

과학데이터의 효율적인 관리 및 활용을 위한 정책방향

윤 중 민*

I. 머리말

최근 과학기술 분야에 있어서 연구개발의 방향이 소위 데이터 중심으로 변화됨에 따라, 종래의 과학이론에 바탕을 둔 실험 중심의 연구개발 활동과는 다르게 과학연구 과정에서 발생하는 방대한 과학데이터를 바탕으로 컴퓨터시뮬레이션을 통한 연구개발 활동이 확대되고 있다. 이와 같은 연구개발 활동의 패러다임 전환에 따라, 연구개발 활동의 과정 및 결과로 발생하는 각종 과학데이터의 수집, 관리 및 활용을 촉진하기 위한 제도정비가 중요한 과제로 대두되고 있다.

주요 외국은 2000년대 중반부터 과학기술 연구개발 활동으로부터 발생된 각종 과학데이터의 공유 및 활용을 촉진하기 위한 체계를 구축하여 연구자간의 정보 교류를 유도하고 있는 한편, 공공 연구개발사업의 성과로 생성된 각종 과학데이터의 국가적 관리를 위한 법제를 정비하고 있다.

우리나라도 과학데이터를 국가적 차원에서 관리 및 활용을 촉진하기 위하여 관련 전략을 수립한 바 있으나,¹⁾ 그 구체적인 실행을 위한 정책수단이나 제도의 정비는 아직 미흡한 실정이다.²⁾ 특히, 국가연구개발 성과로 발생된 공공 과학데이터를 체계적으로 관리하고 이를 효율적으로 공유·활용하도록 하기 위한 법제가 마련되지 않고 있어, 많은 예산이 투자된 과학기술 연구개발의 성과가 사장될 우려가 있으며, 데이터 중심의 연구개발 활동을 효과적으로 지원하기 어려운 실정이다.

이에, 이 논문에서는 국가과학데이터의 공유 및 활용을 촉진하는데 있어서 그 제도적 기반이 되는 관련 법제도의 효율적인 정비방안을 검토하여 제시해 보고자 한다. 특히, 국가연구개발사업의 성과인 과학데이터를 국가적 차원에서 효율적으로 관리하고, 그 공유 및 활용을 촉진하기 위한 주요 정책방향과 관련 입법추진 과제를 도출하여 제안하고자 한다.

II. 과학데이터의 개념과 특성

1. 과학데이터의 의미와 개념 정의

과학데이터(Scientific data)는 과학연구 활동의 과정에서 연구의 결과 또는 부산물로서 생성되는 각종 데이터를 의미한다. 즉, 연구를 수행하는 과정에서 관찰이나 실험과 같은 일련의 과정을 통하여 수집, 측정, 관측되는 기초 데이터를 말한다. 이러한 측면에서 과학데이터를 연구데이터(Research data)라고 부르기도 한다.

과학데이터에 대한 개념정의는 국제적으로나 국내적으로 아직 명확히 정립되지 않고 있다고 할 수 있다. OECD는 “과학연구에서 1차 자료로 사용되는 사실기록(수치, 문자기록, 이미지 및 음향)으로 과학자 사회에서 통상 연구결과를 검증하는데 필요한 데이터로서 인정된 것”으로 정의하고

* 충북대학교 법학전문대학원 교수, 043-261-3592, cmyoon@cbnu.ac.kr.

1) 미래창조과학부·한국과학기술정보연구원, 「국가과학기술 빅데이터 거버넌스 체제구축을 위한 전략계획(ISP)」, 2013.12.

2) 미래창조과학부는 국가 과학기술 데이터의 개방과 공유를 촉진하기 위하여 2014.8.7. ‘국가과학기술데이터 포털’(http://open.ntis.go.kr)을 구축하고 시범서비스를 시작하였으나, 이를 체계적으로 뒷받침할 수 있는 법제도는 아직 정비되지 않았다.

있다. 미국도 이와 유사하게 정의하고 있는데, 미국정부대외관계위원회(Council On Governmental Relations)의 「연구데이터에 대한 접근 및 저장 : 권리와 책임」에서 연구데이터를 “과학자 사회에서 일반적으로 연구결과를 검증하기 위하여 인정된 기록이 된 사실적 자료”로서 정의하고 있다. 독일의 경우, 독일연구재단(DFG)의 「정보기반체제와 과학데이터」(Informationsinfrastrukturen und Forschungsdaten)라는 자료에서 과학데이터(Forschungsdaten)를 “연구를 수행하는 과정에서 기초·원천연구, 실험, 측정, 조사 또는 설문을 통해 생성되는 데이터로서 디지털 및 전자데이터로 저장할 수 있는 데이터”라고 개념정의하고 있다. 호주는 독일과 유사하지만 보다 구체적으로 기술하고 있다. 호주 국가협업연구 기반구조사업인 NCRIS의 협업플랫폼 구축 프로그램인 「호주 국가데이터서비스(ANDS)」는 관련 데이터서비스 지침에서 과학데이터를 “사실, 관찰, 이미지, 컴퓨터프로그램 결과, 기록, 측량 또는 경험(논거, 이론, 테스트 또는 가설 또는 기타 연구물에 기초한)의 형태에서 생성되는 데이터를 의미한다. 데이터의 범위에는 수치, 기술, 시각 또는 청각 데이터가 모두 포함 된다”고 규정하고 있다.

이상에서 보듯이, 과학데이터 또는 연구데이터의 개념은 국가마다 다양하게 정의되고 있는데, 이는 각 국가마다 조금씩 다른 데이터 정책을 취하기 때문인 것으로 생각된다. 다만, 이들의 공통적 개념을 종합하여 판단하면, 과학데이터란 과학연구의 과정 또는 결과로서 생성되는 사실적 기록으로서 수치, 문자기록, 이미지, 음향 등의 형태를 망라한 데이터라고 할 수 있다. 이와 같은 과학데이터에는 가공되지 않은 ‘원시데이터(raw data)’는 물론 원시데이터를 일정하게 가공한 ‘가공데이터(processing data)’가 포함되며, 데이터 형식에 있어서도 시각 또는 청각 등 그 형식과 관계없이 보존이 가능한 모든 데이터가 포함된다.

