

근거 기반 정책 지원을 위한 연구 데이터 관리와 개방



Research Data Management and Openness to
Support Evidence-based Policies

정영철 | 한국보건사회연구원 연구위원

근거 기반 연구, 데이터 중심 연구가 부각하고 정보기술이 발전함에 따라 공적 자금으로 만들어진 연구 성과 과정을 개방하고 공유하여 데이터의 가치 제고와 융복합 연구 성과 창출을 추구하는 ‘오픈 사이언스’ 운동, ‘오픈 사이언스 정책’이 전 세계적으로 확산되고 있다. 우리나라도 이에 부응하여 과학기술 연구에서 데이터 관리 계획을 도입하여 추진하고 있으며, 현 정부의 ‘한국판 뉴딜 정책’에서도 ‘데이터 수집·가공·거래·활용 기반 구축을 통한 경제 가속화(데이터 댐)’ 과제를 10대 대표 과제 중 하나로 선정하는 등 공적 자금으로 만들어진 연구 결과를 활용에 대한 필요성은 더욱더 커지고 있다.

그러나 과학기술계에 비해 경제인문사회계 연구 부문에서는 연구 데이터 개방·공유의 필요성, 중요성이 공론화되지 않고 있으며 그 현황조차 파악이 안 되고 있다.

그러므로 연구 데이터 통합과 재사용에 따라 새로운 지식과 가치가 창출되는 ‘오픈 사이언스’, ‘오픈 데이터’의 의미를 부여하기 위해서는 우선 경제인문사회계 연구기관들의 연구 데이터 관리 및 공유·활용 체계 현황과 이들 기관 연구자들의 인식을 파악하고 이를 토대로 경제인문사회계 연구기관의 실정에 맞는 가이드 라인을 제시할 필요가 있다.

1. 들어가며

정보기술은 나날이 발전하여 데이터 폭발에 이르고 있으며, 방대한 양의 데이터가 축적됨에 따라 이제 데이터는 단순히 현 사회를 파악하고

진단하는 데 사용될 뿐만 아니라 미래 사회를 전망하고 예측하는 데 사용되기도 한다.

축적된 데이터와 이를 활용한 정보기술의 발전은 정책 수립 영역에서도 새로운 부가가치와 수요를 창출하는 원동력으로 작용하고 있다. 이

에 따라 과거의 통념·관행에 의존했던 의견 기반 정책(Opinion-based Policy) 수립에서 나아가 객관적·과학적 증거를 활용한 근거 기반 정책(Evidence-based Policy) 수립의 중요성과 필요성이 점차 커지고 있다(한국정보화진흥원, 2015).

근거 기반 정책은 관련 통계, 연구 및 평가 결과 등을 통해 이루어지는 의사 결정으로서 과정의 투명성과 내용의 신뢰성·설득력을 얻게 된다(오세영, 윤진, 오균, 2017). 근거 기반 사상은 1990년대 초반 의료 영역에서 처음 도입된 이후 사회복지, 교육, 형사 사법 등 여러 분야의 정책 수립 및 실천 분야로 확대되었다(한국정보화진흥원, 2015). 나날이 발전하는 정보기술을 통해 축적된 연구 데이터의 효용성 및 가치 향상과 더불어 연구의 투명성과 재현성, 투자 효율성 확보 등을 위해 연구의 성과와 과정을 개방하고 공유하는 ‘오픈 사이언스(Open Science)’ 운동, ‘오픈 사이언스 정책(Open Science Policies)’이 전세계적으로 확산되고 있는 가운데 객관적이고 과학적인 데이터에 기반한 근거 기반 사상과 이의 핵심 요소 중 하나인 연구 데이터의 투명성과 재현성을 추구하는 ‘오픈 사이언스’ 사상은 일맥상 통한다고 할 수 있다.

한편 정부는 한 나라의 경제, 사회, 과학기술 등 모든 영역에 걸쳐 공공 문제를 해결하기 위한 방침, 즉 ‘정책’을 펼치고 있으며 이를 위한 싱크 탱크가 정부출연연구기관이다. 정부출연연구기관은 정부가 재원을 출연하여 설립한 기관으로, 우리나라에서는 한국개발연구원, 한국보건사회

연구원 등을 시작으로 50여 년 동안 다양한 분야별 정부출연연구기관들이 설립되었다.


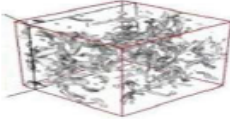

정부출연연구기관은 과학기술계 기관과 경제인문사회계 기관으로 나눌 수 있다. 과학기술계 기관의 기능은 연구·개발이 우선인 데 비해 경제인문사회계 기관의 기능은 정책 연구를 통한 정책 개발이 우선이다(김용훈, 오영균, 2008). 정부출연연구기관의 정책 연구는 근거 기반 정책을 수립하고 분석·평가하기 위한 것으로 근거에 기반한 연구 혹은 데이터 중심 연구가 필수 요소로 자리매김하고 있다. 최근 국가·사회문제가 복잡화, 다양화됨에 따라 이를 해결하기 위한 공공 연구의 역할 확대가 요구되고(이민형, 김태양, 2020) 있는 가운데, 이 글에서는 정부출연연구기관의 핵심 자산이라 할 수 있는 연구 데이터 관리와 개방의 필요성, 그리고 이를 위한 과제들을 살펴보고자 한다.

