

2021년 제9호 (통권16호)



데이터산업 동향 이슈 브리프

ISSUE BRIEF

2021.9

데이터 공유 및 생태계에 관한 국제적 논의

데이터 공유 및 생태계에 관한 국제적 논의

I. 검토 배경	1
II. 데이터 공유 프레임워크 국제 논의	4
III. 데이터 공유 생태계 현황과 협업 장애요인 및 추진 전략	10
IV. 결론 및 시사점	21

요 약

- 데이터를 활용한 디지털화의 가속화는 기존 산업을 새로운 형태로 전환·확장시키며, 기존 및 신규 사업자 모두에게 산업 경계를 초월한 새로운 기회를 제공하는 등의 이점을 제공
 - 동시에 ▲데이터 가용성 및 접근성 제약 ▲낮은 품질의 가용 데이터 ▲데이터셋의 상호운용성 부족 ▲데이터 보호 및 개인정보보호에 관한 규제 불확실성 등 이러한 데이터 경제의 잠재력을 제한하는 요인이 존재
- 데이터 경제의 장애요인을 해소하고, 잠재력을 실현하기 위해서는 데이터 사용·재사용 및 공유를 위한 제도적, 기술적 기반과 기업들의 활발한 데이터 활용이 필요. 이에 따라 현재 관련 법제 및 프레임워크 구축 논의가 활발
 - (EU 데이터 거버넌스 법안 마련) EU는 AI를 비롯한 신제품 및 혁신 창출을 위해 데이터의 가용성 제고를 촉진하는 법안을 마련('20.11.)
 - (WEF의 데이터 교환 프레임워크 논의) 세계경제포럼(WEF)은 공동 목적 이니셔티브를 위한 데이터(DCPI, Data for Common Purpose Initiative) 프로젝트에서 데이터 공유 메커니즘의 한 형태인 데이터 교환¹⁾ 프레임워크를 설정('21.8.)
 - 기업들의 데이터 공유와 관련하여 프랑스 글로벌 컨설팅 회사 캡제미니(Capgemini)는 최근 데이터 공유 생태계 추진을 위한 전략 방안을 제시한 보고서를 발표('21.7.)
- WEF는 데이터 교환을 재화나 서비스가 교환되는 일반적 시장처럼 데이터가 교환되는 시장(Data Marketplace)으로 정의했으며, Capgemini는 데이터 공유 및 데이터 공유 생태계를 사일로(silo)화된 시스템에서는 불가능했던, 새로운 가치를 창출하는 데이터 공유 및 관리를 위한 여러 기관 간의 파트너십으로 정의
 - 그러나 본 고에서는 데이터 공유를 데이터 거래, 교환, 기업 간 제휴 및 협력 등을 포함하는 포괄적인 개념으로 사용하는 우리나라의 현실을 감안하여, '데이터 공유'라는 용어로 통일
- (WEF) 데이터 교환(DEX, Data Exchange)의 공동의 목적을 위한 데이터 활용 측면에서 ▲데이터 발견 가능성 ▲데이터의 결합력 ▲데이터 가용성 ▲데이터 규율 ▲투명성 ▲연구개발 가속화의 이점을 제공
- (Capgemini) 기업들은 데이터 공유 생태계 참여를 통해 고객만족도와 생산성의 향상과 비용 절감의 효과를 얻을 수 있으며, 이를 위해 전략 수립-실행-이점의 유지 및 업그레이드를 단계적으로 추진해야 함
- 우리나라의 경우 데이터산업 활성화 전략('18.6.), 데이터 3법('20.1.), 디지털 뉴딜정책('20.7.), 그리고 최근 데이터 기본법('21.9.) 마련까지 지속적으로 데이터 산업을 위한 법적 제도를 마련하였으며, 데이터 플랫폼 발전전략('21. 6.11) 등을 발표하여 데이터 생태계 조성을 위한 노력을 기울이고 있음

PART

I

검토 배경

▶ (데이터 경제의 잠재력과 실현 장벽) 향후 데이터를 활용한 디지털화의 가속화에 힘입은 데이터 경제의 긍정적 효과가 예상되는 동시에 이러한 잠재력을 제한하는 요인 존재

- (데이터 경제의 산업 파급 효과) 데이터 사용 확산은 기존 산업을 새로운 형태로 전환·확장시키며, 기존 및 신규 사업자 모두에게 산업 경계를 초월한 새로운 기회를 제공
 - (기존 산업) 기존의 제품·서비스를 스마트 제품·솔루션으로 전환할 수 있는 기회를 제공하고, 기존의 사업영역을 확장하거나 전혀 다른 영역으로 진출하게 하는 효과
 - (신규 산업) 디지털 제품, 데이터 기반 제품 및 서비스, 데이터 생태계의 성장은 데이터 수집, 보유, 처리 또는 배포를 위한 맞춤형 서비스를 제공하는 신규진입자 유입을 촉진

표 1 | 데이터 경제의 산업에 대한 긍정적 파급 효과

기존 산업	소매	제조	헬스케어	금융서비스
스마트 제품	- 제품 및 패키지에 대한 즉각적인 맞춤화	- 커넥티드 제품·서비스 - 스마트 부품 및 예비품	- 맞춤형 진단 및 건강 계획 - 블록체인 기반 전자건강기록 (EHR) 저장 및 접근	- 개인투자 활성화 - 맞춤형 자산 관리 - 소셜 뱅킹 및 투자
기존 산업 내 포지셔닝 확장	- 경계가 없는 소매 네트워크	- 제품 추적 가능성 - 연결된 거래 공장 및 신속 서비스 제공	- 지역 단위 건강 허브 관리 네트워크	- 데이터 기반 소매 및 스타트업 액셀러레이터 허브
산업 경계 초월	- 5G 데이터 기반 커넥티드스토어 - 하이퍼 홈, 헬스(hyper home, health), 미디어 허브(media hub)	- 디자인에서 배송까지 디지털 제조 - 드랍shipping(dropshipping) ²⁾	- 커넥티드 메디케어 및 보험 네트워크	- 커넥티드 건강 및 자산 관리
신규 산업	푸드테크(FoodTech)/애그리테크(AgriTech)	핀테크	헬스테크	로지스틱스테크
디지털 제품	- 디지털 식단/식사 계획 - 디지털 애완동물 케어 - 가족의 디지털 생체리듬 관리/디지털 목축업	- 운전 행태 기반 보험 - 소액 예금자 대출 - 디지털 성장 기반 은행	- 맞춤형 건강 비서 - 디지털 간호사/영양사 - 디지털 질병 및 의료 시뮬레이션	- 디지털 모빌리티 사용권 - 디지털 트럭
데이터 기반 신제품 및 신규 서비스	- 예방적 작물 관리 및 재배 - F2M(Farm to Market) 네트워크	- 디지털 BNPL(Buy Now Pay Later) * 선구매 후지불 시스템 - 디지털 통화 및 환전	- 맞춤형 약제	- 커넥티드 물류 네트워크
산업 경계 초월	- 지역사회 시장 및 금융을 주도하는 지역 농업네트워크	- 소액 투자 및 위험 대상 (예: 주택대출 연체가능성이 높은 신용카드 거래자)	- 건강 계획 및 혜택 제공	- 물류 지수 거래소(Logistics index exchange)

출처 : World Economic Forum, Towards a Data Economy: An enabling framework, WHITE PAPER, 2021.8

1) WEF는 데이터 교환을 데이터 시장(Data marketplace)과 같은 데이터 거래의 의미로 사용

2) 판매자가 상품 재고를 두지 않고 주문을 처리하는 유통 방식

- 그러나 디지털 전환을 기반으로 한 데이터 경제의 잠재력 실현에 여전히 장애요인 존재
 - (데이터의 가용성 및 접근성 제약) 생성되는 데이터의 대부분이 사일로(silo)화, 파편화되어 있으며, 공동의 목적으로 사용 불가
 - (낮은 품질의 가용 데이터) 가용 데이터셋이 있더라도 라벨링이 잘못되거나, 구조화되지 않은 포맷(PDF, 종이 문서 등)으로 존재하는 경우, 데이터를 정리 또는 디지털화하는 데에 많은 노력 소모
 - (데이터셋의 상호운용성 부족) 서로 다른 소스의 데이터를 공유, 분석, 통합해야 하는 경우, 동일한 표준 및 프로토콜의 부족으로, 데이터를 이해하는 데 장애요인으로 작용
 - (데이터 보호 및 개인정보보호에 관한 규제 불확실성) 관련 법률이 계속 변화하면서, 데이터 소유권 및 법 준수 문제가 효과적인 데이터 공유를 방해

▶ 데이터 경제의 장애요인을 해소하고, 잠재력을 실현하기 위해서는 데이터 사용·재사용 및 공유를 위한 제도적, 기술적 기반과 기업들의 활발한 데이터 활용이 필요. 이에 따라 현재 관련 법제 및 프레임워크 구축 논의가 활발

- (EU 데이터 거버넌스 법안* 마련) EU는 AI를 비롯한 신제품 및 혁신 창출을 위해 데이터의 가용성 제고를 촉진하는 법안을 마련('20.11.)

* Proposal for a Regulation of the European Data Governance(약칭 Data Governance Act, 이하 DGA), 2020.11.25.

- 최근 의회의 산업연구에너지위원회가 DGA에 대한 입장서를 승인('21.7.16.)했으며, 9월로 예정된 본회의 표결에서 동 법안이 통과되면 데이터 공유의 제도적 기반이 확보될 것으로 전망

| 표 2 | 유럽 데이터 거버넌스 법안의 주요 내용

구분	주요 내용
목적	데이터 중개자에 대한 신뢰 제고 및 EU의 데이터 공유 메커니즘 강화를 통해 데이터의 가용성을 촉진
공공 데이터 재사용	(대상 데이터) 상업 기밀, 통계 기밀, 지식재산권 보호, 개인정보보호를 이유로 공공기관이 보유하고 있는 데이터를 대상으로 하며, 공공기관이 아닌 공공사업체가 보유한 데이터에는 적용되지 않음(제3조) (재사용 원칙) 공공기관이 보유하는 데이터의 재사용을 독점적으로 이용하는 계약 또는 협정은 금지하며(제4조), 데이터 재사용 여부를 판단하는 공공기관은 EU 내에서 해당 데이터의 재사용 허용에 비차별적이어야 함(제5조)
데이터 공유 서비스 사업자(DSSP, Data Sharing Service Provider)	(데이터 공유 서비스 정의) 법인인 데이터 보유자와 데이터 이용자를 중개하는 서비스로, 데이터 교환 서비스 및 데이터를 이용 가능하게 하는 플랫폼과 데이터베이스의 구축 서비스 등(제9조) (DSSP 조건) 중개 이외의 목적 사용 금지, 메타데이터 활용 제한, 투명한 접근절차 제공, 원활한 데이터 교환을 위한 형식 제공, 사기 방지, 보안 및 정보보호 조치, 데이터의 연속성 보장, 경쟁법 준수, 관할 명시 등(제11조)
데이터 이타주의 ³⁾ (data altruism)	데이터 이타주의 조직을 등록하고(제15조), 일반적으로 비영리 기반으로 운영되며, 해당 단체의 다른 활동과는 분리된 법적으로 독립적 조직을 통해 데이터 이타주의에 근거하여 활동할 것을 등록조건으로 규정(제16조)
공동 표준 작업	(유럽 데이터 혁신 위원회 설치) 유럽 공동 데이터 공간의 창출을 촉진하기 위한 공통 표준을 만드는 작업을 수행(제27조)

출처 : Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act). COM/2020/767 final

3) 공익 목적을 위해 개인정보 정보 주체 또는 비개인정보 데이터 보유자가 일정한 보상 없이 개인정보나 비개인정보 사용을 허락하는 것

- (WEF의 데이터 교환 프레임워크* 논의) 세계경제포럼(WEF)은 공동 목적 이니셔티브를 위한 데이터(DCPI, Data for Common Purpose Initiative) 프로젝트⁴⁾에서 데이터 공유 메커니즘의 한 형태인 데이터 교환⁵⁾ 프레임워크를 설정

* WEF, Towards a Data Economy: An enabling framework WHITE PAPER, 2021.8.