2. 과학데이터의 일반적 특성

과학데이터는 과학적 연구 활동의 결과로서 발생하는 동시에, 과학적 연구를 위한 기초자료로 활용된다. 나아가 과학데이터는 과학적 연구 결과의 객관성 및 진실성을 검증하는 자료도 활용되기도 한다. 이와 같은 측면에서 과학데이터는 중요한 연구자원으로서의 성격을 갖는다. 최근에 데이터를 중심으로 하는 제4세대 R&D와 데이터 집중형 R&D가 확대·강화되면서 연구자원으로서의 과학데이터의 중요성은 더욱 증대되고 있다. 또한, 과학데이터는 그 본질적 속성상 공공재로서의 성질을 갖는다. 과학데이터는 연구개발의 성과물이므로 누구나 접근하여 활용할 수 있는 지식정보로서의 성격을 갖는다. 따라서 국가의 공공자원으로서 체계적으로 관리·보존하고, 그 공유·활용을 촉진하는 것이 필요하다.

한편, 과학데이터는 그 관리·활용의 측면에서 다음과 같은 특징을 보인다. 먼저, 과학데이터는 다양한 형식으로 존재한다. 과학데이터는 연구 활동을 통하여 생성된 기록물로서 관측, 감시, 조사, 실험, 분석, 계산 등의 과정을 거쳐 생산된 문자, 이미지, 오디오, 동영상 등의 아날로그 및 디지털 형식의 데이터를 포괄한다. 또, 과학데이터는 재가공이 가능하다. 과학데이터는 특정한 현상을 설명하기 위하여 필요한 경우 데이터를 재가공하여 이용할 수 있다. 예를 들면, 위성을 통해 지구표면에 대기가 분포된 모습을 찍은 사진인 과학데이터를 날씨를 관측 및 예측하는 기상데이터로 재가공할 수 있다. 과학데이터의 재가공을 통하여 다양한 학문·분야 간에 데이터의 공유를 바탕으로 한 융·복합이 이루어질 수 있다. 또한 과학데이터는 그 활용성과 가치를 높이기 위해서는 장기적이고 체계적인 축적이 요구된다.

과학데이터는 특정의 데이터가 시계열적으로 또는 광범위한 범위에 걸쳐 대용량으로 집합되고 축적되어야 데이터 분석의 정확성과 객관성을 높일 수 있다. 이와 같은 측면에서 연구자 개인이나 국가적 차원에서 과학데이터를 관리·활용함에 있어서는 보다 체계화된 정책과 관리방법이 요구된다. 이에 따라 주요 국가들은 과학기술의 연구개발 활동으로부터 생성되는 각종 과학데이터의 수집, 보존 및 관리·활용을 촉진하기 위해 다양한 정책을 추진하고, 관련 법제도를 정비하고 있다.

Ⅲ. 해외의 과학데이터 정책 및 법제 동향

1. 국제기구

과학기술 관련 국제기구들은 오래 전부터 과학데이터의 공유 및 활용을 촉진하기 위한 활동을 수행하여 왔다. 먼저, 국제학술연합(International Council for Science Union, ICSU)이 1966년 설립한 과학기술데이터위원회(Committee on Data for Science and Technology, CODATA)³⁾는 과학기술에 관한 중요 데이터의 축적, 평가 및 보급을 추진하고 있으며, ICSU가 운영하는 WDS(World Data System)는 전 세계 12개국의 52개 센터와 연계하여 과학데이터 정보의 상호 공유 네트워크를 구축·운영하고 있다.

경제협력개발기구(OECD)는 2007년 4월 ‘공공자금으로 발생된 과학데이터에 대한 접근 원칙 및 가이드라인’을 발표하여 공공 과학데이터에 관한 각국 정부의 정책 및 자금지원기관에 대한 폭넓은 정책적 권고안을 제시하였다. 이 원칙 및 가이드라인은 세계 과학시스템의 효율성을 향상시키는 것을 궁극적인 목적으로 하고 있으며, 과학데이터에 대한 개방성·유연성·안전성·지식재산권 보호 등의 13개 원칙을 제시하여 회원국들의 과학데이터 공유 및 활용을 촉진하고 있다.

2. 유럽(EU)

유럽연합은 각 회원 국가들이 개별적으로 운영하는 과학데이터 인프라에 대한 통합 관리시스템의 중요성을 인식하고, 유럽정부 주도로 각종 과학데이터 정책을 추진·지원하고 있다. 특히 FP7(Seventh Framework Programme)에서 정보통신기술 기반의 e-Infrastructure 구축 사업을 통해, 과학데이터와 컴퓨터 자원의 실시간 공유 및 연구자간의 커뮤니케이션을 위한 다양한 정책그룹(e-IRG, ESFRI, ERA, PARADE)을 구성하고, 과학데이터 프로젝트 및 저장소(DRIVER, METAFOR, NMDB, GENESI-DR 등)의 개발을 추진하고 있다.

개별 국가 차원에서 과학데이터의 관리 및 공유체제를 구축·운영하고 있다. 예를 들면, 독일은 막스플랑크연구회(Max Planck Society)와 수학물리정보센터인 FIZ Karlsruhe가 공동으로 개발한 eSciDoc 시스템을 통해 연구결과는 물론, 사전연구문서, 원시 및 실험데이터, 예비출판물(pre-prints), 학습자료 등 모든 연구중간과정을 포함하는 과학데이터의 저장, 처리, 배포, 발간 등을 위한 e-연구플랫폼을 구축하여 운영하고 있다.

3. 미국

미국은 각 분야별로 관리해오던 과학데이터를 국가적 차원에서 통합적으로 관리하기 위하여 2006년 국가과학기술위원회(National Science and Technology Council, NSTC) 산하에 ‘연방기구 간 데이터 워킹 그룹(Interagency Working Group on Digital Data, IWGDD)’을 구성하여 연방기구들에서 생성된 데이터에 대한 공개 접근을 가능하도록 하였다. 특히, 국가연구개발사업의 수행으로 발생된 데이터 공유를 위해 연구비 지원기관인 국립보건원(NIH)이나 국립과학재단(NSF) 등을 통해 과학연구 데이터의 공유 및 활용을 추진하고 있다.