2. 오픈 사이언스와 연구 데이터

연구자들은 연구를 수행할 때 많은 데이터를 수집하고 생산하며, 이를 가공하고 분석하여 정보와 지식을 얻고, 이를 토대로 근거에 기반한 연구 결과를 제시한다.

이러한 연구는 자연현상을 기술하는 실험 중심의 제1세대 연구에서 비롯하여, 수학적 모델과 같은 이론 중심의 제2세대 연구, 복잡한 현상을 모의실험하는 것과 같은 컴퓨팅 자원 중심의 제3세대 연구를 거쳐 오늘날에는 기존의 실험, 이론,

그림 1. 연구 패러다임의 변화

1세대 연구	2세대 연구	3세대 연구	4세대 연구
실험 중심연구	이론 중심연구	컴퓨팅자원 중심연구	데이터 중심연구
	$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$		

자료: Hey, T., Tansley, S., & Tolle, K. (2009). Jim Gray on eScience: A Transformed Scientific Method. The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research. p. xviii. 국가과학기술연구회. (2019c). 출연(연) 연구데이터 관리·활용 방안 연구. p. 1에서 재인용.

모의실험 등을 통합하는 데이터 탐색 중심의 제4세대 연구 패러다임으로 변화하고 있다(그림 1). 이는 데이터를 통해 기존의 이론, 실험, 모델과 다른 새로운 이론과 가설이 만들어지는 것으로 (Hey, Tansley, & Tolle, 2009), 최근의 머신러닝, 인공지능 등과 같은 빅데이터 기술로 실현되고 있다. 특히 정보기술이 발달함에 따라 연구 데이터의 양적 팽창과 다양한 분야에서의 융합 가능성이 점증하고 있으며 이에 따른 데이터 관리와 활용에 대한 이슈가 부각되고 있다.

특히 학문의 자유를 위하여, 공적 자금 투자에 대한 사회경제적 편익을 제고하기 위하여 2000년 전후 학계를 중심으로 세계 각국에서는 연구에 대한 성과 공개에서 시작된 오픈 사이언스 운동이 확산되고 이에 대한 정책이 추진되고 있으며, 이는 민간 부문보다는 공공 정책 영역으로 자리 잡아 가고 있다(신은정 외, 2017). 오픈 사이언스가 처음 등장했을 때는 연구에서의 자율적 규범을 일컫는 개념으로서 기존 지식을 공개·공유하여 검증하고 논리적으로 반박함으로써 과학이 발전하는 것으로 여겨졌다(Merton, 1942/1973;

신은정, 안형준, 정원교, 2016에서 재인용). 이후 정보기술 발달에 따라 인터넷 보급률이 높아지고 디지털 연구 자원이 급격하게 축적되자 2000년대 초반부터 경제협력개발기구(OECD)를 중심으로 글로벌 정책 어젠다로 등장하게 되었다(신은정, 2015).

이후 세계 각국에서 공공 정책 영역으로 다루어지면서 ‘오픈 사이언스’는 “공공 연구의 성과물(출판물 및 데이터)을 디지털 형태로 공개 및 확산시켜 사회경제적 편익을 제고하려는 일련의 노력”, 이에 대한 공적 개입은 ‘오픈 사이언스 정책’이라고 정의되었다. 오픈 사이언스는 정보기술이 발전함에 따라 연구 출판물에 대한 ‘오픈 액세스(Open Access)’, 연구 데이터 및 소재에 대한 접근성을 강조하는 ‘오픈 데이터(Open Data)’, 개방형 연구 모델을 가능하게 하는 ‘오픈 콜라보레이션(Open Collaboration)’까지 포괄하게 되었다(OECD, 2015; 신은정 외, 2016에서 재인용).

