

영국의 연구를 목적으로 한 행정 데이터 연계 활용

Linking and Using Administrative Data for Research in UK

한은희(한국사회보장정보원 부연구위원)

정부 부처 및 공공기관들이 소유한 다양한 행정 데이터는 연계를 통해 근거 기반 정책 수립 및 평가를 위한 연구에 활용될 수 있다. 영국은 개인의 정보인권을 보호하면서 행정 데이터를 연계하여 정책 연구에 활용할 수 있는 안전한 데이터 이용 기반을 만들고자 하는 뚜렷한 정책적 목적을 가지고 범부처적인 노력을 기울여 왔다. 2013년 영국은 잉글랜드, 스코틀랜드, 웨일스, 북아일랜드에 각각 행정 데이터 연구센터(ADRC: Administrative Data Research Center)를 설립하고 거시적인 행정 데이터 연구 네트워크(ADRN: Administrative Data Research Network)를 발족하였다. 이를 통해 안전한 환경에서 행정 데이터를 연계하고 효과적으로 활용할 수 있는 물리적 기반을 마련하고 행정 데이터를 둘러싼 다양한 주체들을 포괄하는 거버넌스 체계를 구축하였다. 현재 행정 데이터 연구 네트워크(ADRN)는 5년간의 활동을 마치고 영국 행정데이터 연구(ARD UK: Administrative Data Research UK)로 대체되었으며, 영국 행정데이터 연구(ADR UK)는 행정 데이터 결합과 활용을 지속 가능한 방식으로 혁신하기 위해 다양한 지원 활동을 하고 있다.

이 글에서는 영국의 행정 데이터 연계 및 활용 체계의 현황과 발전 과정을 검토하고, 우리나라의 보건·복지를 포함한 사회보장 행정 데이터의 연계 및 활용을 위한 시사점을 논하였다.

1. 들어가며

전 세계적으로 코로나바이러스감염증-19 발생 이후 보건복지 정책 연구에서 행정 데이터 활용의 중요성이 더욱 커지고 있다. 특히 대면 설문조사를 통한 데이터 수집이 어려워짐에 따라 행정 데이터는 설문조사 데이터를 대체하거나 보완하는 수단으로서 그 기능과 역할이 더욱 커질 것으로 예상된다. 행정 데이터는 연구 목적으로 만들어진 것이 아니며, 정부 부처 및 공공기관의 서비스 제공 및 운영을 위해 생산되고 축적된 데이터이다. 하나의 행정 데이

터는 그 자체로 정보로서의 가치가 있지만, 여러 정부 부처 또는 기관의 행정 데이터를 연계·결합(linkage, matching 또는 integration)하면 하나의 데이터 소스로는 알 수 없는 풍부한 정보를 생산할 수 있게 되어 연구용 데이터로서의 가치가 배가된다. 하지만 여러 기관이 소유한 행정 데이터를 효과적으로 연계하여 이용하기 위해서는 데이터 연계 및 활용을 위한 물리적 기반 마련과 행정 데이터 연계·활용을 위한 거버넌스 체계 구축 문제가 해결되어야 한다.

영국 정부는 2013년 행정 데이터를 둘러싼 주요 주체들이 참여하는 행정 데이터 연구 네트워크(ADRN: Administrative Data Research Network)를 발족해 여러 정부 부처 및 공공기관이 소유한 행정 데이터를 안전하게 결합하고 연구자들이 행정 데이터에 효율적으로 접근하여 공공의 목적을 위한 연구에 활용할 수 있도록 지원하기 위한 거버넌스 체계를 확립하였다. ADRN의 활동은 2018년 종료되었다. 그 후속으로 행정 데이터 연계 및 활용 방식의 혁신을 위해 영국 행정 데이터 연구(ADR UK: Administrative Data Research UK)를 구성하였는데, 이는 2021년까지 활동을 이어 갈 예정이다(ADR UK, 2019).

이 글에서는 영국의 ADRN과 영국 행정데이터 연구(ARD UK)의 설립 배경 및 역할, 그리고 영국의 행정 데이터 연계 및 활용 체계를 검토하고, 우리나라의 보건·복지 행정 데이터 연계 및 활용을 위한 시사점을 논하고자 한다.

2. 행정 데이터 연구 네트워크(ADRN) 및 영국 행정데이터 연구(ADR-UK)의 설립 배경 및 개요

가. 행정 데이터 연구 네트워크(ADRN) 설립 배경 및 개요

영국은 2010년 이후로 공공 데이터 개방 정책을 적극적으로 추진하여 공공 부문에서 생산된 거의 모든 데이터에 대한 검색과 접근이 가능한 오픈 데이터 포털 'data.gov.uk'를 운영하고 있다. 영국 정부의 공공 데이터 개방 정책은 단순히 정보를 국민에게 개방하는 것뿐만 아니라 공공 데이터를 정책 연구에 활용하여 공공의 이익에 기여하고자 하는 목적이 있다(DBIS, 2013). 하지만 2010년대 초반까지는 행정 데이터를 소유한 정부 부처 또는 기관마다 데이터 개방 원칙이 다르고 연구자가 데이터를 요청하여 받기까지 오랜 시간이 소요되는 데다 연구 목적의 행정 데이터 연계와 개방을 위한 법적 근거가 미비하여 행정 데이터 활용이 효율적이지 못했다. 이에 2011년 경제사회연구위원회(ESRC: Economic and Social Research Council)

