

과학데이터에 관한 입법례와 관리정책 그리고 대응방안

- 호주, 미국, 중국을 중심으로 -

윤종민*·김규빈**·이상환***

I. 서론

21세기 이후 과학연구는 과거보다는 더욱 더 데이터 중심으로, 그리고 융합화 되고 있다. 수집, 분석, 재분석 및 저장되는 데이터의 양 또한 컴퓨터시뮬레이션과 모델링, 자동데이터획득 및 통신 기술로 말미암아 대규모로 증가하고 있다. 방대한 데이터의 생성으로 연구패러다임은 이전의 연구 패러다임, 즉 실험(관측), 이론 및 대규모 컴퓨터시뮬레이션의 연구방식을 통합한 “데이터중심의 과학”(e-Science)이라는 “제4세대 패러다임”으로¹⁾ 전환되고 있다.

과학데이터는 과학기술의 발전에 따른 연구개발과정에서 생성된 사실적 기록으로 집약될 수 있으며, 연구패러다임이 데이터 중심의 연계·융합연구로 전환되면서 과학데이터에 대한 중요성과 그 가치가 높아지고 있다. 수많은 연구개발과정에서 다양한 형태와 규모를 가진 과학데이터는 획득 → 큐레이션(처리, 가공, 저장, 접근) → 분석(모델링, 마이닝, 시뮬레이션) → 가시화(visualization) → 출판 및 인용단계의 과정을 거치게 된다. 이러한 과정에서 생산된 과학데이터가 창의적인 연구개발을 위해 효율적으로 재사용될 수 있기 위해서는 공유와 활용을 위한 관리체제의 구축이 필수적이다. 과학데이터의 공유와 활용을 위한 관리체제의 구축은 국가적 차원에서 이루어져야 하지만, 우리나라의 경우 관리체제의 수준은 호주, 미국, 중국 또는 유럽에 비해서 연계성으로나 효율성으로나 내실을 기하지 못하였다고 볼 수 있다. 실제로 과학데이터는 보고서 또는 논문을 위한 일회소비성 데이터가 되거나, 아예 유실 또는 사장되는 경우가 대부분이었음을 인정하지 않을 수 없다.

이러한 외화내빈(外華內貧)은 여러 가지 원인에 기인하지만 그 중 두 가지 상황이 지배적이라고 볼 수 있다. 그 하나는 연구 분야에 따라 과학데이터의 형태나 규모가 매우 다종다양하기 때문에 데이터를 수집하는 장비 또한 특수할 수밖에 없다. 사정이 이러하므로 장비고유성(instrument-specific)에 맞추어 데이터를 획득하는 과정은 결코 쉬운 일이 아닐 수 없다. 그 결과 대부분의 과학데이터는 데이터의 최종 이용에 초점을 두면서 생산되기 보다는 과학데이터의 생성을 용이하게 하는 형태와 구조 속에서 생산되고 있다.

나머지 하나는 과학데이터의 범위(scale)에 있다. 연구자들은 양적으로나 질적으로 과학데이터의 가치를 극대화하고자 한다. 또한 그들은 이를 실현하기 위해서 데이터를 생성하거나 획득할 때 소요되는 시간·비용을 최소화시킬 수 있는 데이터 수집 장치의 설계능력을 향상시켜야 할 필요성을 가지게 된다. 그렇기 때문에 과학데이터의 활용 면에 있어서는 연구자들은 그 효율성이 떨어질 수밖에 없다. 이와는 달리 전통적인 데이터베이스를 이용한 데이터 수집은 오히려 대부분 획득된 데이터의 질의기반(query-based) 검색의 효율성에 집중되고 있다고 볼 수 있다.

따라서 논의 초점은 효율적이고 공정한 과학데이터의 활용에 모여진다. OECD, CODATA 등의 국

* 윤종민, 충북대학교 법학전문대학원 교수, 010-7412-7421, cmyoon@cbnu.ac.kr.

** 김규빈, 충북대학교 법학연구소 전임연구원, 010-3399-0824, kimkyubin@cbnu.ac.kr.

*** 이상환, 한국과학기술정보연구원 책임연구원, 010-3437-0614, sanglee@kisti.re.kr.

1) Clifford Lynch, “Jim Gray’s Fourth Paradigm and the Construction of the Scientific Record”, in Tony Hey et al. (eds.), *The Fourth Paradigm*, 2009, p. 177.

제기구는 효율적이고 공정한 과학데이터의 활용을 위해 데이터 접근에 대한 개방체제(open access)를 강조하고 있으며, OECD는 이에 관한 가이드라인 또는 지침을 마련하여 회원국들에게 정책수립 방향과 법제도적 근거를 제시하고 있다. 호주, 미국, 중국 등은 국가차원에서 관련 기관을 통해 과학데이터를 수집, 관리 및 유지하는 등 데이터 활용을 적극적으로 추진하기 위하여 중장기적인 정책수립, 법제도 정비, 기반시설에 대한 투자를 지속적으로 확대 해나가고 있다. 이하에서는 효율적이고 공정한 과학데이터의 공유 및 활용을 위한 관리체계구축 및 이를 뒷받침할 수 있는 법제도를 보다 명확히 이해하기 위해서 국외 입법례 및 정책에 관하여 살펴보고 우리나라의 대응방안을 제시해보고자 한다.