- 기업들의 데이터 공유와 관련하여 프랑스 글로벌 컨설팅 회사 캡제미니(Capgemini)는 최근 데이터 공유 생태계 추진을 위한 전략 방안을 제시한 보고서*를 발표

* Capgemini, Data Sharing Masters, '21.7.

4) DCIP는 소비 및 기술 거버넌스, AI와 머신러닝, 데이터 정책, IoT, 도시 전환 플랫폼의 미래를 위한 WEF의 프로젝트이며, 마켓플레이스에 대한 거버넌스(Developing a Responsible and Well-designed Governance Structure for Data Marketplaces), 데이터 경제를 향한 프레임워크(Towards a Data Economy: An enabling framework) 보고서를 발표
5) WEF는 데이터 교환을 데이터 시장(Data marketplace)과 같은 데이터 거래의 의미로 사용



데이터 공유 프레임워크 국제 논의

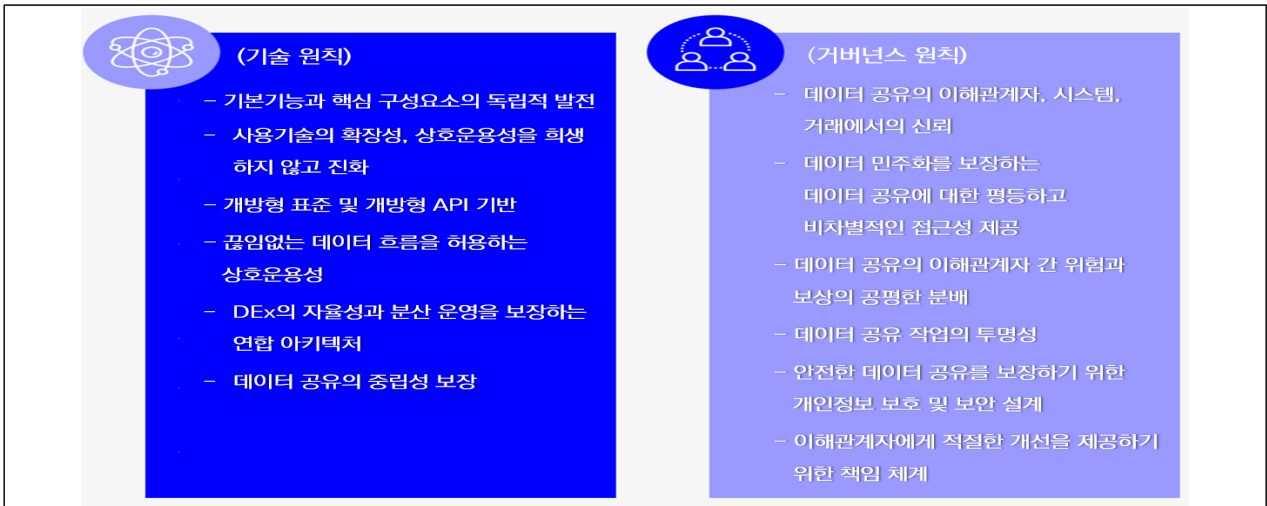
1 데이터 공유의 기대효과

- ▶ (데이터 공유의 개념) 데이터 교환은 가치에 대한 데이터의 원활한 공유가 작동할 수 있는 메커니즘 중 하나로, 이를 통해 기업, 정부 및 시민은 특정 목적을 위해 데이터에 액세스할 수 있음
 - 그러나 이러한 과정에서 모든 이해 관계자와 참가자의 권리를 인정하고 보호하는 것이 필수적
 - ※ WEF의 보고서에서는 데이터 교환(Data Exchange)이라는 용어를 사용하고 있으나, 데이터 공유를 데이터 거래, 교환, 기업 간 제휴 및 협력 등을 포함하는 포괄적인 개념으로 사용하는 우리나라의 현실을 감안하여, 본 고에서는 '데이터 공유'라는 용어로 통일
- ▶ (데이터 공유의 기대 효과) 데이터 공유(DEX, Data Exchange) 공동의 목적을 위한 데이터 활용 측면에서 다음과 같은 이점을 제공
 - (데이터 검색 가능성, Discoverability) DEX 플랫폼은 데이터셋에 데이터 소비자가 접근할 수 있는 검색 가능성을 제공함으로써 효율적인 데이터 공유 거래를 가능하게 함
 - (데이터의 결합력, Combinatorial power) 적절한 데이터셋에 시기적절하고 쉽게 접근함으로써 다양한 소스에서 얻은 여러 데이터셋의 결합력을 사용하여 통합된 종단 간 서비스 제공 가능
 - (데이터 가용성, Availability) 신뢰할 수 있는 양질의 데이터셋에 대한 수요 창출이 데이터 제공자의 공급 증가를 유도함으로써 데이터 가용성 및 사용의 선순환 형성
 - (데이터 규율, Data Discipline) 시장 메커니즘에 의해 부과된 엄격한 품질 요구 사항은 DEX 생태계에 적극적으로 참여하고 이익을 얻으려는 다양한 이해 관계자들에게 데이터 규율 준수 의식을 고취
 - (투명성, Transparency) 다양한 기관의 기능 및 성과에 대한 데이터 분석이 이루어짐에 따라 사회 전반의 투명성 확립이 보다 강화될 것임
 - (연구개발 가속화) 보건, 교육, 농업, 환경 등의 공익 분야의 연구를 가속화

2 데이터 공유 기능의 아키텍처

- ▶ (데이터 공유 원칙) DEX가 미래 지향적인 데이터 경제에 기여하기 위해 핵심적으로 요구되는 사항
 - DEX는 기능상 효율적이고, 기술적으로 견고하고, 법적 규범을 준수하며, 상업적으로 실행 가능성을 갖추어야 함