NSF는 2011년부터 연구관리 지침을 통하여 연구과제신청시 해당 연구개발로부터 발생하는 과학

3) CODATA는 ICSU의 학제간 위원회로서, 세계 물리상수 및 단위 표준을 관장하는 국제기구이다. CODATA의 주요 기능은 과학데이터에 대한 개발과 지식공유, 과학데이터 관련 활동을 통해 과학기술을 육성·발전시키는데 있으나, CODATA가 직접 과학데이터의 관리 및 공유체제를 구축하여 서비스를 제공하지는 않는다.

데이터의 생산 및 관리 등에 관한 계획(DMP)을 반드시 제출하도록 하고, 연구결과 보고 시에 그에 따른 시행여부를 확인하고 있다.⁴⁾ 이에 따르면, NSF의 연구비지원으로 과제를 수행하는 연구자는 과제수행과정에서 생산되거나 수집한 원시데이터, 표본, 물리적 컬렉션, 기타 자료를 추가비용이 발생하지 않는 범위 내에서, 그리고 적절한 시간 내에 다른 연구자들과 공유하도록 하고 있다. 또한, NSF의 산하기관인 ‘사이버기반구조국(Office of Cyber Infrastructure, OCI)’은 2007년에 생산된 정보의 양이 최초로 수용 가능한 저장능력을 초과하는 현실적 문제에 직면하여 관련 기관에서 생산된 과학데이터에 대한 공유를 지원하기 위해 DataNet(Sustainable Digital Data Preservation and Access Network Partners)을 구축·운영하기 시작하였다. DataNet은 과학데이터의 보존과 접근에 대한 장기적인 역량의 강화, 경제적·기술적으로 지속가능한 시스템과 서비스 창출, 과학 중심의 정보의 융합이라는 세 가지 목표를 가지고 추진하고 있다.⁵⁾

4. 중국

중국은 과학데이터의 공유와 활용을 국가 차원에서 적극적으로 추진하고 있다. 중국과학원은 국가 경제와 사회 발전을 위해 과학데이터의 공유 강화 및 과학자원시스템의 통합을 천명하고, 과학데이터의 보존, 과학데이터 기술 공유 및 활용을 위한 법·정책·규범의 확립, 과학데이터 공유 아키텍처 구축 등 과학데이터와 관련된 논의를 활발히 진행하고 있다.

중국 정부는 각종 연구개발사업의 수행으로 생산된 과학데이터의 공유 및 활용을 구체적으로 시행하기 위하여 2002년부터 ‘과학데이터공유공정’(科學數據共享工程)을 추진하였다, 이 계획은 국가 과학기술 발전을 위한 인프라 구축의 중요한 부분으로, 2020년까지 과학기술의 진보와 혁신, 경제 성장, 사회발전 및 국가보안을 위해 과학데이터를 체계적이고 통합적으로 관리하고 효율적으로 이용함을 목적으로 한다. 공정의 주요 임무는 첫째, 국가 과학데이터 관리 및 공유서비스체제 구축, 둘째, 과학데이터 자원의 구축, 셋째, 과학데이터의 규범화 및 표준화 추진, 넷째, 과학데이터 공유 정책 및 법률체계의 구축이다.

중국은 과학데이터의 국가적인 관리와 활용을 제도적으로 지원하기 위하여 2006년 “국가과학기술계획항목과학데이터제출임시법”(國家科技計劃項目科學數據匯交暫行辦法)을 입안하여 자국의 국가과학기술 프로젝트로부터 발생하는 과학데이터의 수집, 제출, 관리 및 활용에 관한 책임과 의무를 규정하는 등 과학데이터 관리정책을 크게 강화해 나가고자 계획하고 있다. 이 법안은 아직 국회를 통과하지 않아 정식 법률로 발효되지는 않았으나, 국가과학기술계획에 따른 각종 사업의 수행으로 생성되는 과학데이터의 체계적인 관리 및 활용을 위해 과학데이터 관리계획의 수립, 과학데이터 제출자의 권리와 의무, 과학데이터의 제출 절차와 방법, 과학데이터 심사 등에 관한 사항을 세부적으로 규정하고 있다.

5. 검토

이상에서 살펴본 바와 같이, 해외에서는 과학데이터를 체계적으로 보존, 관리 및 활용하기 위한 정책을 2000년대 중반부터 체계적으로 추진하고 있으며, 특히, 미국, 호주, 중국의 경우 다소의 차이는 있으나 관련 정책을 제도적으로 뒷받침하기 위한 법제를 동시에 마련해 나가고 있다. 우리나라도 이와 같은 상황을 고려하여 국가적 차원에서 과학데이터 관리 및 활용을 촉진하기 위한 정책과 법제를 조속히 정비할 것이 요청된다.

4) NSF, *The Proposal and Award Policies and Procedures Guide*, Part I - Proposal Preparation & Submission Guidelines GPG, Chapter II, C.2.j, nsf 11-1, II-18 ff.

5) OCI, *Sustainable Digital Data Preservation and Access Network Partners(DataNet)*, NSF, 2007, p. 23.

IV. 과학데이터 관련 현행 법령의 유형과 내용 분석

1. 과학데이터 관리정책 추진 현황

국내외적으로 과학데이터의 중요성 증대 및 국가적 차원에서 과학데이터 관리와 활용을 촉진하기 위한 정책의 필요성이 높아짐에 따라, 우리나라도 관련 정책을 수립하여 추진하고 있으나, 아직은 시작단계라고 할 수 있다. 우리나라에서 과학데이터의 공유와 활용을 촉진하기 위한 정책이 본격 추진된 것은 2010년에 들어와서라고 할 수 있다.