초기 오픈 사이언스 운동의 대상이었던 연구 보고서, 논문 등과 같은 연구 결과물에서 점차 연구 데이터, 소프트웨어, 연구 수단 등과 같이 연

표 1. 주요 선진국들의 초창기 오픈 사이언스 정책 동향

국가 및 기관		오픈 사이언스 정책
유럽연합		- 부다페스트 오픈 액세스 선언(BOAI: Budapest Open Access Initiative)(2002) - 사이언스 2.0(Science 2.0): 과도기 과학(Science in Transition)(2014) - 오픈 이노베이션, 오픈 사이언스, 오픈 세계(Open Innovation, Open Science, Open to the World): 유럽 비전(A Vision for Europe)(2016)
핀란드		- 연구 데이터 이니셔티브(2009) - 핀란드 오픈 사이언스 및 연구 선언(Finish Open Science and Research Initiative(ATT)(2014)
네덜란드		- 오픈 사이언스에 관한 암스테르담 행동 촉구안(Amsterdam Call for Action on Open Science)(2016) - 오픈 사이언스 국가계획(National Plan Open Science)(2017)
미국	연방정부	- 공공 액세스에 관한 과학기술정책국 메모(OSTP Memorandum on Increasing Access to the Results of Federally Funded Scientific Results)(2013) - 미국 오픈 데이터 국가실천계획(US Open Data Action Plan)(2014)
	국립보건원(NIH)	- 공공 액세스 정책(Policy on Enhancing Public Access to Archived Publication Resulting from NIH-Funded Research)(2005), 약칭 Public Access Policy
일본	정부	- 일본 오픈 사이언스 활성화 방안(Promoting Open Science In Japan: Opening Up a New Era for the Advancement of Science)(2015)
	일본과학기술원(JST)	- 공공 자금으로 지원된 연구의 성과에 대한 오픈 액세스 방침(2013) - 오픈 사이언스 추진을 위한 연구 성과 취급에 관한 기본지침(2017)

자료: 신은정, 정원교. (2017. 8. 8.). 오픈 사이언스 정책의 확산과 시사점. STEPI Insight, 제216호, 1-39.

구 과정 중 생산된 성과와 자원에 대한 공개를 요구하는 오픈 데이터 정책으로 확대되었는데 이는 연구 데이터의 공개 자체가 최종 목표가 아니라 준비 과정과 협의를 거쳐 활용·재활용 가능한, 즉 연구 데이터의 부가가치를 창출하는 데 의미를 둔 것이며, 연구 데이터를 생산·가공할 때 사용된 프로그램, 소프트웨어 등에 대한 공유도 수반되어 제공자와 이용자 간 소통과 협력, 홍보 등을 필요로 한다(신은정 외, 2017).

유럽연합, 핀란드, 네덜란드, 미국, 일본 등 주요 선진국에서는 새로운 지식·가치 창출과 공유·확산을 위해 국가, 기관 차원의 오픈 사이언스 정책을 적극적으로 추진하고 있다(신은정, 정원교, 2017).

한편, 이러한 세계적인 기조에 발맞추어 우리

나라에서도 오픈 사이언스, 오픈 데이터가 자리를 잡아 가고 있는 가운데 특히 공적 자금을 활용하여 수행한 정부출연연구기관 연구의 데이터 관리와 개방에 많은 관심을 기울일 필요가 있다.

3. 우리나라의 오픈 사이언스 정책

이 장에서는 우리나라의 오픈 사이언스 정책을 이해하기 위해 관련 법률과 정부출연연구기관의 관련 규정을 살펴보았다.

가. 관련 법률

먼저 흔히 「공공데이터법」이라 불리는 행정안전부의 「공공데이터의 제공 및 이용 활성화에 관

한 법률», 행정안전부의 「공공기관의 정보공개에 관한 법률」(이하, 정보공개법), 문화체육관광부의 「저작권법」에 대해 살펴보고자 한다.

1) 행정안전부의 「공공데이터의 제공 및 이용 활성화에 관한 법률」

2013년 6월 ‘정부3.0 기본계획’에서 공공데이터 개방이 천명되었고 이에 따라 2013년 7월 30일 「공공데이터법」이 제정되었다. 이 법은 공공기관에 공공데이터 제공 의무를 부여하고 공공데이터를 민간에 효과적으로 제공하고 민간의 이용 활성화를 지원할 수 있는 법적 근거를 마련함으로써 국민 삶의 질 향상과 국민 경제 발전에 이바지하는 것을 목적으로 한다.

이 법에서 ‘공공데이터’라 함은 공공기관이 법령 등에서 정한 목적을 위해 생성 또는 취득하여 관리하고 있는 전자적 방식의 자료 또는 정보로서(제2조 정의) 공공기관이 생산한 연구 데이터도 일종의 공공데이터라 할 수 있다. 또한 ‘제공’이란 공공기관이 기계 판독이 가능한 형태의 공공데이터에 접근할 수 있게 하고 이를 전달하는 것을 말한다(제2조 정의).

제3조(기본 원칙)에서는 공공데이터를 제공하는 공공기관과 이를 이용하는 이용자의 의무 사항을 규정하고 있는데, 공공기관이 공공데이터를 제공할 때는 이용의 편리성, 이용권의 보편적 확대, 이용자의 접근성 등을 위해 노력할 것 등을 제시하고 있으며, 이용자에게 대해서는 이용 시 공익이나 타인의 권리를 침해하지 않고 신의에 따

라 성실하게 이용할 것을 제시하고 있다.

제17조(제공 대상 공공데이터의 범위)에서는 국민에게 제공하지 않는 데이터의 범위를 규정하고 있는데, 「공공기관의 정보공개에 관한 법률」 제9조에 따른 비공개 대상 정보, 「저작권법」 등에 따라 이용 허락을 받지 않은 정보 등이다. 제18조(공공데이터 목록의 등록)에서는 공공기관의 소관 공공데이터 목록을 행정안전부 장관에게 등록하도록 하고 있으며, 제24조(공공데이터의 제공 기반 구축)에서는 공공데이터 제공을 위한 공공기관의 역할을 제시하고 있다.

