

# 사회보장 분야 데이터의 활용과 그 미래

**노대명** | 한국사회보장정보원 원장

## 들어가며

**사회보장 행정데이터의 시대** 우리는 과연 수기로 사회보장 급여를 책정하고 집행하던 과거로 돌아갈 수 있을까. 또한 현장 방문을 통해 복지 사각지대를 발굴하고 지원하는 방식만을 고수할 수 있을까. 표본조사를 통해 데이터를 구축하여 문제를 확인하고 다양한 정책 대안의 타당성을 검토하던 방식에만 의존해서 연구를 할 수 있을까. 그렇지 않다. 이미 현재의 사회보장시스템은 다양한 행정데이터가 교환되고 결합되어 새로운 정보를 생산하고, 이를 통해 수급 자격과 급여를 결정하는 방식으로 작동하고 있다. 마찬가지로 사회보장 분야의 학술연구 또한 행정데이터를 적극적으로 활용하는 방식으로 변화하고 있으며, 조사데이터의 오차를 줄이는 방안에도 광범위하

게 활용되고 있다. 현실의 사회보장은 정책과 행정 그리고 학술적인 측면에서 이미 행정데이터 시대로 접어들었다고 해도 과언이 아니다.

## 디지털정부, 데이터경제 그리고 사회보장 혁신

**데이터경제와 디지털정부의 도전** 우리나라에서 4차 산업혁명과 데이터경제는 다양한 모습으로 나타나고 있다. 크게 세 가지 정책 기조를 확인할 수 있다. 먼저, 디지털정부 또는 디지털뉴딜과 같은 정책이다. 여기서 디지털딤이나 디지털 정부는 과거 정부들 또한 표방해 왔던 구호들이다. 다음은 행정데이터 및 빅데이터 생산이 늘어남에 따라 데이터의 결합과 활용 욕구에 대응하기 위한 법제적 대응이다. ‘데이터 3법’<sup>1)</sup> 등을 통해 데이터 생산과 가공(결합) 그리고 활용을 촉

진하는 제도 정비가 그것이다. 끝으로 통계와 데이터의 품질을 높이기 위한 관리체계를 강화하는 것이다. 예를 들면, 다양한 기관이 생산하는 핵심통계에 대한 통계청 승인통계시스템이 그것이다.<sup>2)</sup>

**데이터경제와 사회보장의 혁명적 변화** 세계는 이미 4차 산업혁명의 시대, 데이터의 생산과 유통 그리고 활용이 다른 산업의 발전을 촉진하는 성장 동력이 되는 데이터경제 시대로 진입하였다. 빅데이터를 기반으로 기업은 소비자의 선택을 기다리는 것을 넘어 각 개인의 취향에 맞게 소비를 추천하기에 이르렀고, 국가 또한 국민이 신청하지 않아도 생애주기별로 각 개인의 욕구와 능력에 맞게 급여와 서비스를 제공하는 복지 개혁을 생각할 수 있게 되었다. 새로운 정보시스템이 종종 큰 혼란을 초래하기도 하지만, 21세기 복지국가의 근간이 되었다는 점은 분명하다. 지금 우리는 통합적 데이터에 기반한 ① 정확하고 ② 빠르며 ③ 안전한 정보시스템이 선택이 아닌 필수인 시대를 살고 있는 것이다.

## 사회보장 행정데이터, 어떻게 세상을 바꿀 것인가

**공공 부문의 데이터 결합과 활용의 부진** 우리

나라는 전 세계에서 공공과 민간 부문 정보시스템이 매우 잘 구축된 국가 중 하나이다. 그렇다고 해서 우리가 각종 데이터를 잘 생산, 활용하고 있다는 것은 아니다. 먼저, 공공 부문의 데이터 생산(결합 및 가공) 및 활용이 부진하다. 지난 20년간 공공 부문은 빠르게 정보시스템을 구축해 방대한 행정데이터를 축적해 왔다. 이는 주요 사회보장 집행 기관(즉, 사회보험공단과 한국사회보장정보원 등)의 행정데이터를 결합하면 전 국민을 대상으로 급여와 서비스를 자동으로 보장하는 생애주기별 사회보장체계 도입이 가능하다는 점을 말해 준다.