II. 호주의 책임연구수행법과 과학데이터 관리체계

1. 개 관

호주의 과학데이터 관리체계는 국가차원에서 데이터 수집, 공유, 관리, 서비스 등 과학데이터를 위한 국가 인프라구축이 체계적으로 이루어지고 있다. 2006년부터 국가차원의 과학데이터관리와 관련된 정책수행이 본격화되면서 호주국가데이터서비스(Australian National Data Service, ANDS)가 구축·시행되고 있다. ANDS는 과학데이터·연구자·연구·매체(instruments)·기관간의 협업으로 이루어지는 체제로서 연구자가 데이터에 접속하여 데이터의 정보를 파악하고 공유 가능한 과학데이터수집과 그러한 공유의場に 데이터를 추가하고 활용할 수 있도록 한 호주과학데이터공유체제(Australian Research Data Commons, ARDC)를 구축하여 운영하고 있다.

ARDC의 운용체제는 “Research Data Australia”와 “Cite My Data”등과 같은 포털을 운영함으로써 구체화되고 있으며 데이터캡처 도구, 메타데이터관리 도구, 출판도구, 데이터 활용 도구, 데이터인용 도구 등 다양한 도구들이 활용되고 있다. 이러한 운용체제를 통하여 호주 연구자들은 과학데이터를 용이하게 발표하고, 검색하며, 접근하고 이용할 수 있게 된다.

이러한 ANDS의 운용은 호주국립보건의학연구협의회(Australian National Health and Medical Research Council, NHMRC)가 호주연구협의회(Australian Research Council) 및 대학과 공동으로 전 연구 분야에 걸쳐 적용될 수 있는 연구원칙을 규정한 “책임연구수행법”(Australian Code for the Responsible Conduct of Research, ACRCR)에 그 규범적 근거를 둔다. ACRCR는 연구기관과 연구자가 책임 있는 연구를 수행하고 연구의 진실성(integrity) 확보를 위해 마련된 규범으로 연구자, 정책입안자 및 연구 집단에게 ‘바람직한 연구실천’(best research practices)을 위한 관리방법을 지원하는 역할을 한다. ACRCR는 과학데이터의 관리 및 보급, 연구윤리, 동료심사 등에 관한 연구자와 연구기관의 책임과 의무를 규정하고 있으며 호주의 전 대학이 ACRCR의 이행을 위해 내규지침을 마련·시행하고 있다.

2. 호주책임연구수행법

1) 목적과 구성

호주책임연구수행법(ACRCR)은 연구기관과 연구자가 책임 있는 연구를 수행하고 연구의 진실성(integrity) 확보를 위해 마련된 규범으로 연구자, 정책입안자 및 연구 집단에게 ‘바람직한 연구실천’(best research practices)을 위한 관리방법을 지원하는데 그 목적을 둔다.²⁾

ACRCR는 크게 두 부분으로 구성된다. 제1장은 연구기관 및 연구자의 책임 있는 연구수행을 진작시

키기 위한 일반원칙과 실천규범을 규정하고 있으며 주 내용은 책임 있는 연구를 위한 일반원칙, 과학 데이터와 1차 자료(primary materials)의 관리, 연구연수생(research trainees) 관리 및 감독, 연구 결과 출판 및 보급, 저자표시(authorship), 동료심사, 이해충돌, 연구소간 공동연구(collaborative research)이다. 반면 제2장은 연구책임수행법 위반, 연구부정행위, 혐의조사기준 등을 규정하고 있다.

2) 과학데이터의 규율범위와 대상

ACRCR는 과학데이터와 1차 자료의 관리를 위한 연구기관과 연구자의 책임을 규정하고 있다. 특히 데이터의 안전한 보관, 소유권 확인, 과학데이터의 보안과 비밀보장은 과학데이터를 보존하기 위한 연구기관의 중요한 의무이기도 하다.

ACRCR에는 과학데이터가 무엇인가에 대한 규정은 포함되어 있지 않다. 과학데이터에 어떠한 데이터가 포함되어야 하고 제외되어야 하는가에 대한 기준은 ANDS를 통해서 마련되었다. ANDS에서는 과학데이터의 개념정의를 필요로 하는 대학과 기타 연구기관에게 과학데이터에 관한 ANDS 지침을 마련해 주고 있으며, 이러한 지침은 호주 대학들의 대다수 과학데이터관리정책에 수용되고 있다. 대표적인 예로 Queensland 공과대학의 과학데이터관리정책에서 규정하고 있는 과학데이터에 관한 정의를 들 수 있다. 이에 따르면 과학데이터는 “사실, 관찰, 이미지, 컴퓨터프로그램결과, 기록, 측량 또는 경험(논거, 이론, 테스트 또는 가설 또는 기타 연구물에 기초한)의 형태에서 생성되는 데이터”를 의미한다. 데이터의 범위에는 수치, 기술, 시각 또는 청각 데이터 모두 포함된다.³⁾

많은 경우에 있어서 과학데이터와 1차 자료(primary materials)간의 구분은 명확하지는 않다. ACRCR에 따르면 ‘1차 자료’는 완성된 설문조사와 녹음을 의미하는 반면 과학데이터는 1차 자료에서 파생된 기록물을 의미하기 때문에 서로 다른 관리기준이 적용되어야 한다고 본다.⁴⁾ 그러나 1차 자료가 연구에서 ‘사실적 기록(factual records)’으로 사용된다면 과학데이터와의 구분이 애매 모호해질 수밖에 없다. 따라서 연구 집단에서는 연구 결과물로서 그 타당성이 입증되는 경우에만 과학데이터로 간주한다. 그리고 그러한 데이터는 일정기간동안 보존되어야 한다.⁵⁾