이러한 「공공데이터법」의 입법 취지에 맞게 정부출연연구기관에서는 공공데이터의 일종인 연구 데이터를 이용자의 편리성, 접근성을 고려하여 더욱 적극적으로 개방(제공)하여야 할 것이다.

2) 행정안전부의 「공공기관의 정보공개에 관한 법률」

「정보공개법」은 공공기관이 보유·관리하는 정보에 대한 국민의 공개 청구, 공공기관의 공개 의무에 관한 사항을 정한 법률이다. 제3조(정보공개 원칙)에서는 공공기관이 보유·관리하는 정보를 적극적으로 공개하여야 함을 밝히고 있으며, 제6조(공공기관의 의무)에서는 정보 공개를 위해 공공기관이 정보를 적절히 보존하고 신속한 검색이 가능한 정보 관리 체계 정비에 노력할 것을 규정하고 있다.

이 법에서 ‘공개’라 함은 공공기관이 이 법에 따라 정보를 열람하게 하거나 그 사본·복제물을

표 2. 법률상 '제공'과 '공개'의 차이

구분	공공데이터 '제공'	정보 '공개'
소관 법률	「공공데이터법」	「정보공개법」
대상	전자적 자료 또는 정보	전자적 혹은 비전자적 자료 또는 정보
형태	기계 판독이 가능한 형태	특정한 공개 형태 없음

자료: 행정안전부. (2019). 공공데이터 관리지침. p. 8.

제공하는 것 또는 정보통신망을 통해 정보를 제공하는 것을 말하는데(제2조 정의), 이는 「공공데이터법」에서의 '제공'과 비교해 대상과 형태의 차이가 있다(표 2).

한편 제9조(비공개 대상 정보)에서는 공공기관이 공개하지 않아도 되는 대상 정보로 1) 타 법령에서 비밀이나 비공개 사항으로 규정한 정보, 2) 공개 시 국가의 이익을 해칠 우려가 있는 정보, 3) 공개 시 국민의 생명, 신체 및 재산 보호에 지장을 끼칠 우려가 있는 정보, 4) 공개 시 직무수행이 곤란해지거나 공정한 재판에 영향을 끼칠 우려가 있는 정보, 5) 공개 시 공정한 업무수행이나 연구·개발에 지장을 초래하는 정보, 6) 공개 시 사생활 비밀 또는 자유를 침해할 우려가 있는 정보, 7) 공개 시 법인·단체 또는 개인의 경영상·영업상 비밀에 관한 정보, 8) 공개 시 특정인에게 이익 또는 불이익을 줄 우려가 있는 정보 등을 제시하고 있으며, 이러한 비공개 대상 정보 범위에 관한 세부 기준을 수립하여야 함을 규정하고 있다.

정부출연연구기관에서는 연구 데이터를 적극적으로 공개하고 이를 위해 정보 관리 체계를 정비하여야 하며 정보 공개 시 혹은 공공데이터 제

공 시 특히 비공개 대상 정보에 대해 제9조를 기준 삼아 명확한 기준과 지침을 마련하여야 할 것이다.

3) 문화체육관광부의 「저작권법」

「저작권법」은 학문적 또는 예술적 저작물의 저작자 권리를 보호하고 저작물의 공정한 이용을 도모하기 위한 법률로 “저작물은 인간의 운동 또는 감정을 표현한 창작물”, “저작자는 저작물을 창작한 자”, “공동저작물은 2인 이상이 공동으로 창작한 저작물로서 각자의 이바지한 부분을 분리하여 이용할 수 없으며”, “업무상저작물은 법인·단체 그 밖의 사용자(이하 “법인 등”이라 한다)의 기획하에 법인 등의 업무에 종사하는 자가 업무상 작성하는 저작물을 말한다”고 정의하고 있다(제2조 정의).

정부출연연구기관의 연구보고서나 조사 데이터의 경우, 직원 등이 업무상 작성하여 기관 등의 명의로 공표한 것이므로 원칙적으로 업무상 저작물에 해당된다 할 수 있다. 연구 데이터의 개방 시 업무상 저작물에 대한 저작권, 공동저작물에 대한 저작권 등을 세심하게 고려하여야 한다.

나. 연구사업의 오픈 사이언스 정책

앞에서 살펴본 법률은 공공데이터 혹은 공공기관이 보유·관리하는 정보에 대해 제공 및 공개 의무를 담은 내용으로 오픈 사이언스 정책 정신을 법률적으로 규정하고 있다. 여기에서는 실제 연구사업 수행에서 어떻게 적용하고 있는지 알아보기 위해 과학기술정보통신부의 「국가연구개발사업의 관리 등에 관한 규정」, 국가과학기술연구회의 「연구 데이터 관리 가이드라인」과 「연구 데이터 관리 계획」에 대해 살펴보려고 한다.

