

Insight Report

데이터 캐피탈리즘: 데이터 자본화 이슈





본 저작물은 공공누리 제4유형: 출처표시+상업적이용
금지+변경금지 조건에 따라 이용할 수 있습니다.

↓	요 약	1
	I. 제4차 산업혁명과 데이터 기반 혁신	3
	II. 데이터 기반 혁신의 기존 이슈 분석	8
	III. 데이터 자본화를 위한 기술 및 시장 이슈	12
	참고문헌	19



요 약

▣ 연구의 목적

- 본 보고서에서는 제4차 산업혁명 등 데이터 기반 혁신을 촉진하기 위해 데이터의 가격 평가, 데이터 거래 시장의 구성 가능성에 대해 분석하고자 함

▣ 데이터 자본화의 개념과 필요성

- 데이터 기반 혁신에 관련된 사회경제 및 기술적 이슈들은 다양한데 이러한 문제들의 많은 부분이 데이터가 자본화되지 못하고 있다는 점과 연관이 깊음
 - 본 보고서에서는 ‘데이터 자본화’를 데이터가 독립된 재화로서 시장 참여자들 간에 원활하게 거래되는 것으로 정의
 - 데이터가 정부, 기업, 개인 등 시장 참여자들 간에 적절한 가격으로 원활하게 거래된다면 데이터 기반 혁신의 촉진과 정상적인 시장 작동이 가능할 것으로 전망
 - 데이터의 자본화는 데이터 기반 스타트업의 시장 진입을 촉진할 수 있어 특정 기업의 데이터 독점에 대한 간접적 방어기제로 작동가능

▣ 데이터의 경제적 가치가 가진 특성

- 적시성, 무한가치성, 무형성, 축적효과성, 독점성, 신용가치성 등 데이터 자체의 경제적 가치는 데이터 기반 혁신의 가치와 대립하는 측면이 있어 이에 대한 조정이 데이터 자본화의 핵심적 이슈

데이터가치	설 명	향후 변화 전망
적시성	- 데이터는 사용시점에 민감	- 적시성에 따라 가격 차별화 가능
무한가치성	- 동일한 데이터가 반복적 이용가능	- 복수의 기업이 이용하면 가치 감소
무형성	- 무형자산이기에 가치평가 어려움	- 가치 평가의 난해함은 해결 가능
축적효과성	- 데이터가 연결, 결합되면 가치 증가	- 데이터의 축적효과성 강화
독점성	- 데이터 개방과 가치는 반비례	- 데이터 자본 시장 작동의 장애요인
신용가치성	- 출처의 신뢰도에 따라 가치 변화	- 데이터 양이 신용가치성을 대체, 보완

- 개별 데이터에 대한 가치 및 가격의 결정은 대단히 어려운 과제이며 그 원인은 특허 등 무형자산 가치 평가에서 마주하는 장애요인들과 일정 부분 유사함
- 데이터는 여타 무형자산과 부분적 배타성, 내재적 위험이라는 측면에서 상이하다고 보여지며 이러한 점들이 데이터 자본화에 긍정적으로 작용 가능

경제적 특징	무형자산	데이터
부분적 배타성 및 누출효과	<ul style="list-style-type: none"> - 무형자산은 소유자가 아닌 자가 효익을 누릴 때 배제하기 어려움 - 무형자산의 부분적 배타성으로 인한 누출에서 발생하는 효익이 누출효과 	<ul style="list-style-type: none"> - 데이터는 부분적 배타성, 누출효과는 제한적으로만 나타남
내재적 위험	<ul style="list-style-type: none"> - 무형자산은 획득 과정에서 자산가치 예상이 어려움 - 무형자산 획득비용은 투자가 아닌 비용으로 계상 	<ul style="list-style-type: none"> - 데이터는 활용 목적이 무한히 다양해 가치 평가에 장애요인으로 작용 - 데이터 활용 기술의 혁신으로 가치 평가의 시점이 앞당겨질 수 있음
비시장성	<ul style="list-style-type: none"> - 조직적, 경쟁적 시장 미형성 	<ul style="list-style-type: none"> - 데이터 시장 형성이 미진

▣ 데이터 자본화와 거래 시장 형성을 위한 제언

- 데이터 자산의 배타성과 내재적 위험을 낮추고 다양한 참여자를 유인할 수 있는 데이터 거래 시장의 기술적 구조 마련을 통해 데이터 자본화를 앞당길 수 있음
- 데이터 거래 시장 참여를 촉진할 수 있는 기업으로는 스타트업 등 중소 IT 기업과 금융, 유통, 제조 등 IT 이외의 산업에 속한 업체를 고려할 수 있음
- 데이터 거래 시장은 판매, 대여 등 다양한 거래의 형태를 제공하고 데이터 이전과 활용에 관련된 정보보안, 빅데이터, 기계학습 등의 기술 역량까지 제공해야 할 것
- 이러한 거래 플랫폼은 보안, 클라우드 컴퓨팅, 빅데이터, 기계학습 등의 기술들이 통합적으로 적용되어야 하고 또한 관련 법제도 자체가 따로 마련되어야 할 것

I 제4차 산업혁명과 데이터 기반 혁신

☐ 데이터가 주도하는 제4차 산업혁명

- 2016년 다보스포럼 이후 화제가 된 제4차 산업혁명이라는 개념은 디지털 세계, 생물학적 영역, 물리적 영역에 걸친 기술융합에 기반한 사회경제적 혁신을 예고
 - 제4차 산업혁명의 근간에는 빅데이터, 사물인터넷, 인공지능 등의 IT 기술의 혁신이 추진동력으로 작용하고 있으며 이에 기반한 O2O, 핀테크 등 신서비스 등장과 제조업 등 산업 부문의 효율 개선에 대한 기대가 부각
 - 세 번에 걸친 산업혁명이 인류 발전에 엄청난 기여와 변화를 가져왔듯이 4차 산업혁명 또한 인류가 경험하지 못한 새로운 시대를 가져올 것으로 기대
 - 특히 4차 산업혁명은 세계 주요국이 겪고 있는 경제 성장 정체에 돌파구를 마련해줄 것으로 기대되면서 주요국들의 기술 및 산업정책의 변화가 진행 중
- 제4차 산업혁명, 지능정보사회 등으로 지칭되는 새로운 변화는 데이터를 생산하고 활용하는 IT 기술에 기반하고 있어 향후 사회경제 부문에서 데이터 중요성은 가중
 - 기술적 동인인 빅데이터, 사물인터넷, 인공지능 등은 방대한 데이터를 다루는 기술이며 제품과 서비스의 데이터 기반 혁신이 제4차 산업혁명과 지능정보사회의 실체
 - 알리바바의 마윈은 데이터 자체가 새로운 자원으로 작용하는 DT(Data Technology) 시대가 시작되었음을 주장하며 데이터 기반 혁신을 강조¹⁾
- 제4차 산업혁명이 가져올 주요한 변화 중 하나는 지식창출과 의사결정의 상당 부분이 기계화되고 인간은 보다 고차원적인 전략적 의사결정에 집중하게 된다는 점
 - 센서, 사물인터넷, 3D 프린팅, 의료, 로봇 등 다양한 부문의 기술 혁신과 융합이 전망되는 가운데 가장 두드러진 특징은 빅데이터, 인공지능 등 인간의 사고능력을 일정부분 대체하고 보완할 기술의 혁신과 응용
 - 제4차 산업혁명에서 전망하고 있는 제조업, 의료 등 다양한 부문에서 발생할 성과 향상은 데이터에 기반해 인간 의사결정을 기계화하는 기술적 성과에 의존
 - 지식창출과 의사결정에 직접적으로 기여할 데이터 기반 혁신은 과거 경제·경영 혁신에 비해 효율성보다 효과성 측면에 기여하는 바가 클 것으로 기대²⁾