주도로 행정 데이터 태스크포스(ADT: Administrative Data Taskforce)를 구성하여 부처 간 행정 데이터의 연계와 활용을 활성화하기 위한 방안을 마련하고자 하였다. 2012년 12월 행정 데이터 태스크포스(ADT)는 1년간의 활동을 바탕으로 보고서를 발간하였다. 이 보고서는 잉글랜드·스코틀랜드·웨일스·북아일랜드에 행정 데이터 연구센터(ADRC: Administrative Data Research Center) 설립, 연구자들이 행정 데이터에 접근할 수 있는 단일 창구 개설, 법제 정비, 행정 데이터의 연계 방식 표준화, 연계된 행정 데이터에 대한 접근 절차 및 자격 규정, 그리고 행정 데이터를 활용한 연구에 대한 재정적 투자 및 기술적 지원 등을 제안하였다(ADT, 2012). 행정 데이터 태스크포스(ADT)의 제안에 따라 영국 정부는 2013년 10월 잉글랜드, 스코틀랜드, 웨일스, 북아일랜드에 행정 데이터 연구센터(ADRC)를 설립하겠다고 발표하였으며, 경제사회연구위원회(ESRC)가 행정 데이터를 활용한 연구와 관련된 다양한 주체가 참여하는 거시적인 네트워크인 ADRN의 활동을 지원하기로 결정하였다(DBIS, 2013).

행정 데이터 연구 네트워크(ADRN)는 행정 데이터 사용을 원하는 사회과학 연구자들이 연구용 통합 행정 데이터에 접근할 수 있는 통합 창구 역할을 하며, 안전한 환경에서 행정 데이터가 연계되고 이용될 수 있도록 다양한 지원을 제공하기 위해 구성되었다. 행정 데이터 연구 네트워크(ADRN)는 코디네이터 역할을 하는 행정 데이터 서비스(ADS: Administrative Data Service), 4개의 행정 데이터 연구센터(ADRC), 데이터를 소유한 정부의 각 부처 및 기관들, 국가통계청(ONS: Office of National Statistics), 경제사회연구위원회(ESRC, 기금 제공자) 등으로 구성되었다. 행정 데이터 서비스(ADS)는 엑시식대학에 기반을 두고 연구자들의 행정 데이터 접근 요청에 대한 지원, 연구 제안서 적합성 심사, 데이터 보유 기관과의 협의, 데이터 접근 자격 부여 및 훈련 제공 등의 서비스를 제공하였다. 각 행정 데이터 연구센터(ADRC)는 하나의 파트너 대학¹⁾과 인근 지역 대학 및 연구기관과 지역적 네트워크를 형성하였다. 행정 데이터 연구센터(ADRC)의 역할은 연구자가 행정 데이터에 안전하게 접근할 수 있도록 보안시설을 제공하고 데이터 분석에 필요한 인프라 및 기술적 지원을 하는 것이다.

2018년 행정 데이터 연구 네트워크(ADRN)는 5년간의 활동을 평가하고 개선 사항을 제안하는 보고서를 발간하였다(Jones, Heys, Tingay, Jackson, & Dibben, 2018). 제안된 주요

1) 잉글랜드: 사우샘프턴대학교, 북아일랜드: 퀸스대학 벨파스트, 스코틀랜드: 에든버러대학교, 웨일스: 스와지대학교

개선 사항은 데이터 확보, 승인 과정, 데이터 접근 및 제공에 대한 컨트롤, 데이터 및 메타 데이터 관리, 연구 지원, 그리고 연계된 통합 데이터 유지(retention) 및 재활용 등이다. 이 중 가장 중요한 문제는 행정 데이터 연구 네트워크(ADRN)의 ‘데이터 생성, 파기’ 모델이 경제적으로 지속 가능하지 않다는 것이다. ‘데이터 생성, 파기’ 모델은 승인된 개별 연구를 위해 행정 데이터를 결합하여 통합 행정 데이터 세트를 생성하고, 분석이 끝난 후에는 통합 행정 데이터 세트를 파기하는 방식을 의미한다. 이 보고서는 지속적인 연구가 필요한 연구 주제들을 선정하여 통합 행정 데이터 세트를 구축하고, 안전한 환경에서 이 데이터 세트를 유지하면서 재활용할 것을 제안하였다.

나. 영국 행정데이터 연구(ADR UK)²⁾

영국 행정데이터 연구(ARD UK)는 경제사회연구위원회(ESRC)가 전국생산성투자기금(NPIF: National Productivity Investment Fund)으로부터 4400만 파운드의 기금을 받아 2018년부터 2021년 말까지 3년 동안 운영하는 영국연구혁신(UK Research and Innovation) 사업이다. ADR UK는 정책 연구 목적으로 행정 데이터가 연계되고 활용되는 방식을 혁신하는 것을 목적으로 한다. 행정 데이터 활용의 혁신은 정부의 주요 정책 프로그램들과 협력하여 정책 연구가 필요한 영역 및 주제들을 선정하고 이와 관련된 부처들의 행정 데이터를 결합하여 안전한 환경, 예를 들어 국가통계청(ONS)의 안전한연구서비스(SRS: Secure Research Service) 또는 웨일스의 안전한 가명화된 정보 연계(SAIL: Secure Anonymised Information Linkage) 데이터뱅크³⁾에 저장·유지하면서 승인된 연구자들이 통합 행정 데이터를 지속적으로 활용하는 것을 의미한다. 영국 행정데이터 연구(ARD UK)는 정부의 행정조직을 통해 이미 생성되고 있는 다양한 행정 데이터를 연계하여 상시적인 연구용 데이터 세트를 구축하고 승인된 연구자에게 정책 연구 목적으로 제공함으로써 근거 기반 정책 결정 및 공공서비스 전달체계 개선을 지원하고자 한다.