1) 과학기술정보통신부의 「국가연구개발사업의 관리 등에 관한 규정」

이 규정은 과학기술 분야 국가연구개발사업 추진 시 공통적으로 적용할 수 있는 기본 원칙과 기준을 마련하여 국가연구개발사업이 효율적으로 이루어질 수 있게 하기 위한 것으로, 2019년 개정안(대통령령 제29625호, 2019. 3. 19. 일부개정) 제2조(정의)에서 “연구 데이터란 연구·개발 과제 수행 과정에서 실시하는 각종 실험, 관찰, 조사 및 분석 등을 통하여 산출된 사실 자료로서 연구 결과의 검증에 필수적인 데이터”, “데이터 관리 계획이란 연구 데이터의 생산·보존·관리 및 공동 활용 등에 관한 계획”이라고 규정하고 있다. 또한 제6조 제4항 제9호¹⁾에서는 연구·개발 계획서 작성 시 데이터 관리 계획을 포함토

록 하고 있으며, 제7조 제3항 제10호²⁾에서는 중앙행정기관이 데이터 관리 계획에 따른 연구 데이터 생산·보존·관리의 충실성 및 공동 활용 가능성을 검토하도록 규정하고 있다.

다만 이 규정은 정부출연연구기관 중에서도 과학기술계 연구기관에 적용하는 규정으로서 경제인문사회계 연구기관에서는 적용하지 않고 있다.

2) 국가과학기술연구회의 「연구 데이터 관리 가이드라인」과 「연구 데이터 관리 계획 가이드라인」

「연구 데이터 관리 가이드라인」은 정부출연연구기관들이 국가연구개발사업을 통해 생산된 연구 데이터를 수집, 관리, 공유, 활용할 때의 기준을 제시한 것이며, 「연구 데이터 관리 계획 가이드라인」은 국가연구개발사업에서 생산되는 연구 데이터의 데이터 관리 계획(DMP: Data Management Plan)의 작성 및 점검 기준을 제시한 것이다. 예비분석 결과, 논문이나 저술의 초안, 연구 노트, 보고서 등은 연구 데이터로 간주하지 않는다.

「연구 데이터 관리 가이드라인」에서는 Wilkinson 외(2016)에서 제시한 연구 데이터 관리의 FAIR (Findable, Accessible, Interoperable, Reusable) 원칙인 검색 가능성, 접근 가능성, 상호 운용 가능성, 재사용 가능성을 준용할 것을 요구하고 있으며(표 3), 연구 데이터 관리 절차는 ① 연구 데

1) 국가연구개발사업을 수행하거나 참여하려는 자가 연구·개발 계획서 작성 시 데이터 관리 계획을 포함하도록 한다.
2) 중앙행정기관이 연구·개발 과제 선정 시 데이터 관리 계획에 따른 연구 데이터 생산·보존·관리의 충실성 및 공동 활용 가능성을 검토하도록 한다.

표 3. 연구 데이터에 대한 FAIR 원칙

구분	공공데이터 '제공'
Findable(검색성)	연구 데이터에 메타데이터와 DOI 등 연구식별자를 할당하고 연구 데이터 리포지터리에 등록하여 검색 가능하도록 함
Accessible(접근성)	표준화된 프로토콜을 사용하여 식별자를 통해 연구 데이터(연구 데이터가 이용 가능하지 않을 경우 메타 데이터)에 접근 가능하도록 함
Interoperable(상호 운용성)	연구 데이터는 표준적인 방식과 어휘를 사용하여 상호 운용 가능하도록 해야 하며, 관련된 다른 데이터에 대한 인용을 포함함
Reusable(재사용성)	명확한 데이터 라이선스와 프로비던스 정보를 제공하여 재사용 가능하도록 함

자료: Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A.,...Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. doi:10.1038/sdata.2016.18. http://dx.doi.org/10.1038/sdata.2016.18; 국가과학기술연구회. (2019a). 연구데이터 관리 가이드라인. p. 2에서 재인용.