| 그림 1 | DEx 디자인의 기술 및 거버넌스 원칙

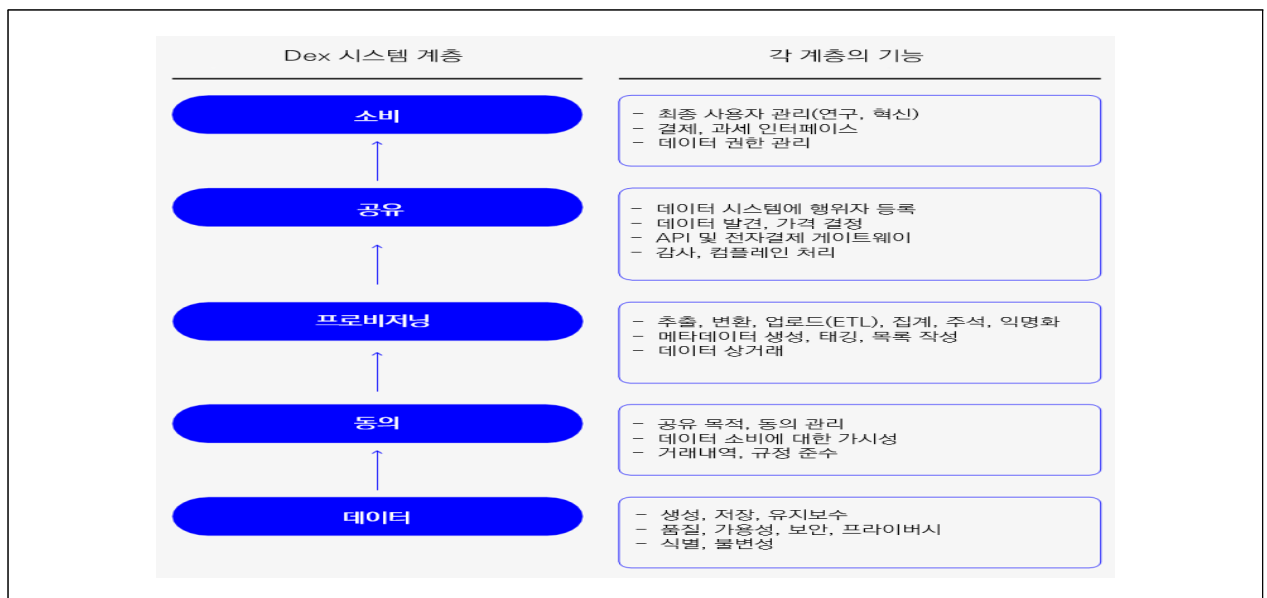


출처 : World Economic Forum(2021.8), Towards a Data Economy: An enabling framework

▶ (DEx 시스템 계층) DEx 시스템은 크게 데이터 계층, 동의 계층, 데이터 프로비저닝 계층, 교환 계층, 소비 계층의 5가지로 구분

● 각각의 계층은 기술원칙에 따라 상호 독립적으로 발전할 수 있도록 설계

| 그림 2 | DEx 디자인의 기술 및 거버넌스 원칙



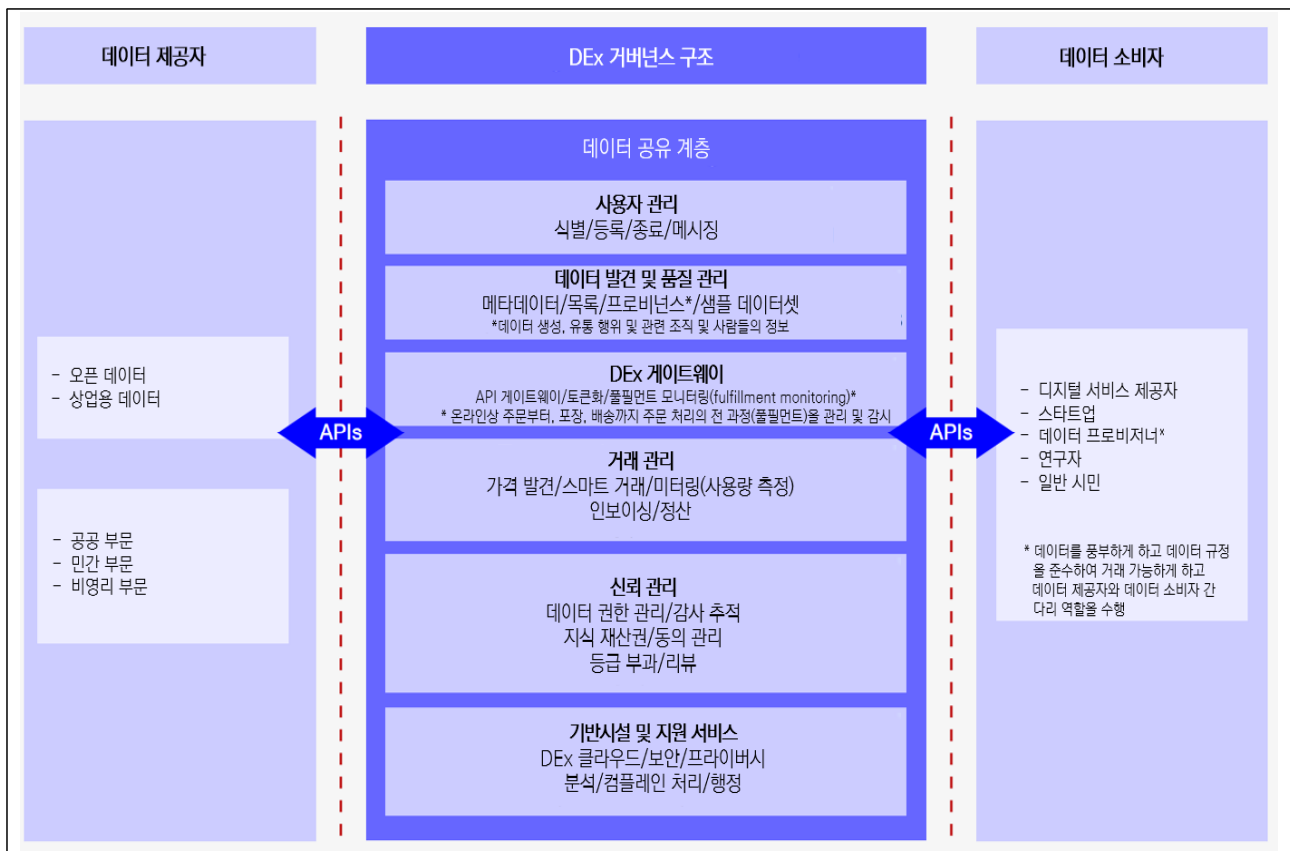
출처: World Economic Forum(2021.8)

▶ DEx 시스템의 중심은 ‘공유’ 계층으로, 기술, 규제 및 관리 측면에서 6가지 기능이 필요

- (ID 관리) 플랫폼에서 거래하려는 데이터 제공자와 소비자의 식별, 인증 및 권한 부여
 - 플랫폼에 대한 신뢰는 DEx가 식별, 인증, 권한 부여 기능을 얼마나 엄격하게 하는가에 달림
- (데이터셋 발견 및 품질 관리) 소비자가 원하는 것을 쉽게 찾을 수 있도록 구조 또는 분류 체계를 갖춰 데이터 제공자와 데이터 소비자 간 데이터의 일 대 일 공유를 용이하게 하는 기능 수행

- (DEx 게이트웨이) DEx 게이트웨이는 데이터 제공자와 소비자 간의 신뢰할 수 있는 데이터셋 공유를 용이하게 하는 것으로서 통합 플랫폼에서 서비스 제공자에게 여러 가지 기회 제공
- (거래 관리) 거래 관리 시스템은 최소한의 단계를 가지고 직관적이며 충돌이 없어야 하며, 특히 사기 거래를 방지하기 위해 안전하고 실수가 없어야 함
- (신뢰 관리) 플랫폼에 대한 신뢰를 위해 무엇보다 데이터의 감사용 기록(Audit trail) 및 데이터 출처 인증이 중요한데, 이를 위해 블록체인 기술 선택 가능
 - 평가 및 리뷰 방식은 가장 인기 있고 신뢰할 수 있는 데이터셋의 식별을 가능하게 함
- (기반 시설 및 지원 서비스) 효과적인 고충 처리 시스템을 포함하는 다양한 기반 시설 및 지원 서비스는 사용자의 신뢰를 높이고 투명성을 높임

| 그림 3 | 공유 계층의 세부 기능



출처: World Economic Forum(2021.8)

3 데이터 공유 촉진 체계

- ▶ 데이터 공유는 세계적으로 아직 초기 단계에 있으며, 성장을 위한 중요한 5가지 요소가 있음
 - 서로 다른 분야에 기반을 둔 공공, 민간 및 공공-민간 파트너십 모두의 다양한 DEx 모델들이 국가나 세계 차원에서 테스트되고 있음

- DEx가 발전하기 위해서는 ① 데이터 가용성, ② 데이터 활용성, ③ 신뢰 구축, ④ 다수 이해관계자 참여, ⑤ 효과적 거버넌스 체계 등 5가지 핵심 요소를 구축하는 것이 중요

(1) 데이터 가용성

- ▶ 데이터의 가용성은 데이터 공유를 위한 전제 조건으로, 대량의 데이터셋에 대한 가용성 제고는 데이터 생태계의 발전을 위한 초기 모멘텀을 제공
- 최고의 데이터 생태계는 여러 종류의 데이터를 통합하고, 분석 과정을 자동화하기 위한 기계학습을 적용하며, 성과 수치를 추적할 수 있게 하는 분석 플랫폼을 중심으로 구축

(2) 데이터 활용성

- ▶ (검색 가능성) 다음의 활동은 일반적으로 DEx에 데이터를 제공하기 전에 수행되는 단계로, 데이터 소비자가 데이터의 존재를 알 수 있게 하는 검색 가능성을 개선
 - (메타데이터 활용) 메타데이터를 통해 데이터셋 검색이 가능하고, 또한 추가로 키워드 정의를 통해 검색 결과 향상
 - (데이터 특성화) 데이터셋의 일반적인 특성 또는 기능을 요약
 - (데이터 구별) 광범위한 검색을 충족하는 다른 데이터셋과의 일반적인 기능을 비교
 - (데이터 분류) 데이터 탐색을 가능하게 하고, 최적의 데이터 유지에 필요
- ▶ (유용성) 데이터가 검색된 후에는 소비자의 사용 목적에 따라 유용성을 지녀야 하는데, 이는 활용성, 품질, 전달성이라는 데이터셋의 필수 속성에 따라 달라짐
 - (활용성, Usability) 활용성 있는 데이터는 예상되는 분석 또는 결과와 연관되고, 소비자가 그것을 사용하여 경제적 가치를 얻을 수 있게 하며, 활용의 비용-편익의 경제성 확보 필요
 - (품질, Quality) 데이터셋은 정확하고, 누락이 없어야 하고, 정보의 출처를 추적해 검증이 가능하며, 포괄적이면서 중복되지 않아야 하고, 신뢰할 수 있어야 함
 - (전달성, Delivery) 데이터의 원활한 소비, 융합, 보급을 위해 시의성, 사용자의 접근성, 사이버 위협 및 개인정보 보호 차원의 보안성 확보가 중요
- ▶ (표준 기반) 데이터는 의미 및 구문 상호운용성 표준을 준수해야 함
 - 표준화된 데이터를 처리하는 도구에 의해 검색 기능을 향상시킬 수 있음
 - 부문별 DEx를 위한 상호운용성 표준이 존재하지 않는 경우, 개발이 이루어져야 하며, 또한 메타데이터, 데이터 특성화, 데이터 식별 기능에 대한 표준도 설계되어야 함

- ▶ (프라이버시 보호) 개인데이터 또는 개인데이터와 비개인데이터가 혼재된 데이터셋에 대한 개인정보보호 규정 준수
 - DEx 설계자의 데이터 보호 준수 보장을 위한 적절한 프레임워크 개발이 필요
- ▶ (동의 기반) 동의 관리(Consent Management)는 개인정보 및 익명정보가 포함된 DEx 에서 발생하는 모든 문제를 처리하는 데 필수적인 요소
 - WEF가 수행하는 공통 목적 이니셔티브(DCPI)는 데이터 주도 경제에 대한 시민 참여 확대를 위한 동의의 중요성을 높이는 데에 집중
- ▶ (접근 권한) 모든 데이터셋의 활용 확대를 위해 데이터 접근 및 해석이 가능해야 하며, 모든 승인된 참여자를 위해 데이터 접근을 보장하는 주요 고려 사항은 다음과 같음
 - 모든 데이터가 한 곳에 집계되어 제공되는 것이 아니기 때문에, DEx가 이루어지는 부문 및 목적에 따라 해당 생태계 이해관계자들이 데이터를 연합하는 방식으로 유지하면서 분산된 방식으로 작업을 수행
 - 데이터 제공자/소유자가 데이터 생애주기 내내 데이터셋에 대한 소유권 및 관련 지식재산권을 보유해야 함

(3) 신뢰 구축

- ▶ 이해관계자들에게 명확성, 통제력 및 안전함을 제공하여 적극적인 참여와 데이터 공유에 확신을 가질 수 있도록 신뢰를 구축하는 것이 중요
 - DEx에 참여하는 이해관계자에 대한 신뢰, 이해관계자 간의 신뢰, DEx에 포함된 시스템에 대한 신뢰, 거래의 정직성에 대한 신뢰가 필요

(4) 다수 이해관계자의 참여

- ▶ 데이터 잠재력이 발휘되기 위해서는 DEx에 참여하는 이해관계자 다수의 이해관계가 균형 있게 조화되도록 운영해야 함
 - * 이해관계자 : 정부, 규제기관, DEx 플랫폼, 데이터 프로비저닝 업체, 데이터 제공자, 기업, 혁신가, 연구개발자, 시민사회
 - DEx는 데이터가 합의된 규범에 따라 참여자의 이익을 위해 운영되어야 하고, 동시에 어떤 특정한 이해관계자의 오용에 의해 DEx에 참여하는 다른 이해관계자에게 해를 입히거나 부당한 이익이 발생하지 않도록 해야 함
 - 여러 이해관계자의 조화를 위해서는 커뮤니티 구축이 필수
 - 커뮤니티는 데이터 소유자의 권리 인식 및 데이터 사용에 발생하는 경제적 이익의 제공, DEx에 참여하기 위한 투명한 규칙, 위험 관리 전략, 포괄성 설정 필요

(5) 효과적 거버넌스 체계

▶ (필요성) 개인, 기업, 공공 분야의 방대한 데이터가 존재하고, 개방되어 있어도 기존의 데이터 거버넌스 모델로 인해 데이터 결합이 용이하지 않음

- 현재 각 지역마다 고려하고 준수해야 하는 특정한 규제 형태가 존재하는데, 이러한 데이터 공유 규제의 필요성과 정도에 대해서 국제적으로 논의가 진행 중
- 즉, 프라이버시, 보안, 상업 기밀 등 민감 데이터, 법적으로 보호된 데이터, 국가 간 데이터 이동을 위한 준수 사항 등 여러 현안들이 존재하는데, 데이터 활용 확대를 위해서는 이러한 문제들을 다루기 위한 적절한 데이터 거버넌스 프레임워크가 필요

▶ (3P 접근 방식) WEF는 데이터 공유 거버넌스 프레임워크의 하나로써 보호(Protect)-예방(Prevent)-증진(Promotion)이라는 3P 접근 방식을 제시

