

KISTEP 수요포럼 브리프

제121회 ‘데이터경제, 빅데이터시대의 R&I 혁신전략과 이슈?’

1. 논의 배경

- 데이터 경제사회의 도래는 코로나 19 사태와 맞물려 모든 사회구조와 경제 시스템의 디지털 전환을 가속화
 - 비대면 사회로의 전환과 맞물려 인공지능(AI), 블록체인, 빅데이터, 클라우드와 같은 디지털 기반 기술을 활용한 기존 산업의 혁신 진행중
 - 산업혁명 과정에서 전기가 대중화되면서 경제적 발전과 생산성 개선을 이뤄냈듯이, 데이터와 디지털 기술이 새로운 생산요소로 등장
- 주요국은 과학기술 및 사회경제의 변화와 발전에 디지털 데이터의 활용에 대한 전략 마련에 고심중
 - 주요국은 데이터 경제를 선도하기 위해 데이터의 확보, 산업·사회에서의 데이터 활용 확대, 데이터 분석 인재 양성 등과 함께 안전한 데이터 활용·제도 정비 등 종합적인 대책 마련 추진
- 우리나라도 데이터 경제시대에 맞는 혁신 성장의 추진 체계 마련을 위해 국정 과제를 비롯한 관련 법·제도의 정비를 추진
 - 과학기술정보통신부는 대통령 업무보고(‘20.1.6.)를 통해 DNA(Digital, Network, AI)를 기반으로 혁신을 선도하는 AI 일등국가와 미래성장을 견인하는 디지털 미디어 강국을 포함하는 3대 전략을 중점 추진
 - 데이터의 안전한 활용과 융합을 촉진할 데이터 3법 국회 통과(‘20.1.9.)

2. 현황 및 이슈

- 데이터 경제는 데이터가 경제활동의 중요한 생산요소로 사용되는 경제 구조를 의미
 - 2011년 데이비드 뉴먼이 가트너 보고서에서 데이터 경제라는 단어를 처음으로 사용하였으며, 빅데이터, 오픈데이터, 연결데이터 등 데이터로 파생되는 경제로 정의
 - 그 외에 데이터의 생성, 수집, 저장, 처리, 분배, 전달 등을 모두 포괄하는 개념(European Commission, 2014; 한국정보화진흥원, 2018) 등의 정의가 있음
- ‘데이터 경제’로의 패러다임 전환에 따른 새로운 국가전략 모색 필요
 - ICT 산업을 기반으로 하는 ‘디지털 경제(digital economy) 시대’를 넘어 데이터가 경제 활동의 주요 생산요소가 되는 소위 ‘데이터 경제(data economy) 시대’로의 패러다임 전환이 촉진됨에 따라,
 - 각종 데이터를 국가 경제발전 및 혁신 성장의 중요 원천으로 인식하고, 이를 체계적으로 관리·활용할 수 있는 새로운 국가전략 수립이 필요
- 국가 데이터 경제의 생산성을 데이터총생산(GDP, Gross Data Product)로 정의하고 4차 산업혁명 시대의 새로운 경제활동 지표로 제안(Chakravort 등 2019)
 - ※ 맥킨지는 데이터 경제가 2030년까지 13조 달러 규모를 형성, 글로벌 경제활동의 새로운 성장세를 가져올 것으로 예측
 - 국가의 경쟁력을 판단할 때, 과학기술이 성장에 기여한 부분이 지표상 드러내지 못하였으며, 이에 GDP(Gross Domestic Product)를 4차 산업혁명 시대의 패러다임 전환을 반영한 데이터총생산(GDP) 제시
 - 데이터총생산은 4개의 구성요소(Volume, Usage, Accessibility, Complexity)를 고려하여 개량화
 - 1) (Volume, 양) 국가에서 사용하는 광대역의 절대량, 생성된 원시

데이터의 프록시(proxy) 개수

- 2) (Usage, 사용) 인터넷에서 활동 중인 사용자의 수로서, 사용 행동의 폭, 요구 및 콘텐츠에 대한 프록시 개수
- 3) (Accessibility, 접근성) 생성된 데이터가 여러 인공지능 연구자, 혁신가 및 응용 프로그램에서 더 넓은 사용성과 접근성을 허용하는 정도를 평가하여 산출
- 4) (Complexity, 복잡성) 디지털 활동의 복잡성을 측정하는 요소로 흔히 1인당 광대역 소비량을 인용함

