보건복지정책에서의 빅데이터 활용 전략과 과제

Strategies and Tasks for Big Data in Health and Welfare

오미애 | 한국보건사회연구원 연구위원

빅데이터는 의료, 금융, 복지, 경제 등 모든 분야에서 적극적으로 활용되고 있으며 국가적 차원에서 이에 대한 관심이 커지고 있다. 정부는 빅데이터를 새로운 가치 창출 수단을 넘어 경제활동의 중요한 생산요소라고 보고 있다. 4차 산업혁명에서는 데이터가 자산인 만큼 데이터를 얼마나 잘 활용하느냐가 국가 경쟁력 확보의 중요한 요소이다. 이미 많은 분야에서 빅데이터 추진 전략을 내세우고 있는데, 이 글에서는 보건복지정책에서의 빅데이터 활용 전략을 정리하고 주요 과제를 살펴보고자 한다.

1. 들어가며

지난 5월 현 정부는 '바이오·헬스산업 혁신 전략'을 발표하였다. 정부는 바이오·헬스 분야를 국민 생명과 건강을 보장하기 위한 핵심 산업으로 보고 차세대 선도 산업으로 육성할 계획이다. 최대 100만 명 규모의 '국가 바이오 빅데이터'를 구축하고 정부 연구·개발(R&D) 투자에 2025년까지 연간 4조 원 이상을 투입할 예정이다. 금융위원회도 지난 6월 '안전한 데이터 활용과 디지털 경쟁·혁신을 위한 금융 빅데이터 인프라 오픈 행사'에서 금융 빅데이터 인프라 구축 계획을 발표하였다. 여기에는 금융 빅데이터 개방 시스템 서비스 개시, 금융 분야 데이터 거래소 오픈 및 시범 서비스 실시, 금융 분야 데이터 전문기관 지정 내용이 포함되어 있다(금융위원회, 2019). 금융 빅데이터 개방 시스템을 통해 우선적으로 일반 신용 데이터베이스(DB)인 200만 명에 대한 차주, 대출, 연체 및 카드 개설 정보 등의 금융 분야 빅데이터가 개방된다. 금융 빅데이터 개방 시

스템 오픈 2주 만에 금융사·핀테크 기업 등 80여 개 기관이 회원 가입(이지윤, 2019)을 하였는데, 이는 많은 기업들이 그동안 빅데이터를 활용하는 데 제약이 있었음을 보여 준다.

빅데이터는 의료, 금융, 복지, 경제 등 모든 분야에서 적극적으로 활용되고 있으며 국가적 차원에서 이에 대한 관심이 커졌는데, 그 배경에는 지난해 6월 4차산업혁명위원회에서의 '데이터산업 활성화 전략' 논의가 있다. 이는 2018년 8월 데이터 경제 활성화 규제 혁신 현장 방문 행사에서의 대통령 연설문 중 "핵심 기술 개발을 지원하고 전문 인력 5만 명, 데이터 강소 기업 100개를 육성하며, 이를 위해 내년 데이터산업에 1조 원을 투자할 것"이라는 내용과 맥락을 같이한다.

이렇듯 정부는 빅데이터를 새로운 가치 창출 수단을 넘어 경제활동의 중요한 생산요소라고 보고 있다. 4차 산업혁명에서는 데이터가 자산인 만큼 데이터를 얼마나 잘 활용하느냐가 국가 경쟁력 확보의 중요한 요소이다. 앞에서 언급한 것처럼 이미 많은 분야에서 빅데이터 추진 전략을 내세우고 있는데, 이 글에서는 보건복지정책에서의 빅데이터 활용 전략을 정리하고 주요 과제를 살펴보고자 한다.

2. 빅데이터 활용 정책 관련 국내외 현황

가. 국외 정책 현황[1)]

데이터 기반 정책 지원을 위해 각국은 데이터 경제를 선도하고자 범국가 차원의 대책을 마련하고 있다.

영국은 오픈 데이터 구축·개방을 통해 공공 서비스를 혁신하여 경제적 가치를 발생시키고 있다. 교통 공공 데이터 개방의 경우, 런던 경제에 연간 최대 1억 3000만 파운드(약 1911억 원)의 가치를 창출하는 데 기여하였다. 데이터 중심의 혁신으로 투명성을 확보하고, 정부와 민간의 혁신이 가능하도록 하며, 광범위한 이해관계자의 참여를 유도하고 있다. 그리고 「디지털 경제법(Digital Economy Act)」을 통해 공공기관 상호간 데이터 공유를 확대하고 비식별화된 개인정보 공개를 추진하고 있다. 에든버러 주정부는 20년 후 에든버러를 '유럽의 데이터 수도'로 만들겠다는 계획을 세우고, 10년간 정부 자금 2500억 원과 민간 펀드 2500억 원을 투자할 계획이다(최병호, 2018).