- 데이터가 갖는 고유의 특성*으로 인해 정부 주도형이든 자율 규제형이든 거버넌스 프레임워크를 설계하기가 매우 어려움

* 많은 양, 빠른 처리 속도, 다양성, 정확성 요구, 사용의 비경합성 및 데이터 활용을 통한 대규모 혁신 창출 등

| 표 3 | DEx 거버넌스 프레임워크 3P 접근방식

구분		주요 내용
보호	개인 및 비개인데이터	- 자연인의 개인식별정보(PII)는 생태계 이해관계자에 의해 데이터 가치 사슬 전체에서 보호 - 모든 디지털 및 데이터 생태계의 설계 및 구현에서 개인정보보호 내재화 원칙 준수
	데이터 권한의 포함 관리	- 데이터의 사용 목적, 기간 및 데이터 수익화 여부를 데이터 소유자가 차등 지정 - 미디어 및 엔터테인먼트 부문의 관리와 유사한 데이터 권한 관리 프레임워크 개발 고려
	효과적 고충처리 메커니즘	- 데이터에 대한 모든 이해 관계자의 이익은 적절한 고충 해결 메커니즘을 통해 보호
예방	데이터의 무단 사용 및 오용	- 데이터의 오용을 방지하기 위해 데이터 공유의 목적을 식별하고 동의 - 데이터 처리 목적은 적절한 표준을 통해 정의 - 프레임워크는 혁신을 저해하지 않고, 자율 규제 형태로 설계
	반경쟁 행위	- 프레임워크는 반경쟁적 관행에 잘 대응할 수 있도록 데이터의 특성을 고려하여 설계 - 규모의 장벽, 독점 기술 등을 통해 신규진입자를 막는 관행에 대한 정보를 사전에 제공
	이해관계자에 대한 피해	- 다양한 데이터셋을 사용하는 애플리케이션에 대해 사용자가 신뢰할 수 있는 인증 메커니즘 수립
증진	역동적이고 경쟁적이며 신뢰할 수 있는 DEx 생태계	- 데이터 생태계 참가자의 등록/탈퇴를 위한 절차 간소화의 수준은 데이터 분야와 목적에 따라 상이 - 데이터의 민감도와 고유한 가치에 따라 적절한 보고 및 규정 준수 메커니즘 개발 - DEx 생태계의 데이터 소유자와 다양한 서비스 제공자 간의 데이터 이동성 권장
	표준 및 프로토콜 개발	- 상호운용성, 스마트 계약 및 데이터셋 발전을 가능하게 하는 표준 및 프로토콜의 개발 및 채택 - DEx 생태계를 쉽게 확장할 수 있는 오픈 소스 도구 및 기술 개발
	오픈 이노베이션	- 혁신 및 연구를 위한 데이터 접근 용이성 촉진 - 스타트업과 중소기업의 데이터에 대한 평등한 접근 보장 - 데이터 접근, 사용 및 평가에 대한 공정하고 합리적이고 비차별적 조건 보장 - 참가자, 부문 및 산업 전반에 걸쳐 확장할 수 있는 파트너십 촉진

출처 : Capgemini Research Institute, Data sharing masters, 2021.07.



데이터 공유 생태계 현황과 협업 장애요인 및 추진 전략

1 개념 및 유형

- ▶ (개요) 캡제미니의 보고서는 미국, 유럽, 아시아 지역 조직의 임원 750명 및 산업 전문가 30명의 설문조사에 기초하여 작성됨
- ▶ (개념) 동 보고서는 데이터 생태계⁶⁾를 이전의 사일로(silo)화된 시스템에서는 불가능했던, 새로운 가치를 창출하는 데이터 공유 및 관리를 위한 여러 기관 간의 파트너십으로 정의
 - 데이터 생태계라는 용어는 보편적인 용어가 아니며, 데이터 공유를 나타내는 데에 데이터 자산(Data estates), 데이터 시장(Data marketplace), 데이터 공간(Data spaces) 등 다양한 표현을 사용
 - ※ 따라서 본 고에서는 보고서의 통일성을 위해 데이터 공유라는 용어를 사용
 - 전통적인 데이터 생태계에서는 여러 조직이 규정 및 지침을 준수하여 데이터를 공유
 - 데이터 소스 조직 등 여러 조직이 최종 소비자 관계에 따라 데이터를 수집
 - 데이터 공유를 용이하게 하는 애그리게이터(Aggregator)나 중개 조직이 하나 이상 존재하거나 없을 수도 있으며, 데이터 소스 조직 중 하나가 애그리게이터의 역할을 수행하기도 함
 - 사용자 조직은 중개자가 통합한 데이터를 사용하는 데 대한 대가를 금전적 또는 비금전적 인센티브(예: 비용 절감, 위험 감소, 고객 경험 개선)의 형태로 지불
 - 데이터 생태계는 조직들이 해당 지역에서 적용되는 규제 하에서 데이터와 인사이트를 공유해 모든 참여자를 위한 새로운 가치 창출을 하는 데에 동의할 때 형성 가능
 - 예를 들어, 물류 파트너 분석을 통한 공급망 최적화, 신제품의 출시 전에 중단 간 관점에서 지속가능성 영향 파악, 제약회사의 약물 발견 데이터 학습으로 보다 안전한 약물을 보다 빠르게 시장에 출시
 - 최근 수년간 기업이나 기관 사이에서 이러한 데이터 생태계를 형성하고자 하는 관심이 증대되는 상황

표 4 | 데이터 공유 생태계에 대한 관심 증대 요인

구분	주요 내용
비즈니스 관점	- '19년 EU 27개국 데이터 경제의 총가치는 3,250억 유로로 추정, '25년까지 5,500억 유로 이상으로 증가하여 전체 EU GDP의 4%를 차지할 전망 - 조사 대상 기업 ⁷⁾ 의 43%가 데이터로 수익을 창출할 수 있고, 데이터의 가치를 정량화할 수 있으며(22%), 수익 창출을 위해 데이터 생태계에 새롭게 참여할 것이라고 응답(54%)
기술적 관점	- '23년까지 1인당 3.6개의 네트워크 장치(연결)를 소유하며 방대한 양의 데이터를 생성할 것으로 예상 - 모바일 네트워크업체인 GSMA는 세계 IoT ⁸⁾ 시장 가치가 '25년까지 9,000억 달러에 달할 것으로 추정했으며, 이러한 가치의 대부분이 데이터를 활용한 애플리케이션, 플랫폼 및 서비스에서 생성될 것이라 예견 - Snowflake, AWS, Dawex 등의 안전한 실시간 클라우드 기반 데이터 공유를 통해 대량의 데이터를 공유
규제 및 주권 관점	- 데이터 주체의 권리와 조직 간 데이터 공유를 규정하는 규제, 법안 채택이 증가하고 있고, 데이터 주권(개인 또는 조직이 개인 및 비즈니스 데이터를 제어할 수 있는 능력)에 대한 요구 증대
지속가능성 관점	- 조사 대상 중 60%는 지속가능한 개발 목표에 대한 진전을 데이터 생태계에 참여하는 핵심 동인으로 언급

출처 : Capgemini Research Institute, Data sharing masters, 2021.07.

6) 캡제미니 보고서에서는 데이터 공유 생태계, 데이터 생태계, 데이터 공유를 유사한 의미로 사용

▶ 데이터 생태계는 실제로 10년 넘게 존재하여 왔지만, 최근의 형태는 과거와 다른 특징을 보이며, 유형은 크게 4가지로 분류

- (최근의 특징) 복잡성(여러 유형의 참여자), 투명성, 개방성과 역동성(용이한 가입, 탈퇴), 상호운용성을 위한 최신 표준 및 기술 준수, 신뢰할 수 있는 소프트웨어 인프라 구축 능력
- (생태계 유형) 생태계를 ①데이터 중개 및 통합 생태계, ②상호 데이터 공유 생태계, ③통합 분석 생태계, ④협업 데이터 공급망 생태계의 4가지 유형으로 구분

| 표 5 | 데이터 생태계 유형

구분	유형 형태
①데이터 중개 및 통합 생태계(55%) - 가치가 있는 데이터를 보유하거나 통합 - 데이터 또는 데이터에서 파생된 아이디어를 상품화 - 데이터는 앱이나 온라인을 통해 직접 수집, 또는 스마트워치 등의 장치를 통해 수집 - 예: 데이터 통합 플랫폼인 Thinknum은 정부 소스, 소셜 미디어 및 회사 웹사이트에서 데이터를 수집하여 예를 들면 소매업체가 수익성 있는 매장 위치를 선택하도록 지원	
②상호 데이터 공유 생태계(20%) - 이 모델에서는 여러 회사가 협력하여 효율성 제고 - 데이터 공유는 협업의 가치를 제고 - 일반적으로 파트너 중 하나가 지배적 위치를 점유(소규모 조직도 데이터를 공유를 통해 가치 확보) - 예: 부품 제조업체는 OEM 업체에게 물류 정보를 제공하고, OEM은 품질 테스트 결과를 알려줌	
③통합 분석 생태계(13%) - 이 모델은 규제, 데이터의 양의 이유로 직접 공유하거나 접근할 수 없는 경우에 다른 조직의 데이터에서 분석을 실행하고 아이디어를 도출하는 것 - 데이터에 대한 접근을 판매하는 것이고, 요금은 제공된 데이터와 분석의 정도에 기반 - 예: NVIDIA의 Clara는 다른 조직과 병원의 데이터 접근을 통해 교육을 수행하고 정확한 진단에 기여	
④협업 데이터 공급망 생태계(11%) - 여러 조직이 단일 고객 또는 시장에 서비스를 제공하는 데이터 생태계 파트너십의 가장 협업적 모델 - 이 모델에서는 조직 간의 협업 자체가 새로운 제품 또는 서비스가 되어 고객을 위한 결과를 도출 - 예: Open Carbon Footprint는 최종 제품의 탄소 발자국을 정확하게 알 수 있도록 가치 사슬의 다양한 파트너를 모음	

출처: Capgemini Research Institute, 2021.07.

주: ()안의 %는 조사대상 기업들의 생태계 형태의 비율임

2 데이터 공유의 진전 현황

(1) 데이터 생태계 참여의 편익

- ▶ 설문조사를 통해 고객만족도 개선, 생산성/효율성 향상, 새로운 제품·서비스·비즈니스 모델 개발, 위험 감소, R&D 속도 단축, 비용 절감의 6가지 측면에서 기업들이 데이터 생태계 참여로 얻은 편익을 추산⁸⁾

※ 캡제미니의 보고서는 미국, 유럽, 아시아 지역 조직의 임원 750명 및 산업 전문가 30명의 설문조사에 기초하여 작성됨

- 설문조사 결과, 데이터 생태계 참여로 인해 지난 2~3년간 고객만족도 개선(15.1%), 생산성/효율성 향상(13.9%)에서 가장 큰 효과를 본 것으로 나타남
- (사례) 해양산업의 위험 모델링 개선 공동 프로젝트인 HiLo는 구명정 사고를 72% 줄였고, 오스트리아 대량포장기업 Greiner Packaging International은 물류 데이터기업인 OBUU와 협력하여 상당한 비용 절감 실현
- ▶ 기업의 데이터 생태계 참여가 수익, 생산성, 비용에 미치는 결과를 보수적, 낙관적 시나리오로 나누어 추산한 결과, 보수적 시나리오에 따르면 5년간 총 2억 3,500만 달러, 낙관적인 시나리오에 따르면 9억 4,000만 달러의 총 재무 효과 예상⁹⁾

[표 6] 데이터 생태계 참여의 재무 효과(5년 누적, 백만 달러)

효과 구분	보수적 시나리오	낙관적 시나리오
추가 매출액	95	378
추가 비용 절감	47	186
추가 생산성 향상	94	377
총 재무 효과 합계	235	940

출처 : Capgemini Research Institute, 2021.07.