영국 행정데이터 연구(ARD UK)는 스코틀랜드, 웨일스, 북아일랜드의 행정데이터 연구(ADR)와 국가통계청(ONS)의 파트너십으로 구성되어 있으며, 경제사회연구위원회(ESRC) 내의 전략허브

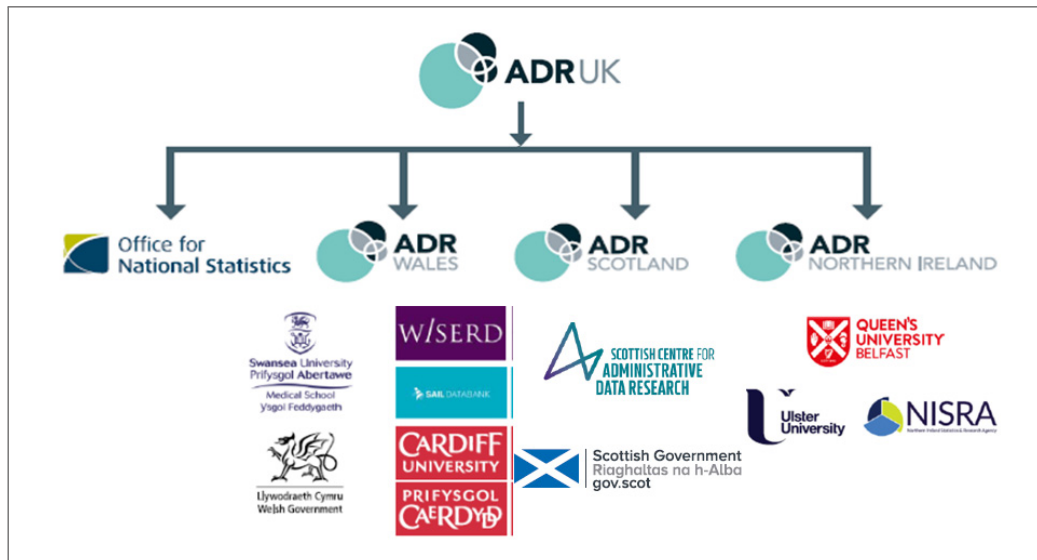
2) 영국 행정데이터 연구(ADR UK)에 대한 자세한 정보는 <https://www.adruk.org/> 참고.

3) 웨일스의 SAIL 데이터뱅크는 보건·의료 관련 연구를 위해 익명화된 개인 단위 데이터의 저장 및 이용을 제공한다(참조: <https://saildatabank.com/>).

팀이 파트너십 조정 역할을 하고 있다. 전략허브팀은 데이터의 안전한 연결을 위해 정부 부처 또는 기관 간의 협의, 행정 데이터를 활용한 연구 결과에 대한 학계 및 대국민 홍보, 연구 자금 관리 등의 사무국 역할도 담당하고 있다.

스코틀랜드·웨일스·북아일랜드의 행정데이터 연구(ADR)는 각각의 대학, 정부의 전문가 및 통계기관 등의 파트너십을 바탕으로 행정 데이터 연계와 데이터 분석을 수행하고 있다. 구체적으로 ADR 스코틀랜드는 정부의 데이터 공유 및 연계부(Data Sharing and Linkage Unit) 전문가들과 스코틀랜드 행정 데이터 연구센터(Centre for Administrative Data Research) 간의 파트너십이며, ADR 웨일스는 스완지대학교의 의과대학과 카디프대학교의 웨일즈 사회경제적 연구, 데이터 및 방법 연구소(WISERD: Wales Institute of Social and Economic Research, Data and Methods), 웨일스 정부의 통계학·경제학·사회과학 연구자들, 그리고 SAIL 데이터뱅크의 파트너십으로 이루어져 있다. 마지막으로 북아일랜드 행정데이터 연구(ADR NI(Northern Ireland))는 북아일랜드 행정 데이터 연구센터(ADR NI)의 퀸스대학 벨파스트 캠퍼스와 얼스터대학, 북아일랜드 통계·연구 에이전시(NISRA: Northern Ireland Statistics and Research Agency) 간의 파트너십으로 이루어져 있다.

그림 1. ADR UK 파트너 및 거버넌스



자료: ADR UK. (2019). Administrative data is an invaluable resource for public good. Let's use it. Annual Report 2018-19. Retrieved from https://www.adruk.org/fileadmin/uploads/adruk/Documents/ADRUK_Annual_Report_2018-19.pdf, /p. 12.

3. 행정 데이터 연계 및 활용과 관련된 법률

영국의 행정 데이터 연계 및 활용을 가능하게 하는 것은 하나의 법률이 아니며 각 영역 혹은 사업별로 다양한 법률들이다. 여기서는 정책 연구 목적을 위한 행정 데이터 연계와 활용에 직접적으로 영향을 미치는 영국의 「정보보호법(Data Protection Act 2018)」과 「디지털 경제법(Digital Economy Act 2017)」에 대해 간략히 살펴보고자 한다.

가. 정보보호법(Data Protection Act 2018)

