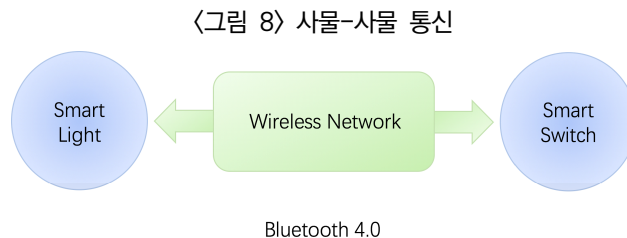
이런 관점에서 보면 많은 서비스 업체가 보다 높은 수준으로 개인정보를 보호하기 위해 엣지 컴퓨팅을 활용하기 시작할 것으로 예측되지만 아직 충분한 성공 사례가 존재하지 않아 사용자로부터 신뢰를 얻기 위해서는 많은 시행착오가 필요할 것으로 예측된다.

4 » 엣지 컴퓨팅의 통신과 네트워크

엣지 컴퓨팅의 기반이 되는 통신과 네트워크를 살펴보고자 한다. 엣지 컴퓨팅에서 통신 모델은 크게 사물-사물 통신, 사물-클라우드 통신, 사물-게이트웨이로 나누어진다.

(1) 사물-사물 통신(Machine-To-Machine Communication)

이 통신 모델은 중간에 다른 하드웨어 지원 없이 서로 직접 연결하고, 정보를 교환할 수 있는 구조를 나타낸다. 이러한 장치는 인터넷의 IP 네트워크로 연결될 수 있고, 그 밖에 다양한 유형의 네트워크를 통해 서로 연결할 수 있다. 예를 들어 아래 그림 8은 스마트 스위치가 Bluetooth 4.0을 통해 스마트 조명기와 통신한다는 것을 보여준다.



이러한 사물 네트워크는 기기들을 직접 연결하여 QoS 요구사항을 만족하는 통신 프로토콜로 정보를 교환할 수 있도록 한다. 이런 방식은 일반적으로 작은 데이터 패킷을 전송하여 통신하고, 상대적으로 낮은 데이터 전송률이 요구되는 스마트 홈, 전기 시스템을 이용한 자동 제어 같은 애플리케이션에 적합하다. 이러한 유형의 대표적인 IoT 장치는 스마트 도어, 스마트 스위치, 스마트 조명 등으로 일반적으로 적은 양의 데이터만 통신한다.

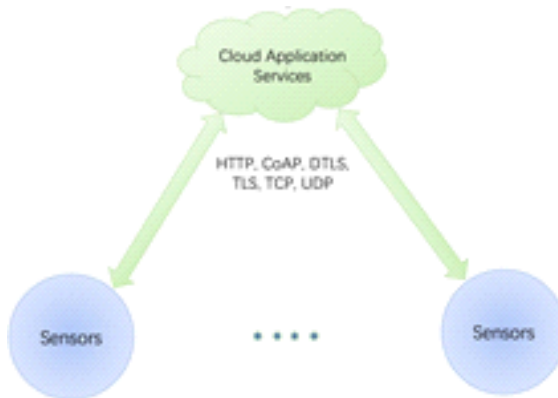
사용자 관점에서 볼 때 기계 간 통신의 문제는 서로 다른 제조업체의 서로 다른 장치가 서로 다른 프로토콜을 사용하여, 호환성이 부족하다는 것이다. 스마트 홈 장치를 예로 들면, Z-Wave 프로토콜 장치는 ZigBee 프로토콜 장치와 통신할 수 없으므로 특정 프로토콜, 심지어 특정 제조사만 연결할 수 있다. 이러한 호환성 문제로 다양한 제품을 혼합하여

네트워크를 구성할 수 없게 된다. 결과적으로 사용자가 구성하려는 컴퓨팅 환경에 대한 선택과 경험을 제한하여 결국 IoT 시장이 성장하는데 장애 요소로 작용하고 있다.

(2) 사물-클라우드 통신(Machine-Cloud Communication)

이 방식은 클라우드 컴퓨팅에서 사용하는 방식이다. 사물-클라우드 통신에서 IoT 장치는 연산 능력과 스토리지 공간의 제약으로 인해 클라우드에 서비스를 요구하거나 클라우드 스토리지에 데이터를 저장한다. 이 방식은 일반적으로 그림 9에서 나타낸 것과 같이 기존의 유선 또는 Wi-Fi 연결을 통해 이루어진다.

〈그림 9〉 사물-클라우드 통신

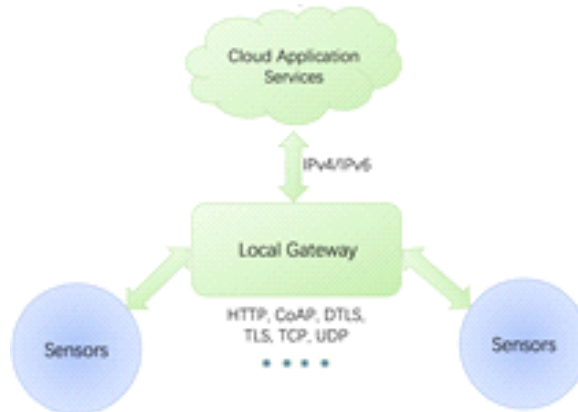


사물-클라우드 통신은 사물 간 통신의 문제점을 해결하지만, 이 모델은 전통적인 네트워크에 연결되어 대역폭과 네트워크 자원을 사용하여 통신 성능이 네트워크에 의해 제한된다. 사물-클라우드 통신의 성능을 향상하려면 네트워크 구조를 최적화해야 한다.

(3) 사물-게이트웨이 통신(Machine-Gateway Communication)

그림 10에서 보듯 사물-게이트웨이 모델에서는 사물과 클라우드를 연결하는 로컬 게이트웨이라는 미들웨어가 존재한다.

〈그림 10〉 사물-게이트웨이 통신



IoT 기기와 클라우드 서비스 간에 중간 매개체를 담당하는 게이트웨이는 다른 네트워크와 보안 체계를 유지하거나, 데이터를 다른 네트워크로 전송할 때에 프로토콜을 변환하는 역할을 한다. 이를 통해 IoT 네트워크의 보안을 강화하고, 다른 사물이나 네트워크의 연결을 자유롭게 한다. 또, 컴퓨팅 작업의 일부를 게이트웨이로 이동(migration)하면 IoT 장치의 전력 소비를 많이 감소시킨다. 예를 들어, 스마트폰은 IoT 기기와 클라우드와 통신하기 위해 일부 애플리케이션을 실행하는 게이트웨이 역할을 한다. 센서가 연결된 스마트폰으로 개인용 건강 서비스를 제공할 때에 센서가 데이터를 생성하고, 스마트폰이 데이터를 암호화하고 클라우드 서비스로 전송하면 센서의 전력 소비를 크게 줄일 수 있다.

사물-게이트웨이 통신 모델은 엣지 컴퓨팅의 기본 통신 모델로 사용된다. 넓은 의미에서 엣지 컴퓨팅의 통신은 앞서 설명한 세 가지 모델, 사물-사물 통신, 사물-클라우드 통신, 사물-게이트웨이 통신 모두를 포함한다.

최근 엣지 컴퓨팅이 주목받는 이유 중 하나는 5세대 이동 통신(5G)의 상용화로 기인한다. 이동 통신 업계는 5G의 중요한 응용으로 엣지 컴퓨팅을 언급하고 있다. 5G 통신을 통해 통신사의 모바일 기지국이나 와이파이(Wi-Fi) AP(Access Point)를 단순히 신호 중계기로만 이용하는 것이 아니라, 핵심 데이터를 실시간 처리하는 엣지 컴퓨팅의 게이트웨이로 용도를 확대하게 된다. 즉, 모든 처리를 통신사의 서버로 보내서 하지 않고, 간단한 서비스를 위한 컴퓨팅 자원과 애플리케이션을 게이트웨이에 배치하여 유연하고 빠른 서비스

를 제공한다. 여기에 5G의 빠른 통신 속도로 엣지 컴퓨팅이 실시간 응용을 지원하게 된다.

2018년 12월부터 SK텔레콤, KT, LG유플러스 등 국내 이동 통신 3사가 5G 전파를 송출하면서 5G 상용화 시대가 시작되었다. 현재 서비스가 이뤄지는 5G는 B2B(기업 간 거래)이고 LTE 네트워크와 5G 네트워크가 혼재하여 있는 NSA(None Stand Alone)이다. 사실 NSA 환경에서는 5G가 목표로 하는 충분한 성능을 내기는 어렵고, 5G 단독모드인 SA(Stand Alone)가 되어야 5G 서비스가 제대로 제공될 수 있다.