1) 2016년 10월 알리바바 항저우 원치 대회에서 5가지 미래 트렌드를 언급하며 “모바일과 인터넷이 만들어내는 신기술 변화와 데이터 자체가 새로운 에너지가 되는 신자원 변화에 주목해야 하며 미래에는 기술과 데이터, 혁신만 이용하면 누구나 기회를 만들어 낼 수 있을 것”이라고 주장

2) 과거의 제품과 서비스, 공정, 마케팅 등에 활용된 정보기술들이 주어진 경영전략을 보다 저비용으로 완수하

☐ 데이터 기반 혁신을 가능하게 하는 변화들

- 데이터 기반 혁신을 가능하게 하는 변화들의 첫 번째는 SNS, 웹페이지, 모바일 인터넷, 사물인터넷 등의 활성화로 인한 데이터의 폭발적인 증가
 - IBM은 매일 250경 바이트의 데이터가 생산³⁾되고 있으며 구글의 CEO인 에릭 슈미트는 2010년 컨퍼런스에서 “최근 이틀 동안 생산된 데이터의 양은 2003년 이전 인류가 생산한 정보의 양만큼 많다” 고 주장⁴⁾
 - 센서, 앱 등 기계가 생산하는 데이터 양이 증가하고 인간의 대화와 글, 이미지, 동영상 등 비정형적인 데이터 비중이 증가한다는 점이 데이터 폭증의 특징
 - 비정형적이고 방대한 양의 데이터가 생산되고 있고 그 양은 앞으로 더욱 증가할 것이라는 점은 데이터 관련 혁신 기술의 등장을 요구하고 촉진하는 요인
- 최근 몇 년간 이루어진 딥러닝으로 대표되는 인공지능 알고리즘의 혁신적 성능 향상은 데이터 기반 혁신의 또 다른 촉진 요인
 - 인공지능 기술은 수 십년간 정체되어 있었으나 2012년 Hinton 등이 딥러닝 알고리즘을 혁신하면서 이미지인식, 자연어 번역 등에 있어서 기술적 돌파구가 형성
 - 딥러닝 등 기계학습 알고리즘 혹은 소프트웨어는 다층적 학습망을 통해 방대한 양의 비정형 데이터를 학습하는 기술로서 빅데이터에 기반한 지식⁵⁾ 창출을 수행
 - 빅데이터 분석 기술의 실체는 인간의 문제해결능력 하에서 컴퓨팅 파워를 활용한 통계적 분석인데 딥러닝 이후 인간의 개입 필요성이 감소
- 컴퓨팅 파워, 데이터 저장용량 등에 관련된 하드웨어들의 성능 향상과 가격 하락이 데이터 기반 혁신을 만들어내는 원동력
 - 초기 컴퓨터들에 비해 천문학적 성능향상이 이루어졌으나 가격은 극적으로 하락한 스마트폰이 수십억 대씩 이용되고 있는 것이 현재의 모습
 - 데이터를 수집하는 센서, 메모리, 네트워크 하드웨어 가격이 꾸준히 하락하고 알고리즘과 소프트웨어를 가동시킬 수 있는 CPU, GPU의 성능 향상

는데 초점을 맞추면서 ‘효율’을 추구하는데 도움을 주었다면 새로운 IT기술로 무장한 제4차 산업혁명에서 벌여질 데이터 기반 혁신은 어떤 제품을 만들 것인가, 누구에게, 언제, 어떤 가격으로 판매할 것인가와 같이 기업이 ‘효과’적인 경영활동을 하는데 도움을 제공한다는 점이 차별화된다고 할 수 있음

3) IBM, “Bringing Big Data to the Enterprise,” 2014.8.16.

4) Marshall Kirkpatrick, “Google, Privacy and the New Explosion of Data,” Techonomy, 2010.8.4.

5) 기계학습 알고리즘에 의한 패턴인식, 자동번역 등을 엄밀히 ‘지식’이라 부르는 것은 논란의 여지가 있을 수 있으나 딥러닝 이전의 알고리즘들이 보였던 인지 성능에 비교할 때 두드러진 패턴인식 성능향상과 향후 기대되는 혁신적 응용을 고려할 때 지식 창출이라 지칭하는 것도 큰 무리는 없을 것으로 보임

데이터 기반 혁신의 사례

- 데이터 폭증, 인공지능 기술 혁신으로 인한 데이터 기반 문제해결의 자동화, 하드웨어 성능향상과 가격 하락은 사회경제 체제의 다양한 부문에서 혁신을 유인
- 구글, 바이두, 페이스북 등 IT 거인들은 각각 딥러닝 등 인공지능 기술 분야의 대표적 학자들을 자사에 영입하면서 다양한 서비스에 데이터 기반 혁신 주도
 - 구글은 Hinton을 영입하여 이미지 인식, 이미지 생성, 자연어 번역, 생각벡터 등에서 데이터 기반 혁신의 기술적 변화를 선도
 - 페이스북 또한 딥러닝 혁신가 중 하나인 Lecun을 통해 이미지 인식 기술을 혁신하고 자연어 처리에 기반한 챗봇 서비스 도입 등의 혁신을 예고
 - 바이두는 Andrew Ng을 자사의 인공지능연구센터장으로 영입하면서 딥러닝 기술력에서 구글을 추격하면서 일부 분야에서 구글에 비교 우위를 주장
- 아마존, 넷플릭스 등은 고객 행태 데이터에 기반한 추천 제품 및 서비스의 정확도를 바탕으로 각각 압도적인 시장 지배력을 유지
 - 아마존이 거대한 종합 온라인쇼핑몰, 미디어 사업자로 성장한 주요한 요인 중 하나는 다양한 제품과 콘텐츠에 대한 고객 맞춤형 서비스 제공능력
 - 넷플릭스는 방대한 영상콘텐츠에 대한 고객 취향을 파악하고 맞춤형 콘텐츠를 제공하는 역량을 통해 고객들의 지속적인 재구매를 유인
- 유통, 증권거래, 엔터테인먼트, 치안 등 사회경제 체제 내 다양한 부문에서 데이터에 기반한 혁신적 변화 사례들이 발생 중
 - 월스트리트에서는 물리학자 2천 명이 알고리즘 트레이딩에 참여하며, 미국 주식 시장의 70%가 알고리즘에 의한 기계적 트레이딩
 - 영국의 데이터 분석 회사인 Epagogix는 대본과 구성, 배우 등의 데이터를 분석하는 알고리즘을 통해 영화의 흥행 성적을 예측하여 판매
 - 온라인 쇼핑몰 Stitch Fix는 이용자 취향, 신체 치수 등과 같은 데이터에 기반해 판매 의류를 선정하고 고객은 배송을 받은 이후에 제품 구매 결정
 - 미국 캘리포니아주 Fresno시의 경찰은 소셜 미디어의 투고를 수집, 해석해서 요주의 인물의 위험도를 평가하는 소셜미디어 감시시스템 이용
 - 미국 캘리포니아, 플로리다주는 범죄발생예측시스템 Huntchlab, PredPol를 이용하여 데이터 기반 범죄발생 예상지역을 특정하고 경찰 인력을 배치