하지만 정작 전 국민 대상 데이터 프레임 구축이나 관련 사회보장 행정데이터의 결합은 아직 논의조차 되지 않고 있다. 더욱이 공공기관 간 행정데이터 결합에 대해서조차 지나치게 엄격한 규제와 복잡한 행정 절차를 적용함에 따라, 연구자는 물론이고 공공기관 간에도 데이터의 결합과 활용이 부진한 상황이다. 이어 민간 금융기관이나 정보포털업체 등의 각종 빅데이터 생산과 가공, 유통 등에 대한 정책이 미진하다. 은행이나 카드사 등 민간 금융기관은 국민의 소득이나 소비 등 정책적 효용 가치가 큰 정보를 보유하고 있다. 이 정보들은 코로나19 확산 단계에서 지역별

1) 데이터 3법이란 「개인정보 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」, 「신용정보의 이용 및 보호에 관한 법률」을 지칭하며, 2020년 1월 9일 이 3법의 개정안이 국회 본회의를 통과하였다.  
 2) 통계청은 2022년 10월 현재 498개에 이르는 정부 부처와 공공기관(통계 작성 기관)이 생산하는 1,298개의 통계를 승인통계로 지정하여 품질을 관리하고 있으며, 그중 75%(979개)가 월간 및 연간 단위 정기 생산 통계이다. 인구와 노동 등을 포함하는 사회 부문 승인통계는 241개이며, 보건과 복지 부문 승인통계는 91개이다(통계청, n.d.).

또는 업종별 매출 변화를 신속하게 파악하는 데 중요한 역할을 하였다. 하지만 이러한 민간 부문의 핵심 빅데이터 또한 생산과 활용, 특히 공공 부문 행정데이터와의 결합 등을 하는 데는 제약이 큰 것이 사실이다.

**행정데이터의 통합(결합)과 활용의 불가역성**  
 데이터의 역사는 반복된다. 역사적으로 통계학이 생성되던 시점의 데이터는 지금의 행정데이터와 유사했을 것으로 판단된다. 물론 데이터의 양과 수집 방법 등에서는 큰 차이가 있었겠지만, 당시의 데이터는 주로 국가 차원의 조세행정과 군사 행정을 통해 집계된 인구데이터와 재정데이터였다. 이는 각 개인에게 세금을 징수하고 병역을 부과하는 것을 목적으로 하였다. 이후 데이터가 축적되면서 분석 방법이 빠르게 발전하였다. 더불어 특정한 연구 목적에 따라 설계된 조사를 통한 데이터 생산 또한 빠르게 증가하였다. 이 조사들은 정치와 사회 부문에서 행정데이터로 파악하기 힘든 각 개인의 가치 선호나 각종 사안에 대한 만족도 등을 파악하는 데 효과적이었다. 하지만 공공 부문 정보시스템의 발달로 개인 단위에서 인구학적 정보, 고용정보, 소득·재산정보, 각종 사회보장 급여 정보 등이 결합되면서 행정데이터의 효용성이 크게 증가하기에 이른다. 21세기의 행정데이터가 전수 데이터의 강점과 함께 조사 데이터처럼 개인 단위에서 다양한 정보를 추가할 수 있다는 강점을 갖기에 가능해진 것이다.

### 보건복지 분야 데이터의 미래와 행정데이터

보건복지 분야에서 각종 조사는 문제를 발견하고 그것을 정책 의제로 만들어 해결을 촉구하는 역할을 해 왔다. 그리고 이렇게 정책화된 내용들은 다양한 형태의 행정데이터로 수집되고 분석된다. 그럼에도 행정데이터가 모든 질문에 답할 수는 없다. 하지만 행정데이터의 효용이 빠르게 커지고 있다는 점을 부인하기는 힘들다. 첫째, 보건복지 분야의 행정데이터는 개인 가구 단위에서 연계 및 통합되면 점점 더 큰 부가가치를 갖는다. 정책의 사각지대(배제 집단)를 해소하고 급여나 서비스의 부적정 수급을 해소하기가 더 쉬워지기 때문이다. 특히 생애주기별 사회보장을 가능하게 하는 중요한 토양이 된다. 둘째, 보건복지 분야의 행정데이터는 근거 기반의 정책 추진을 용이하게 한다. 행정데이터 분석은 전수 데이터에 대한 분석이라는 점에서 추정 과정에서의 오류를 최소화하기 때문이다. 셋째, 행정데이터는 분석 결과와 정책 방안 간의 거리를 좁히는 역할을 한다. 통상적으로 조사 데이터에 대한 분석 결과와 정책 제안은 실제 정책 집행 대상을 특정하지 않는다. 하지만 행정데이터에 대한 분석 결과로 도출된 대상 집단(신규 수급자나 탈락자)은 실제 정책 지원 대상이기 때문이다. 넷째, 보건복지 분야 데이터는 많은 급여 업무의 자동화를 통해 일선 복지 담당 공무원들이 시달리고 있는 서류 행정을 최소화하는 역할을 할 것이다.