미국은 투명한 데이터 거래를 위해 「데이터법(Data Act)」을 제정(2014), 발표하였다. 또 데이터 활용 혜택을 최대화하고 위험을 줄이기 위해 「소비자 프라이버시 권리장전(Consumer Privacy Bill of Rights)」, 「정보유출통지법(National Data Breach Legislation)」을 제정하였다. 또한, 2017년에 국립데이터서비스청(National Secure Data Service)의 신설이 제안된 후 기구가 만들어졌는데, 여러 기관에서 통계 목적으로 보유해 온 자료를 결합하는 업무를 담당하며 자료 보관소(data

1) 한국정보화진흥원(2018), p. 12의 내용을 발췌하여 요약·정리하였다.

warehouse) 역할을 지양한다는 점을 명시하였다(정용찬, 2018).

일본은 '미래 투자 전략 2018(Society 5.0 데이터 추동 사회로의 혁신)' 보고서에서 데이터 추동 사회로의 진입을 선언하며, 세계의 '데이터 패권주의'에 대한 대응 전략으로 행정서비스의 디지털 전환 및 정부 보유 데이터 개방을 통한 혁신과 신규 비스니스 창출 촉진을 강조하고 있다. 또한 '익명 가공 정보 제도[2]'를 도입하여 개인정보 보호와 데이터 활용을 촉진하고 있다.

중국은 산업별 대용량 데이터 활용을 강화하며, 기업별 데이터 활용을 촉진하고, 빅데이터 표준 시스템을 구축하였다. 그리고 보안 역량 강화를 통해 데이터 거래 및 데이터 경제 활성화를 위한 기반을 마련하였다.

나. 국내 정책 현황

우리나라의 경우, 과학기술정보통신부가 '빅데이터 플랫폼 및 센터 구축' 사업으로 분야별 플랫폼 10곳, 이와 연계된 기관별 센터 100곳을 구축하는 데 3년간 총 1516억 원을 투입한다고 발표하였다. 올해에는 총 두 차례에 걸쳐 플랫폼 10곳과 센터 100곳, 총 640억 원 규모로 사업을 추진한다(과학기술정보통신부, 2019a). '빅데이터 플랫폼 및 센터 구축' 사업은 공공과 민간이 협업해 빅데이터센터 등에서 수집된 데이터를 분석·유통할 수 있는 플랫폼을 구축하고, 혁신 서비스를 발굴·확산해 데이터 기반의 가치 창출 생태계를 조성하기 위한 것이다.

과학기술정보통신부는 개인이 자신의 정보를 직접 내려받거나 동의하에 제3자에게 제공해 다양한 분야의 개인 데이터 활용 서비스를 누릴 수 있도록 '본인 정보 활용 지원(MyData) 사업(이하 마이데이터 사업)'도 추진하고 있다. 올해는 의료, 유통, 에너지 등으로 분야를 확장해 총 97억 원 규모로 실증서비스를 추진한다.

강남세브란스병원(의료)은 건강검진 및 처방전 데이터를 개인이 휴대전화 앱에서 직접 내려받아 제3의 기업에 제공함으로써 맞춤형 건강 관리(활동량, 영양 관리 등) 및 식단 추천 서비스에 활용할 수 있도록 할 예정이다. 다음소프트와 세종시(에너지)는 가구별 에너지(상하수도, 전력, 가스 등) 사용량 데이터를 활용해 시간대별 사용량 모니터링 및 시각화 분석, 누진제 적용 시작 구간 알람 등의 에너지 절감 서비스를 제공한다.

보건의료 분야에서도 국민건강보험공단, 건강보험심사평가원, 질병관리본부, 국립암센터 4개 기관의 정보를 연계하여 연구를 지원하는 보건의료 빅데이터 시범사업을 추진 중이다. 시민, 환자단체를 포함한 보건의료 빅데이터 정책심의위원회를 구성하고 시범사업의 추진 방안, 입법 필요 사항, 데이터 활용 방안 등에 대한 논의를 하고 있다.

2) 해당 개인정보 보호법 개정은 2017년 5월 시행되었으며, 의료 분야 데이터 이용을 활성화하기 위해 의료 정보의 '익명 가공 사업자' 인정 제도를 도입하는 '차세대 의료기반법'이 포함된다.