2000년대 후반, 사회적으로 빅데이터 활용이 본격 제기되는 것과 함께, 과학기술분야에 있어서도 과학기술 빅데이터 혹은 과학데이터에 대한 국가적 차원의 공유와 활용을 위한 정책이 필요하다는 요청에 의거, 정부는 국가차원의 과학데이터 관리 및 활용을 촉진하기 위한 방안을 모색하기 시작하였다. 2010년 당시 교육과학기술부는 KISTI를 통해 ‘과학데이터의 체계적 보존 및 활용을 위한 정책연구’를 수행하여 각 분야별 과학데이터 개발 전략을 모색하였으며, 2011년에는 국가 과학데이터의 관리 및 활용을 중심으로 담당할 분야별 ‘국가과학데이터센터 구축에 관한 방안’을 연구하였다. 또한, 미래창조과학부는 2013년 실제 과학데이터의 공유 및 활용을 촉진하기 위한 ‘사업과제를 발굴하기 위한 기획연구’를 수행하고, 이에 기초하여 2013년 말에는 관련 ‘사업의 본격적 추진을 위한 전략계획(ISP)’을 수립하였다. 정부는 수립된 전략계획에 따라, 2014년에는 과학데이터의 관리 및 활용을 위한 ‘시범사업’을 추진하고, 동년 8월 정부 3.0 추진시책과 연계한 ‘국가과학기술데이터포털’을 구축하여 시범서비스에 나서고 있다.

한편, 생물자원이나 화합물 등과 같은 과학적 연구자원이나 데이터에 대하여는 자원관리의 측면에서 접근하여 관리·활용하는 정책을 추진하고 있다. 즉, 국가연구개발사업의 성과물로서의 생물자원 및 화합물, 생물다양성협약(CBD)에 따른 각 분야별 생명자원 등의 경우 해당 부처별로 관련 데이터 및 정보자원을 관리·활용하기 위한 정책을 추진하고 있다. 우리나라의 과학데이터 관리정책은 데이터 자원의 성격에 따라 추진하는 체계라고 할 수 있다.

2. 과학데이터 관련 현행 법령의 종류 및 주요 내용

과학데이터의 수집, 관리 및 활용 등과 관련된 법령은 과학데이터의 유형과 범위를 어떻게 설정하는가에 따라 다를 수 있지만, 과학데이터 자체 또는 과학기술 연구자원으로 간주될 수 있는 데이터자료 등 넓은 의미의 과학데이터로 인식할 경우, 이와 관련된 현행 법령은 크게 과학기술기본법 및 관계 법령과, 분야별 생명(생물)자원의 관리 및 활용 등에 관한 법령 및 공공데이터의 공개 및 활용 촉진 관련 법령 등으로 나누어 볼 수 있다.

1) 과학기술기본법 및 국가연구개발 관련 법령

과학기술기본법은 국가과학기술정책의 기본방향과 목표를 정한 총괄규범으로서, 과학데이터에 관하여는 이를 직접 규정하지 않고 있으나, 과학기술 연구개발 활동에 필요한 각종 과학기술 지식·정보의 관리 및 유통에 관한 정책추진을 규정하고 있어, 과학데이터가 넓은 의미의 과학기술지식·정보에 해당하는 것으로 이해할 경우 간접적으로 규정하고 있다고 해석할 수 있다. 동법은 국가과학기술발전을 위하여 ‘과학기술지식과 정보자원의 확충·관리 및 유통체계의 구축’에 관한 사항을 「과학기술기본계획」으로 수립하도록 하고(제7조 제3항), 이를 구체적으로 실천하기 위하여

6) 미래창조과학부·한국과학기술정보연구원, 「국가과학기술 빅데이터 거버넌스 체제구축을 위한 전략계획(ISP)」, 2013.12.

정부로 하여금 ‘과학기술 및 국가연구개발사업 관련 지식·정보의 생산·유통·관리 및 활용을 촉진할 수 있는 시책’을 세우고 추진하며, 세부시책으로는 i) 과학기술 및 국가연구개발사업 관련 지식·정보의 수집·분석·가공 및 데이터베이스의 구축, ii) 과학기술 및 국가연구개발사업 관련 지식·정보망의 구축 및 운영, iii) 과학기술 및 국가연구개발사업 관련 지식·정보의 관리·유통기관의 육성 등을 규정하고 있다. 그러나 과학데이터 관련 사항은 그 포함여부가 불명확한 실정이다.

한편, 국가연구개발사업의 시행 및 관리 등에 관하여 구체적으로 규정하고 있는 「국가연구개발사업의 관리 등에 관한 규정」은 ‘국가연구개발사업의 연구개발정보를 체계적으로 관리하기 위한 계획(국가연구개발사업 연구개발정보 관리계획)’을 수립하고, 이를 효율적으로 시행할 수 있는 국가과학기술종합정보시스템을 구축·운영하도록 하고 있다(규정 제25조 제1항). 또한, 논문, 특허, 생물자원, 화합물 등 주요 연구개발성과에 대하여는 이를 효율적으로 관리하고 유통하기 위하여 미래창조과학부장관이 지정하는 기관⁷⁾에 등록하거나 기탁하도록 규정하고 있다(동 제13항). 이에 따라, 국가연구개발사업으로 창출된 연구개발성과 중 과학데이터로서 취급될 수 있는 생물자원이나 화합물의 경우 한국생명공학연구원과 한국화학연구원에 각각 등록되거나 기탁된다. 그러나 이 경우에도 연구개발성과로서의 과학데이터의 정의, 과학데이터의 범위와 내용, 과학데이터의 등록 및 기탁절차와 방법 등이 불명확한 실정이다.