- ▶ 또한 캡제미니의 이전 연구¹⁰⁾에서도 데이터 마스터 기업*의 경우 더 높은 재무성과를 보임

* 의사 결정에 데이터를 광범위하게 활용하는 기업으로, 조사 대상 기업의 16%를 차지

- 데이터 마스터 기업은 직원당 매출액이 다른 기업들에 비해 70%, 고정자산 회전율이 245%, 영업이익률이 22% 더 높은 것으로 나타남
- 기업들을 총 14개의 외부 데이터* 중 8개 이상을 사용하는 기업군과 7개 이하를 사용하는 기업군으로 나눠 분석한 결과도 마찬가지로 데이터를 많이 활용하는 기업의 재무성과가 높았음

* 14개 외부데이터: 공급업자, 유통업자, 플랫폼 제공자, 익명 소비자, 식별 소비자, 소비 내역, 소셜미디어, 제품리뷰,

8) 기업들이 데이터 생태계 참여를 통해 얻은 편익을 지난 3년간 실현된 연간 환산 효과와 현재의 수준에서 향후 3년간 실현될 것으로 예상되는 연간 환산 편익의 비율을 백분율로 나타냄

9) 기업의 매출액을 조사 대상 기업의 평균 매출액인 100억 달러로 상정하고, '25년까지 데이터 경제가 EU GDP의 4%를 차지한다는 것을 전제

10) Capgemini(2020.11), The data-powered enterprise

경쟁자, 분석보고서, 주요 플랫폼 기업, 민간 소유, 공개 정보, 기타 외부 데이터

- 외부 데이터 비사용 기업과 비교했을 때, 외부 8개 이상 사용 기업군은 시가총액, 고정자산 회전율, 직원당 매출액이 각각 119%, 1,256%, 37% 더 높았으며, 7개 이하 사용군에서는 각각 56%, 461%, 2% 더 높게 나타남

(2) 데이터 생태계 참여 가속화 및 요인

▶ 가까운 미래에 기업들의 데이터 생태계 참여가 크게 가속화될 것으로 예상됨

- 설문에 응한 조직의 거의 절반(48%)이 새로운 데이터 생태계에 참여하거나 관련 활동을 추진할 방안을 모색 중이며, 3개 중 1개 이상의 조직(36%)이 기존의 참여도를 강화할 방침
- 즉, 설문 응답 조직의 약 84%가 3년 이내에 데이터 생태계의 신규 참여 혹은 활동 강화 추진 예정

▶ 일반적으로 이러한 데이터 생태계 참여 가속화의 동인으로 다음의 사항들이 제시됨

- 조직의 지속 가능성과 사회적 실천 및 규약을 발전시키기 위한 노력
- 데이터 가치에 대한 인식과 데이터를 조직 외부와 공유하려는 강한 의지
- 엄격한 규정을 통해 데이터를 안전하게 공유할 수 있는 기술과 플랫폼

▶ 특히, 이러한 데이터 생태계 참여의 빠른 진전에는 규제 당국에 의한 적절한 제도 확립이 큰 역할을 수행

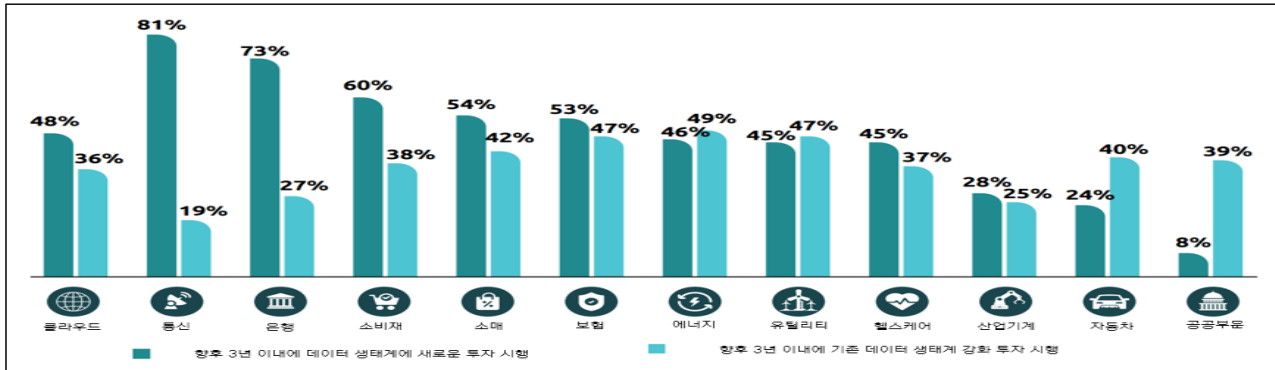
- (EU의 금융 데이터 공유 촉진) EU는 유럽 데이터 공간에서 보다 원활한 금융 데이터 공유를 설정하기 위한 제도를 강력 추진했는데, 이것이 데이터 생태계의 급속한 발전을 초래
- (영국 오픈 뱅킹 생태계 사례) 현재 주요 은행 및 핀테크기업을 포함한 300개 회원사와 250만 이상의 영국 고객으로 구성된 영국의 오픈 뱅킹 생태계가 운영되고 있음
 - 유럽의 새로운 법률인 지불시스템 지침2(PSD2, Payments Services Directive)의 발효와 함께 오픈 뱅킹이 가능('18.1.)
 - 오픈 뱅킹을 통해 고객은 금융 데이터에 대한 접근을 제공자(은행 및 핀테크 등)와 공유하여 새로운 제품 및 서비스 이용, 재정 관리, 신용정보 접근 및 결제 가능
 - 은행 및 기타 금융 기관은 API를 사용하여 소비자의 동의 하에 정보를 공유하여 새로운 서비스 생성
 - 오픈뱅킹 소비자들은 '20년 400만 건 이상의 오픈 뱅킹 결제와 58억 건의 API 호출을 했는데, 이는 '18년 대비 각각 13배, 87배 증가한 수치

(3) 데이터 생태계 투자 계획

▶ (부문별 투자 계획) 통신, 은행, 소비자재 및 소매와 같은 소비자 대면 부문이 향후 데이터 생태계 투자에 적극적인 것으로 나타남

- 설문 조사 결과, 통신 부문 기업의 81%가 향후 3년간 데이터 생태계에 새로운 투자를 추진할 것이라고 응답해 가장 높게 나타났으며, 은행 부문(73%)이 그 뒤를 이음

| 그림 4 | 부문별 데이터 생태계 투자 계획(복수응답, N=750개)

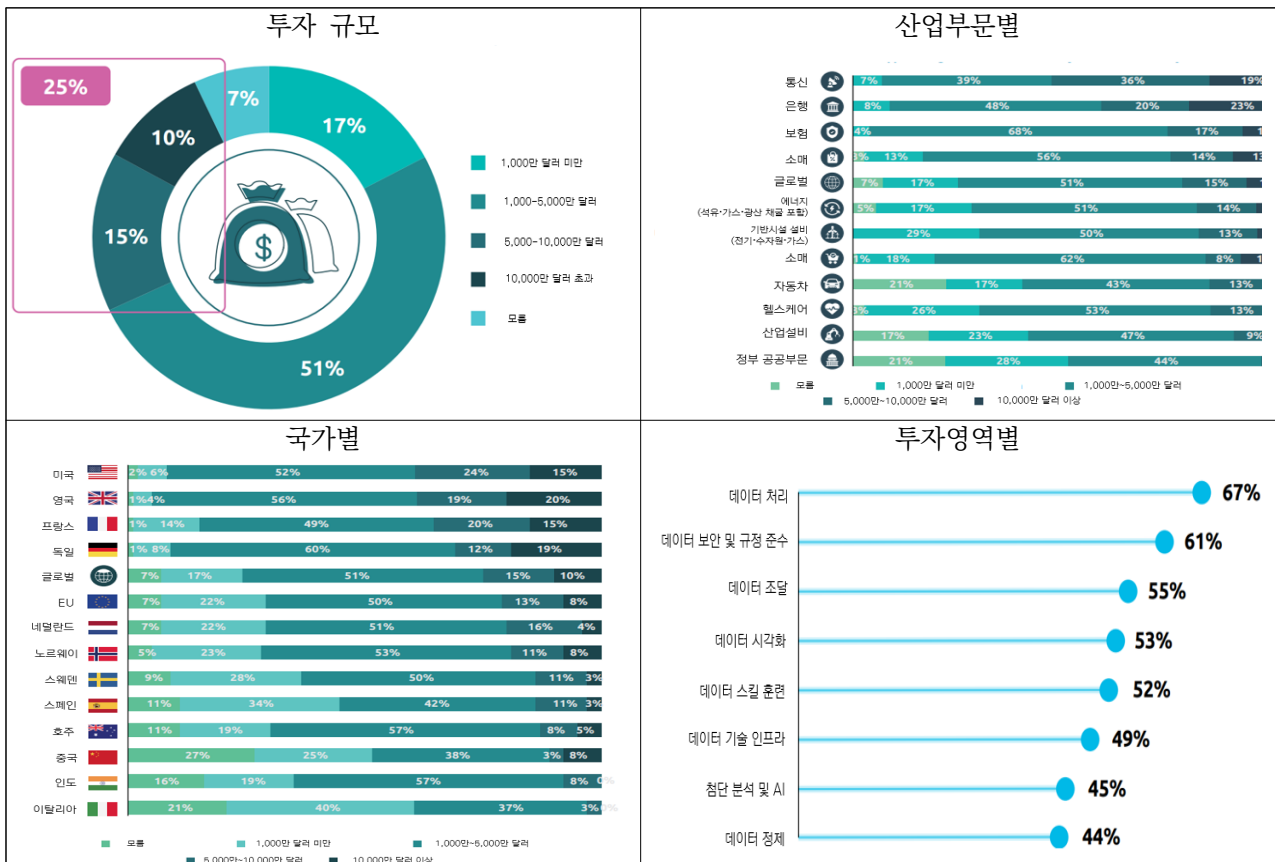


출처 : Capgemini Research Institute, 2021.07.

▶ (투자 규모) 대부분의 조직은 1,000만~5,000만 달러의 투자 규모이고, 응답 조직의 25%가 향후 2~3년 동안 데이터 생태계 활동에 5천만 달러 이상을 투자할 계획(그림5)

- (산업부문별) 통신회사의 55%, 은행 부문의 43%가 향후 2~3년간 5,000만 달러를 투자할 계획
- (국가별) 미국과 영국이 가장 많은 규모로 투자할 계획인데, 두 나라 모두 응답 조직의 39%가 5,000만 달러 이상 투자할 계획
- (투자 영역별) 특히 데이터 기술 인프라, 도구, 인재 및 기술, 프로세스 재설계 역량 확보 등 데이터 생태계 운영을 위한 모든 요소에 대한 투자를 고려하는 것으로 나타남

| 그림 5 | 데이터 생태계 투자 규모(산업부문별·국가별·투자영역별)



출처 : Capgemini Research Institute, 2021.07.