표 4. 연구 데이터 관리 절차

구분	내용	
관리 체계 수립	정책과 절차	관리 기준 정의, 정책 및 절차 수립, 기관 내 공표
	전담 조직 및 인력	전문 인력 총원 및 전담 조직 구성
	요구 사항 정의	기관 내 연구 데이터의 크기, 종류, 생산 주기, 검색 요구 사항 정의 연구 데이터 및 프로세스의 표준화 방안 마련
	연구 데이터 인프라	안정적인 연구 데이터 관리와 장기간 보존을 제공하고 다양한 연구 데이터 공유·활용을 지원하는 데 필요한 정보기술(IT) 인프라 필요
수집 및 저장	지원 서비스	체계적인 연구 데이터 관리·활용을 위해 연구자에게 교육, 자문, 지원 서비스 등을 제공
	연구 데이터 정제 및 표준화	관리·보존을 위해 연구기관의 데이터 리포지터리에 제출하며, 이를 위해 연구기관에서 정의한 연구 데이터 관리 기준에 적합하게 정제되고 표준화된 연구 데이터 준비
	메타데이터 작성	연구기관에서 정의한 메타데이터 스키마를 준용하며, 장기간 보존 및 활용을 위해 풍부하게 기술
	제출	연구 과제 종료 후 연구기관에서 지정한 기간 내에 데이터 관리 계획에서 지정한 데이터 리포지터리에 연구 데이터 제출
승인 및 검토		- 연구기관의 연구 데이터 공개 기준에 따라 비공개, 내부 공개, 대외 공개를 선택하고, 공개의 경우 엠바고 설정을 통해 공개 시한 지정 • 연구 데이터 주요 검토 사항: 데이터 형식, 메타데이터 기술 형식 준용 여부, 연구 데이터 공개 및 라이선스 적합 여부, 민감 정보 포함 여부 등 - 승인 절차와 성과를 관리하기 위해 연구기관의 정보 관리 시스템과 연계 가능
출판	출판 절차	- 다른 연구자들이 (재)활용할 수 있도록 데이터를 공개하기 위해 연구기관 등의 데이터 리포지터리를 통해 온라인으로 접근 가능한 형태로 출판 가능 - 데이터 저널을 통해 연구 데이터 자체에 대해 논문 형태로 출판 가능 - 관련 주제 분야 혹은 데이터 전문가의 피어 리뷰 등 별도의 품질 점검 절차 도입 가능
	데이터 식별자 부여	- 데이터 세트의 식별 및 활용 추적을 위해서 공개된 연구 데이터에 DOI(Digital Object Identifier) 등과 같은 식별자 부여 권장 - DOI는 한국DOI센터, DataCite 등 DOI 등록 기관에서 발급받을 수 있으며, 해당 연구 분야의 식별자 관리 기관이 존재하면 해당 기관 활용
	라이선스 적용	- 연구기관은 연구 데이터의 대내외 공개와 저작권 확보를 위해 라이선스 기준과 지원 서비스를 제공하고, 연구자가 각 연구 데이터에 대해 적절한 라이선스를 선택하여 적용
공유 및 활용	접근 및 공유	- 다른 연구자가 연구 데이터를 쉽게 활용할 수 있도록 표준 프로토콜 및 기술 형식에 따라 연구 데이터 관리 및 배포 - 연구 데이터 리포지터리의 검색에 노출되는 연구 데이터 및 메타데이터의 공개 여부는 이를 제출하는 연구자에 의해 지정 - 해당 연구 분야의 데이터 검색 사이트 또는 외부 연구 데이터 사이트와의 연계를 통해 검색 노출 빈도를 높임으로써 연구 데이터 활용도 증진
	활용	- 연구기관에서는 내외부 연구자에게 보유 중인 연구 데이터의 처리·분석을 위한 다양한 연구 환경 제공 - 타 연구자의 연구 데이터를 활용한 경우 연구논문 등에서 데이터 인용을 통해 출처 표기

자료: 국가과학기술연구회. (2019a). 연구데이터 관리 가이드라인.

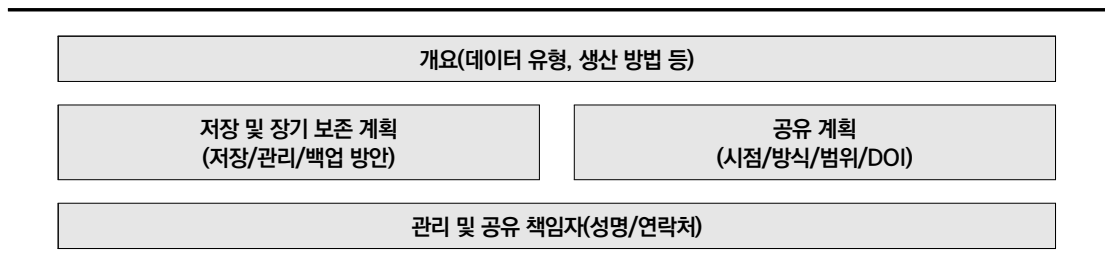
이터 관리 체계 수립 ② 연구 데이터 수집·저장 ③ 연구 데이터 출판 ④ 연구 데이터 공유·활용 등 4단계로 제시하고 있다(표 4).

한편 연구 데이터 관리 계획에는 연구·개발(R&D) 과정에서 생산되는 연구 데이터의 유형, 관리, 공유 계획 내용이 포함되어야 하며(그림 2), 과기정통부에서는 2019년 시범 R&D 과제

를 대상으로 데이터 관리 계획 제출 의무화를 우선 시행하고, 2020년부터 적용 대상을 단계적으로 확대할 예정이다.

이러한 가이드라인도 정부출연연구기관 중 과학기술계 연구기관에 적용되는 것으로 경제인문사회계 연구기관에서는 이를 적용하지 않고 있다.

그림 2. 연구 데이터 관리 계획 포함 내용



자료: 국가과학기술연구회. (2019b). 연구데이터 관리 계획 가이드라인. p. 2.