2018년 6월에 민간표준화기구 3GPP에서 5G SA에 대한 1차 표준(릴리즈 15)이 만들어졌고, 5G SA에 대한 2차 표준(릴리즈 16)이 2019년 말에 이뤄질 예정이다. 5G 표준화는 지난 6월, 릴리즈 15에서 핵심 기능은 마무리되었지만 기존 표준의 성능 향상 부분과 융합 신산업 분야인 고신뢰·초저지연(URLLC), 차량간통신(V2X)에 대한 표준을 릴리즈 16에 포함하게 된다. 트래픽 증가 속도, 네트워크 장비 개발 진행 상황, 정부 규제 상황을 고려하면 5G SA 서비스는 빠르면 2020년 이후에나 상용화될 전망이다. 따라서 그때까지는 진정한 5G 서비스를 제공받기 어려울 것으로 예상된다.

5G 스마트폰이 2019년에 출시되지만, 네트워크 상황은 아직 제약이 많을 것이다. 통신 속도도 그리 빠르지 않을 것이다. 그나마 몇몇 대도시에서, 그것도 일부 지역에서만 가능할 것이다. 그 이유는 5G의 서비스 제공 범위, 기지국 커버리지의 특성과 관계가 있다. 5G는 몇백 미터 수준의 범위를 가질 것이 유력하다. 따라서 충분한 기지국이 설치되지 않으면 창가에 기대어 더 좋은 신호를 찾는 예전의 풍경이 다시 나타날 수도 있다. 상대적으로 4G의 경우, 18km 이상의 커버리지를 가지고 있다.

따라서, 통신사들이 원활한 5G 커버리지를 유지하기 위해서는 이동통신용 초소형 기지국인 펌토셀(Femtocell)을 곳곳에 배치해야 하는데 이를 위해서는 많은 투자가 필요하다. 펌토셀 자체뿐만 아니라 설치하기 위한 장소를 확보하여야 하고, 설치 장소까지 고속 광케이블을 연결해야 기대하는 5G 속도를 유지할 수 있다. 이러한 투자가 충분히 이루어지지 않는다면 결국 5G이지만 충분한 성능을 낼 수 없게 된다. 따라서, 초기 5G의 시장은 기업 내부나 컨벤션 센터와 같이 밀집된 공간에서 활용될 것으로 예측되고, 앞으로 적어도 몇 년 동안은 4G를 유지할 것으로 예상된다.

5 » 엣지 컴퓨팅의 표준

엣지 컴퓨팅이 활성화되기 위해서는 표준 문제가 해결되어야 한다. 그래야 다양한 엣지 디바이스들이 하나의 엣지 컴퓨팅 환경에 통합되고, 서비스될 수 있기 때문이다. 통합된 엣지 컴퓨팅 환경이 만들어져야 의미 있는 비즈니스 모델이 만들어질 수 있다. 반대로 생각해 보면 아직 엣지 컴퓨팅에 킬러 애플리케이션이나 서비스가 없어서 기업들의 참여가 미흡하고, 이로 인해 표준화가 더디다. 이러한 교착상태를 해결하는 것이 엣지 컴퓨팅 표준화의 중요한 숙제이다.

사실 새로운 기술이 도입되었을 때에 이러한 상황은 반복되었다. 인터넷이 도입된 1970년대 후반에서 1980년대 상황이 유사하다. 기업들이 새로운 산업이 시작될 때에 시장을 지배하는 단일 사업자가 없는 상황에서 표준화로 산업이 개방되는 것에 위험 부담을 느끼게 되어 적극적인 표준화를 주저하게 된다. 하지만 시간이 지남에 따라, 인터넷을 활용한 다양한 애플리케이션들이 도입되고, 비즈니스 모델이 만들어지면서 점차 기업들의 참여가 이루어졌다. 1990년대 초 월드와이드웹(WWW)이 킬러 서비스로 등장하면서 시장이 폭발적으로 성장하고, 표준화와 이에 따르는 인터넷 인프라의 구축이 가속도가 붙게 되었다. 엣지 컴퓨팅도 충분히 사용자에게 매력적인 킬러 애플리케이션이 만들어진다면 인터넷만큼 오랜 시간이 걸리지 않고 표준화에 성공할 것으로 예측된다.

엣지 컴퓨팅 관련 표준으로는 산업체 중심의 오픈 포그 컨소시엄, 유럽에는 유럽 전기 통신 표준 협회, 중국은 엣지 컴퓨팅 컨소시엄 등이 적극적으로 활동하고 있다.

오픈 포그 컨소시엄(Open Fog Consortium)은 ARM, Cisco, Dell, Intel, Microsoft, 그리고 프린스턴 대학교가 함께 결성한 표준 단체로 포그 또는 엣지 컴퓨팅 응용을 산업에 확산하기 위해서 노력하고 있다. 이를 위해 서로 다른 업체들이 생산한 디바이스들의 상호 운용성과 확장성을 보장하도록 표준화하고 있다. 또한, 시스템 구축에 필요한 오픈 아키텍처(Open Architecture)를 만들어 산업 혁신을 촉진 시키는 데 노력하고 있다.

유럽에서는 Network Edge Europe과 Cloud for Europe 등 유럽 클라우드 사업자들

이 유럽 전기 통신 표준협회(ETSI : European Telecommunications Standards Institute)를 통해 엣지 컴퓨팅을 표준화하는 계획을 추진하고 있다. 이 프로젝트는 네트워크 기능 가상화(NFV:Network Functions Virtualisation)와 모바일 엣지 컴퓨팅, 자동 구성 네트워크(SON: Self-Organizing Network)를 중심으로 표준화를 진행하고 있다.

중국에서는 엣지 컴퓨팅 컨소시엄(ECC:Edge Computing Consortium)이 설립되었다. 엣지 컴퓨팅 컨소시엄은 화웨이, 중국 과학원 심양 연구소, 중국 정보 통신 기술원(CAICT), Intel, ARM, 그리고 ISoftStone이 공동으로 설립하였다. 엣지 컴퓨팅 컨소시엄은 궁극적으로 디지털과 산업의 통합, 즉 디지털 전환(Digital Transformation)을 목표로 하고 있다. 또한, 중국의 산업 특성상 강력한 표준 정책을 채택하고 있어, 기업들에 동일한 엣지 컴퓨팅 아키텍처를 제시하고 있고, 운영과 보안 관련 지침도 제공하고 있어 표준이 적용되는 속도가 빠를 것으로 예측된다.

일본의 오픈 포그 컨소시엄(Open Fog Consortium)은 미래 인터넷 인프라 구축을 위한 새로운 비즈니스 모델 개발과 표준화 수립을 목표로 하고 있으며 미국 전기 전자학회(IEEE)와 협력을 통해 표준을 개발하고 있다.

산업 분야별 엣지 컴퓨팅을 활용하기 위한 산업체 간 동맹도 본격화됐다. 도요타는 지난 7월 인텔, 에릭슨, NTT, NTT도코모, 덴소, 도요타 산하 인포테크놀로지 센터 등이 참여하는 가운데, 오토모티브 엣지 컴퓨팅 컨소시엄(AECC : Automotive Edge Computing Consortium)을 출범시켰다. 이 컨소시엄은 차량용 엣지 컴퓨팅 기술 개발을 통해 클라우드의 대안을 마련하는 것을 목표로 하였다. 커넥티드카와 자율주행차 확산에 따라 차량과 클라우드 간 오고 가는 데이터 규모는 지금보다 많이 늘어날 것으로 전망된다. 이에 컨소시엄은 엣지 컴퓨팅을 통해 지능형 주행, 실시간 데이터 기반 지도 생성, 운전 지원과 같은 신규 서비스를 지원하는 생태계를 육성해 나간다는 전략이다.



엣지 컴퓨팅의 응용 분야

글로벌 ICT 리서치 업체 가트너(Gartner)는 2017년 글로벌 사물인터넷 디바이스가 지난해보다 31% 증가해 84억 대에 달할 것으로 전망했으며, 2016년부터 2020년까지 연평균 34%로 증가해 2020년에는 204억대에 육박할 것으로 예측했다. 2017년 전체 디바이스와 관련 서비스의 지출 규모는 1조 6,896억 달러(약 1,940조 원)에 달할 것으로 예상하며, 2016년부터 2020년까지 연평균 21%로 성장해 2020년에는 시장규모가 2조 9천억 달러에 이를 것으로 예측했다.

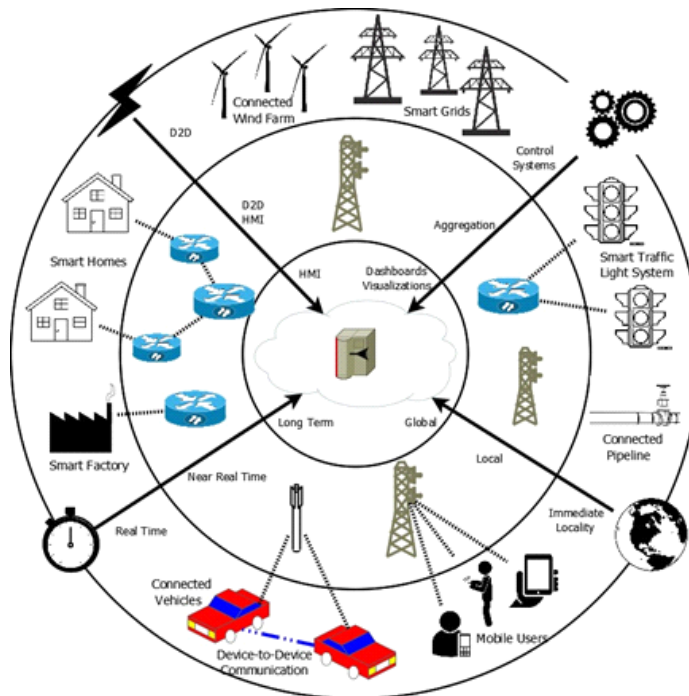
특히, 산업용 사물인터넷의 글로벌 시장규모는 2015년 1,137억 달러를 기록했으며, 연평균 8%씩 성장을 지속해 2022년에는 1,954억 달러의 대규모 시장을 형성할 것으로 전망되었다. 산업용 사물인터넷이 운송, 제조, 에너지, 헬스케어, 소매업 등 산업 전 분야에 적용되면서 4차 산업혁명을 견인할 것으로 예상된다.