표 1. 2019년 본인 정보 활용 지원(MyData) 실증서비스 선정 과제 목록

분야	과제명	수행 기관		서비스
		주관	참여	
의료	의료 마이데이터 플랫폼 및 검진 데이터를 활용한 건강 관리 서비스	강남세브란스병원 (연세대학교 산학협력단)	(주)아롬정보기술 CJ프레시웨이 에쓰푸드(주)	• 건강검진·처방전 등의 데이터를 이용한 영양 건강 식단 추천 서비스
의료	응급 상황을 위한 개인건강 지갑 서비스	(주)브이티더블유	삼성서울병원 서울아산병원 동아대학교병원	• 응급 환자가 응급 진료 기록 및 일상생활 속 건강 기록을 보관하고, 진료와 처방에 활용할 수 있는 '개인건강 지갑' 서비스
의료	마이 헬스 데이터 플랫폼 및 서비스 실증	서울대학교병원	차의과대학교 산학협력단, 메디블록, 웰트, 삼성화재	• 환자가 동의한 개인 의료 정보 기반의 건강 정보 교류 플랫폼 개발, 라이프로그 데이터와 융합한 개인 맞춤 코칭 서비스
금융	본인 정보 통합 조회 및 생애 주기별 맞춤형 금융 상품 추천 서비스	엔에이치엔 페이코(주)	(주)하나은행 한화생명보험(주) 한화투자증권(주) 한화손해보험(주) 신한금융투자(주) 웰컴저축은행(주)	• 금융·비금융 정보를 융합한 빅데이터 기반 맞춤형 금융 상품 추천 서비스
에너지	에너지 마이데이터를 활용한 사용자 맞춤형 절감 서비스	(주)다음소프트	세종시 (주)에이엠에이닷컴 (주)유디아이	• 세종시 주민 대상 에너지 데이터 수집 및 에너지 수급 예측, 맞춤형 서비스 구현을 통한 에너지 소비 효율화
유통	개인 데이터 저장소 기반 소상공인 마케팅 관리 서비스	(주)한국신용데이터	신한카드	• 결제 데이터와 가맹점 방문객 정보, 사업장 정보를 활용하여 소상공인 대상 경영 개선 및 개인 대상 할인·이벤트 정보 제공 서비스
유통	소상공인 성장을 돕는 문서·자금 플랫폼 서비스	한국기업데이터(주)	한국전자 세제협회 (주)비즈니스온 커뮤니케이션 기웅정보통신(주)	• 소상공인(개인사업자)을 위한 문서 자금 플랫폼을 통해 간편 본인 정보 전송, 계산서 발급 및 정책 자금 매칭 등 경영 지원 서비스
기타	연구자를 위한 매칭 및 분석 서비스	(주)코난테크놀로지	국회도서관	• 국회도서관으로부터 수집된 학술연구자 정보를 활용하여 연구자를 위한 전문 서비스 '스칼라뱅크' 구축

자료: 과학기술정보통신부. (2019. 5. 17.). 의료·금융·에너지 등 마이데이터 서비스 8개 과제 선정, 붙임 1. 보도자료.

각 지방자치단체도 최근 3년 동안 빅데이터 분석 사업에 집중하고 있는데, 서울시의 경우 인프라 조성 중심의 사업이 많았고, 그 외의 지자체는 공공 행정, 관광 문화, 재해 안전 사업 연구가 많았다(강민선, 2018).

3. 빅데이터 활용 전략별 과제

보건복지정책에서 빅데이터가 잘 활용될 수 있으려면 표준화, 플랫폼 구축, 법·제도 검토, 대국민 서비스 방안, 이렇게 4가지 전략이 필요하다. 이는 예전부터 언급된 전략으로 새로운 것이

아니다. 하지만 데이터를 둘러싼 주변 환경과 여건이 변화하기에 4가지 전략을 되짚어 볼 필요가 있다. 표준화는 데이터 포맷, 용어, 생성 주기 등을 검토해야 하는 부분이고, 플랫폼 구축은 정형 빅데이터와 비정형 빅데이터를 고려해야 한다. 법·제도 관련 사항으로는 2017년 말에 '데이터 기반 행정 활성화에 관한 법률' 제정안이 국무회의를 통과하여 국회에 상정 중이다. 이 법률은 각 기관이 데이터 기반 행정업무를 발굴할 수 있도록 하고, 기관 간 데이터 제공, 데이터 관리 체계 마련 등의 사항을 규정하여 데이터가 공유·연계될 수 있는 근거가 되기 때문에 법률이 통과되어 법제화되기 전까지 검토할 수 있는 부분을 고려해 보아야 한다. 대국민 서비스 방안은 보건복지 정책에서의 빅데이터 활용이 국민의 지지를 받기 위해 필요한 부분으로, 빅데이터 관련 추진 사업에서 공청회나 시범사업 등을 고려해야 한다.