3 데이터 공유의 파급효과

(1) 데이터 생태계의 사회경제적 파급효과

▶ 동 보고서는 데이터 생태계가 사회경제적 측면에 이점을 제공하는 3가지 예시를 제시

● 시민 서비스 증진과 소비자 영역에서의 수익원 창출

- 예를 들어, IoT 등 새로운 기술을 통해 획득한 차량 운행 데이터(목적지, 운행시간, 빈도, 평균 속도 등에 대한 정보)는 스마트 시티, 광고회사와 연계하여 다음과 같은 파급효과 발생
 - (스마트시티와 연계) 교통 데이터의 상세한 흐름은 도시의 교통 흐름을 계획하고, 혼잡 및 교통 움직임에 대한 데이터는 주차 공간이 필요한 지역을 식별하며, 교통 체증 지역과 인기 있는 방문지는 시민서비스를 계획하는 데 도움이 됨
 - (광고회사와 연계) 교통 흐름 데이터를 활용하여 도시 내 어디가 가장 좋은 광고 지점인지를 식별함으로써, 소비자 회사가 보다 효과적인 마케팅 캠페인을 추진해 수익을 얻을 수 있도록 함

● 다양한 부문에 걸친 협업 및 데이터를 공유를 통한 수익 창출

- (소매업체) 통신이용 데이터를 통해 소매업체의 매장을 방문하기 전후에 고객이 방문한 매장/장소를 파악하여 보다 정교한 고객의 활동지도 작성
 - 인구 통계 데이터와 연계함으로써 소매업체가 매장 레이아웃을 계획하고, 효과적인 마케팅 결정을 내리는 데 활용
 - 나아가 이러한 소매 데이터를 소비자 판매(CPG, Consumer packaged goods) 기업과 공유하면 제품에 대한 더 나은 예측 가능성을 제공하여 재고 관리 최적화 가능
- (관광업계) 통신이용 데이터를 통해 관광객이 가장 많이 방문한 식당 및 소매점을 식별할 수 있어 향후 관광객에 대한 맞춤서비스 제공 가능
 - 또한 현지 레스토랑이나 호텔의 효율적인 자원 계획과 관리에 유용하고, 관광업체는 이 데이터를 사용하여 평균 관광객 흐름, 체류 기간 등을 보다 정확하게 파악할 수 있음
- (공공 부문) 집계되고 익명화된 통신 데이터는 정부 기관에서 시민의 이동, 이주 패턴, 비상사태 대응 등을 추적하는 데 활용
 - 실제로, Vodafone Analytics Platform은 코로나19로 인한 봉쇄 전후의 인구이동 변화를 보여주는 히트 맵(Heat map)을 이탈리아 롬바르디아 지방 정부에 제공

● 공공 서비스 보급의 효율성과 시민 참여 촉진

- 공공기관 간의 원활한 데이터 공유를 통해 안전, 비즈니스 및 금융 관련 시민 서비스를 데이터 공유 플랫폼에 통합하여 시민들에게 제공
- 이는 스마트 도시 전환에 도움이 되고, 시민들이 언제 어디서나 자신의 의견, 질문 및 우려 사항을 제시하는 등 도시 개발에 직접 참여할 수 있는 기회를 제공

(2) 가치사슬 부문에 따른 데이터 공유 주요 사례

- ▶ 설문조사를 통해 조직들이 흔히 참여하는 37개의 데이터 공유 활용 사례(붙임 참고)를 도출했으며, 그 중에서도 가장 일반적인 분야는 광고, R&D, 에너지, 스마트시티, 온라인 구매로 나타남

[표 7] 가장 일반적인 데이터 공유 분야 및 주요 사례

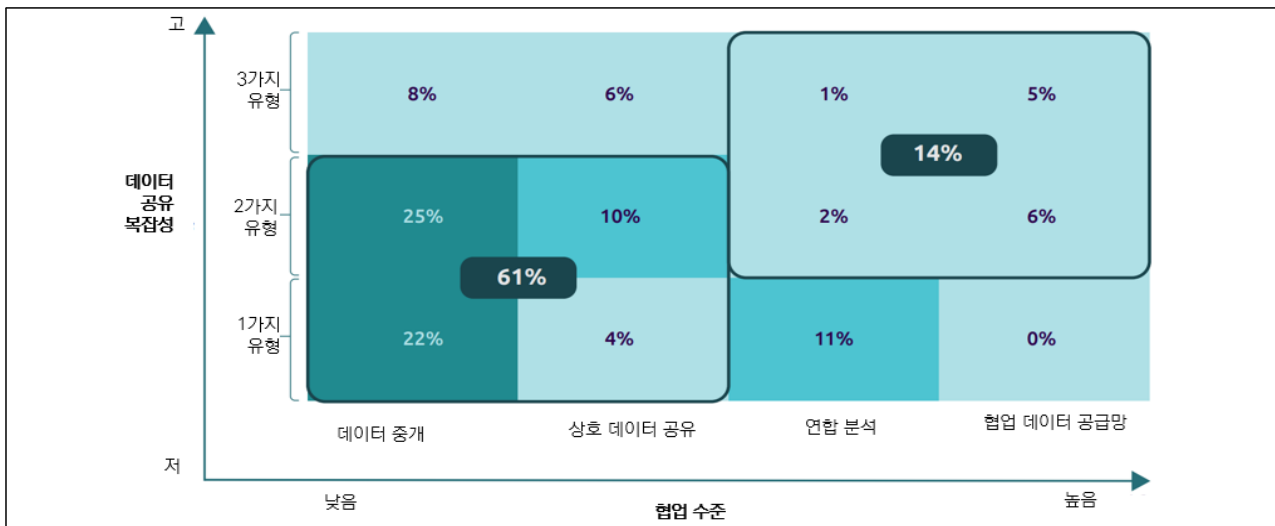
분야	참여 내용	사례
광고효율성 및 타겟팅 모니터링	- 생태계 내 여러 부문의 협업은 기업들이 고객의 기호 및 일상 활동을 더욱 명확하게 파악할 수 있는 기회를 제공 - 소비자 대면 부문인 소비재(48%), 소매(51%), 통신(87%)에서 가장 활발	- (Starbucks)데이터 공유를 통해 고객 구매 행동 분석, 개인맞춤 향상을 위한 고객 세분화, 미디어 자출의 효과 제고에 활용 - (인공지능 빅데이터 분석 회사 LUCA)통행자수 정보, 고객 인구 정보를 제공하여 이상적인 소매점 위치를 설정하고, 나아가 영업시간을 고려하는 데 활용
새로운 연구 개발에 활용	- 전체 조사 대상 조직 중 41%가 시범적으로 데이터 생태계에 참여하고 있고, 20%는 본격적 적용 단계 - 특히, 의료 및 제조 부문은 데이터 공유를 통한 효과적인 R&D에 크게 의존	- Airbus의 Skywise는 세계 100개 이상의 항공사와 파트너 생태계를 구축하여 9,000대 이상의 연결된 항공기와 18,000명 이상의 사용자 정보 확보 - Skywise 프로그램의 데이터에 대한 분석을 활용하여 미국 대표 항공사들은 연간 연료비용을 1,300만 달러 절약 - EBI(European Bioinformatics Institute)는 유럽 위원회, European Open Science Cloud, ELIXIR 유럽, 대학 및 의료 센터와 협력하여 COVID-19 데이터 플랫폼 구축 - COVID-19 전염병과 관련된 데이터셋에 신속하게 접근할 수 있게 하여 진단, 치료제 및 백신에 대한 연구 가속화에 기여
실시간 에너지/수도 공급-소비 데이터 공유	- 에너지 및 유틸리티 부문은 에너지 낭비를 줄이고 이익을 늘리기 위해 매우 효율적인 예측 방법 필요 - 데이터 생태계의 세분화된 데이터를 수집하는 방식을 통해 에너지 소비 패턴과 부하 변동에 대한 명확한 파악 가능 - 에너지 및 유틸리티 부문의 73%가 시험 및 개념 증명을 실행하고 있고, 16%는 완전한 규모의 생태계 구축	- (청정에너지 소비 축진을 추구하는 인도 스타트업 Bidgely) 500억 개 이상의 계량기 판독 값에서 수집된 데이터를 사용하여 가전제품의 에너지 소비 패턴에 대한 정확한 정보 파악 - 이 데이터는 더 정확한 고객 사용 패턴 파악과 계획 수요를 도출하기 위해 유틸리티 회사와 공유됨 - (영국 스마트 에너지 플랫폼 회사 PassivSystems) Sheffield University 및 비영리기관인 Open Climate Fix와 협력하여 태양 에너지 발전에 대한 보다 정확한 예측을 수행 - 전력의 수요와 공급의 균형을 유지하고, 그리드 운영과 소비자의 비용을 절감하고, 환경 편익을 가져오는 재생에너지 활용 확대에 기여
스마트시티 시민서비스 관련 데이터 공유	- 스마트시티, 주차, 맵핑 등의 서비스를 위한 모빌리티 통합 데이터 공유로서 자동차, 보험 및 공공 부문 전반에 적용 - 자동차 부문의 81%가 시험 및 개념증명을 실행하고 있고, 7%는 완전한 규모의 생태계 구축	- (모빌리티 솔루션 분야의 선도적인 글로벌업체 Avis Budget Group) 자동차 데이터 서비스 플랫폼인 Otonomo와 제휴하여 생태계 구축 - 다양한 브랜드와 모델을 포함하는 대규모 커넥티드 차량의 데이터('20년에 약 70억 마일에 해당하는 주행 데이터)를 활용 - 차량 데이터 분석을 통해 Avis의 운영 간소화, 비용 절감 및 예측 유지 보수 솔루션에 의한 고객 경험 개선에 적용 - JD Power를 첫 번째 고객 중 하나로 확보한 이 데이터 생태계는 다른 여러 스마트시티 사업에 서비스 제공 확대를 모색
온라인 구매 및 거래 데이터 공유	- 은행 데이터는 일반적으로 규제 제약장벽이 높지만 활용시 비즈니스나 소매고객을 위한 유용한 아이디어 제공 - 온라인 구매 및 거래 데이터 공유는 은행 부문에서 가장 활발 - 은행 부문 조직의 45%가 이를 전면적으로 구현하고, 42%가 시범 형태로 실행	- (Visa) Bank of America Merrill Lynch 및 Citi와 같은 생태계 참여자들에게 Amazon Business에서 Visa 카드를 사용하여 이루어진 모든 구매에 대한 데이터 공유를 통해 제공 - 데이터 구축 과정은 기업을 추적하고 기록하는 데 시간이 많이 소요되는데, Visa, Amazon 및 참여사들 간의 협력을 통해 원활하게 이루어짐

출처 : Capgemini Research Institute, 2021.07.