▣ 주요 국가의 데이터 기반 혁신 대응 정책

- 글로벌 저성장 기조, 생산성 하락 등을 겪고 있는 주요국들은 제조업 등 주요 산업 성장 전략, 공공부문 운영효율 향상 정책으로서 데이터 기반 혁신 정책을 마련
 - 미국은 첨단제조업파트너십(AMP) 프로그램, 독일은 Industrie 4.0, 일본은 산업부흥전략, 중국은 중국제조 2025를 발표하는 등 국가별 대응 전략을 추진
- 데이터 기반 혁신을 위한 미국의 정책적 대응은 제조업의 부흥에 초점을 맞추고 있으며 이 과정에서 GE 등의 민간기업과의 공조가 두드러짐
 - 기업 수준에서 발생하고 있는 데이터 기반 혁신 사례의 대부분이 미국에 집중된 만큼 관련 기술과 인력 측면에서 미국의 역량이 가장 앞선 것이 현실
 - 사실상 제4차 산업혁명이라는 개념 자체가 미국이 선도하는 데이터 기반 혁신을 추격하려는 유럽 경제권역의 자각이라 할 수 있음
 - 구글, 페이스북, 아마존 등 인터넷 기반 기업의 기술 및 서비스 역량이 두드러진 가운데 GE가 주도하는 제조업 부문 데이터 기반 혁신 움직임이 부각
 - data.gov로 대표되는 공공데이터 개방 움직임을 선도한 것 역시 미국으로서 공공과 민간 양 부문에서 미국의 데이터 기반 혁신 능력이 독보적
- 유럽에서 데이터기반 혁신의 중요성을 빠르게 인지하고 정책적으로 대응한 것은 독일로서 industrie 4.0을 중심으로 제조업의 데이터 기반 혁신을 준비
 - 독일은 제4차 산업혁명, industrie 4.0 등의 개념을 유럽의 아젠다로 확장해가면서 빅데이터, 사물인터넷, CPS⁶⁾ 등 IT 기술에 기반한 제조업 혁신을 추진
 - industrie 4.0은 현실과 동떨어진 표준 선정과 경직된 협력체계 등으로 난항을 겪었으나 2015년 업계 의견을 받아들여 보다 유연한 제조업 혁신 전략으로 재탄생
- 중국은 제조업 부문에서 저임금 기반 성장의 한계를 돌파하기 위한 대책으로서 중국제조 2025 등의 전략을 표방
 - 데이터 기반 혁신의 일부 개념을 포함하고 있으나 현실적으로는 저임금 기반 단순 조립을 넘어선 고부가가치 제조업 역량을 확보하기 위한 정책으로 판단됨
- 일본의 산업부흥전략은 데이터 기반 혁신을 제조업, 농업, 의료, 금융 등 전 산업부문에서 활용하자는 전방위적 산업 전략으로서 로봇에 대한 강조가 두드러짐

6) CPS는 Cyber Physical System의 약자로 사물인터넷, 로봇 등의 기술을 통해 디지털 정보 기술과 물리적 세계의 결합을 시도하는 기술적 노력

- 기술 전략보다는 산업 응용 범위에 대한 강조가 눈에 띄는 산업 전략으로서 일본 정부의 통상적인 산업 전략에 IT 기술과 데이터에 대한 강조가 추가된 형태
- 딥러닝, 빅데이터 등 IT기술에 있어서 일본의 혁신 역량이 눈에 띄지 않는 것이 현실이며 다른 주요 국가들과의 정책적 차별점은 로봇 기술에 대한 강조
- 로봇 기술 강조의 이유는 알고리즘, 소프트웨어 보다는 정밀 기계 및 부품 분야의 장점을 데이터 기반 혁신에 접목하려는 현실적 전략으로 판단됨

II 데이터 기반 혁신의 기존 이슈 분석

▣ 혁신을 둘러싼 다양한 논점들 : 기술 혁신, 위험과 효용, 시장 왜곡

- 데이터 기반 혁신은 새로운 지식과 가치의 창출, 새로운 제품과 서비스 탄생, 생산과정과 조직 운영효율의 혁신 등 다양한 영향력을 가진 만큼 관련된 문제도 다양
 - 딥러닝, 빅데이터 등 데이터 수집 및 처리를 위한 IT기술 혁신이 선결과제로서 미국 중심으로 세계적 기술개발 경쟁이 진행 중
 - 기계화된 지식 창출과 의사결정이라는 새로운 가능성의 이면에는 보안, 프라이버시, 인간이 배제된 의사결정 등의 문제점들이 존재
 - 기업과 기업, 개인과 기업 간 균형발전의 측면에서 거대 기업의 시장지배력 강화, 정보 비대칭 등 시장 왜곡의 가능성이 대두
 - 기술 혁신 성과에 대한 합리적 기대수준과 위험 조정에 대한 의견이 각계에서 쏟아져 나왔고 비교적 늦게 시장 왜곡에 대한 경각심이 제기되는 추세

▣ 알고리즘, 알고리즘, 알고리즘, 그리고 인간

- 데이터 기반 혁신의 전제조건인 딥러닝, 빅데이터 등 기술 혁신 경쟁이 미국 중심으로 진행
 - 딥러닝, 빅데이터 기술에 있어서 기초 기술 혁신 역량, 산업 내 성공사례가 미국의 거대 IT 기업에 집중되어 있어 타 국가들 및 기업들이 경쟁적으로 추격
 - 바이두가 실리콘 벨리 내에 인공지능 연구거점을 마련하는 등 미국 외 기업과 학계에서 기술 추격 노력을 계속하고 있으나 여전히 기술 격차는 존재
 - ※ 2016년 바이두가 이미지 인식 성공률에서 구글을 앞질렀다고 발표했으나 구글은 이미지와 상황의 매칭, 다국어 번역 등에서 새로운 딥러닝 기술 혁신을 선보여 기술 격차는 여전
 - ※ ETRI 등 국내 연구자들도 이미지 인식, 인간과 알고리즘의 퀴즈 대결 등에서 일정 수준의 성과를 보이고 있으나 세계 최고 수준에는 미달
 - 2015년 Hinton은 인간의 생각 자체를 모델링하는 생각벡터(thought vectors) 개념을 기계학습에 접목함으로써 보다 인간에 가까운 인공지능의 가능성을 제기
 - ※ 기존의 딥러닝은 하나의 단어, 이미지, 상황에 대해 잘 훈련된 기계적 반응을 도출하는데 성공한 것이며, 생각벡터는 동일한 단어, 이미지, 상황에 대해서도 다양한 맥락을 종합적으로 인지하는 인간의 두뇌 활동을 기계적으로 모사하려는 시도로서 과거 context awareness라 칭했던 기술 부문을 포함한 딥러닝 기술의 활용 가능성을 대폭 증가시킬 것으로 기대