보건복지 분야에서 4가지 추진 전략이 필요한 이유는 다음과 같다.

표준화는 용어 및 기준 통일로 명확한 의사소통이 가능하게 한다. 그리고 필요한 데이터 소재 파악에 소요되는 시간이나 노력이 감소되며, 일관된 데이터 형식·규칙의 적용으로 데이터 품질 향상을 기대할 수 있다. 또한 정보 시스템 간 데이터 결합 시 데이터 변환 및 정제 비용을 감소시킬 수 있다. 의료 데이터의 경우 타 기관과의 자료 연계 시 진료 기록부 표준화는 필수적이라고 할 수 있다.

플랫폼 구축은 보건복지 분야의 다양한 데이터를 활용할 수 있는 환경 기반을 마련하는 데 필요하다. 보건복지 분야 정책을 수립하고 실행 과정을 모니터링할 수 있는 체계 마련과 빅데이터 분석을 할 수 있는 분석 환경 마련, 객관적이고 과학적인 분석에 따른 정책 의사 결정 체계 마련 모두 플랫폼 구축하에서 이루어질 수 있다.

법·제도 검토는 데이터 활용과 개인정보 보호

그림 1. 빅데이터 활용 전략 4과제

자료: 저자 작성.

가치 간 조화의 기반을 마련하기 위해 꼭 필요하다. 행정안전부 공공데이터정책과는 2018년까지 빅데이터 활용과 개인정보 보호 간의 사회적 합의를 도출하고 법적 효력을 지닌 법령, 가이드라인 등 관련 규정을 개선하여 2019년부터 빅데이터 활용 촉진을 위한 제반 시스템 구축을 추진 중이다(행정안전부 공공데이터정책과, 2018).

대국민 서비스 방안은 국민의 알 권리를 보장하고 정보 접근성의 비대칭성을 제거하며, 서비스 제공 피드백을 통해 대국민 서비스를 고도화함으로써 서비스 질을 제고할 수 있다. 보건복지 정책에 대한 국민의 공감대 형성과 정부에 대한 신뢰도 제고에 꼭 필요한 요소로, 정책 수요자의 적극적인 정책 참여를 유도할 수 있다.

4. 빅데이터 활용을 위한 전략별 실행 과제

앞서 언급한 4가지 추진 전략을 위한 구체적인 실행 과제는 다음과 같다.

가. 표준화 방안

표준화는 시스템 구축 단계부터 중요하게 검토되어야 함에도 불구하고, 데이터 표준 정책 없이 단위 시스템 위주로 표준 정책을 수립한 채 프로젝트들이 진행되고 있다. 기관마다 상황은 다르지만 데이터 표준화를 지원하는 관리 도구를 가지고 있는 기관이 많지 않고, 데이터 관리 전담 인력도 부족하며, 데이터 관리의 중요성을 크게 인지하고 있지 않은 것이 현실이다.

이러한 한계를 극복하기 위해서는 데이터 관리 교육을 해야 한다. 데이터의 질 확보를 위해 표준화 단계가 선행되어야 함을 인지하고, 데이터 정제 절차 및 프로세스를 마련하여 교육을 해야 한다.

그리고 분야에 적합한 표준화 지침을 작성해야 한다. 공공 데이터 관리 지침, 행정 정보 데이터베이스 표준화 지침이 존재하나 각 분야의 특성을 반영하지 못한 일반적인 양식으로 구성돼 있어 분야별 특성에 맞는 표준화 지침 작성이 필요하다. 보건의료 분야의 경우, 진료 기록 자료에 대해 데이터 생성부터 관리할 수 있는 표준화 지침을 작성한다면 의료 데이터의 활용성을 높일 수 있다.

또한 보건복지 분야의 용어 표준화 작업도 필요하다. 복지 분야에서 많이 쓰이는 '노인 가구'의 개념에 대한 정의를 보면 국어사전에는 "65세 이상의 노인만으로 이루어진 가구"로 되어 있다. 반면 통계청 2017 등록 센서스에 정의된 노인 가구는 "65세 이상 가구원 포함 가구"이다. 이러한 개념 정의에 대한 차이는 통계 수치에 혼란을 줄 수 있기 때문에 하나의 개념으로 통일하여 사용할 필요가 있다.