3) 과학데이터 관계자의 책임과 의무

(1) 책임 있는 연구수행

ACRCR에서의 기본원칙은 책임 있는 연구수행으로 조직의 연구문화를 조성하는 것이다. 이를 위해서는 연구에 대한 정직성과 진실성이 필수적이며 연구수행 시 인체연구참여자, 동물, 환경을 존중해야 하며 연구수행에 사용되는 공공재를 좋은 상태로 관리하고 다른 연구수행자의 역할에 대한 적절한 배려와 연구결과에 대한 책임 있는 보고가 중요한 요소로 작용한다.⁶⁾

(2) 연구 부정행위

연구부정행위에 관해서는 여러 가지 절차를 규정하고 있는데 이는 연구부정행위가 잘못된 정보의

2) NHMRC/ARC, Australian Code for the Responsible Conduct of Research, Australian Government 2007, p. 1, [http://www.nhmrc.gov.au/guidelines/publications/r39\(2012.8.18\)](http://www.nhmrc.gov.au/guidelines/publications/r39(2012.8.18)). 방문).

3) ANDS, “ANDS Guide: What is Reserach Data?”, 2011.8, ANDS, p. 1, <http://ands.org.au/guides/what-is-research-data.pdf> (2012.8.29. 방문).

4) Ibid., p. 3.

5) Ibid.

6) NHMRC. ARC, Australian Code for the Responsible Conduct of Research, Part A, Sec. 1, p. 1.3.

공식적 기록화 또는 연구 참여자, 동물 및 환경에 좋지 않은 영향을 미치게 되는 심각한 결과를 초래할 위험성이 현저하기 때문에 이러한 행위가 발생했을 경우에는 엄격한 대처를 하겠다는 의지표명이기도 하다. ACRCR에 따른 연구부정행위에는 ACRCR 위반의혹뿐만 아니라 위조, 변조 및 표절(FFP)이 포함된다. 여기에는 의도적인 행위이든 부주의로 야기된 행위이든 모두 포함된다.⁷⁾

4) 과학데이터 공유정책의 수립과 추진

ACRCR에서는 연구 수행과정에서 생산된 과학데이터와 1차 자료에 관하여는 보존할 것을 원칙으로 한다. 따라서 이러한 데이터를 안전하게 보존하고 관리할 수 있는 체계구축이 중요함을 명확히 하고 있다. 또한 과학데이터와 1차 자료에 대한 소유권과 비밀정보와 관련된 사항에 대해서는 법적인 보호가 이루어져야 함을 명시하고 있다.⁸⁾

연구결과의 출판 및 보급과 관련하여 연구기관의 경우 연구결과물의 책임 있는 출판 및 발표를 촉진시켜야 하며 비밀보호와 지적재산에 대한 관리체계를 마련해야 한다. 반면 연구자는 모든 연구결과물을 발표해야 할 의무가 있으며 출판과 발표 시 연구결과의 정확성을 유지해야 한다. 그렇기 때문에 다른 저자의 연구를 완전하고 정확하게 인용해야 하며 연구결과는 학술대회를 통해 발표해야 한다.

5) 과학데이터 관리·운영 관련 규범

ACRCR에서는 책임 있는 연구수행을 위해 과학데이터의 적절한 관리와 보존의무를 규정하고 있다. 과학데이터보존은 연구과제에서 생산된 모든 데이터를 대상으로 하기 때문에 중요하다. 현실적으로 광석, 생물, 설문조사, 녹음과 같은 1차 자료는 보존되기가 어려운 반면 분석시험, 테스트 결과, 기록, 실험실과 현장기록과 같은 1차 자료로부터 파생된 과학데이터는 보존되어야 하고 접근이 가능해야 한다.⁹⁾

연구기관은 과학데이터의 보존에 관한 정책을 수립해야 하며 관련 법령, 지침, 원칙이 적용되는 관행과 일치해야 한다. 과학데이터의 보존기간은 출판일로부터 최소 5년이다. 그러나 과학데이터의 특성에 따라 보존기간은 달라질 수 있다. 또한 연구기관은 과학데이터를 안전하고 보안이 잘 갖춰진 장소에 보관할 수 있는 보관시설을 설치해야 하며 아울러 과학데이터 소유권과 보관에 관한 정책을 수립해야 한다.

연구자는 연구기관의 정책에 따라 과학데이터를 관리해야 하며 다음의 사항들을 준수해야 한다. 첫째, 연구방법, 데이터출처 등에 관한 명확하고 정확한 기록을 해야 하며, 둘째, 과학데이터의 안전과 보안을 유지해야 하고, 셋째, 1차 자료를 실험실기록, 과학데이터 분석 자료와 동일한 수준으로 보호하고, 넷째, 과학데이터를 내구성이 강한 형태, 색인목록, 회복 가능한 형태로 보관해야 하며, 다섯째, 과학데이터목록을 접근 가능한 형태로 보존해야 하며, 여섯째, 윤리강령과 관련 법률에 따라 과학데이터를 관리해야 한다.¹⁰⁾

6) 과학데이터 공유절차와 방법

ACRCR는 호주정부에서 개발해온 과학데이터에 대한 접근성(Accessibility) 프레임워크를 반영하

7) Ibid., Part B, Sec. 10, p. 10.1-10.4.

8) Ibid., Part A, Sec. 2, p. 2.1-2.3.

9) Ibid., Part A, Sec. 2, p. 2.1.

10) Ibid., Part A, Sec. 2. 2.6, p. 2.3.