이러한 산업 전망에 따라 사물인터넷에 기반한 엣지 컴퓨팅의 장래는 밝다고 할 수 있다. 특히 최근 인공지능기반의 사물인터넷 시대가 열리게 되어 통신이 가능한 다양한 IoT 디바이스로부터 수많은 데이터가 수집되기 때문에, 데이터 기반으로 머신러닝을 통해 해당 분야의 산업을 개선하거나 비즈니스 혁신을 가져올 것이다.

인공지능 기반의 사물인터넷이 구현되기 위해서는 이를 지원하는 컴퓨팅 환경이 구축되어야 하고, 이러한 측면에서 엣지와 게이트웨이에 있는 연산과 저장 기능으로 다양한 서비스를 실시간으로 제공할 수 있는 엣지 컴퓨팅이 그 해결책으로 제시되고 있다. 엣지 컴퓨팅은 데이터가 수집되는 끝단인 엣지에서 데이터를 즉시 분석하고 현장에 적용하기 때문에 클라우드를 통해 데이터 분석 결과를 가져오는 것보다 즉시성과 안정성이 담보되는 컴퓨팅 기술이라고 할 수 있다.

이번 장에서는 엣지 컴퓨팅의 장점을 살릴 수 있는 다양한 응용 분야를 살펴보고자 한다. 아래 그림 11에서 보듯이 엣지 컴퓨팅의 응용 분야는 매우 다양하다. 이들 중 대표적인 분야를 하나씩 살펴본다.

〈그림 11〉 엣지 컴퓨팅의 응용 분야



출처 : Internet of Things: Principles and Paradigms 2016, Pages 61–75 Chapter 4: Fog Computing: Principles, Architectures, and Applications

1 » Smart Factory

4차 산업혁명, 제조업 4.0으로 인해 스마트 팩토리(Smart Factory)가 주목받고 있다. 스마트 팩토리는 국가 차원에서 산업 혁신을 추진하고 있는 독일과 뒤를 이어 미국, 일본이 시장에 뛰어들면서 활발히 성장하고 있다. 이 중에서도 스마트 팩토리 분야에 한발 늦게 진입한 일본 기업들이 엣지 컴퓨팅을 차별화된 개념으로 강조하며 새로운 사업 기회를 모색하고 있다.

엣지 컴퓨팅이 적용된 스마트 팩토리를 살펴보면 공장 내 온도, 습도 조절이나 공장의 가동 현황, 로봇이나 기계에서 발생한 데이터로 실시간 생산현황을 파악하여 생산량을 효율적으로 관리한다. 이때 단순 데이터 처리는 클라우드에서 처리하는 대신에 엣지 영역인 공장 내에서 데이터를 분석하고, 이를 기반으로 제어하여 실시간으로 기기가 지능적으로 작동할 수 있도록 한다. 반면에 클라우드에서는 기계 수명 관리나 사고 위험 예측 같은 고차원 분석을 수행하고, 이에 필요한 데이터는 중앙 데이터센터에 저장하여 처리한다.

결과적으로 제조업체들은 신속한 의사결정 지원을 위해 매일 수많은 데이터가 발생하는 상황에서 실시간 데이터 분석이 가능한 엣지 컴퓨팅을 채용하게 된다. 엣지 컴퓨팅은 생산 현장에서 실시간 모니터링과 제어, 예측, 정비를 가능케 해 전반적인 생산공정을 최적화시켜, 생산성 향상과 비용 절감을 기대할 수 있다.

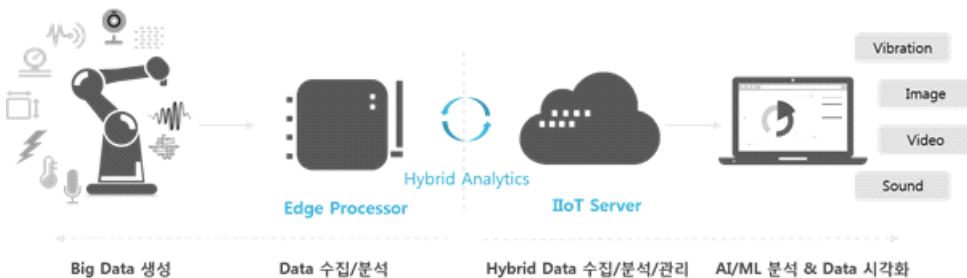
엣지 컴퓨팅에 인공지능을 결합한 엣지 인텔리전스는 데이터 전송비용 절감, 보안 개선, 실시간 의사결정에 따른 생산설비 최적화 등의 지능형 처리가 가능하다. 엣지 인텔리전스는 기계 부품들이 기대하는 성능 목표를 충족하는지 확인할 수 있도록 생산 현장에서 데이터를 수집하고 분석한다. 이로 인해 엣지 컴퓨팅은 원격으로 장애를 예방하고 장비의 보수에도 기여하여 생산설비의 중단 시간을 최소화하고, 노동 생산성을 높이고 에너지 소비를 줄이는 효과가 있다.

엣지 컴퓨팅을 활용하여 단순히 제조 생산성을 높이는 것을 넘어 사업 분야를 서비스로 확장할 수도 있다. 하나의 예로, 상황을 예측하여 유지 관리 비용을 획기적으로 줄이는

스마트 엘리베이터가 있다. 스마트 엘리베이터는 소음, 진동, 온도 등에 대한 데이터를 수집하기 위해 많은 센서를 사용한다. 이때 엘리베이터의 운전상태는 수집된 데이터 분석에서 파악이 된다. 이때 엣지 디바이스는 적외선 센서, 중량 센서, 연기 감지기, 진동 센서, 노이즈 센서, 카메라 등이 될 수 있고, 엣지 노드는 엘리베이터 운영 센터 또는 엘리베이터 자체가 된다.

이 엘리베이터를 엣지 컴퓨팅에 연결하고 수집된 데이터를 클라우드로 연동하여 엘리베이터 운영자는 모든 엘리베이터의 작동 상태를 파악할 수 있다. 엘리베이터 기술자는 엣지 컴퓨팅 데이터와 클라우드의 데이터를 사용하여 분석에 따라 예측 유지보수를 수행하여 고장 가능성이 큰 엘리베이터를 선택적으로 미리 점검하고 유지 관리할 수 있다. 예측 가능한 유지보수는 장애 예방과 예측하지 못하는 정지 시간 방지를 통해 유지보수 비용을 절감하는 동시에 장비의 운영 효율성을 향상시킨다.

〈그림 12〉 엣지 컴퓨팅을 활용한 Smart Factory



출처: <http://www.alooh.io/platform/>

2 >> Smart City

도시에서 공공 자원을 효과적이고 효율적으로 사용하고, 시민들의 생활 수준을 높이기 위해, 스마트 시티가 제안되고 많은 도시가 구축되고 있다. 스마트 시티는 주로 안전이나 시설물 관리에 적용되나, 스마트 신호등과 연동하여 스마트 교통관제에도 적용된다. 예를 들어, 스마트 교통 신호등은 사고 발생 시 인근 차량에 경고 메시지를 보내고 신호등끼리 서로 연동이 된다. 또, WiFi, 이동통신망, 도로변 장치를 통해 스마트 교통 신호등과 차량이 상호 연동된다.

엣지 컴퓨팅은 데이터 소스에 최대한 가깝게 처리가 이루어지고, 분산처리가 가능하므로 엣지 컴퓨팅이 스마트 시티에 적용되면 가정에서 지역사회로, 더 나아가 도시 규모로 유연하게 확장될 수 있다. 엣지 컴퓨팅은 다음과 같은 특성으로 인해 스마트 시티에 이상적인 플랫폼이 된다.