2) 생명(생물)자원의 관리 및 활용 등에 관한 법령

생명(생물)자원의 관리 및 활용 등에 관한 법령은 “생물다양성협약(CBD)”의 체결에 따른 협약조건의 이행과 함께, 국가의 중요 생명(생물)자원을 체계적으로 보존·관리하기 위해 제정된 법령으로서, 2009년 5월 제정된 「생명연구자원의 확보·관리 및 활용에 관한 법률」을 비롯하여, 2011년 7월 전부 개정된 「농수산생명자원의 보존·관리 및 이용에 관한 법률」, 2011년 7월 일부 개정된 「야생생물 보호 및 관리에 관한 법률」, 2012년 2월 제정된 「생물다양성 보전 및 이용에 관한 법률」 및 2012년 6월 제정된 「해양생명자원의 확보·관리 및 이용 등에 관한 법률」 등과 그 하위규정들을 말한다. 이들 법령은 국가생명(생물)자원이 생명공학연구의 기반이 될 뿐만 아니라 육지, 연안, 해양 등 다양하게 존재하는 각종 생명(생물)자원이 관련 연구 또는 산업을 위해 실체적 또는 잠재적 가치가 있는 자원으로서 기능하므로 국가적인 차원에서 해당 생명(생물)자원을 체계적으로 보존·관리 및 활용을 도모하기 위한 것이다.

이 법령들은 각종 생명(생물)자원의 보존·관리의무와 함께, 일정한 자원의 기탁·등록, 해당 자원의 유통과 보호 등에 관한 사항을 규정하고 있다. 예를 들면, 「생명연구자원의 확보·관리 및 활용에 관한 법률」은 국가 및 지방자치단체로 하여금 생명연구자원의 효율적인 확보·관리 및 활용을 위해 노력하고(동법 제3조), 미래창조과학부를 비롯한 농림축산식품부, 환경부, 해양수산부, 산업통상자원부 등 관계부처는 생명연구자원의 확보·관리 및 활용체제의 조성과 이를 지원하기 위한 소관분야의 세부 시책을 마련하도록 규정하고 있다(동 제6조). 또한, 생명연구자원의 기탁·등록·보존 기관의 지정(제8조), 생명연구자원의 기탁등록보존 절차와 방법(제9조), 생명연구자원 정보센터의 지정·운영(제10조) 등에 관한 사항을 규정하고 있다.

그러나 이와 같은 법령의 규정으로 해당 생명(생물)자원과 관련된 일부 과학데이터의 수집·관리 및 공유·활용은 부분적으로 이루어질 수 있으나, 이는 연구자원의 보전 및 관리의 측면에서 다루어지는 것으로 연구 성과물로서의 과학데이터의 특성을 고려한 데이터 관리 혹은 데이터 자체의 공유와 활용 촉진을 직접 목적으로 하는 법령 규정으로서의 의미는 부족하다고 할 것이다.

7) 중요 국가연구개발 성과 및 동 등록·기탁기관으로는 한국과학기술정보연구원(논문, 보고서원문), 한국지식재산전략원(특허), 한국기초과학지원연구원(연구시설·장비), 한국산업기술진흥원(기술요약정보), 한국생명공학연구원(생물자원)(생물자원 중 신물질은 농림축산식품교육문화정보원/농업유전자원센터), 한국화학연구원(화합물), 한국저작권위원회/정보통신산업진흥원(소프트웨어) 등이다.

3) 공공데이터의 공개 및 활용 촉진 관련 법령

공공데이터의 공개 및 활용 촉진 관련 법령은 국가나 지방자치단체 등 공공기관이 생성 또는 취득하여 관리하고 있는 각종 공공데이터를 일반 국민들에게 널리 공개하여 활용을 촉진하기 위한 것으로서, 대표적으로 2013년 7월 제정된 「공공데이터의 제공 및 이용 활성화에 관한 법률」을 들 수 있다.⁸⁾ 이 법률은 공공데이터를 국민이 최우선적으로 이용할 수 있도록 보장하고, 공공기관에 공공데이터 제공의무를 부여하여 효과적인 민간제공과 이용 활성화를 지원할 수 있는 법적 근거를 마련하는데 있다.

‘공공데이터’라 함은 데이터베이스나 전자파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보를 말하는 것으로서(동법 제2조), 이를 보유한 공공기관⁹⁾은 원칙적으로 비공개대상이거나 제3자의 권리를 침해하는 정보를 제외한 모든 공공데이터를 국민에게 제공하여야 한다(동법 제17조). 공공기관의 장은 공공데이터의 제공을 효율적으로 시행하기 위하여 소관 공공데이터 목록을 안전행정부장관에게 등록하고(동 제18조), 이용자가 온라인을 통해 접근·활용할 수 있도록 해당 공공기관 또는 정부가 운영하는 공공데이터 포털을 통해 제공하여야 하며(동 제26조), 특별한 사정이 없는 한 이를 중단할 수 없다(동 제28조).

이와 같은 규정에 따라, 공공기관이 보유하고 있는 과학데이터가 공공데이터에 해당하는 경우에는 일반에게 공개·제공될 수 있으므로, 이 법률은 과학데이터의 관리 및 활용에 관한 법령으로서 기능할 수 있다. 그러나 이들 법령은 공공데이터의 제공대상을 일정내용으로 제한하고 있는 바, 국가R&D사업의 수행으로 생성되는 과학데이터의 비공개성 및 지식재산으로서의 성격 등을 고려할 때 해당 과학데이터가 원활히 공개 및 활용될 수 있을지에 대하여는 일정한 한계가 있다.

3. 검토 및 소결

이상에서 살펴본 바와 같이, 현재 우리나라의 과학데이터 수집, 관리 및 활용 촉진을 위한 법제는 미흡한 실정이라고 할 수 있다. 즉, 과학기술 연구개발 관련 법령의 경우 특정 과학데이터 정보의 등록 및 유통 등에 관하여 규정하고 있으나 그 규율내용이 간접적·부분적이어서 일정한 한계가 있으며, 생명(생물)자원의 관리 및 활용 등에 관한 법령도 해당 자원의 보존 및 관리에 중점을 두고 있어 과학데이터의 체계적인 공개 및 유통과는 직접적인 관계가 없다. 공공데이터의 공개 및 활용 촉진에 관한 법령은 그 성질상 과학데이터의 관리 및 활용의 법제로서 일정 기능을 수행할 수 있으나, 공개대상 데이터의 범위가 제한되는 등 연구개발 성과정보인 과학데이터를 종합적·체계적으로 수집, 관리 및 활용할 수 있는 직접적인 근거법제로서는 일정한 한계가 있다.