마지막으로, 데이터 관리를 전담할 수 있는 인력을 양성할 필요가 있다. 행정안전부는 공직 분류 체계 내 데이터직류 신설 및 데이터 전문 인력 충원을 계획 중이다. 실효성 있는 정책 지원이 이뤄지기 위해서는 형식적인 차원을 넘어 데이터 관리 부서를 실질적으로 지원하고 타 부서와 협

력할 수 있는 인재 양성이 우선되어야 할 것이다.

나. 플랫폼 구축

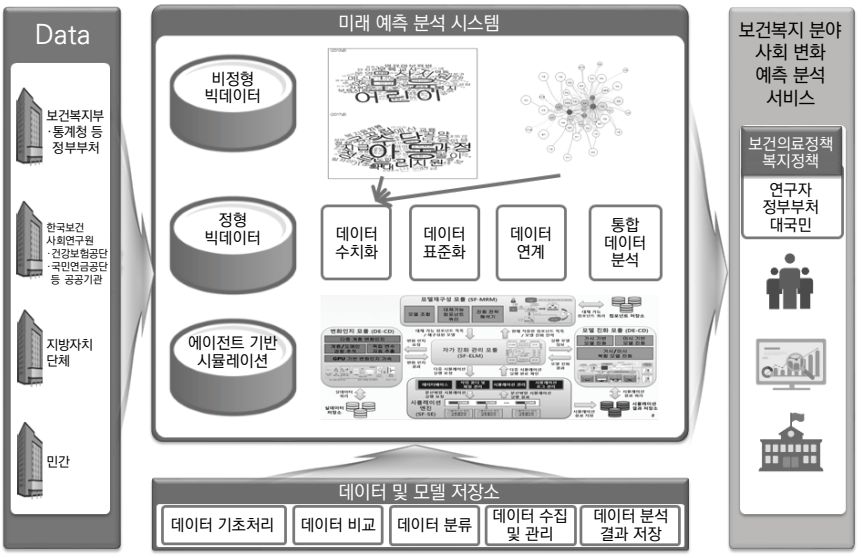
플랫폼 구축은 연구 및 정책 지원 환경의 기반을 마련하는 데 필요한 인프라이다. 플랫폼은 각 목적에 맞게 기능과 구성이 이루어져야 하기 때문에 보건복지 분야에 하나의 완성된 단일 플랫폼을 구축하기보다는 기존의 플랫폼을 유지·활용하면서 타 시스템과 연계 가능한 유연한 플랫폼을 구축하는 방안이 필요하다. 데이터 기반 정책 연구의 경우, 수집·연계 데이터는 보건복지 데이터, 공공 데이터, 외부 데이터 등으로 나뉠 수 있고, 데이터 분석 플랫폼 안에서는 데이터 수집 및 연계 영역, 데이터 분석 영역, 포털 서비스 영역, 주제별 데이터 영역으로 나뉠 수 있다. 이러한 플랫폼은 보건복지 분야에서의 선제적인 정책 대안 제시 및 타 분야와의 융합 연구에 활용될 수 있다. 민간 영역에서는 사회 변화에 대한 실시간 분석이나 맞춤형 정보 서비스 제공이 가능할 것이다. 국제 비교 연구에서는 주요 국가 통계를 연계하거나 해외 연구기관과의 협업 체계를 확대할 수 있는 기반이 마련될 수 있다.

플랫폼 구축에서는 빅데이터의 분석 기술 발달로 사회 현안을 실시간으로 분석할 수 있는 비정형 빅데이터도 분석 시스템에 포함된다. 이는 분석 결과를 바탕으로 정형 빅데이터와 연계하여 다양한 정책 시뮬레이션을 하는 데 활용될 수 있다.

그림 2. 데이터 기반 정책 지원 시스템 예시

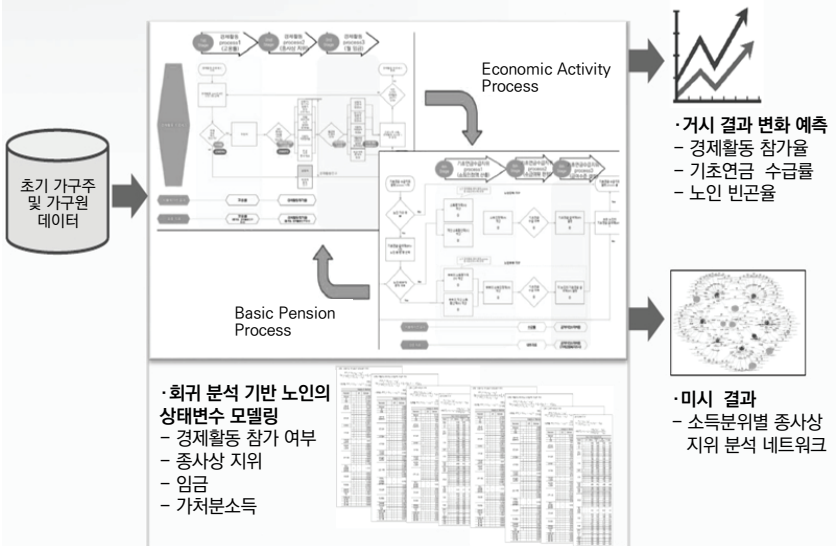
자료: 경제·인문사회연구회. (2019. 5. 15.). 데이터기반 미래예측정책지원시스템. 데이터 기반 정책시뮬레이션 컨퍼런스의 일부 내용 수정하여 제시.