- (1) 대규모 데이터: 인구 100만 명이 거주하는 도시는 매일 180PB의 데이터를 생산하게 되며, 공공 안전, 보건, 유틸리티 및 교통 등에 활용된다. 대규모의 데이터가 생성되기 때문에 모든 데이터를 처리할 수 있는 중앙 클라우드 데이터센터를 구축하는 것은 비현실적이다. 따라서 엣지 컴퓨팅을 활용하여 각 네트워크 엣지에서 데이터를 처리함으로써 효율적인 솔루션이 된다.
- (2) 짧은 대기 시간: 스마트 시티의 교통이나 안전에 관련된 비상 상황이 발생할 때 이에 대한 처리는 예측할 수 있고 대기 시간이 짧아야 한다. 엣지 컴퓨팅은 데이터 전송 시간을 단축하고 네트워크 구조를 단순화할 수 있으므로 스마트 시티에 적절한 환경이다.
- (3) 위치 인식: 교통 및 유틸리티 관리와 같은 위치 기반 애플리케이션의 경우, 위치 인식으로 인해 엣지 컴퓨팅이 클라우드 컴퓨팅보다 효과적이다. 엣지 컴퓨팅에서 데이터는 클라우드로 전송되지 않고 지역에 따라 분리되어 수집되고 처리될 수 있다.

스마트 시티의 예로 실시간 비디오 분석을 활용한 보안이 있다. 휴대전화와 네트워크 카메라가 널리 보급되고, 인공지능 기술이 발전하면서 비디오 분석이 새로운 응용으로 떠오르고 있다. 클라우드 컴퓨팅은 긴 데이터 전송 지연 시간과 개인정보 보호 문제로 인해 비디오 분석을 처리하는 데 적합하지 않다.

도시에서 길을 잃어버린 아이, 미아를 찾는 예를 살펴보자. 다양한 종류의 카메라가 도시 지역과 각 차량에 널리 배치되고 있다. 아이가 실종되면, 이 아이가 카메라에 포착될 가능성이 매우 크다. 그러나 다양한 기기들에서 포착된 카메라 영상은 개인정보 보호 문제나 통신 비용 때문에 클라우드에 업로드하는데 문제점이 많다. 클라우드에 데이터를 보내 광역 영상 데이터를 구축하더라도 대량의 데이터에서 특정 인물을 검색하는 데 시간이 오래 걸릴 수 있으며, 개인정보 이슈로 인해 인물 검색을 허용되지 않을 수도 있다.

엣지 컴퓨팅에서는 길을 잃은 어린이를 검색해 달라는 요청을 생성하여 해당 지역에 있는 모든 디바이스에 보낼 수 있다. 예를 들어, 스마트폰으로 요청을 수락하여 로컬 카메라로 검색할 수 있으며, 미아의 인식 여부만 클라우드에 보고할 수 있다. 이렇게 엣지 컴퓨팅에서 엣지 디바이스의 컴퓨팅 성능을 활용하는 것이 클라우드 컴퓨팅에 처리하는 것에 비해 훨씬 빠를 뿐만 아니라 개인정보 문제도 해결할 수 있다.

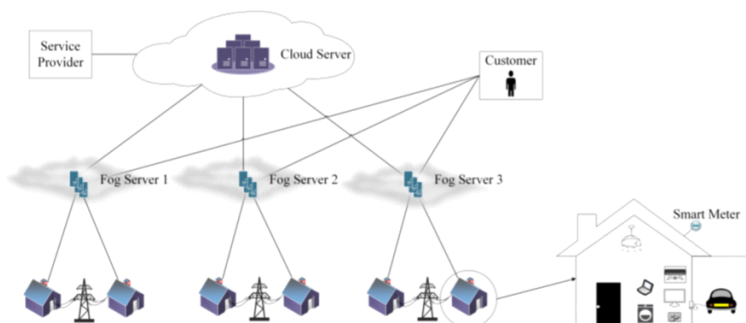
3 » Smart Grid

스마트 그리드는 기존의 전력망에 정보통신기술을 접목해 에너지 네트워크와 통신 네트워크가 합쳐진 지능형 전력망으로 전력공급자와 소비자가 실시간으로 전기사용 관련 정보를 주고받음으로써 에너지 사용을 최적화할 수 있는 서비스이다.

스마트 그리드가 제공하는 이점은 안전을 높이고, 보안을 강화하고, 효율적인 에너지 사용을 가능하게 한다. 이를 위해서는 스마트 그리드에서 데이터를 수집하고, 서로 데이터를 교환하기 위해 다수의 스마트 미터, 센서 및 액추레이터가 설치되어 있어야 한다. 또한, 전력의 제어 지연이나 전력 중단을 방지하기 위해 데이터 수집원에서 바로 동작이 처리되어야 해야 한다. 사실 다양한 지역에 분포하고 있는 계량기와 센서로부터 전력 관련 데이터를 수집 처리하고, 전력 관리를 위한 결정을 적시에 최적화하기 위한 시스템을 구축하는 것은 매우 어려운 문제이다.

엣지 컴퓨팅은 현재 중앙집중형 시스템에 비해 빠른 속도로 스마트 그리드가 전력 관련 의사결정과 전력의 전송과 관리를 할 수 있는 기능들을 데이터가 발생하는 근접한 엣지에서 수행할 수 있다. 엣지 컴퓨팅은 기존의 자동화 시스템보다 더 높은 유연성을 제공하는 것뿐만 아니라, 최소 비용으로 전체 전력 네트워크에 많은 노드를 추가할 수 있는 확장성을 제공한다.

〈그림 13〉 엣지 컴퓨팅을 활용한 Smart Grid



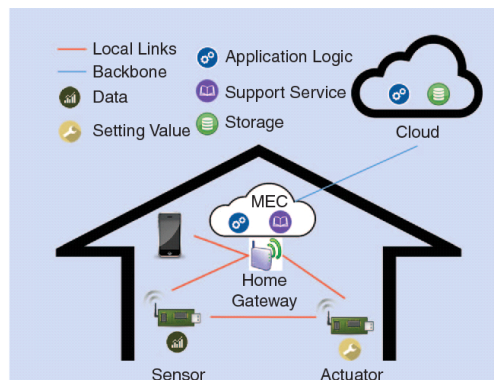
출처 : F. Y. Okay and S. Ozdemir, "A fog computing based smart grid model," 2016 International Symposium on Networks, Computers and Communications, 2016.

4 >> Smart Home

IoT는 스마트 홈 제품을 통해 가사를 포함한 개인에게 큰 도움을 준다. 스마트 라이트, 스마트 TV, 로봇 청소기 등 일부 제품이 개발되어 시판되고 있다. 그러나 홈 어플라이언스 제품에 Wi-Fi 모듈을 추가하고 클라우드에 연결하는 것만으로는 스마트 홈에 충분하지 않다. 스마트 홈 환경에서는 연결된 기기 외에도 무선 센서와 컨트롤러를 방, 파이프, 심지어 바닥과 벽에 설치된 센서로부터 수집된 데이터와 연동하여 지능적인 서비스를 제공해야 한다. 따라서, 스마트 홈 환경은 대규모의 데이터를 생성하게 되고, 이들 데이터를 클라우드에 전송하여 처리하게 된다. 또한, 이들 데이터는 가정이라는 환경에서 개인정보를 포함하고 있어 보호와 관리가 필요하다.

이러한 스마트 홈 기능을 클라우드 컴퓨팅에 구축하는 것은 적합하지 않다. 그 이유는 이 장치들이 서로 다른 업체로부터 공급되고 플랫폼이 다르므로 함께 작동하기 어렵기 때문이다. 즉, 상호 운용성이 보장되지 않는다. 엣지 컴퓨팅은 서로 다른 기기를 통합하기 위한 통합 인터페이스를 제공한다. 이로 인해 가정에서 모든 것을 쉽게 연결하고 관리할 수 있고, 인터넷 대역폭에 대한 부담을 해소하기 위해 데이터를 로컬로 처리할 수 있고, 가정에 고유한 서비스도 구축할 수도 있다. 결과적으로 엣지 컴퓨팅은 다양한 플랫폼을 통합하고 유연한 연결로 스마트 홈 애플리케이션을 강화한다.

〈그림 14〉 엣지 컴퓨팅을 활용한 Smart Home



출처: Carlo Vallati et al., "Mobile-Edge Computing Come Home Connecting things in future smart homes using LTE device-to-device communications", IEEE Consumer Electronics Magazine, 2016.

5 Connected Car

커넥티드카는 다른 차량이나 교통 및 통신 인프라, 보행자 단말 등과 실시간으로 통신하며 운전자의 편의와 교통안전을 돕고 인터넷의 다양한 서비스를 제공한다. 특히 커넥티드카는 반자동 및 자율주행 자동차로 발전하면서 차량에서 생성되는 데이터가 폭증하고 있다.