영국의 개인정보 처리에 관한 기본 법률은 「정보보호법(Data Protection Act) 2018」이다. 「정보보호법」은 1984년 처음 제정되었으며, 이후 유럽연합(EU)의 개인정보보호지침(Data Protection Directive 95)이 제정되자 영국 내 이행 입법으로 1998년 개정되었다. 이후 2018년 5월 25일부터 EU의 개인정보보호지침을 대체하는 일반정보보호규정(GDPR: General Data Protection Regulation)⁴⁾이 시행됨에 따라, 영국은 「정보보호법 2018」을 제정하여 시행하고 있다. 「정보보호법 2018」은 일반정보보호규정(GDPR)의 기준을 대부분 그대로 적용하고 있기 때문에 영국이 EU를 탈퇴한 후에도 영국 내에서 일반정보보호규정(GDPR)이 실질적으로 적용될 것이다. 일반정보보호규정(GDPR)은 개인정보의 연계·결합은 정보 주체가 본인의 개인정보 처리를 동의하였거나 정보처리자의 법적 의무를 준수하기 위해 개인정보 처리가 필요하거나, 공익상 이유 또는 정보처리자의 공식 권한을 수행하기 위한 업무에만 가능하도록 규정하고 있다. 단, 공익적 기록 보존, 과학적·역사적 연구, 통계적 목적의 연계·결합은 예외 항목으로 두고 있으며, 이 경우에도 개인정보는 가명 처리 등 안전조치를 취하도록 하고 있다(89조 1항). 일반정보보호규정(GDPR)의 개인정보 처리 기본 원칙 6가지, 즉 ① 적법성, 공정성, 투명성 원칙(Lawfulness, fairness and transparency), ② 수집 목적 제한의 원칙(Purpose limitation Storage limitation), ③ 개인정보 최소화 원칙(Data minimization), ④ 정확성 원칙(Accuracy), ⑤ 저장 제한 원칙(Storage limitation), ⑥ 무결성 및 기밀성 원칙(Integrity and confidentiality)은 영국의 「정보보호법 2018」에 그대로 적용되고 있다.

4) 일반정보보호규정(GDPR)의 주요 항목은 사용자가 본인의 데이터 처리 관련 사항을 제공받을 권리, 열람 요청 권리, 정정 요청 권리, 삭제 요청 권리, 처리 제한 요청 권리, 데이터 이동 권리, 처리 거부 요청 권리, 개인정보의 자동 프로파일링 및 활용에 대한 결정 권리 등이다.

나. 디지털경제법(Digital Economy Act 2017)

「디지털경제법(Digital Economy Act 2017)」은 영국의 디지털 인프라 및 서비스에 대한 규정을 담고 있는 법률이다. 이 법의 제5부 전자정부의 5장은 행정 데이터의 ‘연구 목적의 공유’를 규정하고 있으며, 행정 데이터 연계 및 제공의 법률적 근거가 되고 있다. 구체적으로, 정부 기관이 보유한 행정 데이터는 연구 목적으로 제3자에게 제공될 수 있으며 그 정보가 개인정보일 경우, 다음과 같은 조건을 만족해야 한다. 첫째, 개인을 식별하는 정보가 포함된 경우 데이터가 공개되기 전에 개인의 신원을 추론하는 것이 합리적으로 가능하지 않도록 처리되어야 한다. 둘째, 데이터 처리에 관여하는 모든 사람은 개인을 식별할 수 있는 위험을 최소화하고, 의도적인 개인정보 유출을 방지하기 위한 합리적인 조치를 취해야 한다. 데이터를 활용한 연구는 (71조에 따라) 승인을 받아야 하며 데이터 처리에 관여한 사람, 데이터를 제공 받는 사람과 연구 목적으로 데이터를 이용하는 사람도 (71조에 따라) 승인을 받아야 한다. 71조는 행정 데이터 활용 승인에 대해 규정하고 있는데, 통계위원회(the Statistics Board)에서 승인을 담당하고 승인을 위한 조건을 수립하고 공개하도록 규정하고 있다.

4. 행정 데이터 이용 절차 및 연계 방식

가. 행정 데이터 이용 절차

영국 행정데이터 연구(ARD UK)와 스코틀랜드·웨일스·북아일랜드의 각 행정데이터 연구(ADR)는 행정 데이터 접근 및 이용에 대한 절차를 규정하고 있다. 이용하고자 하는 데이터의 소유 또는 제공 기관에 따라 세부 절차는 다를 수 있지만, 일반적으로 다음과 같은 절차를 따른다.

1단계: 연구 제안서 제출 및 승인

행정 데이터를 연구에 활용하고자 하는 연구자는 연구 주제와 내용을 논리적으로 요약·정리한 연구 제안서를 제출해야만 한다. 행정 데이터 이용 승인을 얻기 위해서는 연구의 결과가 어떻게 공익에 기여할 것인지를 명확히 기술해야 한다. 예를 들어, 제안하는 연구가 정부 정책을 어떻게 지원할 것인지, 또는 제안하는 연구가 일반 국민 또는 현장 실무자들의 이익에 어떻게 기여하는지 증명해야 한다(ADR Scotland, 2019).

연구 제안서 제출 외에도 연구윤리위원회(ethical review board)의 승인 및 개인정보영향평가(DPIA:

Data Protection Impact Assessment) 등 여러 표준화된 절차를 거쳐야 할 수 있다. 또한 접근하고자 하는 데이터 및 데이터가 저장되어 있는 장소에 따라 각기 다른 절차를 거쳐야 할 수 있다. 만약, ADR 웨일스에서 연계하여 안전한 가명화된 정보 연계(SAIL) 데이터뱅크에서 보유하고 있는 데이터에 접근하고 싶다면, 먼저 예비 신청서를 제출하고 안전한 가명화된 정보 연계(SAIL) 데이터뱅크 분석가와 협의하는 절차를 거친 후 정보 거버넌스 심의 패널(Information Governance Review Panel)에 신청서를 제출해야 한다. 스코틀랜드에서 보건의료 데이터에 접근하고자 한다면 보건 분야 공공이익과 정보보호 패널(PBPP: Public Benefit and Privacy Panel)에 연구 제안서를 제출하여 하며, 연구 제안서와 함께 필수 정보 보안 서류들, 데이터 공유 합의서, 그리고 데이터 관리자(controller)와 처리자(processor) 간의 계약서 등도 첨부해야 한다(ADR Scotland, 2019).