그림 3. 보건복지 분야 미래 예측 분석 시스템 예시



자료: 저자 작성.

그림 4. 시스템 기반 거시 및 미시 결과 예측 예시



자료: 경제·인문사회연구회. (2019. 5. 15.). 데이터기반 미래예측정책지원시스템. 데이터 기반 정책시뮬레이션 컨퍼런스. ETRI 발표자료.

이러한 플랫폼이 잘 구축된다면 연구자가 본인이 설정한 연구 영역에서, 예를 들면 '인구 고령화에 대응한 노후소득보장정책 및 노인고용정책 변화가 노인 빈곤과 불평등 변화에 미치는 영향 예측·분석'과 같은 다양한 시나리오 분석이 가능할 것이다. 분석 결과는 거시변화로 경제활동 참가율, 기초연금 수급률, 노인 빈곤율을 예측할 수 있고, 미시 결과로 소득분위별 종사상 지위 네트워크 분석이 가능할 것이다.

다. 법·제도 검토

빅데이터와 관련된 법령은 빅데이터 활용을 촉진하는 것과 빅데이터 활용을 제한하는 것으로 나눌 수 있다. 빅데이터 활용을 촉진하는 법으로는 「공공데이터의 제공 및 이용 활성화에 관한 법률」, 「공공기관의 정보공개에 관한 법률」 등이 있으며, 빅데이터 활용을 제한하는 법으로는 「저작권법」, 「의료법」 등이 있다.

이러한 법률의 상충 관계를 해결하기 위해서는 빅데이터 활용의 명암을 면밀히 검토한 뒤, 관련 법률의 일괄 개정이 필요하다. 현실적으로 가능한 안은 '데이터 기반 행정 활성화에 관한 법률' 제정안이 법제화되는 것인데, 그 전까지 우선적으로 「통계법」 24조하에서 빅데이터 활용 방안을 검토할 수 있다. 통계청은 통계 작성 목적으로 행정 자료 제공을 요청할 수 있다. 통계청 통계데이터허브국의 올해 업무 추진 계획에는 국가정책 수립 및 신규 통계 작성 등에 필요한 통계

표 2. 빅데이터 관련 법령 현황

구분	법령
빅데이터 활용 촉진	범부처별 •공공데이터의 제공 및 이용 활성화에 관한 법률 •공공기관의 정보공개에 관한 법률 •전자정부법 •국가정보화기본법 •민원사무 처리에 관한 법률 •행정절차법 등 분야별 •기상법/기상산업진흥법(기상 정보) •통계법(통계 정보) •건설기술 관리법(건설기술 정보) •발명진흥법(특허 정보) •공간정보산업 진흥법(공간 정보) 등
빅데이터 활용 제한	•공공기관의 정보공개에 관한 법률(비공개 대상 정보) •행정조사기본법(행정조사 관련 정보 활용 제한) •국유재산법(국가 재산 처분 제한, 비용 산정 등) •공유재산 및 물품 관리법(지자체 재산의 처분 제한 등) •저작권법(권리 처리 등) •의료법 등

자료: 이상윤. (2017. 10.). 빅데이터 관련 개인정보 보호법제 개선방안 연구. 법제처, 한국법제연구원. pp. 29-30.