자율주행차는 엣지 컴퓨팅이 가장 유용하게 활용될 수 있는 분야라고 볼 수 있다. 자율주행차는 차량에 부착된 센서에서 실시간으로 데이터를 수집해 앞차와 거리 유지, 주변 도로 상황, 차량흐름 등을 파악하고 주행 중 돌발상황이 발생했을 때 신속하게 대처하여 사고를 피할 수 있다. 예를 들어 차량이 주행하고 있는데 전방에 아이들이 무단횡단을 한다면, 사람이 운전하고 있다면 급정차를 할지, 빠르게 지나칠지를 수 밀리 초 내에 결정할 수 있다. 반면 자율주행차는 차량에 설치된 레이더와 카메라, 초음파 그리고 위성통신 등으로부터 식별한 물체의 종류와 차량 흐름, 자동차간 거리, 주변 도로 상황 등을 클라우드로 전송한 뒤 차량의 진행 방향을 전달받게 된다. 이때 걸리는 시간은 사람이 판단을 내리는 시간보다 꽤 늦을 것이며 주행 중 네트워크 지연이나 데이터 전송 오류가 있을 수 있다. 이는 큰 사고로 이어질 수 있어 돌발상황에 유연하게 대처하지 못할 확률이 높다.

엣지 컴퓨팅 환경은 엣지(차량 내)는 물론 포그 노드(제조업체의 관제센터)에서 데이터 처리가 가능하다. 엣지 컴퓨팅을 활용하면 차량이 독립적으로 주변 환경, 주행 상태, 주행 방향 등의 데이터를 실시간으로 로컬에서 분석하여 주행한다. 기타 차량과 관련된 데이터는 제조업체에 전송해 차량 유지보수 개선이나 차량의 사용 패턴을 파악하는데 활용할 수 있다.

커넥티드카 관련하여 엣지 컴퓨팅이 활용되는 대표적인 응용 시나리오는 아래와 같다.

① 자율 지도 생성

엣지 컴퓨팅은 실시간 데이터를 활용한 자율주행에 필요한 지도 정보를 생성한다. 자율주행 자동차가 센서를 통해 수집한 차량의 위치와 주변 정보를 처리하여 클라우드에 업로드한다. 클라우드는 이들 정보를 가공하여 실시간 지도 데이터에 반영한다. 갱신된 지도

데이터는 다시 해당 지역의 차들에 배포되어 자율주행을 위해 사용된다.

② 실시간 보행자 안전

엣지 컴퓨팅에 기반한 보행자 안전 지원 시스템은 도로의 차량과 보행자 간의 충돌 방지를 통한 안전을 제공한다. 엣지 디바이스는 보행자의 스마트폰과 주행 중인 차량이다. 스마트폰과 차량이 현재 위치와 이동 방향을 포그 노드에게 전송하면 포그 노드는 보행자와 차량 간의 거리를 실시간으로 분석한다. 보행자 주변에 접근하는 차량이 있으면, 위험 상황으로 인지하고 차량에 경고를 보내 충돌을 방지한다.

③ 군집 주행

트럭 운송 회사는 차량을 안전하고 효율적으로 운영해야 한다. 배달 경로에서 최적의 경로설정 외에, 각 차량의 기계적 상태를 모니터링하여 적절한 시기에 유지보수를 수행하여 운전 효율성과 안전을 개선할 수 있다. 자율주행 기술을 결합하여 여러 대의 트럭을 묶어 집단을 형성해 군집 주행을 통해 운전 인력 감소, 에너지 절감 같은 운영 효율성을 높일 수 있다.

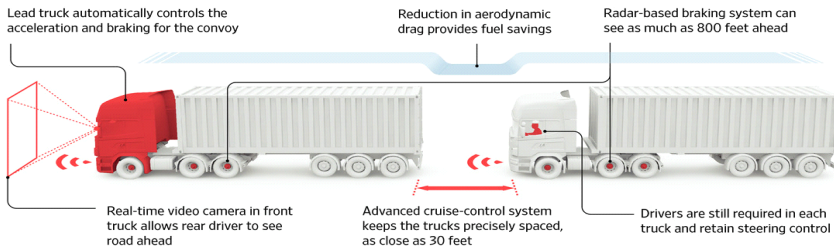
군집 주행은 일반적으로 자율주행차에 사용된다. 선두 차량은 차선 및 교통 조건을 감지하여 최적의 주행 속도를 결정한다. 이 정보는 소대의 다음 차량에 전달된다. 각 차량은 차량 간 통신 기능과 운전 안전을 위한 제어 장치를 갖추고 있다. 각 차량은 편대에서 위치와 각 차량의 운전상태 같은 정보를 관제센터로 전송한다.

이 경우, 엣지 디바이스는 차량 상태 모니터링 센서(엔진 속도, 온도, 브레이크 패드 두께 및 유압 서브시스템, 변속기 서브시스템, 중량 및 충격 상태, 타이어 압력, 연료 레벨 등), GPS 트래커, 환경 센서(외부 온도, 차선 감지, 주변 차량 감지) 등이 되고, 엣지 노드는 집단 주행을 하는 소대의 유도 차량이 된다.

〈그림 15〉 엣지 컴퓨팅을 응용한 군집 주행

How 'Platooning' Lowers Fuel Expenses

New automated-driving technologies allow two or three trucks to closely follow each other. Peloton Technology claims its system allows the rear trucks to use 10% less fuel. Combined savings for the front and rear trucks are 7% at 65 mph.



Source: Peloton Technology

Graphic by Alberto Cervantes/ THE WALL STREET JOURNAL.

6 » 의료 서비스

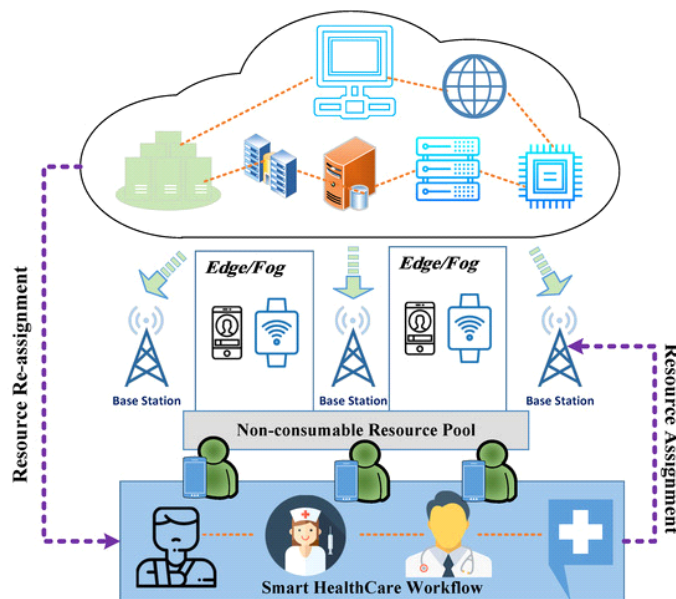
그동안 의료 서비스는 임상에서 의사결정을 지원하는 데이터 분석에 클라우드 솔루션을 사용해 왔다. 엣지 컴퓨팅은 진보한 사물인터넷 인프라를 활용하여 의료 서비스 수준은 물론, 실시간 모니터링과 관리를 할 수 있도록 하여 임상 업무를 개선한다.

머지않은 미래에 엣지 컴퓨팅을 의료 서비스에 사용할 수 있는 애플리케이션 중 하나는 의료 빅데이터 시스템이다. 지리적으로 분산된 데이터 처리로 의료 서비스를 제공하기 위해서는 다양한 데이터의 공유와 통합이 필요하다. 이러한 문제를 해결하기 위해 협업을 위한 분산 컴퓨팅 기술로 엣지 컴퓨팅이 사용된다. 엣지 컴퓨팅에 저장된 가상의 공유 데이터는 미리 정의된 서비스 인터페이스를 통해 다양한 서비스를 제공할 수 있게 된다. 특히, 개인정보 보호를 위해 연산은 엣지나 포그가 있는 로컬 영역에서만 이루어진다.

엣지 컴퓨팅을 활용한 건강관리의 예로 독감 추적이 있다. 환자가 병원을 이동하며 환자의 전자 진료기록(EMR: Electronic Medical Record)이 업데이트된다. 병원은 환자의 증상, 치료 환자의 규모 등과 같은 독감 발생에 대한 정보를 공유하고, 환자는 약국에서 처방전에 따라 약을 구매할 것이다.

이제 약국은 환자의 구매 기록을 병원에 제공하면서 병원들이 제공하는 엣지 서비스를

이용하여 독감 발생 인구를 파악한다. 이로 인해 약국들은 충분한 재고를 확보하여 매출을 증대시킬 수 있다. 또한, 제약회사에서 제공하는 데이터를 활용해 약품 판매점의 위치, 가격, 재고 등을 파악할 수 있다. 그러면 약국은 검색된 정보에 따라 주문 계획을 세울 수 있다. 제약회사들은 약국으로부터 독감 약품 주문을 받는다. 이에 따라 제약회사는 생산 계획을 재조정하고 창고 재고를 재조정할 수 있다. 한편, 정부의 질병 통제와 예방을 위한 센터는 광범위한 지역에서 증가하는 독감을 감시하고 적절한 지역의 주민에게 독감 경보를 발령할 수 있고, 독감의 확산을 막기 위한 추가적인 조치를 할 수 있다.



〈그림 16〉 엣지 컴퓨팅을 활용한 의료 서비스

출처: Soraia Oueida et al., "An Edge Computing Based Smart Healthcare Framework for Resource Management," Sensors, 2018.