- 데이터 기반 혁신에서 알고리즘에 대한 강조는 지나칠 수 없으나 기업과 조직 측면에서는 새로운 정보 및 지식의 생산, 소비 체계에 적합한 인적 자원의 확보에 대한 강조가 꾸준히 제기
 - 데이터 사이언티스트라 불리우는 새로운 인적 역량 확보에 대한 요구는 빅데이터, 딥러닝 기술의 확산과 더불어 지속적으로 강화
 - 기업과 조직의 문제에 빅데이터, 딥러닝 기술을 활용하기 위한 필수적 자원으로 통계학, 코딩에 관한 지식과 종합적 문제해결 역량을 함께 갖추어야 함
 - 근래에는 조직 전체가 데이터 기반 혁신을 수용하기 위한 준비로서 데이터 기술에 대한 이해 역량을 갖춘 중간관리자 계층에 대한 요구 또한 제기

▣ 위험과 효용, 대립에서 조화로

- 데이터 기반 혁신이 가져올 효율과 효과 증가에 대한 기대의 이면에는 프라이버시, 보안, 기계화된 지식 창출과 의사결정에 내재된 위험에 대한 우려가 존재
- 프라이버시와 정보보안에 대한 경고는 데이터 기반 혁신에 관련되어 제기되는 가장 빈번한 쟁점으로 개인과 기업, 정보보호와 정보효용에 걸친 근본적 대립
 - 기업이 개인 정보를 수집, 분석함으로써 얻는 수익이 개인에게 지불되는 대가 또는 혜택에 비해 지나치게 크다는 인식이 데이터 수집 및 활용 확산에 걸림돌
 - 옵트인, 옵트아웃 등과 같은 개인정보 제공 선택권이 마련되어 있으나 기업이 복잡한 약관 등을 통해 사실상 개인정보를 제공을 강제하고 있는 것이 현실
 - 프라이버시 보호를 위한 데이터 비식별화 등의 기술적 대안이 제안되고 있지만 빅데이터에서 나타나는 모자이크 효과⁸⁾로 인해 비식별화의 효용이 손상
 - 해킹 등 정보누출은 데이터 기반 혁신의 가치를 손상시킬 수 있는 위협이며 개별 데이터 누출이 모자이크 효과로 인해 증폭된 위험을 낳을 가능성 존재⁹⁾
 - 각 조직별 보안 대책뿐만 아니라 모자이크 효과 발생 위협에 대한 종합적 보안 대책이 정보 생태계의 인프라로서 마련되어야 할 것¹⁰⁾

7) 본 저자의 개인적인 추측으로는 생각벡터, 특히 확장 가능한 생각벡터의 도입은 메타 학습, 즉 기존 학습 내용에 기반한 학습, 학습과정 자체에 대한 학습이 가능한 딥러닝으로의 확장 가능성을 여는 중요한 기술적 시도로 보임

8) 모자이크 효과는 개인정보가 없거나 비식별화 처리를 거친 데이터라 하더라도 충분히 많은 양의 데이터가 통합될 경우 다시 개인정보 추출이 가능해지는 현상을 가리키는 용어. 미국 data.gov에 공개된 공공정보들을 통합, 분석했을 때 주요 정부기관 간 물류량과 같은 민감한 정보가 파악가능해지는 현상이 발견되어 모자이크 효과는 개인정보 뿐만 아니라 국가 안보에도 영향을 미치고 있음

9) 금융회사 A에서 누출된 정보와 온라인 쇼핑몰 B에서 누출된 정보를 온라인 상에 공개된 정보와 결합하여 분석을 시도할 경우 모자이크 효과로 인해 대단히 민감한 정보의 유출이 가능해질 위험이 존재

10) 최근 국내에서 시행되고 있는 개인 PC 내 주민등록번호, 신용카드정보 등의 저장 금지 규제가 초보적인 정

- 기계화된 지식 창출과 의사결정에서 생산되는 결과물의 신뢰성, 활용에 따른 위험 발생에 대한 인간의 판단을 가능하게하기 위해 시스템의 투명성 확보가 필요
 - 딥러닝 학습 과정은 사실상 인간이 이해하는 것이 불가능하며 학습된 신경망이 산출하는 결과의 전체 집합을 검증하는 것 또한 불가능한 것이 현실¹¹⁾
 - 구글은 2015년 이미지 인식에 사용된 딥러닝 알고리즘의 중간 결과를 살펴보기 위해 시도했으나 그 결과 관측된 것은 이해불가능한 몽환적 이미지의 집합¹²⁾
 - 이해 불가능한 과정을 거쳐 산출되는 지식과 의사결정은 경험적으로 유용하더라도 기업과 조직의 생존, 인명 등이 달린 문제에 활용하는 것은 위험
 - 강한 인공지능에 대한 각계 전문가들의 우려도 큰 맥락에서 이해 불가능한 지식 시스템에 내재된 위험을 경고
- 데이터 기반 혁신을 둘러싼 위험과 효용의 대립은 다양한 법제도와 정보기술을 통해 해소되어야 할 것이며 기본적 해결 방향은 민주적 절차에 따른 사회경제 체제 전체의 효용 최대화

■ 시장을 변화시키되 왜곡시키지 않는 데이터 기반 혁신

- 데이터 기반 혁신에 따른 시장 왜곡의 위험은 크게 거대 데이터 기업에 의한 시장 독점과 소비자 종속, 정보비대칭으로 인한 시장 효율 저해 등
- 방대한 데이터를 축적한 기업에 의해 시장 진입장벽이 형성되면서 사실상의 시장독점이 발생하고 이용자가 종속될 위험이 존재
 - Google은 검색이력 데이터를 활용하여 검색 정확성을 30% 이상 증가시키고 있다고 알려져 있는데 이는 Bing 등 후발 서비스가 극복하기 어려운 경쟁 우위¹³⁾
- ※ Google의 PageRank 알고리즘 등은 학술 논문을 통해 공개되어 있으며, 딥러닝 알고리즘, 개발 툴킷 등도 논문과 공개 SW 등의 형태로 제공되고 있으나 검색이력 데이터는 비공개¹⁴⁾
- 독점적 데이터와 지식창출 알고리즘을 동시에 소유한 기업은 데이터-지식 독점자로서 강력한 시장지배력을 획득 가능
 - Google의 경우 이용자 유입, 데이터 축적, 검색엔진 알고리즘 진화 등이 일련의