2단계: 트레이닝

행정 데이터에 접근하여 분석하기 위해 연구자는 데이터 보호/정보 거버넌스 트레이닝(data protection/information governance training)을 이수하고 이수 증명서를 제출해야 한다. 데이터 제공 기관에 따라 연구자에게 특정한 트레이닝 코스를 이수할 것을 요구할 수 있다. 예를 들어, 국가통계청(ONS)이 제공하는 데이터에 접근하기 위해서는 국가통계청(ONS)의 연구자 정보보안훈련(Safe Researcher Training)을 이수해야 한다. 스코틀랜드에서 보건·의료 데이터에 접근하기 위해서는 국가보건의료서비스 스코틀랜드 정보 거버넌스(NHS Scotland Information Governance) 온라인 코스 또는 의료연구위원회(Medical Research Council)의 연구의 일반정보보호규정과 비밀유지(Research, GDPR and Confidentiality)* 온라인 코스를 이수해야 한다(ADR Scotland, 2019).

3단계: 데이터 확정과 관련 데이터 센터와의 협의

어떤 데이터를 연구에 활용할 것인지 탐색하고 확정하는 것은 시간이 많이 걸리는 작업이다. 연구자는 연구 제안서를 제출하기 전에 데이터에 대한 탐색해 볼 수 있다. 국가의료서비스 스코틀랜드(NHS NSS: NHS National Service Scotland)의 국가 데이터 카탈로그(National Data Catalogue), 인구 센서스 데이터 및 센서스 기반의 횡단 데이터, 아동 및 사회적 돌봄 관련 데이터들을 사전에 탐색해 볼 수 있다. 만약 행정 데이터 연계·결합 프로젝트를 기획 중이라면, UK 데이터 서비스(Data Service), 국가통계청(ONS), 안전한 가명화된 정보 연계(SAIL) 데이터뱅크, 북아일랜드 통계·

연구 에이전시(NISRA) 등 연계 서비스를 제공하는 기관들의 정보를 검색해 볼 수 있다(ADR Scotland, 2019).

또한 개인 또는 가구 기반의 행정 데이터를 연계하여 사용하고자 한다면, 일반정보보호규정(GDPR)의 규정에 따라 각 데이터 소유 기관의 관리자에게 개별적으로 사용 승인을 얻어야 한다. 예를 들어 스코틀랜드의 경우, 데이터 관리자들은 국가의료서비스 스코틀랜드(NHS NSS), 국가기록스코틀랜드(NRS: National Records Scotland), 교육분석서비스(EAS: Education Analytical Services), 영국 데이터 서비스(UK Data Service), 스코틀랜드 경찰청(Police Scotland/Scottish Police Authority) 그리고 지방정부들이다(ADR Scotland, 2019).

4단계: 데이터 접근과 분석

연구용으로 결합된 행정 데이터는 대개 안전한 환경, 소위 정보 보안 시설인 세이프 헤이븐(safe haven)이라고 불리는 곳을 통해 제공된다. 세이프 헤이븐(safe haven)은 보안 설비가 되어 있으며 외부와 인터넷이 연결되어 있지 않고 오직 승인된 데이터에 대한 접근과 분석만이 가능한 플랫폼을 의미한다. 예를 들어, 스코틀랜드에는 5개의 지역 세이프 헤이븐(safe haven)이 있으며, 이 중 에든버그의 국가세이프헤이븐(National Safe Haven)은 원격 접속이 가능하다. 하지만 접근하려는 데이터 또는 데이터 제공자에 따라 원격 접속이 제한될 수도 있다. 이런 경우에는 물리적으로 외부와 차단된 공간에 직접 가서 분석을 수행해야 한다(ADR Scotland, 2019).

5단계: 결과물에 대한 점검

데이터 분석의 최종 결과물은 개인을 식별할 수 있는 정보를 포함하고 있는지 검증 절차를 걸친다. 이 과정을 통계적 노출관리기법(SDC: Statistical Disclosure Control)이라고 하며, 통계적 기법을 사용하여 분석 결과물에서 개인을 식별할 수 없게 만드는 것이다. 데이터 관리자와 처리자는 개인정보를 활용하여 연계된 데이터의 분석 결과물에서 개인을 식별하지 못하도록 조치를 해야 하는 의무가 있기 때문에 스코틀랜드의 국가세이프헤이븐(National Safe Haven)은 전자정보연구혁신서비스(eDRIS)팀⁵⁾의 두 명이 모든 행정 데이터를 활용한 분석의 최종 결과물을 체크하

5) 전자정보연구혁신서비스(eDRIS: electronic Data Research and Innovation Service)팀은 국가의료서비스 스코틀랜드(NHS NSS)의 정보서비스국에 소속되어 있으며, 스코틀랜드 내에서 수행되는 대부분의 행정 데이터를 활용한 연구를 지원하고 있다.

도록 하고 있다(ADR Scotland, 2019).