7 » 증강현실

증강현실(AR: Augmented Reality)은 현실 세계에 가상 정보를 추가하는 시스템을 의미한다. AR은 컴퓨팅 기기가 소형화되고, 강력한 컴퓨팅 기능, 휴대성이 강화되면서 중요한 애플리케이션이 되었다. AR 애플리케이션은 사용자 요청에서 응답까지 약간의 시간 지연이라도 생기면 사용자 경험을 손상시키기 때문에 실시간 서비스가 매우 중요하다. 즉, 시간 지연이 발생하면 사용자는 입력에 반응하지 않는 것으로 인식될 수 있다. 엣지 컴퓨팅의 장점인 실시간 응답성을 활용하여 증강현실이나 실시간 게임에 엣지 컴퓨팅이 사용된다.



엣지 컴퓨팅에 대한 기업의 대응

초기에는 센서 데이터를 수집하고 전달하는 게이트웨이의 역할만 했던 엣지 컴퓨팅이 이제는 점차 그 활용 범위를 넓혀나가며 위상을 강화하고 있다. 지금까지는 클라우드가 산업의 트렌드로 자리 잡아가고 있지만, 대규모 데이터 수집과 빠른 분석을 통한 사업적 활용을 위해 엣지 컴퓨팅도 점차 주목받고 있다.

이 같은 엣지 컴퓨팅의 성장에는 네트워크 속도 개선과 장비 자체의 제품규격이 높아져 일정 수준 이상의 처리가 가능해진 데 그 이유가 있다. 점차 데이터 전송속도가 빨라지면서 대용량 데이터의 전송 시간도 짧아지고, CPU의 성능도 향상돼 많은 데이터를 좀 더 빠르게 처리할 수 있게 되었다.

실제로 통신서비스 제공업체들은 소비자용 이동통신 서비스 속도 향상과 더불어 IoT 전용망을 구축해 서비스에 나서고 있다. 국내에서도 이통3사는 IoT 전용망에서 경쟁 우위를 점하려 경쟁하고 있다. 기존의 장비를 제조하는 업체도 엣지 컴퓨팅 시대를 맞아 이에 대응하는 제품을 속속 출시하고 있다.

본 장에서는 엣지 컴퓨팅에 대한 기업의 대응을 살펴보고자 한다.

1 Amazon

아마존은 2016년 12월 ‘re:Invent’에서 Greengrass를 소개함으로써 엣지 게이트웨이에 대응하는 소프트웨어를 출시하였다. AWS Greengrass에서는 인터넷에 연결되어 있지 않더라도 디바이스에서 AWS 람다 함수를 실행하고, 디바이스의 데이터를 동기화된 상태로 유지하고, 다른 디바이스와 안전하게 통신할 수 있다. 또한, IoT 디바이스들을 Greengrass 그룹으로 구성해 서로 통신할 수 있고, 이들 디바이스들은 클라우드와 연결이 끊어지더라도 Greengrass 그룹에 속한 디바이스끼리는 로컬 네트워크를 통해 서로 계속 통신할 수 있다.

〈그림 17〉 아마존의 AWS Greengrass



아마존 Greengrass의 초기 고객과 파트너는 인텔, 퀄컴, 테크니칼라, 삼성, 에이스, 노키아, QNAP 등이다. 이후 시스템 통합 기업, ISV, 전자 회사, 하드웨어와 반도체 회사 등과 협력하고 있다. Greengrass는 라즈베리 파이나 ARM 프로세서로 지원되는 시스템 온칩(System On Chip) 기기에서 동작할 정도로 사이즈가 작다.

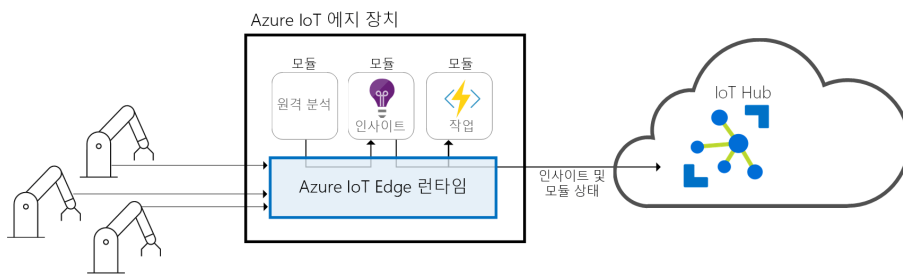
2 >> Microsoft

마이크로소프트는 인공지능의 발전과 스마트 기기의 확대에 따라 클라우드로 전략을 크게 바꾸었다. 마이크로소프트는 2017년 빌드 개발자 컨퍼런스에서 인공지능과 연계하는 새로운 클라우드 환경을 제시했다. 다양한 디바이스와 인공지능 그리고 서버 중심에서 벗어난 새로운 컴퓨팅 환경을 소개하였다.

클라우드 환경에서 연동되는 엣지 디바이스에서 동작하는 애플리케이션 개발환경을 제공하기 위해 인텔리전트 클라우드와 인텔리전트 엣지의 개념을 제시했다.

또, 마이크로소프트는 2017년 Edge 솔루션인 ‘Azure IoT Edge’를 발표하여 클라우드에 엣지 디바이스를 연결할 수 있도록 하였다. Azure IoT Edge는 IoT 허브(Hub)에서 동작하는데 클라우드가 아닌 디바이스에서 데이터를 분석할 수 있는 환경을 제공한다.

〈그림 18〉 Azure IoT Edge (출처: Microsoft Azure IoT Edge)



기업은 이들 서비스를 사용해 요구조건에 맞는 적합한 클라우드 및 엣지 IoT 솔루션을 구현할 수 있으며, Azure IoT Edge는 클라우드와 엣지 사이에서 안전하게 데이터가 송수신되고 클라우드 플랫폼에서 운영되는 다양한 서비스와 기능을 IoT 디바이스에 사용할 수 있는 환경을 제공한다.

3 Dell

Dell은 빌딩과 공장 자동화 용으로 데이터센터 솔루션과 분석 기능이 포함된 엣지 게이트웨이(Edge Gateway) 솔루션을 제공한다. Edge Gateway는 데이터를 수집, 집계, 분석하여 가치 있는 데이터만 데이터센터 또는 클라우드로 전달하는 로컬 분석과 기타 미들웨어를 제공하고 있어, 클라우드로 전송되는 데이터를 최소화할 수 있다.

〈그림 19〉 Dell Edge Gateway



Categories	Edge Gateway 3001 General-Purpose Automation	Edge Gateway 3002 Transportation/Logistics	Edge Gateway 3003 Media/Retail Kiosks
Use Cases	Factory floors, remote energy locations like windmills, etc.	Interstate trucks, trains, planes, marine. Cold-chain logistics in warehouses, shipping containers.	Streaming media (audio & video), ticket kiosks, other retail
Features	- Intel Atom E3805 2GB RAM 8GB eMMC(32GB w/WWAN) - uSDcard slot	- Intel Atom E3805 2GB RAM 8GB eMMC(32GB w/WWAN) - uSDcard slot	- Intel Atom E3815 –has 400 MHz GPU 2GB RAM 8GB eMMC(32GB w/WWAN) - uSDcard slot
I/O Ports	- USB (1 x 3.0, 1 x 2.0) 2x 10/100 Fast Ethernet - Multi Function IO connection (0-5V, 8 channel, independently programmable, DAC, ADC, GPIO). 1KV isolation. 2x Programmable Serial (RS232, 422, 485) - Accelerometer & atmospheric pressure sensors	- USB (1 x 3.0, 1 x 2.0) 2x 10/100 Fast Ethernet - CAN bus interface (supports CAN and J1939) - Supports CAN 2.0a/b/FD 2.5KV of isolation. - Accelerometer & atmospheric pressure sensors	- USB (1 x 3.0, 1 x 2.0) 2x 10/100 Fast Ethernet - DisplayPort1.1 –Supports up to 2560x1600@60Hz; Daisy-chaining of displays not supported - Audio –3.5mm Line In/Line Out - Accelerometer & atmospheric pressure sensors
Wireless	- WiFi / BLE –802.11n, RS9113 - GPS - Optional: 3G and 4G LTE in select countries - Optional: zigbee(dongle required)	- WiFi / BLE –802.11n, RS9113 - GPS - Optional: 3G and 4G LTE in select countries - Optional: zigbee	- WiFi / BLE –802.11n, RS9113 - GPS - Optional: 3G and 4G LTE in select countries - Optional: zigbee(dongle required)
OS	- Ubuntu Core Series 16 - Win 10 IoT Enterprise LTSC 2016 (WWAN Required)	- Ubuntu Core Series 16 - Win 10 IoT Enterprise LTSC 2016 (WWAN Required)	- Ubuntu Core Series 16 - Win 10 IoT Enterprise LTSC 2016 (WWAN Required)

Dell은 여러 IoT 애플리케이션에 사용할 수 있는 Edge Gateway 3000 서버를 선보였다. 데이터 수집 시점에 분석을 수행하고 필요한 정보를 데이터센터로 전송한다. 이 서버에는 코어가 최대 16개인 인텔의 아톰(Atom) 칩이 장착돼 있다. 모델 3001은 산업 자동화와 에너지 관리용이고, 모델 3002는 운송 및 물류용으로 트럭 같은 큰 자동차에 설치할

수 있다. 모델 3003은 비디오 스트리밍 및 POS(Point-Of-Sale) 애플리케이션용이다.