보생태계 내 정보누출 위험에 대한 방어책으로 볼 수 있음

11) 예를 들어 구글이 선보인 자연어 번역 알고리즘에 대해 모든 조합 가능한 문장을 입력하고 그 결과를 검증하는 것은 사실상 불가능

12) 구글은 이러한 몽환적 이미지들이 기계학습이 새로운 패턴을 생성해 낼 가능성으로 해석했으나 여전히 인간이 딥러닝 학습 과정을 이해하거나 간섭하는 것은 어렵다고 판단됨

13) Kira Radinsky, "Data Monopolists Like Google Are Threatening the Economy", 2015.03

14) http://deeplearning.net/software_links

선순환 구조를 형성함으로써 지구적 시장지배력을 확보¹⁵⁾

- Watson 기반 컨설팅 서비스를 판매 중인 IBM, 정밀농업 컨설팅 서비스를 시작한 GE 등도 산업 데이터 축적에 따라 데이터-지식 독점자로 변화 가능
- 유럽 각국은 Google을 데이터 독점자라 공격하고 있는데 이는 검색엔진의 시장 지배력이 온라인 광고 등 서비스로 전이하고 있다는 해석에 기반¹⁶⁾
- 특정 업체가 축적한 데이터와 지식창출 알고리즘을 독점적 지위의 원천으로 보고 이를 규제할 것인지에 대해서는 논의 필요
- 데이터를 축적한 기업이 개인 고객을 대상으로 정보비대칭에 기반한 가격 차별화, 서비스 차별화를 시행할 경우 이에 대한 방어 기제가 미비
 - 금융회사가 고객의 대출 이력, 공공데이터 등에 기반하여 고객의 미래 재무 상태에 대한 예측정보를 축적할 경우 개인에 대한 강력한 red-lining¹⁷⁾ 발생
 - 이러한 정보비대칭 우려는 데이터 기반 혁신이 진일보한 상황에서 발생 가능하지만 발생하기 시작한다면 그 시점에서는 실태파악이 어려워 규제 마련이 곤란
- 데이터 기반 기업의 시장 독점, 정보비대칭 등에 대한 법제도와 규제는 기술과 기업 변화 속도에 뒤쳐질 것으로 전망되기에 현 시점에서 규제 마련을 위한 공감대 형성이 중요

15) Google은 한국 등 특정 지역을 제외하면 검색엔진과 인터넷 광고 부문에서 세계적 시장지배자

16) The Huffington Post, "Google's remarkable monopoly on data intrigues Europeans", 2015.07

17) 은행이 대출 신청자의 재무 상태를 평가하고 예측하는 것은 자연스러운 경영활동이지만 데이터에 기반한 미래 예측의 신뢰성이 비약적으로 향상된다면 특정한 고객집단은 대출 등 금융서비스에 대해 원천적으로 이용이 불가능해지는 사태가 벌어질 수 있으며, 이 때 고객은 자신이 어떤 이유로 대출심사에서 탈락되었는지를 파악하기조차 어려울 수 있음. 예를 들어, 흑인에 대한 대출 거부율이 두드러진다면 금융규제당국이 간섭할 명분이 만들어질 수 있지만 표면적으로 쉽게 파악하기 어려운 데이터 패턴에 의해 대출여부를 결정한다면 특정 고객그룹에 대한 원천적인 대출 불허가 시행되지만 규제당국과 개인고객은 실태 파악이 곤란

III 데이터 자본화를 위한 기술 및 시장 이슈

III 데이터 자본화의 개념과 필요성

- 앞서 언급된 데이터 기반 혁신 관련 이슈들은 대단히 다양하지만 이러한 문제들의 많은 부분이 데이터가 자본화되지 못하고 있다는 점과 연관이 깊음
 - 본 보고서에서는 ‘데이터 자본화’를 데이터가 독립된 재화로서 시장 참여자들 간에 원활하게 거래되는 것으로 정의
 - 데이터가 정부, 기업, 개인 등 시장 참여자들 간에 적절한 가격으로 원활하게 거래된다면 데이터 기반 혁신의 촉진과 정상적인 시장 작동이 가능할 것으로 전망
- 데이터 자본화의 핵심적인 요소는 데이터라는 새로운 무형 자산에 대한 적정 가치 평가 체계와 가격 결정이 가능한 거래 시장의 형성
 - 데이터 가치 평가와 가격 결정이 어려운 상황에서는 데이터 거래가 일부 시장 참여자 간에 이루어지게 되고 데이터 활용에 따른 혁신이 제한될 것으로 예상
 - 또한 데이터 가치와 가격에 대한 시장 내 시그널이 부재하면 데이터 생산, 수집, 저장, 처리, 분석, 활용을 담당하는 개인, 기업의 소유권과 통제권에 대한 쟁점 해결이 복잡한 규제에 의존하게 될 것
 - ※ 데이터 혁신 관련 프라이버시와 보안 문제, 위험과 효용의 가치 평가 문제는 개인과 기업의 데이터 소유권과 통제권 문제로 데이터 가치와 가격이 명확하게 결정될수록 완화될 것임¹⁸⁾
- 데이터의 자본화는 데이터 기반 스타트업의 시장 진입을 촉진할 수 있어 특정 기업의 데이터 독점에 대한 간접적 방어기제로 작동가능
 - 구글 등이 가진 데이터 관련 서비스 시장에 대한 강력한 지배력은 데이터 확보에 있어서 선순환 구조를 구축하여 사실상의 데이터 독점력을 가졌다는 점에 기반¹⁹⁾
 - 선발 업자가 구축한 데이터 확보의 선순환 구조에 의해 형성된 시장 진입장벽은 후발 업자들이 자본화된 데이터 거래를 통해 일정 부분 극복가능해질 수 있음
 - 데이터 기반 혁신 시장에서 나타나는 기업의 시장 독점은 시장 확정에 대한 쟁점 정리가 어려워 규제가 곤란한데 데이터 자본화는 독점력에 대한 시장 내 대항수단을 마련할 수 있다는 점에서 의의를 가짐

18) 데이터의 가격이 명확해지면 프라이버시, 보안, 기계화된 지식창출의 효용과 위험이 시장 내 가격이라는 동일한 기준으로 판단될 가능성이 높아지면서 시장 논리에 의해 쟁점 해결이 가능해질 것

19) 강력한 자본력과 인적 자산을 가진 Microsoft의 검색엔진 Bing이 구글의 검색엔진 시장 점유율을 뺏지 못하는 현상은 데이터 기반 혁신 시장에서 데이터 확보 역량이 실질적인 시장 지배력의 근간이라는 점에 대한 간접적 증거