고 있다. 접근성 프레임워크는 호주 혁신산업과학연구고등교육부(Department of Innovation, Industry, Science, Research and Tertiary Education, DIISRTE)에서 추진해 온 NCRIS(National Collaborative Research Infrastructure Strategy)에 기반을 둔다. NCRIS는 연구자에게 접근가능하고 그들의 장기적 필요성에 상응할 수 있는 연구시설, 네트워크 및 기반시설구축을 위한 호주정부의 투자전략사업으로 특히 연구 집단의 우선과제를 지원하기 위해 개발된 eResearch전략이며, 각 분야별 연구의 우수성(excellence) 지원, 공동연구 육성, 국가이익창출에 목표를 두고 있다.¹¹⁾

ANDS는 데이터 공유를 위한 컬렉션의 근원을 미국 NSB에서 구별한 데이터수집의 세 가지 카테고리에 기반을 두면서 공유체제의 개념도를 발전시키고 있다.¹²⁾ 여기에서 세 가지 카테고리란 첫째, 과학데이터 수집은 연구과제 중심의 하나 또는 그 이상의 성과물이며 일반적으로 제한된 처리 또는 큐레이션의 대상이 되는 데이터를 포함한다. 둘째, 자원 또는 공동체 데이터 수집은 단일 과학 또는 공학집단이 제공한다. 셋째, 참조데이터 수집은 과학과 교육집단의 광범위한 세그먼트를 제공하는데 있다.¹³⁾ ARDC 데이터의 근원은 첫째, 연구기관, 연구자가 소유하고 있는 과학데이터, 둘째, 학술단체, 또는 학문분야에서 관리하는 과학데이터, 셋째, 비즈니스와 사회로부터 합법적으로 수집된 과학데이터, 다섯째, 정부기관에서 획득한 운영데이터라고 볼 수 있다.

ARDC의 운용을 위해서 다음과 같은 요소들을 포함하고 있는 바, 이는 ① 연구비가 지원되는 프로젝트와 공동참여(collaborative engagements)를 통해 연구와 데이터를 생산하는 조직과의 제휴, ② “Research Data Australia”와 “Cite My Data”와 같은 국가서비스 제공, ③ 공유 가능한 데이터수집, ④ 데이터컬렉션에 대한 정보, ⑤ 데이터캡처 도구, ⑥ 메타데이터관리 도구, ⑦ ANDS 포털: Research Data Australia, ⑧ 훈련과 기관포털, ⑨ 출판도구, ⑩ 데이터 활용 도구, ⑪ 데이터인용 도구에 관한 것이다.

III. 미국의 국립과학재단의 과학데이터 관리지침과 관리체제

1. 개 관

미 의회는 1950년 5월 과학의 진전, 국민보건 및 안녕·복지의 증진, 국가안전보장 등의 분야와 관련된 연구개발을 지원할 목적으로 국립과학재단(National Science Foundation, NSF)을 설립하였다.¹⁴⁾ NSF는 연방의 독립기관으로서 수학, 물리, 의료, 생물, 공학 등의 기초과학 분야 연구를 지원하는 핵심기관으로 기초과학 연구지원기관의 패러다임이 되고 있다. 또한 국가안보와 관련된 특수 분야의 과학연구 활동지원과 관련 인적자원 교육 및 양성을 위해 다양한 과학교육프로그램을 진행하고 있으며 국내외 과학자들 간에 과학데이터(scientific information)를 상호 교환할 수 있도록 그 기반을 조성하는 등 여러 가지 기능을 담당하고 있다.¹⁵⁾

11) Rhys Francis, “eResearch: from research by the knowledge poor to research by the knowledge rich”, International Association of Technological University Libraries 29th Annual Conference: Digital Discovery: Strategies & Solutions Auckland New Zealand, 21-24 April 2008, Australian eResearch Infrastructure Council, Department of Innovation, Industry Science and Research, p. 3.

12) ANDS Technical Working Group, “Towards the Australian Data Commons: A proposal for an Australian National Data Service”, DEST, 2007, p. 11.

13) National Science Board(NSB), “Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century”, NSF, 2005, p. 37 ff.

14) NSF의 설립은 1950년 제정된 “National Science Foundation Act of 1950”에 근거를 두고 있다, www.nsf.gov/about/history/legislation.pdf(2012.6.22. 방문).

15) NSF Act of 1950, Section 3, Public Law 81-507, 1950. 5. 10.

NSF는 데이터 중심으로 변화하고 있는 연구 환경에 맞춰 2007년부터 디지털데이터, 데이터분석 및 가시화(visualization)가 과학, 공학 그리고 교육 분야의 진보를 위해 필수적임을 강조하면서 과학데이터에 관한 기본 틀을 다음과 같이 제시하고 있으며, 주요내용은 ① 기록된 형태로 저장·보관된 데이터, ② 전문가를 불문하고 토론 및 분석 가능한 데이터, ③ 적절한 보호체제 하에 서의 데이터에 대한 개방된 접근, ④ 데이터의 신뢰할 만한 보존이다.¹⁶⁾

2011년부터 NSF는 NSF 신규과제신청 시 데이터관리계획(Data Management Plan, DMP)을 필수적으로 요구하고 있다. DMP는 2010년 발표한 “연구 성과물의 보급과 공유”(Dissemination and Sharing of Research Results)에¹⁷⁾ 관한 NSF정책을 달성하는데 핵심요소라 할 수 있다.¹⁸⁾