## 데이터경제와 개인정보 보호

**개인정보 보호, 어디를 어떻게 통제해야 하는가**  
 21세기 국민들은 여전히 빅브러더(Big Brother)를 경계하고 있다. 국가가 개인정보를 과도하게 취합하여 사찰하고, 개인이나 기업 등이 개인정보를 탈취하여 신체에 위해를 가하거나 경제적 손실을 입히는 것에 대한 불안감이 있다. 그렇다고 디지털경제 하에서 행정데이터 활용의 경제적, 행정적, 사회적, 학문적 효용을 무시하기도 힘든 실정이다. 이 점을 고려할 때, 행정데이터의 이용을 활성화하는 동시에 개인정보를 효과적으로 보호할 수 있는 다른 대안이 필요하다. 그것은 다음 두 가지 주장을 어떻게 효과적으로 결합시킬 것인가 하는 점이다.

먼저, 개인정보 보호를 중심으로 행정데이터의 결합과 활용을 통제해야 한다는 주장이다. 행정데이터 생산부터 활용에 이르는 모든 단계에서 개인정보 유출을 예방하는 안전장치를 우선적으로 고려해야 한다는 것이다. 예를 들면, 데이터 생산 단계에서 과도한 정보 취합을 억제하도록 생명윤리규정을 적용하고, 데이터 결합 단계에서 개인정보를 엄격하게 비식별화하며, 데이터 유통 및 활용 단계에서 개인정보 제공을 금지하고 유

출 사고에 대해 엄하게 처벌하는 방식이다. 이는 현재의 데이터 관리 방식에 근접하는 것이다. 하지만 이 방식이 행정데이터의 결합과 활용을 촉진하는데 한계가 있다는 점도 분명해 보인다.<sup>3)</sup>

다음으로, 행정데이터의 활용에 초점을 두어 개인정보 보호 방안을 마련해야 한다는 주장이다. 정부와 이미 주민번호를 기반으로 행정을 집행하는 공공기관에 대해서는 엄격한 승인 장치를 전제로 주민번호에 기반한 데이터 결합과 활용을 허용하되, 데이터 개방 단계에서 개인정보를 비식별화하는 등 정보 보호 조치를 강화해야 한다는 것이다. 예를 들면, 정부 부처나 사회보험공단 또는 사회보장 집행 기관처럼 주민번호를 기반으로 행정을 해야 하는 기관에 대해서는 행정데이터의 결합과 가공, 정책연구를 위한 데이터 개방 측면에서 데이터 결합과 활용을 용이하게 하고<sup>4)</sup> 외부에 데이터를 개방하는 단계에서 개인정보 삭제나 비식별화를 강화하는 방식을 취하는 것이다.

행정데이터의 결합과 활용을 둘러싼 최근의 논의를 살펴보면, 생명윤리규정을 준수하며 개인정보를 효과적으로 보호할 수 있으며, 그와 더불어 데이터의 결합과 활용을 촉진할 수 있는 보다 실질적인 대안이 필요하다.

3) 행정데이터의 가명처리와 결합과정에서 행정적으로 너무 복잡하고, 긴 시간이 소요되며, 데이터 결합과정에서 적지 않은 소실이 발생하고 있으며, 이것이 데이터의 활용을 가로막는 장애요인이 되고 있다.

4) 사회보장 분야에서는 정부 부처(보건복지부), 사회보험공단, 한국사회보장정보원 등 주민번호를 기반으로 외부 데이터를 연계하는 업무를 수행하는 기관 간에는 데이터를 결합할 때 주민번호를 연계키로 활용하게 하고 외부에 데이터를 반출할 때는 모든 개인정보를 비식별화하도록 하는 내용의 별도 법률 제정이 필요하다.

## 맺으며

사회보장 데이터의 개방과 활용 문제는 향후 한국 사회보장의 혁신에 결정적인 영향을 미칠 것이다. 이는 사회보장정보시스템을 효율화하고, 복지 사각지대를 더욱 쉽게 발굴하며, 사회보장과 관련된 정책연구를 활성화하는 데 매우 중요한 사안임을 의미한다. 개인정보 보호에 대한 정부 정책이 어떤 방향으로 전개되는가에 따라 향후 사회보장 데이터와 통계의 발전은 다른 길을 걷게 될 개연성이 높다. 이러한 맥락에서 정책 결정에 앞서 개인정보 유출 등 침해 사건이 어떻게 발생하는지에 대한 현실적이고 객관적인 진단이 전제되어야 한다는 점을 강조하고자 한다. 지나친 이상주의는 때로 기대했던 것과 정반대의 결과를 초래할 수 있다. 罍

---

## 참고문헌

통계청. (n.d.). 승인통계 현황. [https://kostat.go.kr/portal/korea/kor\\_pi/2/1/index.action](https://kostat.go.kr/portal/korea/kor_pi/2/1/index.action)에서 2022. 10. 1. 인출.