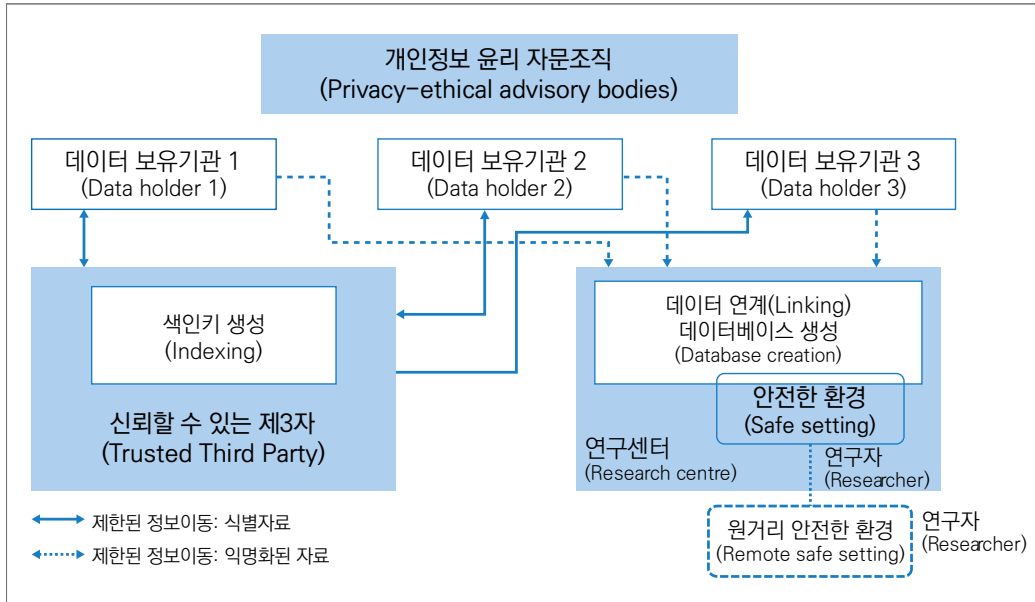
나. 행정 데이터 연계 방식

영국의 행정 데이터 연계에서 최우선적으로 고려되었던 것은 데이터의 유출이 일어나거나 개인의 신상이 노출되지 않도록 하는 것이다. 서로 다른 정부 기관들이 보유하고 있는 행정 데이터를 결합하기 위해서는 개인식별정보를 활용한 매칭이 불가피하지만, 매칭된 정보는 개인식별정보 제거 후 연구자에게 제공된다. 2011~12년 활동했던 행정 데이터 태스크포스(ADT)는 2012년 보고서에서 다양한 방식의 행정 데이터 연계를 검토한 후 신뢰할 수 있는 제3자(TTP: Trusted Third Party)를 통한 연계 방식을 제안하였다(ADT, 2012).

신뢰할 수 있는 제3자(TTP)를 활용한 데이터 연계 방식은 [그림 2]와 같다. 먼저 개별 행정 데이터 보유 기관들은 개인식별정보와 연구용 정보를 분리한 후 각 레코드에 고유참조번호를 부여한다. 각 데이터 보유 기관들은 고유참조번호와 개인식별정보만을 신뢰할 수 있는 제3자(TTP)에 보내고, 신뢰할 수 있는 제3자(TTP)는 각기 다른 기관으로부터 받은 개인식별정보를 활용하여 공통적인 색인 키(Index Key)를 생성한다. 신뢰할 수 있는 제3자(TTP)는 개인식별정보를 삭제하고 각 기관이 보내 온 고유참조번호와 신뢰할 수 있는 제3자(TTP)가 생성한 색인 키만을 각 데이터 보유 기관으로 보낸다. 데이터 보유 기관은 색인 키와 연구용 정보를 결합하여 데이터분석센터에 보낸다. 데이터분석센터의 연구자는 각 보유 기관으로부터 받은 데이터 세트를 색인 키를 활용하여 결합할 수 있다.

신뢰할 수 있는 제3자(TTP)를 통한 결합 방식은 결합 과정에 관여하는 모든 주체들이 부분 정보만을 다루게 함으로써 개인정보가 식별될 수 있는 위험을 최소화하고 프라이버시 침해 를 방지할 수 있다. 신뢰할 수 있는 제3자(TTP)는 데이터 연계·결합을 위한 보안시설을 완비하고 있는 국가통계청(ONS), 스코틀랜드의 안전한 가명화된 정보 연계(SAIL) 데이터뱅크 또는 북 아일랜드 통계·연구 에이전시(NISRA)가 될 수 있다.

그림 2. 신뢰할 수 있는 제3자(TTP: Trusted Third Party)를 활용한 데이터 연계 방식



자료: ADT(Administrative Data Taskforce). (2012). The UK Administrative Data Research Network: Improving access for research and policy. Economic and Social Research Council, Medical Research Council, and Wellcome Trust. Retrieved from <https://esrc.ukri.org/files/research/administrative-data-taskforce-adt/improving-access-for-research-and-policy/p.39>

5. 영국 행정데이터 연구(ADR UK)의 행정 데이터 연계 및 활용 사례

영국 행정데이터 연구(ARD UK)는 고용, 주거, 보건, 아동, 노인, 불평등 및 사회적 배제(social exclusion), 기후와 환경, 범죄 등 다양한 영역의 행정 데이터를 연계하는 데이터 연계 사업과 이미 구축된 통합 행정 데이터 세트를 활용하는 연구를 지원하고 있다. 여기서는 영국 행정데이터 연구(ARD UK) 및 스코틀랜드·웨일스·북아일랜드 행정데이터 연구에서 지원하거나 수행 중인 보건·복지 관련 행정 데이터 연계 및 분석 프로젝트 중 일부를 발췌·요약하여 소개하고자 한다.⁶⁾

6) <https://www.adruk.org/our-work/> 참고.