4 공유 생태계 협업 장애요인

- ▶ 데이터 생태계 참여 조직을 대상으로 데이터 공유의 복잡성 유형과 앞서 제시한 4가지 데이터 생태계 유형¹¹⁾을 기준으로 참여 형태를 분석한 결과, 향후 보다 발전된 형태로 데이터 생태계를 진전시켜야 할 것으로 보여짐
 - 데이터 공유의 복잡성은 다음의 3가지 유형 중 몇 가지를 사용하였는가로 구분
 - 직접 데이터 공유(파일 전송, API 등 활용)
 - 집계 기반 데이터 공유(중개자/통합자가 여러 소스에서 데이터를 수집하여 고객과 공유)
 - 교환 또는 시장(예: AWS, Snowflake, Verrisk)을 통한 데이터 공유
 - 또한 앞서 제시한 데이터 생태계 유형은 데이터 중개→상호 데이터 공유→연합 분석→ 협업 데이터 공급망으로 갈수록 데이터 협업 수준이 높아짐
 - 조사 대상 조직들을 분석한 결과, 대다수의 조직(61%)이 데이터 공유 유형 중 한두 가지만 사용하고 데이터 중개 및 상호 협업 모델에만 집중하는, 복잡성이 낮고 협업도가 낮은 생태계 활동을 하는 것으로 나타남(그림6의 좌하단)
 - 응답 조직의 14%만이 가장 협업적인 생태계 모델과 복잡한 유형의 데이터 공유를 채택(그림6의 우상단)
 - 협업도와 복잡성이 높은 데이터 공유에 관여하는 14%의 조직은 단순한 데이터 공유 및 낮은 협업 모델에 의존하는 조직보다 10% 포인트 이상의 재무 성과(수익, 비용절감, 생산성 향상 효과)를 나타냄

| 그림 6 | 데이터 생태계 참여 형태 분포(N=698개)



출처 : Capgemini Research Institute, 2021.07.

- ▶ 조사 대상 기업들은 데이터 생태계 참여 확대 과정의 다양한 장애 요인들을 지적
 - 대다수의 조직이 이러한 과제들에 직면해 있기 때문에, 더 큰 규모의 데이터 공유 추진이 어려운 상태

11) 데이터 중개 및 통합 생태계, 상호 데이터 공유 생태계, 통합 분석 생태계, 협업 데이터 공급망 생태계

[표 8] 데이터 생태계 참여 확대 장애요인 (5년 누적)

구분	내용	최대과제라고 응답한 조직의 비율 (복수응답)
거버넌스, 리스크, 규정 준수	사이버 보안 위협	62%
	지식재산권 및 민감데이터 문제	57%
재무 요인	인재와 스킬에 대한 투자 부족	56%
	데이터 생태계 투자에 대한 수익의 불확실성	29%
신뢰	데이터 공유에 대한 소비자 저항	56%
	외부 데이터에 대한 신뢰 부족	46%
기술 요인	접근 권한을 통제할 수 있는 데이터 공유 플랫폼의 부족	56%
	낮은 데이터 품질 (데이터 가용성 부족 포함)	53%

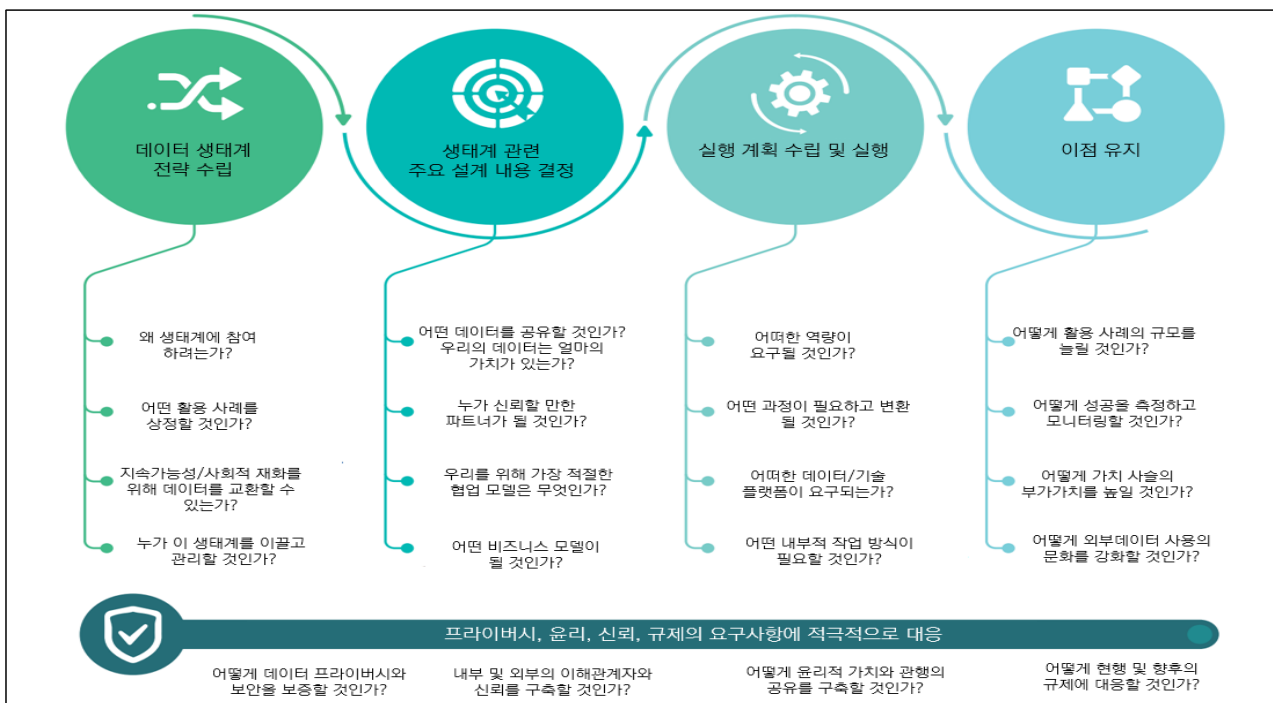
출처 : Capgemini Research Institute, 2021.07.

2 추진 전략

▶ 조직이 데이터 생태계로부터 지속적으로 이점을 얻을 수 있도록 하는 데이터 생태계 추진 로드맵을 제시

- 동 로드맵의 핵심은 전략 수립, 생태계 주요 설계 내용 결정, 실행 및 성장이라는 참여 이전, 참여 도중, 참여 이후의 단계에서 꼭 필요한 사항들을 점검하여 대응방안을 제시

[그림 7] 데이터 생태계 추진전략의 주요 구성요소 및 관련 주요 질문



출처 : Capgemini Research Institute, 2021.07.

(1) 데이터 생태계 전략 수립

▶ 데이터 생태계에 참여하려는 조직은 사전에 명확한 전략을 수립하는 것이 중요

- 고객만족도 개선, 운영 효율성 및 고객 증대 등 생태계에 참여하려는 목적을 명확하게 설정
- 조직과 데이터 공유 파트너의 가치를 창출할 수 있는 데이터 공유 사례를 파악
 - 여러 생태계에 참여할 가능성이 많다면, 가급적 단기적으로 빠르게 가치를 창출할 수 있는 생태계를 고려하는 것이 중요하며, 여러 부문과 가치사슬에서 가장 흔한 활용 사례를 출발점으로 삼는 것이 나음
- 사회적 이익(지속가능한 개발, 사회통합 등)에 기여할 수 있는 기회를 모색
 - 설문조사에서 많은 조직이 지속가능한 개발(73%)과 사회적 통합(65%)을 중시
- 외부 데이터 공유의 감독, 지원의 책임과 효율적 운영을 보장하는 리더와 팀을 지정

(2) 생태계 관련 주요 설계 내용 결정

▶ 다양한 생태계 모델들의 주요 특성을 파악하여 조직에 적합한지를 판단하고, 공유 데이터의 내용, 조직과 데이터 공유 파트너의 역할 등 생태계의 주요 설계 사항을 결정

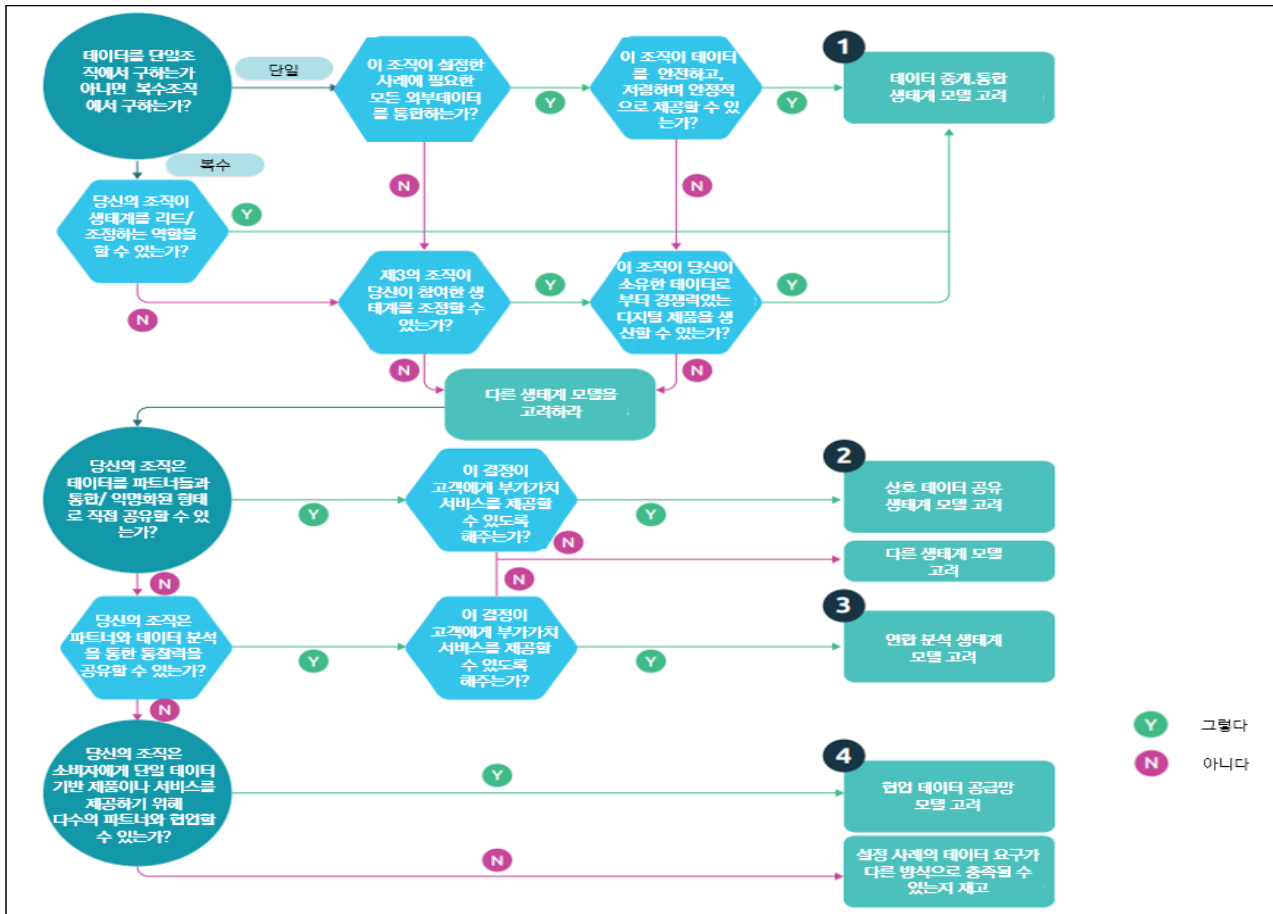
- 외부 데이터의 활용도에 맞추어 조직 내에 존재하는 모든 데이터와 특성을 파악하고, 해당 데이터의 목록 작성 및 품질 평가 수행
- ▲외부 데이터의 가치 평가 및 그에 대한 지불 용의(Willingness to pay), ▲외부에 공유 가능한 데이터 유형과 공유 가능한 외부데이터의 출처 유형 ▲데이터 분석 모델, 정의, 설명 등을 통한 데이터의 가용성 및 공유 가능성 여부 ▲데이터 접근 가격 설정 등의 설계 사항에 대한 결정
- 데이터 생태계 모델 결정과 생태계 운영에서의 조직과 데이터 공유 파트너의 역할 설정
 - 의사결정 트리(그림8)를 통한 적합한 모델 결정 및 혁신, 그리고 운영효율성 차원의 역할 설정
- 생태계 파트너간에 윈-윈 파트너십을 구축하려면 어떤 데이터와 아이디어를 공유할지를 고려하는 것이 중요