표 3. 통계법 24조

제24조(행정자료의 제공)
① 중앙행정기관의 장 또는 지방자치단체의 장은 통계의 작성을 위하여 필요한 경우에는 공공기관의 장에게 행정자료의 제공을 요청할 수 있다. 〈개정 2009. 4. 1., 2012. 12. 18.〉
② 공공기관의 장은 제1항에 따라 행정자료의 제공을 요청받은 때에는 국가기밀, 개인과 기업의 중대한 비밀의 침해 등 대통령령으로 정하는 정당한 사유가 없는 한 이에 응하여야 한다.
③ 제2항에 따라 행정자료를 제공하는 경우 그 제공 범위 및 방법 등에 관하여는 요청기관의 장과 제공기관의 장이 협의하여 결정하되, 제공기관의 장은 요청기관의 장에게 행정자료에 포함되어 있는 개인이나 법인 또는 단체 등의 정보를 보호하기 위하여 사용 방법·사용 부서나 그 밖에 필요한 사항에 대하여 제한을 하거나 행정자료의 안전성 확보를 위하여 필요한 조치(이하 "정보보호조치"라 한다)를 강구하도록 요청할 수 있다.
④ 제2항에 따라 공공기관으로부터 제공받은 행정자료는 이를 통계 작성 외의 목적으로 사용하거나 다른 자에게 제공하여서는 아니 된다.
⑤ 행정자료 제공기관의 장은 요청기관의 장이 제3항에 따라 요청한 정보 보호 조치를 하지 아니하거나 제4항을 위반하는 경우에는 행정자료의 제공을 중지 또는 제한할 수 있다.

자료: 통계법, 법률 제14843호 (2017).

생산용 DB를 구축·제공하고 종합통계등록부를 고도화한다는 내용과, 국가승인통계 작성 시 공공기관이 보유한 행정 자료의 활용 가능성을 검토하고 관련 조사에 활용토록 지원하는 활동을 강화하겠다는 내용이 담겨 있다. 따라서 각 기관에서 통계 작성 목적으로 국가승인통계 작성을 검토하고 있다면, 통계청과의 협의하에 행정 빅데이터 연계 가능성도 검토해 볼 수 있다.

라. 대국민 서비스 방안

빅데이터 활용의 대국민 서비스 방안과 관련해서는 먼저, 데이터 기반 정책 의사 결정에 대한 국민 인식을 살펴볼 필요가 있다. 정부의 빅데이터 활용에 대한 국민의 신뢰도가 낮다면, 국민의 신뢰도를 높일 수 있는 방안을 우선적으로 검토해야 하고, 반대로 국민의 신뢰도가 높다면 대국민 서비스의 질을 어떻게 높일 것인지를 고민해야 할 것이다.

일반 국민의 경우, 정보 사각지대가 존재하며 데이터센터에 대한 접근성이 낮기 때문에 국민의 이용도 제고를 위한 교육과 홍보가 필요하다. 또한 정책 수요자 관점의 서비스 사용자 인터페이스(UI)를 제공함과 동시에 대국민 아이디어 공유 기반을 마련하여 국민 개개인이 생각하는 빅데이터 활용 방안이 보건복지 빅데이터 정책에 반영될 수 있는 기회를 제공할 수 있어야 한다. 그리고 데이터 기반 정책 의사 결정에 대한 국민 인식 제고를 위한 공청회를 마련하여 국민의 생활에 도움이 되는 빅데이터 기반 정책 활용 사례를 공유하고, 이를 통해 개인을 식별할 수 없는 방법으로 개인정보가 보호되고 각 개인의 데이터가 정책 결정에 긍정적으로 활용될 수 있음을 이해시킬 필요가 있다.

5. 나가며

지금까지 빅데이터의 활용도 제고를 위한 표준화, 플랫폼 구축, 법·제도 검토, 대국민 서비스

방안의 4가지 전략을 살펴보았다. 전략별 추진 과제도 제시하여 고려해야 할 점에 대해 다시 한 번 짚어 보았다.

보건복지정책에서 빅데이터는 행정 자료를 포함하기도 하고 텍스트 자료, 이미지 자료를 포함하기도 한다. 공공 데이터 역시 빅데이터로 혼용하여 쓰이기도 한다. 이렇듯 최근에는 빅데이터를 광의의 개념으로 사용하며 정책에 다각도로 활용하고 있다. 이러한 현상은 사회적으로빅데이터에 대한 관심이 그만큼 높아진 것으로 해석할 수 있다.

보건복지 분야에서 정책 지원을 위한 빅데이터 활용이 원활히 이루어지기 위해서는 정부의 빅데이터 활용과 개인정보 보호에 대한 국민의 신뢰도 제고가 우선되어야 할 것이다. 보건의료 빅데이터 사업의 경우 2017년 말부터 추진되었으나, 정부가 국민의 건강 정보를 상업적 목적으로 기업에 제공할 수 있다는 우려 때문에 시민사회단체에서 반대했다. 그 후 정부는 설득 작업을 거쳐 시민단체, 전문가가 참여하는 보건의료정책심의위원회를 구성하고 여러 차례 회의를 진행하여, 안전장치를 마련한 후 시범사업을 추진해 나가는 것으로 의견을 모았다. 개인 건강 정보는 민감 정보 중 하나로, 유출이나 오용될 경우 사회적으로 큰 파장을 일으킬 수 있기에 빅데이터 활용의 우려 사항을 충분히 검토해야 하며, 공청회 등을 통해 국민들의 의견을 수렴하는 자리가 마련되어야 할 것이다.