가. 올 이어즈(All Years) 데이터 세트

‘올 이어즈(All Years) 데이터 세트’는 영국 행정데이터 연구(ARD UK)와 국가통계청(ONS), 유니버시티 칼리지 런던(UCL: University College London), 국가의료서비스 디지털(NHS Digital), 그리고 교육부가 합동으로 수행 중인 프로젝트이며, 교육부 시스템에 포함된 아동의 정보와 병원 기록을 연계하여 약 1600만 명을 포함하는 데이터 세트를 구축하고자 한다. 이 프로젝트는 국가 학생 데이터(NPD: National Pupil Database)와 병원 에피소드 통계(HES: Hospital Episode Statistics)의 데이터를 연계한다. NPD는 잉글랜드 내의 학교와 대학에 소속된 학생들에 대한 다양한 정보를 포함하고 있으며, 병원 에피소드 통계(HES)는 잉글랜드 국가의료서비스(NHS) 병원의 외래 기록, 사고 및 응급 서비스(Accident and Emergency Attendance) 기록, 입원 기록을 포함하고 있다. 국가 학생 데이터(NPD)와 병원 에피소드 통계(HES)의 데이터 연계는 사회과학 분야와 보건 분야의 융합 연구를 가능하게 할 것이며 아동의 건강과 교육, 그리고 삶의 질 향상을 위한 정책 지원을 가능하게 할 것으로 기대된다. ‘올 이어즈(All Years) 데이터 세트’는 연구용 데이터 세트로 만들어져 국가통계청(ONS) 안전한 연구서비스(SRS)를 통해 개방될 예정이며, 이 데이터를 이용하기 위해서는 연구 제안서를 제출하고 승인을 받아야 한다.

나. 노숙인 데이터 잉글랜드(Homelessness Data England)

‘노숙인 데이터 잉글랜드(Homelessness Data England)’ 프로젝트는 지방정부가 가지고 있는 노숙인 케이스 레벨 정보(H-CLIC: Homelessness Case Level Information Collection)와 고용, 교육, 사회복지, 보건·의료 등 여러 분야의 행정 데이터를 연계하여 노숙인에 대한 통합 행정 데이터를 구축하는 것이다. 2017년 노숙인감소법(2017 Homelessness Reduction Act)은 노숙자 방지를 위해 지방정부들의 초기 개입을 의무화하였으며, H-CLIC는 이 법에 따른 지방정부의 여러 조치들에 대한 정보를 포함하고 있다. 노숙인 케이스 레벨 정보(H-CLIC)와 지방정부 내 타 기관의 데이터를 연계함으로써 노숙인 현황을 포괄적으로 이해하고 어떤 개입이 효과적이었는지 평가할 수 있게 됨에 따라 장기적으로 노숙인 문제를 해결하기 위한 시사점을 얻을 수 있을 것으로 기대된다. 영국의 통계청(ONS)과 주택·지역사회·지방정부부(MHCLG: Ministry of Housing, Communities & Local Government)의 협업을 통해 통합 행정 데이터가 구축되고 있으며, 개인정보가 제거된 연구용 데이터는 국가통계청(ONS)의 안전한 연구서비스(SRS)를 통해 승인된 연구 프로젝트와 승인된 연

구자에게 제공될 예정이다.

다. 데이터퍼스트(Data First)

‘데이터 퍼스트(Data First)’는 영국의 법무부 내 여러 행정 데이터를 연계할 뿐 아니라 다른 정부 부처 데이터의 연계와 활용을 통해 교정 정책을 개선하기 위한 프로젝트이다. 법무부(MoJ: Ministry of Justice) 내부 데이터 연계는 왕립교정보호청(HMPPS: Her Majesty’s Prison and Probation Service)와 법원행정처(HMCTS: HM Courts and Tribunals Service)의 행정 데이터를 포함하여 민사, 가족, 형사 소송과 관련된 모든 데이터를 연계하는 것이다. 또한 법무부 데이터와 교육부 및 다른 행정 부처의 데이터 연계를 통해 법무부 행정 데이터의 활용 범위를 확대하고자 한다. ‘데이터퍼스트(Data First)’ 프로젝트를 통해 민사소송과 형사재판의 교집합에 속하는 사람들은 어떤 인구학적 특징을 가지는지, 민사재판을 반복적으로 이용하는 사람들은 어떤 특징이 있는지, 교육 정도와 범 죄는 어떤 관계가 있는지, 가정폭력 가해자 중 아동의 면접교섭 관련 가정법원 케이스에 관여되어 있는 사람은 얼마나 되는지와 같은 다양한 주제의 연구가 가능해질 것으로 기대된다. 개인정보가 제거된 연구용 데이터 세트는 국가통계청(ONS)의 안전한 연구서비스(SRS)를 통해 제공될 예정이며 데이터를 활용하기 위해서는 연구 제안서를 제출하고 승인을 받아야 한다.

라. 북아일랜드 아동보호 서비스 성과(Outcomes for looked after children in Northern Ireland)

이 프로젝트는 북아일랜드 행정데이터 연구(ADR)와 퀸스대학 벨파스트 소속의 에이든 맥과이어 박사가 수행하며 아동기의 아동보호 서비스 경험이 성인기의 건강 및 사회적 결과에 어떤 영향을 미치는지 탐구하고자 하는 목적을 가진다. 이 연구 프로젝트는 지난 30년간 북아일랜드의 사회서비스 데이터(SOSCARE: Social Services Care Administrative and Records Environment, 1985-2015)와 센서스 데이터, 약물 처방 데이터, 병원 데이터, 자해 등록 데이터, 사망 정보를 결합하여 분석에 활용할 예정이다. 주요 분석 과제는 ① 지난 30년간 아동보호 서비스 대상의 특성 변화, ② 아동보호 서비스 대상의 신체적·정신적 건강을 포함한 다양한 장기적 결과(예를 들어, 향정신성 의약품 사용, 정신병원 입원, 자해, 고용

상태, 교육 정도, 자살로 인한 사망 등), 그리고 서비스 이용 사유와 서비스 종류에 따른 차이, ③ 서비스 기간 및 이행(transition)의 효과 등이다.