3. 정책 제언

- 과학기술 분야의 국가적 데이터 관리·활용 체계의 조속한 구축 시급
 - 과학기술정보통신부는 데이터 활용을 통한 경제·사회 전반의 혁신 창출을 위한 다양한 시책을 수립하여 추진 중
 - * 데이터 바우처, 마이데이터 실증서비스, 데이터 플래그십, 빅데이터 분석지원 등 4개 사업에 약 730여 억원 지원(2020 주요 데이터 활용지원 사업내용 참조)
 - 향후, 과학기술 분야에서도 '데이터중심 R&D(data intensive R&D)'가 급속히 발전할 것으로 예상되는 바, 이에 대한 적극적 대처가 요구
 - * 과학기술 연구데이터를 국가적 차원에서 공유·유통·활용할 수 있는 플랫폼 구축
 - 특히, 데이터 경제 실현을 위한 '데이터 3법'의 정비에 따라, 데이터의 수집, 축적 및 이용과정에서 발생하는 개인정보 보호 문제를 해소할 수 있는 구체적인 제도적 장치를 마련하는 것이 무엇보다 중요
 - 이를 위해, KISTEP 등 출연(연)이 중심이 되어 '과학데이터총생산'에 대한 정책개념의 적극적인 설정과 수립, 관리체계를 정비할 필요
- 주요국은 데이터 경제시대의 생존전략 차원에서 데이터 역량 강화에 주력하고 있음

- 영국의 정부혁신전략(Government Transformation Strategy, 2017), 디지털경제법(Digital Economy Act, 2017), 미국의 연방통계 혁신(2017), 증거기반 정책수립(2017), 일본의 통계개혁추진회의 설치(2017)와 미래투자전략(Society 5.0 데이터추동 사회로의 혁신, 2018) 등은 데이터 전략이 정부 혁신의 핵심 사안임을 시사
- Chakravorti 외의 데이터총생산(GDP, Gross Data Product) 개념은 4차 산업혁명 시대의 새로운 경제활동 지표로 의미 있음
 - Volume, Usage, Accessibility, Complexity의 관점에서 한국이 미국, 영국, 중국, 스위스에 이어 5위에 오른 점은 설득력이 있음
 - * 인터넷 이용률과 스마트폰 보급률 세계 1위(Pew Research, 2018)
 - 양적인 측면의 경쟁력 우위에서 질적 경쟁력 강화로 전환이 필요
- 이러한 관점에서 국가데이터전략의 확립이 절실
 - 데이터 전략이란 데이터가 비즈니스 전략을 활성화하고 촉진시키는 방법을 설명하는 핵심적인 통합 개념(MIT CISR Data Research Advisory Board)
 - 영국 정부는 국가 데이터 전략(NDS, National Data Strategy)의 목적을 세계 최고의 데이터 경제를 구축할 수 있도록 지원하는 비전을 추진하는 것으로 정의
 - 10대 원칙, 40대 실천 전략, 20대 실행 과제로 구성된 미국의 연방 데이터 전략(Federal Data Strategy)의 배경에는 증거 기반 행정의 확립을 위해 행정자료, 민간 자료 등 다양한 데이터 활용 전략을 위한 고민이 담겨 있음
- 우리나라의 경우 통계 데이터와 행정 데이터를 포괄하는 통합 데이터 거버넌스 관점에서 국가 데이터 전략 수립이 필요
 - 통계조직의 독립성과 전문성, 통계데이터와 행정데이터의 유기적 연계 측면에서 주요국에 비해 취약