이에 따라, 과학기술 연구개발 활동을 통해 생산되거나 처리·가공되는 중요 과학데이터에 대한 종합적인 관리정책을 조속히 수립함은 물론, 그 제도적 기반으로서 관련 법제를 체계적으로 정비하는 것이 필요하다. 특히, 국가연구개발사업의 성과인 과학데이터에 대한 체계적인 보존·관리 및 원활한 공개와 활용을 촉진할 수 있는 제도를 마련하여야 할 것이다.

8) 공공데이터를 포함하는 ‘공공정보’의 민간제공에 관하여는 「국가정보화기본법」(중요지식정보자원), 「전자정부법」(행정기관 정보), 「기상산업진흥법」(기상정보), 「공간정보산업진흥법」(공간정보), 「통계법」(통계정보) 및 「발명진흥법」(특허정보) 등이 있으나, 이들 법령의 규정은 선언적인 내용에 그치고 있어 근거법률로서의 실효성이 미흡하고, 또한 민간제공 대상 정보의 범위와 제공수단 및 이용자를 제한하는 등 제한적인 범위에서 정보가 제공되고 있는 실정이다.

9) 이 법에서 ‘공공기관’이라 함은 국가기관, 지방자치단체 및 「국가정보화 기본법」 제3조 제10호에 따른 공공기관 즉, 「공공기관의 운영에 관한 법률」에 따른 공공기관, 「지방공기업법」에 따른 지방공사 및 지방공단, 특별법에 따라 설립된 특수법인, 「초·중등교육법」, 「고등교육법」 및 그 밖의 다른 법률에 따라 설치된 각급 학교, 그 밖에 대통령령으로 정하는 법인·기관 및 단체 등을 말한다.

V. 과학데이터 공유 및 활용 촉진을 위한 법적 과제

1. 과학데이터 공유 · 활용을 둘러싼 법적 쟁점

과학데이터를 국가적 차원에서 체계적으로 생산 · 수집 · 관리하고, 연구자 등에게 효율적으로 공유 및 활용되도록 하기 위해서는 그 제도적 기반으로서의 관련 법령을 적절히 정비하는 것이 필요하다. 특히, 공공자금으로 지원되어 수행되는 국가연구개발사업의 성과로서 발생하는 과학데이터는 국가의 공공자산이자 연구개발 혁신을 위한 첨단정보자원으로 관리되어야 하므로, 이를 위한 체계적인 공유 · 활용체계를 구축하는 것이 중요하다.

그런데, 국가연구개발 과학데이터의 공유 및 활용체계가 원활히 구축 · 운영되도록 하기 위해서는 무엇보다 과학데이터를 중요한 국가적 자산으로서 인식하고 이를 보존 · 관리하도록 규정하는 것과 함께, 과학데이터의 생산자와 이용자 등 과학데이터 관계자간의 이해관계를 적절히 조절할 수 있는 법적 매카니즘을 구축하는 것이 필요하다고 할 것이다. 이러한 측면에서 국가연구개발 과학데이터의 공유 및 활용과 관련하여 제기되는 주요 법적 쟁점 및 제도화 과제로는 i) 과학데이터의 권리 및 그 귀속관계, ii) 과학데이터 공유 · 활용 촉진을 위한 정보의 등록 및 관리제도, iii) 과학데이터 소유자와 이용자 간의 사용관계, iv) 과학데이터 생산 · 제공자 등에 대한 평가 및 권리보호제도 등을 들 수 있다.

2. 과학데이터에 대한 소유 등 권리관계의 명확화

국가연구개발사업의 성과로서 발생된 과학데이터를 효율적으로 보존 · 관리 및 활용되도록 하기 위해서는 그에 대한 소유관계를 명확히 할 필요가 있다. 과학데이터에 대한 권리를 누가 소유하는가를 명확히 하여야 그에 대한 권리귀속 및 사용관계를 규율할 수 있으며, 국가연구개발 성과관리에 따른 일정한 권리의무의 주체를 확정할 수 있기 때문이다.

구체적으로는 과학데이터도 연구 성과의 하나로 명시하고, 일반 연구 성과와 같이 해당 국가연구개발사업을 주관하여 수행하는 주관연구기관 등이 그에 대한 권리를 보유하도록 하는 것이 필요하다. 이를 통해, 과학데이터에 대한 권리관계를 국가연구개발 성과관리의 일반원칙에 부합되면서, 다른 연구 성과의 권리관계와 일원화하는 것이 바람직하다.

또한, 과학데이터에 대한 권리주체(소유자)는 해당 과학데이터를 일정한 방법에 따라 보존하도록 하여 과학데이터를 중요 국가연구자원으로서 관리하도록 하는 것도 필요하다. 과학데이터의 보존 및 관리는 주관연구기관 등에게 일정한 준수 의무로 규정하는 방안도 고려할 수 있으나, 연구사업의 신청서 및 관련 연구계획서를 제출할 때 과학데이터의 보존 및 관리에 대한 자체계획을 수립하여 제시하도록 함으로써 자율적인 참여를 유도하는 것도 검토할 수 있다.

3. 과학데이터 등록 및 관리제도의 활성화

과학데이터를 산 · 학 · 연 등의 연구자가 원활히 공유 및 활용하도록 하기 위해서는 관련 정보에 상시 접근하여 그 소재와 이용방법 등을 파악할 수 있도록 하는 것이 필요하다. 또, 중요 과학데이터는 국가적 차원에서 안정적으로 보존되도록 관리체계를 정비하는 것이 바람직하다.

구체적으로는 과학데이터의 소유주체로 하여금 연구사업의 종료 후 또는 과학데이터를 생성 · 취득한 후 일정기간 이내에 관련 정보를 공개하고, 이를 성과관리기관에 등록하거나 기탁하도록 규정하는 것이 필요하다. 등록 및 기탁의 절차와 방법은 원칙적으로 다른 연구 성과와 같은 방식으로 하되, 과학데이터의 특성을 고려하도록 하여야 한다. 다만, 모든 국가연구개발 과학데이터를 등

록하여 관리하는 것이 바람직할 수도 있으나 방대한 과학데이터의 수량과 유형에 따른 현실적인 관리 가능성 등을 고려하여 등록 및 기탁대상은 연구과정 및 결과를 검증할 수 있는 데이터 자료나 기록, 또는 해당 분야의 연구계속성 등을 위해 연구자 사회에서 가치가 있다고 인정되는 등 국가적으로 보존·관리할 필요가 있다고 인정되는 과학데이터 자료나 정보파일로 한정하는 것이 바람직하다.