또한 ‘Embedded Box PC’는 산업 현장의 환경과 요구조건에 맞춰 개발한 임베디드 솔루션이다. 별도의 냉각 팬이 없는 팬리스(Fanless) PC로, 내부 케이블을 최소화한 온보드 구성과 방열 본체로 모델에 따라 최대 섭씨 70도까지 견딜 수 있다. 극단적인 환경에서도 안정적인 성능을 유지하며, 우수한 내구성과 편리한 설치로 제조업 등 다양한 산업 시설에서 활용된다.

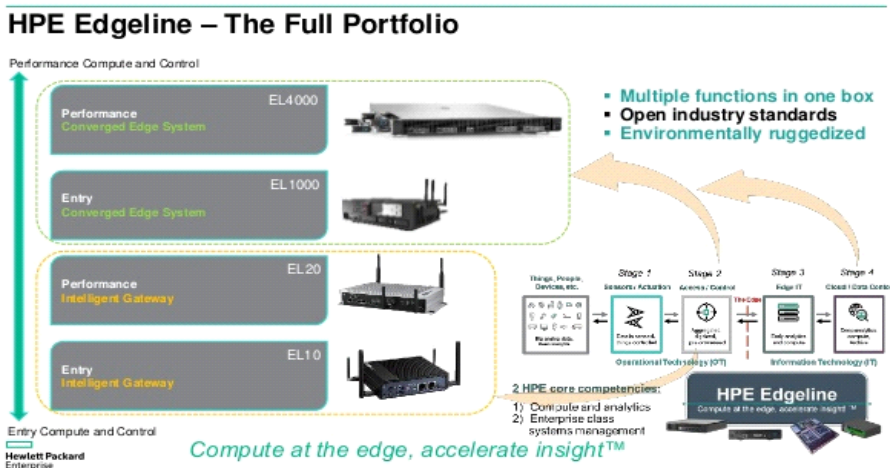
4 HPE

HPE는 아마존과 같은 대형 클라우드 서비스 회사와 경쟁할 분야로 엣지 컴퓨팅을 선정하였다. 사실 HPE는 직접 클라우드 서비스를 내놓고 클라우드 시장에 진출하였으나, 경쟁에서 밀려 사업을 철수했다. 이번에는 HPE가 엣지 컴퓨팅에서 새로운 도전을 한다.

HPE는 엣지 컴퓨팅 인프라로 엣지라인(Edgeline)이라는 IoT 서버를 공개했다. 이 서버는 엣지에 위치하고 데이터센터로 데이터를 보내기 전에 분석을 수행한다. 또 엣지에서 데이터 패킷을 가상화할 수 있어 컴퓨팅 자원을 더 효과적으로 활용할 수 있다.

예를 들어, Edgeline 서버는 페이스북 게시물 응답이나 검색 요청 등 모바일 컴퓨팅 작업을 엣지에서 처리한다. 반면에 이미지 인식 등 자원을 더 많이 소비하는 작업은 중앙의 데이터센터로 전송한다.

〈그림 20〉 HPE Edgeline



HPE는 사물인터넷을 위해 현장 환경에 적합한 서버도 출시하였다. 이러한 시스템들은 충격, 진동, 극한 기온을 견뎌낼 수 있도록 만들어진다. 최근 슈나이더일렉트릭은 HPE와 공동으로 개발한 HPE 마이크로 데이터센터(HPE Micro Datacenter)를 출시했다. 이 제

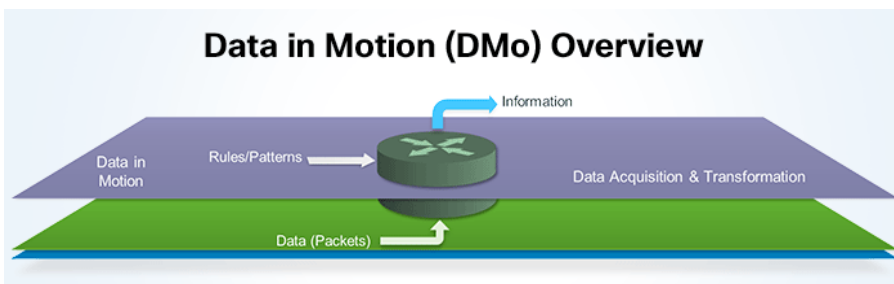
품은 엣지 컴퓨팅을 기반으로 간편한 설치, 원격 관리, 보안, 모니터링 등 통합적인 솔루션을 제공하며, 방대한 데이터를 관리한다. 이 솔루션은 데이터센터에 필요한 공간이 부족하거나 물리적 보안이 중요한 지역에 적합하다. 무정전 전원 장치(UPS), 전력 공급, 냉각과 모니터링을 통합적으로 제공하여 엣지 환경을 구축하도록 지원한다.

HPE Micro Datacenter는 클라우드 기반의 HPE 컨버지드 시스템(Converged Systems), 하이퍼컨버지드(HyperConverged)와 IoT HPE 엣지라인 솔루션(IoT HPE Edgeline Solutions)을 지원한다. 운영에 최적화된 솔루션을 구현해 지연 속도를 줄이면서 데이터센터의 용량을 신속하게 조절할 수 있고, 사용자가 데이터센터를 원격으로 안전하고 간편하게 관리할 수 있는 특징이 있다.

5 >> CISCO

CISCO는 포그 컴퓨팅이라는 용어를 만든 회사이다. CISCO 엣지 플랫폼으로 'IOx'를 제공하는데 이 플랫폼은 라우터가 데이터 분석과 처리를 가능하게 한다. 또한, CISCO는 데이터 처리 모델인 '데이터 인 모션(Data in Motion)'을 제공하고 있어 사용자는 실제 데이터를 받아들이고 처리하는 일련의 동작 유형을 정의하고, 직접 프로그래밍할 수 있다. 프로그래밍이 된 규칙들은 라우터와 스위치에 탑재되고 네트워크에 연결된 모든 디바이스에 적용된다. CISCO가 제공하는 분산형 엔진은 네트워크 장비에 탑재된 규칙을 해석하여 최적화된 검색이 가능하고, 개발자는 이를 사용하여 데이터 수집과 분석을 할 수 있는 IoT 엣지 디바이스용 애플리케이션을 쉽게 프로그래밍할 수 있다.

〈그림 21〉 CISCO의 Data in Motion

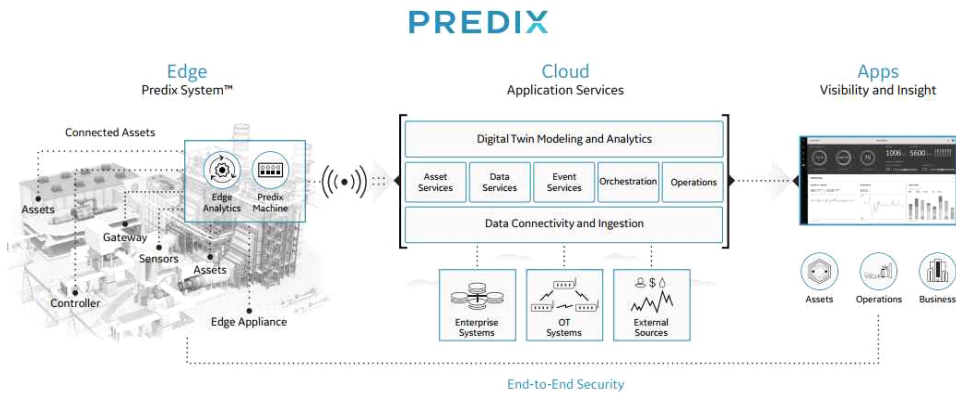


6 » GE

GE는 2015년 산업 인터넷 소프트웨어 플랫폼인 프리덱스(Predix)를 발표하였다. Predix는 산업용 장비를 연결하고, 데이터를 분석하고, 실시간으로 처리한다. 예를 들어, 터빈, 엔진 등 산업용 중대형 장비나 부품에 부착된 센서를 통해 축적되는 데이터를 분석해 현장에서 발생하는 각종 문제를 해결한다.

또, Predix 기반의 애플리케이션을 개발할 수 있다. 이를 통해 계속해서 개발자 풀을 확대하면서 GE는 Predix를 통해 자사의 소프트웨어 비즈니스 역량을 강화하는 동시에 새롭게 떠오르고 있는 데이터 분석 시장에도 도전하고 있다. 이를 위해 GE는 2016년 Predix를 클라우드 서비스에 접목하기 위해 마이크로소프트와 전략적 제휴를 발표하고, GE의 Predix를 MS의 클라우드 서비스 Azure에서 활용할 수 있도록 하였다. GE는 소프트웨어 부문에서의 매출을 계속해서 확대하는 계획을 하고 있다.