- 주요국들이 공공 데이터 개방을 앞 다투어 시행하고 있기에 데이터 자본화 쟁점은 민간 데이터의 거래에 집중
 - 현재 사회적 효용의 제고, 혁신 유인 등의 명분에 따라 공공데이터는 무료로 개방되는 추세이며 이러한 현상은 데이터 거래 시장 형성의 씨앗으로 작용
 - 영국의 컨설팅 회사인 Lateral Economics는 영국의 공공 데이터 개방 전략과 관련하여 2016년 공공 데이터가 무료로 개방될 때 국가 총생산에 기여하는 부분이 최대화된다고 보고²⁰⁾
 - ※ 무료에서 유료로의 전환, 프리미엄(freemium) 가격 체계의 적용, 데이터 사전 처리 수준에 따른 가격 차별화 등에 대한 검토에 기반하여 완전 무료 개방 원칙이 공공의 효용 증대에 가장 효과적이라는 결론을 제시
 - 민간 데이터의 거래는 M&A, 기업 간의 포괄적 거래 계약의 일부로서 발생하는 경우가 대부분이며 최근 데이터 브로커에 의한 데이터 거래가 차츰 증가하는 추세
 - 민간 데이터의 거래는 사례가 제한적이며 가격과 가치의 평가가 불투명한 경우가 대부분이어서 데이터 자본화 수준이 아직은 낮은 상태

▣ 데이터의 경제적 가치와 자본화 가능성

- 적시성, 무한가치성, 무형성, 축적효과성, 독점성, 신용가치성 등 데이터 자체의 경제적 가치는 데이터 기반 혁신의 가치와 대립하는 측면이 있어 이에 대한 조정이 데이터 자본화의 핵심적 이슈
 - 일반 재화와 비교할 때 데이터 자체는 재화로서 적시성, 무한가치성, 무형성, 축적효과성, 독점성, 신용가치성 등의 특징을 보유
 - 빅데이터, 기계학습 등의 기술 발전과 더불어 적시성, 축적효과성 등이 주요한 데이터 가격 결정 요인으로 작용할 것으로 보임
 - 데이터의 폭발적 증가와 비정제 데이터에서 지식을 생산할 수 있는 기술적 역량의 혁신 추세를 볼 때 향후 신용가치성이 가진 중요도는 일정 부분 감소할 것
 - 데이터는 무형자산이지만 데이터로부터 지식을 생산하는 기술이 자동화되어갈 수록 데이터의 가치에 대한 예측은 더욱 쉬워질 것으로 보임
 - 데이터가 기업 내부에서 활용되는 경우에는 무한가치성이 유지될 수 있으나 데이터 거래를 통해 데이터에 기반한 유사 제품, 서비스가 경쟁적으로 출현할 때는 데이터의 무한가치성은 손상될 것

20) Lateral Economics, "Permission granted: The economic value of data assets under alternative policy regimes", 2016.3.

표 1 | 데이터의 경제적 가치요인과 데이터 자본화로 인한 변화 전망

가치요인	설 명	데이터 자본화로 인한 변화 전망
적시성	<ul style="list-style-type: none"> - 데이터는 일반 재화에 비해 사용시점에 민감 - 데이터의 경제적 가치는 특정 시기를 지나면 현격하게 감소 	<ul style="list-style-type: none"> - 실시간 데이터 흐름이 강화되는 기술 발전 추세에서 데이터 적시성의 중요성은 계속 증가 - 자본화된 데이터 시장에서 적시성에 따라 가격 차별화 가능
무한가치성	<ul style="list-style-type: none"> - 일반 재화는 소유의 개념이 분명하고 단일 가치를 지닐 수 있으며 사용 이후에는 소멸되거나 다른 재화의 가치로 전환 - 데이터는 사용의 개념으로 동일한 데이터가 가치 소멸없이 무한히 반복적으로 이용가능하며 이 과정에서 경제적 가치는 무한히 증가 가능 	<ul style="list-style-type: none"> - 기술 발전은 동일한 데이터의 활용을 확대함으로써 데이터 가치를 증가시킴 - 데이터 자본화는 복수의 기업이 동일 데이터로 유사하게 이용할 경우 데이터의 가치는 현격하게 감소
무형성	<ul style="list-style-type: none"> - 데이터는 무형자산으로서 실제 가치에 대한 인지가 어렵거나 평가 절하되는 경우가 빈번 	<ul style="list-style-type: none"> - 데이터 자체는 무형성을 유지 - 기술 발전에 따라 데이터 활용 과정의 기계화 비중이 높아지고 데이터가 특정 제품, 서비스로 전환되는 과정이 짧아지면 데이터 무형성으로 인한 가치 평가의 난해함은 일정 부분 완화될 수 있음
축적효과성	<ul style="list-style-type: none"> - 데이터는 소량으로 산재되어 있을 때보다 동일 저장소에 대규모로 축적되어 있을 때 가치 극대화 - 다양한 데이터가 연결, 결합되면서 개개의 가치 합 이상의 새로운 가치 창출이 가능 	<ul style="list-style-type: none"> - 기술 발전과 더불어 데이터의 축적효과성 강화
독점성	<ul style="list-style-type: none"> - 동일한 내용을 가진 데이터가 복수로 존재할 경우 데이터 가치는 감소 - 데이터 가치는 데이터 개방 정도에 반비례 	<ul style="list-style-type: none"> - 데이터 생태계 전체의 성장과 데이터 독점자의 수익 증대 간에 대립 관계가 심화
신용가치성	<ul style="list-style-type: none"> - 동일 정보를 가진 데이터라 할지라도 출처 신뢰도에 따라 가치 차이 존재 - 개인이 측정한 온도와 기상청이 측정한 온도가 비록 같더라도 사용자는 기상청 데이터의 가치를 높게 평가 	<ul style="list-style-type: none"> - 기술 발전과 데이터 수요 폭증 추세로 인해 충분한 양의 데이터는 데이터 자체가 가진 신용가치성을 일정 부분 보완

- 데이터의 독점성은 개별 기업 입장에서는 강화해야 할 가치이지만 데이터 기반 혁신 생태계 전반의 성장을 고려하는 시장 정책의 관점에서는 독점성 완화에 대한 요구가 존재
- 최근 기술혁신이 빠르게 진행되고 있는 빅데이터, 기계학습 기술은 데이터의 양이 충분히 방대하다면 데이터의 신용가치를 일정 부분 보완할 수 있음을 시사