개인정보 노출의 연장선상에서 살펴보면, 자료 공개 시 개인을 식별할 수 있는 정보를 삭제한 후 공개하더라도 외부 자료와의 연계를 통해 민감한 정보가 유출될 수 있다. 미국의 보안회사는 뉴욕시의 택시 운행 기록과 연예인 가십 기사를 분석하여 특정 연예인들의 택시 승하차 장소와 지불한 요금 정보를 알아낼 수 있었다(박민정, 이용희, 권성훈, 2018). 서로 다른 자료가 연계될수록 개인의 민감한 정보가 유출될 가능성도 커질 수 있다. 따라서 차등정보보호와 같은 새로운 시각의 개인정보 보호 기법을 고려해 보아야 한다.

마지막으로, 보건복지 분야의 빅데이터를 분석·활용할 때는 알고리즘의 편향성도 고려해야 한다. 빅데이터 분석에서 알고리즘은 중요한 부분을 차지하는데, 알고리즘의 불공정성·편향성 등의 윤리적 문제에 대한 고려는 모든 사람에게 공정하고 포용적인 시스템을 만들기 위해 꼭 필요한 일이기 때문이다. 구글의 인공지능(AI) 공정성을 연구하는 연구자는 불공정성은 데이터 수집, 분류, 이용 과정에서 항상 개입할 수 있는 요소라고 하였다. 그리고 단일 원인을 추적하기 어렵고 복합적으로 편향성 문제가 발생하기 때문에 이에 대한 해결책도 복합적일 수밖에 없다고 설명하였다(이기범, 2019). 그렇기 때문에 빅데이터 분석 시, 공정성 개념을 여러 시각에서 바라볼 필요가 있다.

빅데이터 활용의 장점을 극대화하고 단점을 최소화하도록 민관이 협력하여 노력해 나간다면 우리나라도 데이터 경제의 선두 국가로 자리매김할 수 있을 것으로 기대한다.

참고문헌

강민선. (2018). 지방자치단체 빅데이터 분석 사업 추진 현황 부록. 2018년 지역정보화 이슈리포트. 한국지역정보개발원. pp. 1-35.

과학기술정보통신부. (2019a. 5. 13.). 빅데이터 플랫폼 및 센터 구축 사업. 보도자료.

과학기술정보통신부. (2019b. 5. 17.). 의료·금융·에너지 등 마이데이터 서비스 8개 과제 선정. 보도자료.

경제·인문사회연구회. (2019. 5. 15.) 데이터기반 미래예측정책지원시스템. 데이터 기반 정책시뮬레이션 컨퍼런스.

금융위원회. (2019. 6.). 금융분야 빅데이터 인프라 추진방안. 보도자료.

박민정, 이용희, 권성훈. (2018). 차등정보보호에 관한 연구. 통계개발원 2018년 하반기 연구보고서.

이기범. (2019. 6. 25.). 구글이 말하는 머신러닝의 공정성. 블로터. http://www.bloter.net/archives/344120에서 2019. 7. 1. 인출.

이상윤. (2017. 10.). 빅데이터 관련 개인정보 보호법제 개선방안 연구. 법제처, 한국법제연구원.

이지윤. (2019. 6. 17.). 금융 빅데이터 개방' 2주 만에 80개사 몰려. 서울경제. https://www.sedaily.com/NewsView/1VKGCT017K에서 2019. 7. 1. 인출.

정용찬. (2018. 11. 8.). 4차 산업혁명 대응 데이터 거버넌스 국가 전략. ICT 산업전망 컨퍼런스 발표자료.

최병호. (2018. 10. 15.). (국정기획2050)⑥미래 기술과 산업·고용·불평등의 위험과 해결방안. 뉴스토마토. http://www.newstomato.com/ReadNews.aspx?no=851119에서 2019. 7. 1. 인출.

통계법, 법률 제14843호 (2017).

한국정보화진흥원. (2018). 데이터 경제의 부상과 사회경제적 영향. IT & Future Strategy 보고서.

행정안전부 공공데이터정책과. (2018. 3. 22.). 정부의 빅데이터 추진체계 및 정책방향. 국회입법조사처세미나.