〈그림 22〉 GE의 Predix





결론

본 보고서는 엣지 컴퓨팅의 개념과 탄생 배경을 살펴보면서 엣지 컴퓨팅이 클라우드 컴퓨팅을 보완하는 새로운 컴퓨팅 환경을 제공하는 것을 설명하였다. 또한, 엣지 컴퓨팅과 관련된 기술을 살펴보고, 엣지 컴퓨팅이 활용될 분야와 활용 방안을 설명하였다. 그리고 마지막으로 엣지 컴퓨팅과 관련한 기업들의 동향을 살펴보았다.

본 보고서를 요약하면 다음과 같다.

- 사물인터넷이 증가함에 따라 수많은 디바이스와 센서가 인터넷에 연결되어 엣지 컴퓨팅에 대한 필요성이 증가하고 있다. 엣지 컴퓨팅은 센서와 디바이스가 분산된 엣지와 중앙 집중식 클라우드를 결합하여 구축한다.
- 엣지 컴퓨팅은 중앙 집중화된 클라우드 컴퓨팅을 보완하여 분산된 컴퓨팅 환경을 제공하고, 네트워크 가장자리인 엣지에서 연산과 저장, 그리고 분석을 지원하는 패러다임이다. 특히, 엣지 컴퓨팅은 실시간 처리, 대규모 분산처리, 확장성, 상호 운용성의 강점이 있어 IoT 애플리케이션의 성능을 향상시킨다.
- 기업은 엣지 컴퓨팅으로 인해 애플리케이션의 성능 향상과 실시간 응용 지원으로 기존 서비스의 개선 또는 신규 서비스의 도입이 가능하며, 클라우드 중심의 컴퓨팅 환경 대비 운영비용을 절감하고, 에너지 절감 효과도 있다. 특히, 분산된 컴퓨팅 환경으로 인해 기술적으로는 난점은 존재하나 보안을 강화할 수 있는 환경을 제공한다. 이로 인해 기업은 운영 효율성을 높이고 고객에게 개선된 서비스를 제공하는 데 도움을 준다.
- 엣지 컴퓨팅 산업은 기술이 성숙해 짐에 따라 제조, 스마트 시티, 에너지 관리, 스마트

홈, 커넥티드카, 의료 서비스, 증강현실 등 다양한 응용 분야에 의미 있는 사례가 만들어질 것으로 예상된다.

- 많은 글로벌 기업들이 엣지 컴퓨팅 시장을 선점하기 위해서 다양한 제품과 서비스를 출시하고 있다. 특히, 클라우드 시대를 주도하지 못한 기업들은 엣지 컴퓨팅을 통해 새로운 도약을 준비하고 있다. 클라우드 시장의 강자인 아마존과 마이크로소프트뿐만 아니라, Dell, HPE, CISCO, GE는 엣지 컴퓨팅을 위한 하드웨어와 소프트웨어 시장을 선점하려고 노력하고 있다.

지금까지 엣지 컴퓨팅에 대해서 개념과 상황을 설명하였다. 그러면 엣지 컴퓨팅을 도입하여 실제 시스템을 구축하고자 하는 기업들의 관점에서 시사점을 살펴보고자 한다. 앞서 설명한 것처럼 사실 엣지 컴퓨팅 구축에 있어 아직 성숙한 표준이 없고, 참고할만한 사례나 아키텍처가 부족하여 다양한 IoT 애플리케이션들을 여러 분야에 적용하기는 어려움이 있다. 현재로는 엣지 컴퓨팅을 구축하는 것은 다양한 하드웨어, 소프트웨어 및 네트워킹 옵션 중에서 몇몇을 선택하고 이들을 통합하여 구축하는 수준이다.

또한 사물인터넷이 결합된 다양한 스마트 서비스의 엣지 컴퓨팅이 적용되고 있고, 이러한 경향은 계속 증가할 것이 확실하나 아직은 전문화된 하드웨어와 소프트웨어, 그리고 이를 통합하는 기술이 부족하여 시도되는 프로젝트들의 사업적인 성공 여부를 예단하기는 이른 상황이다.

따라서 기업들의 엣지 컴퓨팅 적용은 파일럿 형태의 시도를 통해 가능성을 검증하고, 이를 확대하는 Start Small, Finish Big의 전략을 취하는 것이 타당할 것으로 생각된다.

참 | 고 | 문 | 헌

- [1] B. Varghese, N. Wang, S. Barbhuiya, P. Kilpatrick and D. S. Nikolopoulos, "Challenges and Opportunities in Edge Computing," 2016 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2016.
- [2] CIO Korea, "'엣지 컴퓨팅 + 네트워크' 포그 컴퓨팅의 정의와 응용 분야," 2018.1.
- [3] CISCO, "Fog Computing and the Internet of Things: Extend the Cloud to Where the Things Are," White Paper, 2015.
- [4] Forbes, "Amazon Makes Foray Into Edge Computing With AWS Greengrass," 2017. 6.
- [5] forcepoint.com, "2019 년 포스포인트 정보 보안 예측 보고서" 2018.
- [6] GE, "Predix: The Industrial Internet Platform," 2016. 11.
- [7] GE리포트 코리아, "엣지 컴퓨팅, 사물인터넷 데이터 왕국으로 통하는 관문?," 2016. 10.
- [8] Hany F. Atlam et al., "Fog Computing and the Internet of Things: A Review," Big Data Cogn. Comput. 2018.
- [9] Industrial Internet Consortium, "Introduction to Edge Computing in IIoT," <https://www.iiconsortium.org/>
- [10] Mahadev Satyanarayanan, "The Emergence of Edge Computing," IEEE Computer, 2017.1.
- [11] Mahmud R., Kotagiri R., Buyya R., "Fog Computing: A Taxonomy, Survey and Future Directions," Internet of Everything. Internet of Things Springer, Singapore, 2018.
- [12] Microsoft Blog, "In a new era of intelligent cloud and intelligent edge, Microsoft aims to empower every developer," 2017. 5.
- [13] Microsoft, "Microsoft Azure IoT Edge Extending cloud intelligence to edge devices," 2017. 5.
- [14] OpenFog, "OpenFog Reference Architecture for Fog Computing," 2017. 2.
- [15] Openstack, "Cloud Edge Computing: Beyond the Data Center," openstack.org.
- [16] Pedro Garcia Lopez et al., "Edge-centric Computing : Vision and Challenges,"

- ACM SIGCOMM Computer Communication Review, 2015.
- [17] SK텔레콤 네트워크 기술원, " SK Telecom 5G White Paper," 2014.
- [18] Sophia I. Vargas et al., "Edge Computing: IoT Will Spawn A New Infrastructure Market," Forrester Research, 2017. 3.
- [19] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, "Edge Computing: Vision and Challenges," in IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, Oct. 2016.
- [20] Wei Yu, et al., "A Survey on the Edge Computing for the Internet of Things," IEEE Access, 2018.
- [21] 방성수, "시스코의 야심작 '포그 컴퓨팅' ... 사물인터넷과 빅데이터 시대의 새 패러다임," http://biz.chosun.com/site/data/html_dir/2016/10/21/2016102102501.html, 2016.
- [22] 배명남, 이강복, 방효찬, "인지 IoT 컴퓨팅 기술동향," 전자통신동향분석, 2017.
- [23] 손가령, 이은민, "IoT 생태계 확산과 엣지 컴퓨팅의 역할," 정보통신방송정책, 29권 16호 2017.
- [24] 윤용익, 김스베틀라나, "모바일 클라우드 컴퓨팅 기술 동향" 정보통신산업진흥원 주간기술 동향 2010.
- [25] 이강찬, "모바일 클라우드 개념과 기술 동향," TTA Journal Vol. 139, 2012.
- [26] 이지현, "Streaming analytics from the center to the edge," <http://www.bloter.net/archives/250015>, 2016.
- [27] 최계영, 4차 산업혁명과 ICT, KISDI Premium Report, 2017.5.31.
- [28] 한상기, "엣지 컴퓨팅의 현황과 주요 기업의 전략," 한국정보화진흥원 이슈리포트, 2018. 3.

데이터로 연결하는 사람·세상·미래 Bridging Data

발 행 : 2019년 3월

발 행 인 : 문 용 식

발 행 처 : 한국정보화진흥원

대구광역시 동구 침단로 53

Tel. 053-230-1114

인 쇄 처 : (사)한국장애인유권자연맹인쇄사업부

Tel. 02-325-1585

1. 본 보고서는 방송통신발전기금으로 수행한 정보통신·방송 연구개발 사업의 결과입니다.
2. 본 보고서의 내용을 발표할 때는 반드시 과학기술정보통신부 정보통신·방송 연구개발 사업의 연구결과임을 밝혀야 합니다.
3. 본 보고서의 내용은 한국정보화진흥원의 공식 견해와 다를 수 있습니다.
4. 본 보고서의 내용에 대해 무단전재를 금하며, 가공·인용할 때는 반드시 「한국정보화진흥원」이라고 출처를 밝혀 주시기 바랍니다.