(3) 명확한 실행계획 수립 및 실행

▶ 조직이 전략 실행 단계에 들어서게 되면 실행의 성공을 위한 명확한 계획 수립이 요구됨

- 앞서 제시한 3가지 유형(직접 데이터 공유, 집계 기반 데이터 공유, 교환 또는 시장을 통한 데이터 공유)에서 적합한 기술 플랫폼의 결정
 - 대다수의 조직(56%)이 가장 큰 기술적 과제로 데이터 공유 플랫폼의 부족을 지적
- 내부에 갖추고 있지 못한 기능에 대해 자체 구축 방안이나 파트너로부터 아웃소싱 하는 방안 마련
- 외부 데이터의 수집, 저장, 관리를 위한 기존 조직의 새로운 프로세스와 역할이 요구되고, 또한 규제 사항에 잘 대응하도록 파트너간의 기존 역할을 전환하는 것 필요(그림8)

| 그림 8 | 의사결정 트리 과정을 통한 적합한 생태계 모델 결정



출처 : Capgemini Research Institute, 2021.07.

(4) 데이터 생태계의 이점 유지와 성장

▶ 전략 실행의 토대가 마련되면 실제 데이터를 시험적으로 활용하여 수익 여부를 판단

- 데이터 생태계 활동의 우선순위를 명확하게 정의하고, 데이터 소유자, 데이터 관리자, 데이터 보호 담당자 등의 주요 역할을 설정
- 데이터 생태계의 산업화를 통해 가치 창출을 유지하면서 시간이 지남에 따라 부가가치가 보다 높은 쪽으로 전환해가는 시도 필요

(5) 프라이버시, 윤리, 신뢰, 규제에 적극 대응

▶ 데이터 생태계를 시작하기 전에 규제 및 윤리 요구 사항에 대한 선제적 대응 중요

- 모든 생태계 파트너들과 신뢰의 토대를 마련하기 위해 규제 사항 등에 대한 가이드라인과 절차 구축 필요
- 혹은 생태계 내의 어느 한 조직에서 윤리 지침이 잘 정의되어 있을 경우, 모든 파트너들이 그 원칙을 공유하고 실행 가능

PART IV

결론 및 시사점

- ▶ 데이터 경제의 큰 잠재력 실현을 위해서는 적절한 데이터 거버넌스 체계 구축이 중요
 - 데이터 경제의 번영은 효율성과 투명성 향상에서 혁신을 통한 새로운 경제적 가치 창출에 이르기까지 다양한 방식으로 정부, 시민, 산업, 기업, 혁신가나 연구자에게 혜택을 제공
 - 개방적이고, 통합되었으며, 확장 가능한 데이터 공유 체계 구축을 촉진함으로써 데이터 경제의 큰 잠재력을 실현할 수 있다는 시각이 일반적
 - 데이터 공유에서의 형평성과 투명성을 보장하고, 모든 이해 관계자의 권리와 이익을 보호하며, 개별 당사자의 권리를 보호하기 위해서는 적절한 거버넌스 프레임워크 구축이 요구됨
 - 이러한 맥락에서 프라이버시 보호, 보안, 혁신 촉진, 반경쟁 관행 방지, 기술 법적 규제 프레임워크 및 인센티브 메커니즘 등의 문제에 대한 합의할 수 있는 체계 마련 필요
 - 동 보고서의 내용은 데이터 경제 잠재력의 발현을 위한 기반으로, 데이터 공유 생태계를 개발할 때 발생하는 다양한 도전 과제에 대한 공동의 이해 발견과 바람직한 해결 방향을 제시하려는 논의의 출발점 역할을 수행
- ▶ 한편, 실제로 데이터 경제에서 진행되고 있는 데이터 공유의 상황을 살펴보고, 데이터 경제의 완전한 발현을 위해 더욱 진전되어야 할 과제를 발견하여, 이를 기반으로 향후 관련 조직들의 추진전략을 모색하는 것이 중요
 - 데이터 공유 생태계는 고객만족도와 생산성의 향상과 비용 절감의 효과를 제공
 - 데이터 생태계에 깊게 참여하는 조직들은 향후 5년 동안 그렇지 못한 조직들에 비해 고객만족도 및 생산성 향상과 비용절감 측면에서 연간 수익의 최대 9억 4,000만 달러의 총 재무 효과 발생
 - 데이터 공유 생태계에 참여하기 위해 조직들은 참여 전략 수립-실행-이점의 유지 및 업그레이드의 단계적 추진과 이 과정에서 규제 제도에 대한 선제적 대응 필요
- ▶ 우리나라의 경우 데이터산업 활성화 전략('18.6.), 데이터 3법('20.1.), 디지털 뉴딜정책('20.7.), 그리고 최근 데이터 기본법('21.9.) 마련까지 지속적으로 데이터 산업을 위한 법적 제도를 마련하였으며, 데이터 플랫폼 발전전략('21.6.11) 등을 발표하여 데이터 생태계 조성을 위한 노력을 기울이고 있음
 - 데이터 기본법(데이터산업 진흥 및 이용촉진에 관한 기본법안)을 통해 ▲데이터산업 진흥 기본계획 수립 ▲국무총리 소속 국가데이터정책위원회 설치 ▲데이터 생산·결합 촉진 등을 위한 시책 추진 ▲데이터 품질관리 및 가치평가체계 마련 ▲데이터 유통·거래 체계 구축 ▲데이터 전문인력 양성 ▲중소사업자 대상 지원 등의 주요과제 도출을 통한 데이터 산업 기틀 마련

- 데이터 플랫폼 발전전략 발표하여 분야별 대표 데이터 플랫폼 확충 및 통합연계, 데이터 거래·유통기반강화, 데이터 분석·활용 생태계 조성, 데이터 통합·관리 체계 마련 등 데이터 생태계 조성 방향 제시
- 그러나 국제적으로도 데이터 공유 프레임워크에 대한 현안들이 완전히 정비되지 않았고, 데이터의 생산, 거래 및 활용 촉진을 위해 풀어야 할 과제가 많아, 향후 동 보고서의 내용을 기반으로 그 논의를 더욱 발전시켜 나가야 할 것으로 보임

〈불임〉 부문별 37개 데이터 공유 활용 사례

부문	사례
소매	· 회사 내에서 생성된 거래 및 판매 데이터를 수익화
	· 소비자 거래 및 수수료 데이터 공유
	· 다양한 부문의 파트너와 소비자 행동 데이터 공유
	· 광고 효율성 및 타겟팅 모니터링
소비재	· 브랜드별 판매 데이터 공유
	· 소셜 미디어 및 유사 브랜드 타사 데이터 공유
	· 다른 재배자 및 정부와 농장 농산물 데이터 집계 및 공유
	· 광고 효율성 및 타겟팅 추적
금융서비스	· 사내에서 생성된 거래 및 판매 데이터를 수익화
	· 온라인 구매 및 거래 데이터 공유
	· 조직간 고객 데이터의 원활한 흐름을 위한 동료와의 파트너십
	· 규정 준수/규제 목적을 위한 데이터 공유(예: 신용 위험 데이터)
자동차	· 유지보수 로그, 연비 등 차량의 운행 데이터를 종합적으로 공유
	· 차량의 식명화 및 집계된 텔레마틱 데이터를 지자체, 서비스 제공업체 등과 공유
	· 보험 및 금융 기관과 차량 사용 데이터 공유
	· 브랜드 및 유형별 판매 데이터 공유
	· 모든 공장의 데이터를 결합하여 모든 공급망 파트너 및 통합 파트너가 사용하도록 하는 산업용 클라우드
	· 스마트시티, 주차, 매핑 등의 서비스를 위한 종합 모빌리티 데이터 공유
	· 웹사이트 위험을 줄이기 위한 모범 사례를 정의하기 위해 동료와 데이터 공유
에너지/유틸리티	· 실시간 에너지/물 공급 및 소비 데이터를 공유하여 수요 예측 개선
	· 집계 및 식명화된 고객 데이터 공유
	· 플랜트 및 시설 관리자, 시스템 통합업체, 기술 제공업체간 데이터 공유
	· 빠른 진단을 위한 환자 데이터 공유/접근
헬스케어	· 건강 및 환자 이력 공동화를 위한 데이터 공유/접근
	· 연구 및 개발을 가능하게 하는 데이터 공유/접근
	· 회사 내에서 생성된 거래 및 판매 데이터를 수익화
통신	· 타겟 홍보를 위해 소비재 브랜드와 통화 수익 공유
	· 지역 디렉토리 서비스 제공업체 및 판매자와 집계 트래픽 데이터 공유
	· 소비자 거래 및 수수료 데이터 공유
	· 다양한 부문의 파트너와 소비자 행동 데이터 공유
	· 광고 효율성 및 타겟팅 추적
	· 상품의 모든 단계 모니터링
	· 운영 데이터(자산 성능, 고장 정보, 유지 보수 및 효율성 등) 공유
제조/공급망	· 인바운드 물류 관리(화물 데이터, 공급 일정, 배송 시간 등) 데이터 공유
	· 설비 및 개별 기계의 데이터를 공유하여 생산/입고 설비 관리
	· 새로운 연구 및 신제품 개발에 대한 데이터 및 아이디어 공유
혁신	· 지속 가능한 운영에 대한 데이터(모범 사례, 담수 소비, 누출, 유출 등)의 부품업체/유통업체와 공유
지속가능성	

출처 : Capgemini Research Institute, 2021.07.

참 고 문 헌

- World Economic Forum, Developing a Responsible and Well-designed Governance Structure for Data Marketplaces, 2021.8.
- World Economic Forum, Towards a Data Economy: An enabling framework, WHITE PAPER, 2021.8.
- Capemini Research Institute, Data sharing masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge, 2021.7.



데이터산업 동향 이슈 브리프

| 발행일 2021년 9월 30일

| 발행처 **Kdata** 한국데이터산업진흥원

서울시 중구 세종대로 9길 42, 부영빌딩 8층

| 기획 및 편집 데이터산업본부 산업기획팀

| 문의처 Tel: 02-3708-5363, 5364

ISSUE BRIEF

* 본 지에 실린 내용은 한국데이터산업진흥원의 공식 의견과 다를 수 있습니다.
본 내용은 무단전재를 금하여, 가공/인용할 경우 반드시 출처를 밝혀주시기 바랍니다.