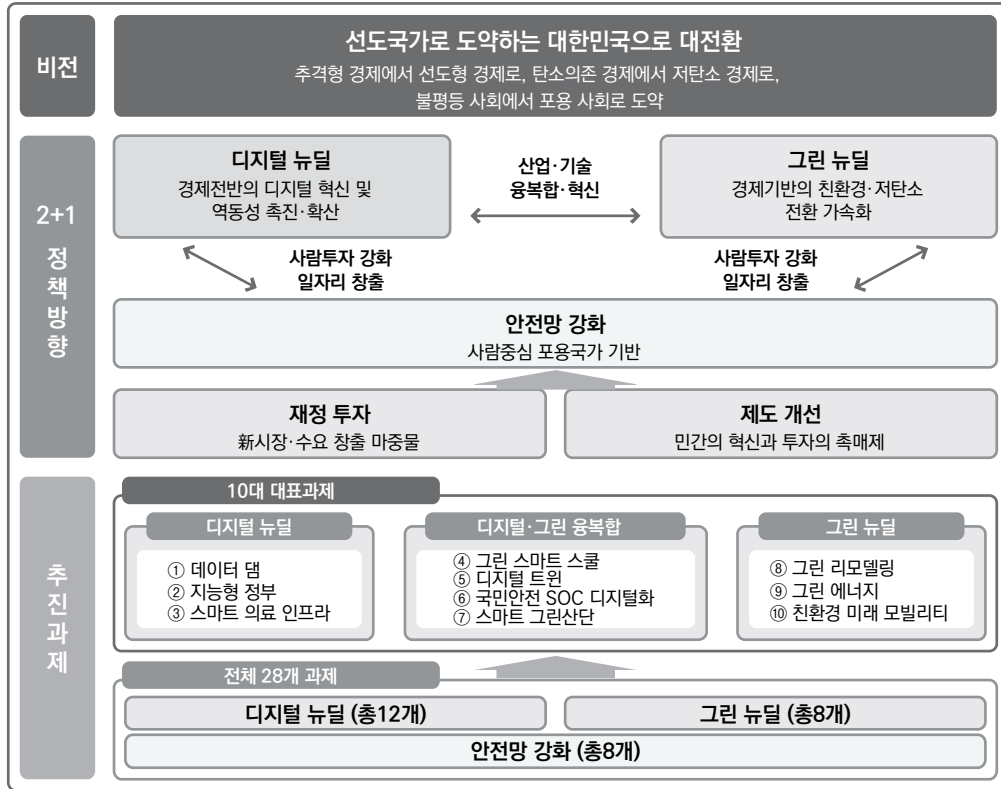
다. 정부출연연구기관과 연구 데이터

전 세계적으로 연구 성과와 과정을 개방하고 공유하려는 오픈 사이언스 움직임이 확산되고 있는 가운데 우리나라에서도 과학기술연구 부문에서 연구 데이터 관리와 개방을 시작하여 점차 그 범위를 확장하고 있다. 현 정부의 ‘한국판 뉴딜 정책’에서도 ‘데이터 수집·가공·거래·활용 기반 구축을 통한 경제 가속화(데이터 댐)’를 10대 대표 과제 중 하나로 선정하여(관계부처 합동, 2020, 그림 3) 공공데이터 개방에 많은 노력을 기울이고 있어 공적 자금으로 만들어진 연구 결과물 활용에 대한 필요성은 더욱더 커질 것이다.

그러나 아직까지 우리나라 연구자들은 연구 결과물 공유와 재이용에 대한 우려와 제한적인 공유, 연구 데이터 관리에서 미흡함을 나타내고 있다(김지현, 2015). 연구 데이터 공유에 대해서는 약 82%가 필요하다고 생각하는 반면, 약 9%만이 연구 데이터를 외부에 공유한 경험이 있는 것으로 나타나(국가과학기술연구회, 한국과학기술정보연구원, 2019, 그림 4) 연구자들의 연구 데이터 개방·공유에 대한 인식은 개선의 여지가 많은 것으로 파악되었다.

우리나라에서 오픈 사이언스에 대한 논의는 과학기술계 연구기관 중심으로 펼쳐져 왔으며 경제인문사회계 연구기관에서는 사실상 적극적인

그림 3. 한국판 뉴딜 종합 계획 비전 및 추진 계획



자료: 관계부처 합동. (2020. 7. 14.). 「한국판 뉴딜」 종합계획. p. 5.

그림 4. 국가과학기술연구회 소속 출연(연) 과제 책임자들에 대한 연구 데이터 개방·공유 수준

(단위: 명)



자료: 국가과학기술연구회. (2019b). 연구데이터 관리 계획 가이드라인. p. 2.

표 5. 한국보건사회연구원이 최근 3년간 수행한 연구 중 조사사업 포함 연구 현황

(단위: %)

구분	조사사업 포함 연구	수행 조사 방법		
		양적 조사	질적 조사	양적+질적
평균	39.6	65.9	21.1	13.0
2017년	38.8	74.4	16.7	9.0
2018년	40.8	63.4	22.0	14.6
2019년	39.1	60.5	24.4	15.1

자료: 저자 작성

논의가 나타나지 않고 있다. 경제인문사회계 연구기관에서 정책 연구를 수행할 때 대표적인 방법론 중 하나는 사회조사로서, 조사 결과가 생산되는 조사 데이터는 중요한 연구 데이터 중 하나이다.