2010 미국 경쟁력강화법 제104조는 연방의 과학컬렉션(scientific collection)에 관한 내용을 규정하고 있다. 과학컬렉션을 관리하는 연방기관은 과학기술정책사무국(Office of Science and Technology Policy, OSTP)이다. OSTP가 주관하는 업무는 과학컬렉션의 장기보존과 온라인상의 접근, 품질, 조직, 접근을 개선하기 위하여 연방과학컬렉션의 관리와 이용정책을 정비·발전시키는 것이다.¹⁹⁾ 경쟁력강화법에서 규정하고 있는 “과학컬렉션”이란 수집품 또는 역사적, 예술적, 또는 문화적으로 의미가 있는, 그리고 경제성이 높은 표본데이터나 자료로서 그들이 지닌 시장 가치 보다는 장기간 연구자산(research asset)으로 활용될 수 있고, 과학을 지원할 목적으로 생산되는 일련의 물리적 표본, 생물 또는 무생물을 의미한다.²⁰⁾ OSTP는 다른 관련 연방기관과 협의 하에 연방과학컬렉션에 대한 접근과 콘텐츠에 관한 정보를 관리하는 정보센터(clearinghouse)를 마련해야 한다.²¹⁾

2. 미국 국립과학재단의 과제신청 및 선정절차 안내지침

1) 목적과 구성

과제신청 및 선정절차지침(The Proposal & Award Policies & Procedures Guide, PAPPG)은 NSF 지원프로그램의 이용, NSF 계약과 관련된 정보, NSF 계약대상자 선정 매뉴얼을 위해 마련된 규정이며 과학데이터와 관련된 규정들도 포함한다. PAPPG는 NSF의 연구과제 신청 및 선정절차와 관련된 내용을 규율하고 있으며 크게 두 부분으로 나누어진다. 첫 번째 부분은 NSF 과제신청안내지침(Grant Proposal Guide, GPG)으로 NSF의 과제신청 준비 및 제출과 관련된 규정이다. 그러나 NSF의 몇몇 프로그램은 GPG의 일반 규정을 바꿀 수 있는 권고조항이 있으며 그러한 경우 권고 규정을 따라야 한다. 두 번째 부분은 NSF의 일반연구지원(Grants), 공동협약(Cooperative agreements)에 관한 선정절차, 관리 및 모니터 등으로 구성되어 있다.

2) 과학데이터의 규율범위와 대상

NSF의 과학데이터관리계획에는 NSF 분과별로 지원된 과제에서 도출된 전체 또는 부분적 성과로

16) National Science Foundation(NSF), “Cyberinfrastructure Vision for 21st Century Discovery”, NSF Cyberinfrastructure Council, 2007, p. 2.

17) NSF, “Proposal and Award Policies and Procedures Guide, Part II - Award & Administration Guide(AAG)”, Chapter VI, D.4, NSF 11-1, 2010, p. VI-8 f.

18) NSF, “Proposal and Award Policies and Procedures Guide, Part I - Proposal Preparation & Submission Guidelines GPG, Chapter II, C.2.j, NSF 11-1, 2010, p. II-18 ff.

19) America Competes Reauthorization Act of 2010, Sec. 104 (a).

20) Ibid., Sec. 104 (e).

21) Ibid., Sec. 104 (b).

표현될 수 있는 모든 과학데이터가 포함된다. NSF는 미국 예산관리국(Office of Management and Budget, OMB)의 정의를 근거로 삼는다. 즉 과학데이터를 “과학자사회에서 일반적으로 연구결과를 검증하기 위하여 인정된, 기록이 된 사실적 자료”로 규정하고 있는 것이다. 기록보관이 가능하고 활용할 수 있는 기본적 수준의 과학데이터에는 ‘분석된 데이터’(analyzed data)와 ‘메타데이터’(metadata)가 포함된다. 이러한 유형의 데이터는 학위논문, 논문, 학술지논문, 원고 및 책(章 포함)부록데이터 그리고 인쇄 또는 전자출판물 형태로 발표되어야 한다.²²⁾

3) 과학데이터 관계자의 책임과 의무

과학데이터와 관련하여 연구기관, 연구자, NSF 및 기타 관계자에 대한 책임과 의무는 “책임 있는 연구수행”(Responsible Conduct of Research, RCR)과 “연구 부정행위”(Research Misconduct)의 범주에서 고려해 볼 수 있다. 미국은 소극적 의미의 ‘연구진실성 확보’인 연구 부정행위 발생 시의 공정하고 투명한 처리를 넘어 적극적이고 미래지향적인 노력으로 “책임 있는 연구수행”에 관해서 활발한 교육활동과 다양한 프로그램을 진행 중이다.²³⁾ NSF는 “책임 있는 윤리적 연구수행”(Responsible and Ethical Conduct of Research, RCR)이 연구결과의 탁월성과 공공의 신뢰를 위해 매우 중요함을 강조한다.

NSF의 “책임 있는 윤리적인 연구수행”은 정직성과 진실성(integrity)을 바탕으로 연구자가 연구결과를 발표하고 평가받으며 후속 연구에 긍정적인 영향을 줄 수 있도록 하는데 초점을 둔다. 이를 정착시키기 위한 조치로 NSF는 “연구 부정행위”(Research Misconduct)에 관한 엄격한 대응을 하고 있다. NSF의 “과제선정관리지침”(AAG)에는 “연구 부정행위”에 대한 NSF정책과 책임, 그리고 연구수행자의 역할에 관하여 규정하고 있다.