과학데이터 관리기관은 정보 및 데이터 자료의 관리·유통서비스 등을 고려하여 과학기술정보 유통기관으로 지정하는 것이 합리적이며,¹⁰⁾ 특히 이들 관리기관은 등록 또는 기탁 받은 과학데이터에 대한 관리의무와 대외적 서비스의무를 규정하여 과학데이터가 널리 공유 및 활용되도록 제도화할 필요가 있다.

4. 과학데이터 보유자와 이용자 간의 사용관계 정립

과학데이터의 공유 및 활용을 촉진하기 위해서는 앞서 검토한 국가적 차원의 관리체계를 정비하는 것도 필요하지만, 궁극적으로는 과학데이터를 보유한 자와 이를 이용하는 자 간의 자율적인 제공 및 사용이 장려되도록 하는 것이 중요하다. 과학데이터는 그 속성상 단순히 데이터 정보를 공유하는 것에 그치지 않고 데이터 자원의 실질적인 제공과 데이터 사용에 필요한 분석 도구나 방법론 등이 동시에 공유되어야 하며, 또한 과학데이터의 질적 수준과 내용 및 이용가치를 제고하는 등 실효성 있는 과학데이터를 확보하고 공유하도록 하기 위해서는 실제 연구를 담당할 연구자의 협조 없이는 원활히 운영되기 어렵기 때문이다.

구체적으로 과학데이터 보유자와 이용자 간의 ‘과학데이터 이용계약’을 통해 과학데이터의 사용 목적, 사용범위와 내용, 데이터 제공절차와 방법, 과학데이터 사용에 따른 성과의 소유 및 활용, 과학데이터 사용에 따른 대가의 지불 등에 관한 사항들을 자율적으로 협의·결정할 수 있도록 유도하며, 이를 위한 표준계약서나 약관 등을 제정하여 이용이 활성화되도록 하는 것이 필요하다.

과학데이터 이용계약은 데이터의 성격 및 특성에 따라 민법상의 계약이나 국공유재산법상의 대부계약, 저작권법상의 이용허락계약 등의 방식으로 다양하게 체결할 수 있으며, 계약에 따른 세부적인 사항은 관련 법령에 따라 당사자 간의 개별적인 합의사항을 반영할 수 있을 것으로 생각된다. 다만, 과학데이터 이용계약을 체결함에 있어서 국가연구개발 과학데이터의 경우 그 공공성을 고려하여 이용조건 등에 있어서 일정한 공적 제한을 가하는 것은 가능할 것이다.

5. 과학데이터 생산자 등에 대한 권리보호 강화

과학데이터의 공유와 활용을 촉진하기 위해서는 무엇보다 과학데이터를 생산하거나 제공하는 자에 대하여 그 성과에 대한 적절한 보상체계를 마련하는 것이 필요하다. 과학데이터의 공유 및 활용은 해당 과학데이터의 질적 수준과 제공되는 수단과 방법 등 이용편의성에 따라 좌우되며, 이와 같은 데이터의 신뢰성 및 사용성의 정도는 과학데이터의 생산자 또는 데이터 제공자의 역량과 의지에 달려 있기 때문이다.

구체적으로는 과학데이터를 생산한 연구자와, 과학데이터를 소유하는 연구기관 등에 대하여 과학데이터의 생산 및 제공 실적에 따라 일정한 성과평가를 실시하고, 특히 과학데이터의 활용으로 일정한 경제적 이익이 발생하는 경우 관련자에게 적절한 보상체계를 마련하는 것이 필요하다. 즉, 국가연구개발 사업이나 과제 평가 시 논문, 특허 등과 같이 과학데이터도 하나의 연구 성과로서 인정하고, 연구기관이나 연구자에 대한 평가에 있어서도 중요 과학데이터의 생산 및 개발실적을

10) 국가 과학데이터 관리기관은 과학데이터의 다양성 및 관리상의 특수성 등을 고려하여 과학데이터의 통일적인 관리와 표준화 등을 담당 지원하는 국가종합과학데이터센터와, 각 기술 분야별 과학데이터의 관리와 유통을 담당하는 분야별 과학데이터센터로 구분하여 지정하는 것이 바람직하다.

독립된 연구 성과로서 반영하여야 한다.

과학데이터의 활용에 따라 경제적 이익이 발생한 경우의 보상은 현재 특허 등 국가연구개발 성과물의 실시에 따라 징수하는 기술료 등 수입의 배분원칙과 연계하여, 과학데이터의 개발·생산자에게도 그 이익에 상응하는 보상금을 지급할 수 있도록 하는 방안을 검토할 수 있다. 즉, 과학데이터를 판매하거나 사용계약을 통해 이용료 등의 수입이 발생하는 경우에는 해당 과학데이터의 개발 또는 생산에 기여한 자에게 일정한 보상을 할 수 있도록 제도화하는 것이다.

VI. 맺음말

최근 각 분야에서 데이터를 이용한 현상 분석과 정책수립 등 데이터 활용이 크게 늘어나고 있는 경향에 맞추어, 과학기술 분야에 있어서도 연구개발 활동이 가상화되고 첨단기술 간의 융·복합 현상이 확대됨에 따라 과학데이터의 중요성이 크게 증가하고 있다.

그런데, 우리나라의 경우 국가R&D 성과로서의 과학데이터를 체계적이고 종합적으로 관리 및 활용하는 체계가 크게 미흡한 실정이다. 국가연구개발 성과관리 법령을 비롯하여 관계 법령상 과학데이터 규정이 불명확·불충분한 실정이고, 과학데이터의 자율적인 공개 및 공유에 대한 연구자들의 인식과 노력도 부족하며, 정부의 관련 정책도 아직 본격적으로 추진되지 않고 있다.