- 데이터 시대에 부합하는 데이터 거버넌스 체계 강화를 위해서는 국가 핵심 데이터 자원인 국가 통계 거버넌스 강화를 출발로 행정 데이터와 통계 생산 시스템의 유기적 결합, 민간 데이터 활용을 위한 거버넌스 체계 정립이 단계적으로 이루어져야 함
- * 인공지능(AI) 국가 전략(2019. 12)을 통해 교육시스템 개선, 디지털 정부 구현 실천 과제 제시
- 국가 데이터 전략의 핵심 요소로는 시민과 공무원의 데이터 역량 강화를 위한 교육 및 제도 개혁, 데이터(특히 정부 데이터) 공유와 활용 활성화를 위한 법제 정비, 행안부 중심의 공공데이터 활성화 기능과 통계청 중심의 국가통계 기능의 효율적 연계를 위한 정부 조직 개편, 행정데이터와 통계 생산 시스템의 유기적인 연계 체계 정립, 데이터와 알고리즘의 공정성·편향성 대응 방안 등이 필요
- 데이터에 대한 기대와 요구가 어느 때보다 큰 시점이며 코로나 19로 인해 디지털화 및 데이터와 관련된 과감한 실행이 민간과 공공 모두 진행중이므로 데이터 총생산 관점의 개념과 국가 데이터 전략의 논의가 필요한 시점임.
 - 코로나 사태로 인해 경제사회 전반에 걸친 큰 변화가 촉발되었고 세계 2차 대전 이후 유례없는 대규모 광범위에 걸친 진행임.
 - 불확실한 미래 시나리오에서 디지털이 중심에 있을 것이며 디지털의 핵심은 데이터임.
- 데이터 수집과 관리, 그리고 활용에 있어서 총량도 중요하지만 데이터가 유효한 범위의 문제도 매우 중요하며 해당 이슈는 2가지 질문에 대한 답변을 얻기 위한 것으로 볼 수 있음.
 - 지금 어떠한 일들이 일어나고 있는가? 즉 정밀분석의 문제와 미래에 무슨 일이 발생할 것인가? 즉, 예측력 확보 문제이며 해결을 위한 다양한 기술 개발과 활용 노력이 전개되고 있음.
- 연구혁신 정책의 전주기(예측-기획-분석-평가 등)에서 데이터 사용이

활성화되고 실질적 기여를 높이는 데에 다양한 장애물들이 있음.

- 유럽집행위원회에서 현재 회람중인 EU 데이터 전략계획을 보면 데이터 경제의 잠재력 실현을 저해하는 요인으로 8가지*를 제시하였으며 거버넌스 법제화, 인프라투자, 정보주체 역량강화, 공동 데이터 공간 구축을 전략의 축으로 제시

* Availability of data, Imbalances in market power, Data interoperability and quality, Data governance, Data infrastructures and technologies, Empowering individuals to exercise their rights, Skills and data literacy, Cybersecurity

- 데이터와 관련된 연구혁신 정책 부문의 최근 동향을 살펴보면 세 가지 격차 극복이 필요함.

- 데이터 생산량과 분석가능량 간의 격차, 데이터 생산자와 이용자의 데이터 품질 격차, 데이터 관련자들의 배경 지식과 경험의 격차를 좁히기 위한 지속적이고 다양한 노력이 필요

※ R&D와 혁신정책의 경우 정부관료, 정책분석가 이외에 다양한 R&D분야별 전문가 그룹간의 효과적이고 지속적인 논의와 공감대 형성이 중요하나 각 그룹간 배경지식과 관점차이로 인해 데이터 분석 결과의 의미 전달과 효과적인 의사소통 난점 존재

- 데이터 품질을 높이기 위한 노력과 함께 전체 데이터량의 증가는 반 정보 또는 오정보* 유통량 증가도 수반되므로 데이터 총량 내부에 대한 측정과 검토가 필요하고 지속적인 점검과 견제 활동이 필수적임.

* disinformation or misinformation(유럽집행위원회, 2019)

- 데이터 유통 활성화를 통한 양의 효과 증가도 중요하지만 가짜정보의 생성과 유통으로 인한 사회적인 비용 발생도 해법 마련이 시급하며 인공지능과 기계학습 등을 통한 기술적 접근이 진행중

- 대용량이며 복잡한 구조의 데이터들을 검토하고 확인하는 활동의 경우 다양한 데이터군들의 연계 또는 경계면에 대한 분석이 중요하므로 이를 담당할 수 있는 융합 전문가 그룹 또는 경계조직(Boundary organization) 활성을 통한 적극적인 대응이 필요함.

- 공공과 민간의 협력과 적절한 역할분담을 통해 지속가능한 데이터 생태계 구축이 중요
- ※ 최근 급부상하는 민간주체의 인터넷 플랫폼은 큰 역할을 하고 있으나 공공정책 부문의 경우 정부 관심과 지향점에 차이가 있다는 문제 제기가 있음(JRC, 2019).
- ※ 연구혁신 정책과 관련된 데이터들의 데이터원천(행정데이터, 상용데이터, 설문데이터, 비정형데이터, 파생데이터)과 관계, 데이터간 상호운용성 확보 정도, 데이터 활용과 참여주체에 대한 종합적인 현황 파악 프로젝트 진행이 디지털화에 대한 정교한 고민의 시발점이 될 수 있음.