마. 약물 관련 사망과 응급 서비스 이용

이 연구는 행정데이터 연구(ADR) 스코틀랜드와 에든버러 대학의 수전 맥비 교수가 수행하며 국가 약물 관련 사망 데이터베이스(NDRDD: National Drug-Related Deaths Database)와 앰블런스 호출 정보 그리고 경찰청 정보를 연계하여 약물로 인한 사망 이전 응급실 사용 패턴을 분석하고자 한다. 스코틀랜드는 약물 관련 사망자 수가 급속히 증가해 왔을 뿐 아니라 현재 EU 국가 중 약물 관련 사망자 수가 가장 높다. 이에 스코틀랜드 정부는 2019년 9월 이 문제를 해결하기 위한 태스크포스를 구성하였으며, 정책 지원을 위해 약물 관련 사망자 통합 행정 데이터를 구축 중이다. 국가 약물 관련 사망 데이터베이스(NDRDD)로부터 추출한 2015~2018년 코호트와 국가의료서비스 스코틀랜드(NHS NSS)의 비예약진료데이터마트(Unscheduled Care Datamart)로부터 추출한 앰블런스 호출 및 긴급 서비스 데이터, 스코틀랜드 경찰청으로부터 추출한 사건 기록 데이터를 연계하여 분석할 예정이다.

6. 나가며

최근 우리나라에서도 근거 기반 정책 수립과 평가를 위한 행정 데이터 연계·활용의 필요성과 중요성에 대한 인식이 확산되고 있다. 하지만 각기 다른 정부 기관이 소유하거나 관리하고 있는 행정 데이터를 연계하고 활용하기 위해서는 안전한 환경의 구축뿐 아니라 각기 다른 목적과 데이터 활용의 근거를 가진 행정부처와 기관들 간의 이해관계를 조정하여 협조를 이끌어 내기 위한 거버넌스 구축이 필요하다. 이를 통해 연구자들이 데이터에 효율적으로 접근하여 효과적으로 데이터를 활용할 수 있도록 절차 및 규정을 마련하고, 데이터 접근과 활용을 활성화하기 위한 재정적·기술적 지원이 중장기적으로 제공되어야 할 것이다. 이런 점에서 행정 데이터를 둘러싼 주요 주체들이 참여하는 거시적 네트워크인 ADRN과 영국 행정데이터 연구(ARD UK)를 구축하고 이를 통해 신뢰할 수 있는 행정 데이터 연계와 활용 체계를 구축한 영국의 사례는 우리에게 많은 시사점을 준다.

우리나라는 최근 「개인정보 보호법」 개정을 통해 통계 작성·과학적 연구·공익적 기록 보

존 등의 목적을 위해 정보 주체의 동의 없이 개인정보를 가명 처리하고 가명 처리된 데이터를 결합할 수 있는 법적 근거를 마련하였다. 하지만 여전히 보건·복지를 포괄하는 사회보장 분야의 여러 부처 및 공공기관이 소유하고 있는 행정 데이터를 결합하여 활용하기 위한 국가적인 로드맵은 뚜렷하게 보이지 않고 있으며, 연구자들이 정책 연구를 목적으로 행정 데이터에 접근하거나 활용할 수 있는 기회는 매우 제한적이다. 영국은 정부 및 공공기관이 보유하고 있는 행정 데이터를 정책 연구에 활용하여 공공의 이익에 기여하고자 하는 뚜렷한 비전을 가지고 있었으며, 데이터의 연계와 활용 과정에서 개인정보 노출이나 사생활 침해가 일어나지 않거나 최소화되어야 한다는 원칙하에 데이터의 연계와 활용을 활성화하기 위한 로드맵을 마련했다. 그리고 지난 10년간 그 로드맵에 따라 체계적이고 범부처적인 노력을 기울여왔다. 우리 정부도 사회보장 행정 데이터 연계와 활용에 대한 국가적 비전을 확립하고 사회보장 행정 데이터 연계와 활용에 관한 로드맵을 완성해 나가야 할 것이다. 또한 사회보장 행정 데이터 연계와 활용을 위한 거버넌스 체계 구축, 법·제도 정비, 물리적 인프라 구축, 연계된 통합 행정 데이터의 유지 및 활용을 위한 행정적·재정적 지원 등이 실질적으로 뒤따라야 할 것이다. 마지막으로 정보 주체인 국민에게 모든 절차와 과정, 결과를 투명하게 공개하고 국민의 정보인권이 침해되지 않도록 해야 할 것이며 데이터 활용을 통해 실질적으로 국민의 삶의 질 향상에 기여할 수 있어야 할 것이다.

참고문헌

- ADR Scotland. (2019). Research handbook: Data linkage and administrative data research in Scotland. Scottish Center for Administrative Data Research. Retrieved from <http://www.scad.ac.uk/sites/default/files/Researcher%20access%20handbook%20Nov%202019%20-%20SCADR.pdf>
- ADR UK. (2019). Administrative data is an invaluable resource for public good. Let's use it. Annual Report 2018-19. Retrieved from https://www.adruk.org/fileadmin/uploads/adruk/Documents/ADRUK_Annual_Report_2018-19.pdf
- ADT(Administrative Data Taskforce). (2012). The UK Administrative Data Research Network: Improving access for research and policy. Economic and Social Research Council, Medical Research Council, and Wellcome Trust. Retrieved from <https://esrc.ukri.org/files/research/administrative-data-taskforce-adt/improving-access-for-research-and-policy/>
- DBIS(Department for Business, Innovation and Skills). (2013). Seizing the data opportunity: A strategy for UK data capability. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/254136/bis-13-1250-strategy-for-uk-data-capability-v4.pdf
- Jones, K. H., Heys, S., Tingay, K. S., Jackson, P., & Dibben, C. (2018). The Good, the Bad, the Clunky: Improving the Use of Administrative Data for Research. *International Journal of Population Data Science*, 4(1):586. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7479922/>