4) 과학데이터 공유정책의 수립과 추진

NSF의 과학데이터관리 및 연구 성과물의 공유계획은 GPG의 “특수정보와 부속서류”(Special Information and Supplementary Documentation) 규정에 근거하여 작성해야 한다. 여기에는 과학데이터의 보존, 문서화, 데이터의 공유, 표본, 물리적 수집, 교육과정자료, 기타 관련된 연구, 교육성과물도 포함된다.²⁴⁾ GPG의 “특수정보와 부속서류”에는 NSF의 다년간에 걸친 과학데이터정책을 명료하게 규정하고 있다. 모든 과제신청에는 데이터관리계획과 연구 성과물의 공유계획을 기술해야 하며, 만약 그러한 계획이 부재한 경우에는 그에 해당하는 합당한 근거를 제시해야 한다. 과학데이터관리계획은 신청과제의 지적 탁월성(intellectual merit)과 신청과제의 광범위한 영향력(broader impacts) 중 하나 또는 두 개 기준에 의해 평가된다.²⁵⁾

과제신청 시, 제출하는 “데이터관리계획”에는 연구 성과물의 보급과 공유에 관한 NSF의 정책에 어떻게 부합할 것인가에 대한 구체적 사항이 포함되어야 하며, 그 내용에는 아래와 같은 요소들이 포함될 수 있다.

- 데이터, 표본, 물리적 수집, 소프트웨어, 교육과정자료, 기타 과제수행 중 생산되는 자료의 유형
- 데이터, 메타데이터 포맷 및 내용에 사용되는 기준

22) NSF, Data Management for NSF Engineering Directorate Proposals and Awards, 2011.1.18., http://nsf.gov/eng/general/ENG_DMP_Policy.pdf(2012.8.1. 방문).

23) 한국학술진흥재단, 연구윤리 확립을 위한 정부 정책 방안 연구, 2007, 33면.

24) NSF, “Proposal and Award Policies and Procedures Guide, Part I - Proposal Preparation & Submission Guidelines GPG, Chapter II, B.2.d.(i).

25) NSF, *ibid.*, Chapter II, C.2.j.

- 개인정보, 비밀정보, 보안, 지적재산권, 또는 다른 권리나 요건의 적절한 보호규정을 포함한 데이터접근 및 공유정책
- 과학데이터의 재활용, 재배포(re-distribution), 파생물(derivatives) 생산에 관한 정책 및 규정
- 데이터, 표본, 기타 연구 성과물의 보관 및 접근유지계획

5) 과학데이터 공유절차와 방법

NSF는 과학과 공학 연구 집단의 이익을 위해 디지털과학데이터를 공유하고 관리하기 위한 강력하고 지속 가능한 기반을 마련할 수 있는 정책을 개선하고 시행하는데 중점을 두고 있다. 따라서 디지털과학데이터의 확산이 NSF의 중요한 과학데이터정책에 영향을 미칠 수 있다는 인식하에 국가과학위원회(National Science Board, NSB)를 통해 전략예산위원회(Committee on Strategy and Budget, CSB)와 위원회 산하 데이터정책프로젝트팀(Task Force on Data Policies, TFDP)을 설립하여 NSF의 연구지원으로 수집된 과학데이터가 적절하고 장기적으로 연구 집단에게 유용하게 이용될 수 있도록 공유 및 관리방안을 모색하였다.²⁶⁾ 이를 위해 TFDP가 2010년 4월부터 2011년 5월까지 준비한 보고서인 “디지털과학데이터 공유 및 관리”(Digital Research Data Sharing and Management)에서는 NSF 투자 사업으로 생성된 디지털과학데이터의 공유 및 관리와 관련된 중요한 문제와 권고사항을 제시하고 있다. 여기에서 디지털과학데이터는 대규모 협업프로젝트에서 수집되고 처리된 데이터를 의미하며 NSF의 지원을 받는 연구 결과물로서 디지털 형식으로 생산된 모든 과학데이터가 포함된다.²⁷⁾

데이터 공유 및 분석을 위한 새로운 접근 방법은 대규모 데이터셋으로 그 분야에서의 연구수행을 변환시키는 것이다. 보고서에서는 여섯 개 분야에서 10 개의 과제를 제시하고 있다. 6개 분야는 공유실천(commitment to sharing), 재현성(reproducibility), 교육, 인력개발 및 양성, 사이버인프라구조, 보존기한(longevity)과 지속가능성, 윤리적, 법적 함의가 포함된다.

CSB는 NSF와 관련하여 연구 집단을 위한 다섯 가지 디지털과학데이터 정책에 관한 권고사항(Recommendations)을 제안하고 있다. 권고는 공유실천, 재현성, 교육, 인력개발 및 양성, 사이버인프라구조, 보존기한과 지속가능성, 윤리적, 법적 함의에 관한 분야이다. 비록 이 5개 분야에서의 권고가 디지털과학데이터 공유 및 관리와 관련된 모든 핵심문제를 포함하고 있지는 않지만 과학과 공학 연구 집단에 유익한 디지털과학데이터에 관한 NSF의 정책이 시행되는 발전적인 단계를 나타내고 있다고 볼 수 있으며, 권고사항을 살펴보면 다음과 같다.²⁸⁾

- ① NSF 과학데이터정책에 부합되는 과학데이터공유 및 관리체제를 구축하기 위해 개별적인 과학 집단의 욕구를 포함한 디지털과학데이터정책의 개발과 시행을 주관하는 연방기관 및 기타 국내외 이해관계자에게 리더십을 제공한다.
- ② 연구 프로젝트에서 생성된 디지털과학데이터와 일치되도록 연구비수혜기관에게 상호 검토된 출판물의 그림, 표, 결과 및 결론의 체계구축 또는 검증을 목적으로 접근할 수 있는 데이터의 생산과 분석을 위해 사용되는 데이터, 방법 및 기술을 만들도록 요구해야 한다. 과학데이터는 저자와 협약과제지원의 자동속성을 인식할 수 있는 영속적 전자식별자를 사용하여 공유해야 한다. 인간, 독점콘텐츠 및 국가 안보와 관련된 과학데이터는 국내 및 국제규정에 부합하도록 적절한 조치를 필요로 한다.