- 개별 데이터에 대한 가치 및 가격의 결정은 대단히 어려운 과제이며 그 원인은 특허 등 무형자산 가치 평가에서 마주하는 장애요인들과 일정 부분 유사함
 - 데이터는 특허, 상표권, 노하우 등과 같이 무형자산으로 분류되는데 무형자산 가격 결정은 공정거래감시, 조세 등의 관점에서 주요 연구대상 중 하나이지만 현재까지 명확한 가치 결정 기준은 미비
 - ※ 특허, 상표권 등의 무형자산에 대한 국제 거래가 빈번해지면서 OECD, EU 집행부, 미국 정부 뿐만 아니라 국내에서도 무형자산 가치 평가 및 정상가격 결정에 대한 법제도, 재무회계기준 등의 마련을 위한 노력이 진행 중이지만 2016년 현재까지 명확한 기준을 설정하는데 실패
 - 특허, 상표권, 노하우 등의 무형자산은 재무적 가치가 평가되는 자산으로 인지되지 않으며 획득에 투입된 자원은 기업의 비용으로 분류하는 것이 국제적 현실
 - 무형자산은 부분적 배타성과 누출효과, 내재적 위험, 비시장성 등의 특징으로 인해 재무적 가치 평가가 어려운데 이러한 특징은 데이터도 일정 부분 공유

표 2 | 무형자산과 데이터의 경제적 가치 특징 비교

특징	무형자산	데이터
부분적 배타성 ²¹⁾ 및 누출효과	<ul style="list-style-type: none"> - 무형자산은 소유자가 아닌 자가 효익을 누릴 때 배제하기 어려움 - 노하우를 갖춘 피고용인이 퇴직하여 다른 기업에서 과거에 습득한 노하우를 활용할 때 이를 배제하는 것에는 현실적 한계가 존재 - 무형자산의 부분적 배타성으로 인한 누출에서 발생하는 효익이 누출효과 	<ul style="list-style-type: none"> - 디지털데이터는 실물 저장기기를 가지며 정보보안 기준이 있음 - 따라서 노하우 등 무형자산에서 발생하는 부분적 배타성, 누출효과는 제한적으로만 나타남
내재적 위험	<ul style="list-style-type: none"> - R&D 활동, 특허 획득, 브랜드 구축 등과 같은 무형자산의 획득 과정에서 자산으로서의 가치를 예상하기 어려움 - 따라서 무형자산의 획득 과정에서 소요된 자원은 기업의 비용으로 처리하는 것이 국제적인 재무 기준 	<ul style="list-style-type: none"> - 데이터도 생성 초기에 미래 가치를 예상하기 어렵고, 처리·통합·분석을 거치면서 가치가 점진적으로 명확해짐 - 특허, 노하우, 상표 등이 생산활동, 비즈니스 모델 등 특정 가치창출 활동에 직접적으로 연계되어 있는 반면, 데이터는 활용 목적이 거의 무한히 다양해질 수 있다는 점에서 차이가 있음 - 따라서 데이터는 무형자산으로서의 가치가 명확해지는데 있어서 타 무형자산에 비해 제약이 있음
비시장성	<ul style="list-style-type: none"> - 노하우, 상표, 특허 등 무형자산에 관한 조직적이고 경쟁적인 시장이 형성되기 어려움 	<ul style="list-style-type: none"> - 데이터 또한 현재로서는 시장 형성이 미진

21) 유가증권, 부동산 등 유형자산은 재산권이 명확하고 소유자의 배타적 효익 보장

- 데이터는 여타 무형자산과 부분적 배타성, 내재적 위험이라는 측면에서 상이한 점이 있다고 보여지는데 이러한 점들이 데이터 자본화에 긍정적으로 작용 가능
- 정보보안 기술을 통해 데이터의 부분적 배타성을 감소시킬 수 있고 이에 따른 데이터 자산의 배타적 효익 보장은 데이터 거래 가능성이 증가할 수 있음
- 무형자산의 내재적 위험은 자산 획득 노력 자체에 대한 유인을 줄이기도 하는데 데이터의 경우 생산 및 저장 비용이 지속적으로 감소하고 있어 최소한 데이터 획득에 대한 유인의 감소는 크지 않음
- 또한 데이터의 활용 부문이 데이터 자체에 의해 크게 제약받지 않는다는 점은 데이터 자산의 범용성을 강화해 거래 가능성을 높인다고 할 수 있음
- 경제적 가치로서의 무형자산과 데이터의 특징 비교를 종합해보면 시장 형성이 어렵다는 점은 공통적이나 부분적 배타성과 내재적 위험 측면에서는 데이터가 여타 무형자산에 비해 자본화 가능성이 높다고 할 수 있음

■ 데이터 거래의 실제 사례와 향후 변화 전망

- 데이터 가치 평가와 가격 결정의 어려움에도 불구하고 민간 데이터의 거래가 제한적이거나 M&A, 포괄적 계약 등의 형태로 발생하고 있음
- 민간 데이터 거래의 대표적 사례는 인터넷 서비스 기업들의 M&A, 개인의 IT 서비스 이용에 따라 부수적으로 발생하는 개인정보 제공 등
 - ISP, 구글 등의 서비스를 이용하는 개인이 명시적으로 혹은 암묵적으로 개인신상 정보와 이용패턴 등의 데이터를 기업에 제공
 - 서비스 이용에 따른 부수적 개인정보의 제공은 대부분 명시적 대가 지불이 없으며 이는 거대 기업에 의한 프라이버시 침해 우려로 전환
 - ※ 미국 FCC는 2015년 Communication Act의 섹션 221에서 ISP가 개인의 명시적 허락없이 개인정보를 수집하는 것을 금지하는 조항을 명문화했으나 ISP를 제외한 구글 등 서비스 기업들의 정보 수집을 제한하는 법적 움직임이 아직 없음
 - 페이스북은 왓츠앱을 190억 달러에 인수했는데 이는 왓츠앱 가입자 당 약 30달러의 대가를 지불한 것으로 해석가능하며 모든 유형, 무형자산에 대한 포괄적 매입이기는 하지만 데이터에 대한 거래로 볼 수 있음
 - ※ 성장 잠재력이 확인된 신생 기업을 인수하는 거대 IT 기업들은 평균적으로 가입자 당 30\$ 수준의 인수비용을 지불하는 것이 통상적

- 인터넷 서비스 기업들에게 있어서 고객 데이터, 서비스 시행에 따라 발생한 데이터는 그 자체로 소중한 자산이기는 하지만 데이터와 비즈니스 모델이 유기적으로 얽혀있기에 데이터 거래보다는 기업 자체에 대한 M&A가 빈번
- 최근 점차 사례가 늘어나고 있는 데이터 브로커를 통한 데이터 거래는 데이터의 품질, 데이터의 다양성 측면에서 아직까지는 미성숙
 - 성별, 나이, 거주지 등 간단한 신상정보와 특정 제품, 서비스에 대한 취향 정보를 담은 데이터가 주로 거래되며 그 가격은 개인 데이터 당 0.0007달러 수준²²⁾
 - 거래 데이터 중에서 최고 가치를 가진 것 중 하나인 개인의 병력 데이터는 일인당 30센트 수준으로 알려짐²³⁾
 - 데이터 브로커에 의한 데이터 거래는 대개 판매 형태로 이루어지며 주요 데이터 공급자는 데이터 전문 수집 업체들
 - 데이터 브로커 시장에서 거래되는 데이터의 다양성, 신뢰성, 적시성, 규모는 구글 등 거대 IT 기업이 수집, 보유하고 있다고 보여지는 데이터에 비해 극히 취약
- 향후 데이터 거래는 현재 양상과 크게 달라지기 어려운 것이 현실이기에 데이터 기반 혁신의 촉진을 위해서는 데이터 거래 구조의 혁신이 절실
 - 개인 수준에서는 현재의 IT 서비스 이용행태를 변화시킬 유인이 미미하고 거대 IT 기업들이 내부 데이터를 거래 시장에 내놓을 유인이 없음
 - 데이터 브로커 시장은 데이터 공급자들의 성장이 전제되어야 하는데 현재 시장에 참여중인 데이터 공급자들만으로는 방대한 데이터 공급이 어려움
 - 따라서, 다양한 데이터 공급자들이 데이터 거래 시장에 참여하도록 하는 촉진책이 요구