최근 공공데이터의 개방 및 이용 활성화를 위한 법령이 정비되어 공공기관 등이 보유한 과학데이터의 경우 부분적으로 공유 및 활용할 수 있는 체계가 일부 갖추어지기는 하였으나 기술개발 관련 내용이나 영업정보 등이 포함된 경우 데이터 공개가 제한되는 등 그 범위와 내용이 한정적이어서 과학데이터의 종합적이고 지속적인 관리 및 활용체계를 구축하는 데는 한계가 있다. 따라서 과학데이터가 사장되지 않고 연구자들 간에 서로 공유 및 활용이 촉진되도록 하기 위해서는 관련 법령의 정비를 통해 다음과 같은 제도적 장치를 마련하는 것이 필요하다.

첫째, 과학데이터의 소유 등 권리관계를 명확히 하여 데이터의 공개 및 활용과 관련한 법률문제를 효과적으로 해결할 필요가 있다. 즉, 과학데이터도 연구 성과의 하나로서 규정하고, 다른 연구개발 성과와 마찬가지로 주관연구기관 등이 소유함을 명확히 하여야 한다.

둘째, 과학데이터의 국가적 관리를 효율적으로 수행하기 위하여 과학데이터의 등록 및 기탁제도를 활성화할 필요가 있다. 현행 국가연구개발 성과의 등록·기탁제도를 확대하여 중요 과학데이터에 대해서는 이를 등록하고, 특히 핵심적인 과학데이터는 성과관리기관에 기탁하도록 하여 체계적인 관리를 도모하고, 과학데이터 정보의 적극적인 공개와 유통을 촉진하여야 한다.

셋째, 과학데이터 소유자와 이용자 간의 사용관계를 합리적으로 정립하여 자율적인 협력을 바탕으로 한 공유와 활용을 도모할 필요가 있다. 즉, 과학데이터 이용계약을 통해 과학데이터의 이용조건과 범위, 데이터 사용에 따른 이용료의 지불 등이 원활히 이루어지도록 함으로써 과학데이터 생산과 이용의 선순환구조가 정착되도록 하여야 한다.

넷째, 과학데이터 생산자 등에 대한 평가와 권리보호를 강화하여 고품질의 과학데이터가 생성되도록 하고, 과학데이터 활용에 따른 적절한 인센티브의 제공을 통해 연구자들의 참여의욕을 제고하는 것이 필요하다. 즉, 현행 국가연구개발 성과에 대한 평가 및 보상제도와 연계하여 그 생산자 및 제공자 등에게 인센티브를 제공하는 방안을 적극 검토하여야 한다.

마지막으로, 위에서 언급한 여러 가지 법적·제도적 장치의 마련과 함께, 무엇보다 현장 연구자들의 과학데이터에 대한 중요성 인식과, 과학데이터를 상호 공유하여 활용하고자 하는 참여의지가 필요하다는 점을 강조하고자 한다.

[참고문헌]

- 교육과학기술부·한국생명공학연구원, 「국가생명연구자원 기반구축 기획연구」, 2010. 12.
- 교육과학기술부·한국과학기술정보연구원, 「과학데이터의 체계적 보존/활용을 위한 큐레이션 정책 연구 및 시범 구축」, 2010. 12.
- 교육과학기술부·한국과학기술정보연구원, 「국가 과학데이터센터 구축을 위한 과학데이터 관리 및 활용체계 연구」, 2011. 12.
- 미래창조과학부·한국과학기술정보연구원, 「과학기술 빅데이터 추진과제 발굴 및 활용 극대화를 위한 추진전략 마련 기획연구」, 2013. 3.
- 미래창조과학부·한국과학기술정보연구원, 「국가과학기술 빅데이터 거버넌스 체제구축을 위한 전략계획(ISP)」, 2013. 12.
- 박동진, “오스트레일리아의 과학데이터 서비스체제(ANDS) 분석과 시사점”, 디지털정책연구 제9권 제3호, 2011.
- 윤종민·김규빈, “과학데이터에 관한 입법례와 관리정책 그리고 대응방안 - 호주, 미국, 중국을 중심으로 -”, 기술혁신학회지 제16권 제1호, 2013. 3.
- 윤종민 외, 「국가과학데이터의 효율적인 관리 및 활용을 위한 법제도 기본연구」, 한국과학기술정보연구원, 2012. 9.
- 윤종민 외, 「국가연구개발 결과물의 공동 활용을 위한 권리보호 방안 연구」, 한국과학기술정보연구원, 2014. 11.
- 한국과학기술정보연구원, 「국가과학데이터 공유·융합체제 구축에 관한 연구」, 2011.
- 행정안전부, 「스마트 행정 구현을 위한 빅데이터 마스터플랜 추진현황 및 향후 계획」, 2012. 6.
- 張先恩, “國家科學數據共享工程”, 科學中國人, 2004,
- 中國 國家科技計劃項目科學數據匯交暫行辦法案, 2006.
- Alex Szalay, “Data Driven Discovery in Science : The Fourth Paradigm”, NITRD(the Federal Networking and Information Technology Research and Development) Symposium, 2010.
- Australian National Data Service(ANDS), *ANDS Guide: What is Research Data*, 2011.
- Clifford Lynch, “Jim Gray’s Fourth Paradigm and the Construction of the Scientific Record”, in Tony Hey et al. (eds.), *The Fourth Paradigm*, 2009.
- NHMRC/ARC, *Australian Code for the Responsible Conduct of Research*, Part I.
- NSF, *The Proposal and Award Policies and Procedures Guide*, Part I - Proposal Preparation & Submission Guidelines GPG, Chapter II.
- OECD, *OECD Principles and Guidelines for Access to Research Data from Public Funding*, 2006. 12.
- OCI, *Sustainable Digital Data Preservation and Access Network Partners(DataNet)*, NSF, 2007.
- Trever Gale, Rob Gilbert, Terri Seddon, “The RQF and educational research in Australia: Implications and responses across the research community”, Annual Conference of the Australian Association for Research in Education, Sydney, 2005.