26) NSF, “Digital Research Data Sharing and Management”, 2011, p. 1, <http://www.nsf.gov/nsb/publications/2011/nsb01211.pdf>(2012.9.8. 방문).

27) Ibid.

28) Ibid., p. 7-10.

- ③ 공유되고 접근할 수 있는 과학데이터를 활용하기 위해 컴퓨터와 데이터를 활용하는 과학·공학 연구자 및 사이버인프라구조 전문가에 대한 지원을 확대한다.
- ④ 디지털과학데이터의 유지와 관련된 문제와 장기사업 모델을 탐색, 개발하기 위한, 그리고 행동에 대한 핵심권고사항을 마련하는데 필요한 이해관계자의 패널을 개최한다.
- ⑤ 데이터공유 기반구조와 연구 집단, 연구기관 및 연구과제(research grants)의 범위를 포괄하는 장기적인 사업모델의 개발 및 시행을 장려하여 새로 생성되고 장기 보존되는 과학데이터의 보존과 큐레이션을 포함한 지속가능한 과학데이터 관리를 확장한다.

IV. 중국의 과학데이터입법례(안)와 관리체계

1. 개 관

중국 정부는 오랜 기간에 걸쳐 다양한 연구 분야에서 관측, 탐측, 조사, 실험 등을 통하여 대량의 과학데이터를 수집·생산하였으나, 공유체제가 구축되지 않아 과학데이터의 활용 면에서는 만족할 만한 성과를 거두지 못하였다. 중국에서 과학데이터와 관련된 정책의 수립은 ‘제10차 5개년계획(10.5계획, 2001-2005)’ 부터이다. 여기에는 ‘국가과학기술기초조건인프라건설(国家科技基础条件平台建设)’도 포함된다.

2002년부터 중국과학원에서 시행하고 있는 과학데이터공유공정은 국가과학기술기초조건 인프라 구축의 중요한 부분으로서 2020년까지 과학기술진보와 혁신, 경제성장, 사회발전 및 국가 보안을 위하여 통합계획, 정책규범 및 해당 규정, 현대 정보기술의 적용, 사회 전체의 네트워크 등을 구축하여, 과학데이터를 체계적이고 통합적으로 관리하고 효율적으로 이용함을 목적으로 한다. 그 주요 내용은 데이터베이스 구축, 과학데이터센터 구축, 인터넷포털사이트 구축으로 이를 위해서 국가과학데이터관리 및 공유서비스체계 구축, 과학데이터자원의 구축, 과학데이터의 규범화 및 표준화, 과학데이터 공유정책 및 법규체계의 구축이 선행되어야 함을 제시하고 있다.

중국에서 과학데이터와 관련한 논의들은 상당한 기간 동안 지속적으로 논의되어 왔으며, 과학데이터와 관련한 정책 및 법규 등은 국가차원에서 오랜기간 동안 체계적으로 추진하여왔다. 과학데이터 관련 현행 법률은 우리나라의 「과학기술기본법」과 비슷한 지위에 있는 「중화인민공화국 과학기술진보법」과 「중화인민공화국과학기술성과전환촉진법」에 일부 규정들이 연구성과 관리 차원과 공유인프라 구축차원에서 존재하고 있으며, 과학데이터 관련 단일법령은 현재까지는 제정되지 못하였다. 현행 관련 법률들에서 과학데이터의 공유가 법문에 명시되어 있지만 과학데이터에 대한 정의 및 범위, 과학데이터 관리 및 공유의 구체적인 내용이 명확하지 못하여 과학데이터를 효율적으로 관리·활용하기에는 미흡한 점이 있어 독립법제를 운영하기 위해 추진하고 있는 것으로 파악된다.

현재 중국에서는 과학데이터 공유와 관련한 입법안인 「국가과학기술계획항목 과학데이터 제출 실시법」, 「중화인민공화국 과학데이터공유조례」가 입안되어 입법추진 중에 있다. 동 법안들의 입법형식은 법률과 행정법규이며 과학데이터의 공유와 관리, 활용에 관한 문제들을 종합적이고 체계적으로 규정하고 있다.

2. 과학데이터 공유 법률(안)

1) 목적과 구성

과학데이터공유공정의 일환으로 과학데이터 공유인프라 구축에 대한 법적 근거 및 규제 지침을 마련하기 위해서, 2003년 9월 과학데이터공유공정판공실(科学数据共享工程办公室)에서 「국가과학기술계획항목 과학데이터 제출 임시법(国家科技计划项目科学数据汇交暂行办法)」을 입안하여 발표했다.²⁹⁾ 동 법안은 총칙, 제출 과학데이터의 종류 및 범위, 과학데이터 제출 계획, 과학데이터 제출 의무자의 권리와 의무, 과학데이터 제출의 순서, 과학데이터의 심사, 과학데이터의 관리, 과학데이터 제출 관리, 장려 및 처벌 등 총 9장 63조와 부칙으로 구성되어 있다.

2) 주요 내용

(1) 제출 과학데이터의 종류 및 범위

법안에 의하여 제출되는 과학데이터는 과학가치와 실용가치가 있어야 하며, 제출할 과학데이터의 범위는 국가과학기술계획항목 실시과정에서 생성된 초기관측데이터, 탐측데이터, 시험데이터, 실험데이터, 조사데이터, 정밀관찰데이터, 통계데이터 및 메타 데이터 등을 가리킨다(법안 제7조, 제8조). 과학데이터의 제출은 과학데이터 관련 관리 기구에서 규정한 양식에 따라 표준에 부합하게 전자문서로 제출하여야 하며, 과학데이터를 제출할 때에는 계약서, 과학데이터 제출계획서 등과 함께 제출하여야 한다(법안 제9조, 제11조).