■ 데이터 자본화를 위한 준비

- 데이터 자산의 배타성과 내재적 위험성을 낮춤과 동시에 다양한 참여자를 유인할 수 있는 데이터 거래 시장의 기술적 구조 마련을 통해 데이터 자본화를 앞당길 수 있음
 - 데이터 자산의 가치 평가 가능성을 높이기 위해서는 데이터 자산의 배타성과 내재적 위험성을 낮추기 위한 정보보안, 빅데이터, 딥러닝 등의 기술 혁신에 집중
 - 구글 등 거대 IT 기업은 데이터 거래 시장에 참여하기에는 자신들의 독점적 데

22) Tech Crunch, "What's The Value of Your Data?", 2015.10.13.

23) Tech Crunch, "What's The Value of Your Data?", 2015.10.13.

- 이터 가치에 비해 데이터 거래로 얻을 수 있는 효용이 적음
- 데이터 거래 시장 참여를 촉진할 수 있는 기업으로는 스타트업 등 중소 IT 기업과 금융, 유통, 제조 등 IT 이외의 산업에 속한 업체를 고려할 수 있음
 - 중소 규모의 IT 기업들은 데이터 독점보다 데이터 거래를 통해 얻을 수 있는 효익이 더 클 수 있고 특히 이들이 연합해 거대 IT 기업과 유사한 수준의 서비스를 부분적으로나마 제공할 가능성이 있음
 - IT 이외 산업에 속한 업체들 또한 최근의 데이터 기반 혁신 성과를 활용하기 위해서는 다른 업체의 데이터를 통합해 활용할 필요가 있기에 시장 참여 가능
- 데이터 거래 시장은 판매, 대여 등 다양한 거래의 형태를 제공하고 데이터 이전과 활용에 관련된 정보보안, 빅데이터, 기계학습 등의 기술 역량까지 제공해야 할 것
- 적시에 다양한 데이터에 접근, 활용하기 위해서는 다양한 라이선싱 계약이 필요하고 거래 지불 형태 또한 사전지불, 사후지불, 성과비례 가격, 프리미엄(freemium) 가격 등의 대가 지불 구조 지원이 필요
 - 데이터 자산의 내재적 위험을 줄이고, 적시성, 추적효과성을 극대화하기 위해서는 거래 중개뿐만 아니라 데이터 저장, 큐레이션, 분석 소프트웨어 등이 통합된 거래 플랫폼²⁴⁾이 요구
 - ※ 다양한 데이터를 실시간으로 통합하고 거래할 수 있을 때 데이터의 적시성, 추적효과성이 증가하는데 이를 위해서는 방대한 데이터 저장용량, 실시간 큐레이션 역량이 기본적으로 요구
 - ※ 데이터의 내재적 위험을 줄이기 위해서는 데이터 가치에 대해 거래 참여자들이 신뢰할 수 있는 평가가 이루어져야 하는데, 이를 위해서는 데이터 분석과 활용의 일정 부분이 데이터 거래 플랫폼 내에서 이루어지는 것이 이상적
 - ※ 데이터 분석, 활용 과정에서 사용된 데이터 목록, 분석 및 활용의 결과 등은 데이터 공급자, 이용자, 거래 플랫폼의 데이터 소유권, 통제권에 따라 다르게 공개되어야 할 것
 - 이러한 거래 플랫폼은 보안, 클라우드 컴퓨팅, 빅데이터, 기계학습 등의 기술들이 통합적으로 적용되어야 하고 또한 관련 법제도 자체가 별도로 마련되어야 할 것

24) 이러한 거래 플랫폼의 유사한 예로서 증권거래시장의 정보시스템을 들 수 있음. 증권거래 중앙 시스템, 홈 트레이딩 시스템 등은 거래 관련 뉴스, 기업재무제표 등 각종 가격 결정 정보가 통합적으로 제공되며 각 시장 참여자들의 권한에 따라 서로 다른 정보를 공개.

※ | 참고문헌

- [1] Chris Higon and Dave Waltho, “Valuing Information as an Asset”, Harvard Business School Press, 2007.
- [2] Kira Radinsky, “Data Monopolists Like Google Are Threatening the Economy”, 2015.03
- [3] Lateral Economics, “Permission granted: The economic value of data assets under alternative policy regimes”, 2016.3.
- [4] Steve Lohr, Data-ism; The Revolution Transforming Decision Making, Consumer Behavior, and Almost Everything Else, HarperBusiness, 2015.03
- [5] Tech Crunch, “What’ s The Value of Your Data?” , 2015.10.13.
- [6] The Huffington Post, “Google’ s remarkable monopoly on data intrigues Europeans” , 2015.07
- [7] 소프트웨어정책연구소, “개인정보 비식별화기술의 쟁점 연구” , 2016.8.
- [8] 유럽연합 집행위원회 (한국지식재산연구원 번역), “지식재산 가치평가에 대한 전문가 그룹의 최종보고서” , 2014.3.
- [9] ETRI 창의경제연구부, “지능정보사회 특징 및 변화(시대관점)”, 내부 작성 자료, 2015.
- [10] 한국국제조세협회, “무형자산이 포함된 국제거래의 정상가격 산정에 관한 합리적인 기준 마련” , 2012.9.
- [11] 한국정보산업연합회, “제4차 산업혁명의 핵심은 데이터다” , 2016.11.
- [12] 한국정보화진흥원, “공공데이터 민간개방의 경제적 파급효과 분석 연구” , 2012.11.

저자소개

정지형 ETRI 미래전략연구소 기술경제연구본부 미래사회연구실 선임연구원
e-mail: jhc123@etri.re.kr Tel. 042-860-1775

데이터 캐피탈리즘

- 데이터 자본화 이슈 -

발행인 : 한성수

발행처 : 한국전자통신연구원 미래전략연구소 기술경제연구본부

발행일 : 2016년 11월 30일

ETRI 한국전자통신연구원
미래전략연구소

305-700 대전광역시 유성구 가정로 218
전화 : (042) 860-3874, 팩스 : (042) 860-6504

* 주의 : 본서의 일부 또는 전부를 무단으로 전재하거나 복사하는 것은
저작권 및 출판권을 침해하게 되오니 유의하시기 바랍니다.