실례로, 한국보건사회연구원은 대표적인 경제인문사회계 연구기관으로 최근 3년간 수행한 과제 중 조사를 수행한 연구가 평균 40% 정도를 차지할 만큼 근거 기반 연구에서 사회조사가 중요함을 잘 보여 주고 있으나(표 5) 과학기술계에서 제시하는 데이터 수집·관리·공유·활용에 대한 기준을 준용하고 있지는 않다.

4. 나가며

데이터를 관리하고 공개하는 것은 그 자체가 목적이 아니라 공개 후 지식을 발견하여 혁신으로 이끌고 데이터와 지식을 통합하고 재사용하는 단초가 되는(Wilkinson et al., 2016) 것이 의미가 있는 만큼 오픈 사이언스, 오픈 데이터 운동에서 정부출연연구기관의 역할은 매우 크다.

그러나 앞에서 살펴본 바와 같이 정부출연연구기관 중에서도 과학기술계 기관에 비해 상대적으로 경제인문사회계 연구기관에서는 아직까지 이러한 연구 데이터 관리와 개방의 필요성, 중요성이 공론화되지 않고 있으며 그 현황조차 파악이 안 되고 있다.

그러므로 연구 데이터 통합과 재사용에 따라 새로운 지식과 가치가 창출되는 오픈 사이언스, 오픈 데이터의 의미를 부여하기 위해서는 우선 경제인문사회계 연구기관들의 연구 데이터 관리 및 공유·활용 체계 현황과 이들 기관 연구자들의 인식을 파악하는 것이 필요하며, 이를 토대로 경제인문사회계 연구기관의 실정에 맞는 가이드라인을 제시하고 적용할 필요가 있다. ■

참고문헌

국가과학기술연구회. (2019a). 연구데이터 관리 가이드라인.
 국가과학기술연구회. (2019b). 연구데이터 관리 계획 가이드라인.

- 국가과학기술연구회. (2019c). 출연(연) 연구데이터 관리·활용 방안 연구.
- 국가과학기술연구회, 한국과학기술정보연구원. (2019). 출연(연) 연구데이터 관리·활용 방안 연구.
- 관계부처 합동. (2020. 7. 14.). 「한국판 뉴딜」 종합계획.
- 김용훈, 오영균. (2008). 정부출연연구기관의 지배구조에 관한 연구. 한국행정연구, 17(3), 279-297.
- 김응렬. (2001). 사회조사방법론의 이해. 고려대학교출판부.
- 김지현. (2015). 데이터 관리와 공유에 대한 대학 연구자들의 인식에 관한 연구. 한국문헌정보학회지, 49(3), 413-436.
- 신은정. (2015). 오픈 사이언스(Open Science)에 관한 OECD 논의 동향과 시사점. 동향과 이슈(22), 과학기술정책연구원, 1-21.
- 신은정, 안형준, 정원교. (2016). 오픈 사이언스를 위한 연구 성과물 공개 정책과 과제. 과학기술정책연구원.
- 신은정, 안형준, 양현채, 최병삼, 양승우, 정원교, 김수연. (2017). 오픈사이언스 정책의 도입 및 추진 방안. 과학기술정책연구원.
- 신은정, 정원교. (2017. 8. 8.). 오픈 사이언스 정책의 확산과 시사점. STEPI Insight, 제 216호, 1-39.
- 오세영, 윤건, 오균. (2017). 증거기반 정책을 위한 정부의 통계 구축 및 활용에 대한 현황 조사. 한국행정연구원 사회조사센터.
- 이민형, 김태양. (2020. 4. 10.). 공공연구기관 R&D 협력체계 구축을 위한 주요 혁신방안. STEPI Insight 제252호. 과학기술정책연구원.
- 한국정보화진흥원. (2015. 9. 15.). 데이터 증거기반(Evidence-Based)의 과학적 정책 수립 방안. IT & Future Strategy 제6호.
- 행정안전부. (2019). 공공데이터 관리지침.
- Merton, R. K. (1973). The sociology of science: Theoretical and empirical investigations. Chicago: University of Chicago Press. (원서출판 1942)
- OECD. (2015. 10. 15.). Making open science a reality. OECD science, technology and industry policy papers, No. 25, OECD publishing, Paris.
<http://dx.doi.org/10.1787/5jrs2f963zs1-en>
- Hey, T., Tansley, S., & Tolle, K. (2009). Jim Gray on eScience: A transformed scientific method. The fourth paradigm: Data-intensive scientific discovery. Microsoft Research.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A.,...Mons, B. (2016). The fair guiding principles for scientific data management and stewardship. Scientific Data, 3(1), 160018. doi:10.1038/sdata.2016.18.
<http://dx.doi.org/10.1038/sdata.2016.18>.

관련 법률

- 「공공기관의 정보공개에 관한 법률」
- 「공공데이터의 제공 및 이용 활성화에 관한 법률」
- 「국가연구개발사업의 관리등에 관한 규정」
- 「저작권법」
- 「국가연구개발사업의 관리 등에 관한 규정」