(2) 과학데이터 제출 계획

과학데이터 제출 계획은 연구 신청 시 반드시 연구신청서와 연구계획서에 포함되어야 한다. 제출 계획의 내용으로는 제출 의무자, 과학데이터의 생성방식, 제출 종류와 범위, 제출 방식과 기한, 제출한 과학데이터의 관리기구, 과학데이터의 사용기한, 과학데이터의 품질요구, 기타 반드시 설명해야 하는 사항 등이다(법안 제12조, 제15조).

(3) 제출의무자의 권리와 의무

과학데이터 제출의무자는 국가과학기술계획항목의 담당자 혹은 그가 지정한 대리인이 과학데이터 제출의무를 부담하게 된다. 제출의무자의 소속기관은 과학데이터 제출에 협조할 의무가 있으며, 제출의무자가 확정되지 않았을 때에는 과기계획항목 담당기관의 책임자가 제출의무자가 된다. 과학데이터 제출의무자는 제출 계획에 따라 과학데이터를 제출하여야 하며, 제출규정, 순서, 기한에 따라 과학데이터를 제출하여야 한다. 또한 제출 과정 중에 과학데이터의 가치가 훼손되지 않도록 적절한 보관을 하여야 하며, 제출한 과학데이터의 진실성에 대해 책임을 지어야 한다. 제출의무자는 제출한 과학데이터에 대한 발표권리, 서명권리, 수정권리, 과학데이터 완성 보호권리, 사용권리 등을 가진다(법안 제21조-제25조).

(4) 과학데이터의 심사

제출된 과학데이터는 과학적인 의미와 실용가치, 제출계획의 요구조건에 부합되는지 심사받는다. 심사대상에는 과학데이터의 완전성과 진실성도 포함된다. 소속기관의 심사를 거친 과학데이터는

29) 동 법안의 원문은 科學數據共享工程 홈페이지(<http://www.sciencedata.cn/fagui.php>)에서 확인할 수 있음.

과학데이터 관리 기구의 전문가로 구성된 위원들의 심사를 받게 된다(법안 제38조).

(5) 과학데이터의 관리

과학데이터 제출의무자, 그 소속기관, 과학데이터 관리 기구는 과학데이터의 수집, 제출, 심사과정에서 과학데이터를 적절하게 관리하여야 하며, 과학데이터 관리 기구는 과학데이터 관리에 필요한 규정을 제정하고, 전문적인 관리인원을 배치하여 현대적으로 과학데이터를 보존하는 수단을 취해야 한다. 그리고 비밀이 요구되는 과학데이터는 특수한 조치를 취하여야 하며, 보존가치를 잃은 과학데이터는 적절하게 처리하여야 한다(법안 제45조). 과학데이터 관리 기구는 과학데이터의 합리적인 이용을 보장하고, 과학데이터의 공유를 추진하여야 한다(법안 제47조).

과학데이터 관리 기구에 제출된 과학데이터는 30일 이내에 공개하도록 규정하고 있으며, 과학데이터 제출의무자가 보호기간 내에 공개하는 것을 동의하였다면 30일 이전이라도 공개할 수 있다(법안 제49조-제50조).

3. 과학데이터 공유 행정법규(안)

1) 목적과 구성

「중화인민공화국 과학데이터공유조례(中华人民共和国科学数据共享条例)」에서는³⁰⁾ 과학기술의 진보와 사회발전을 위하여 사회적 이익과 경제적 이익에 합치해야 한다는 과학데이터 공유의 기본원칙을 천명하였으며, 과학데이터와 관련한 문제에 대하여 종합적이고 구체적인 행정법규로서 총칙, 과학데이터 공유관리, 과학데이터 제출, 과학데이터 공유서비스, 과학데이터 사용, 과학데이터 공유보장, 법률책임 등 총 7장 53조와 부칙으로 구성되어 있다.

2) 주요 내용

(1) 제출 과학데이터의 정의 및 적용범위

동 조례에서의 과학데이터는 국가재정 지원에 의한 국가과기계획항목에서 생성된 원시관측데이터, 탐측데이터, 실험데이터, 조사데이터, 고찰데이터, 원격탐지데이터, 통계데이터, 원시데이터 및 체계적으로 가공된 데이터 등이 포함된다(제2조-제3조).

(2) 과학데이터의 공유 관리

과학데이터 공유는 정부차원에서 각 부서가 역할분담을 하여 과학데이터의 생산, 보관, 사용에 이르기까지 통일된 관리 체계를 구축하도록 규정하고 있으며(제9조, 제11조). 국무원 과기행정 주관 부서가 관련 부서와 함께 ‘국가과학데이터공유센터’를 구축하도록 하고 있다(제12조).

(3) 과학데이터의 소유권 귀속

과학데이터는 과학기술창조와 국가발전의 중요한 자원으로, 국가재정 지원에 의한 과학기술활동에 의하여 생성된 경우, 과학데이터의 소유권은 원칙적으로 국가가 소유한다. 다만 일정한 경우에는 계약에 의하여 정할 수 있다고 규정하고 있다(제13조).

30) 동 조례의 원문은 중국법학망(<http://www.iolaw.org.cn/showNews.asp?id=22236>)에서 확인할 수 있음.

(4) 과학데이터 제출의무자의 권리와 의무

과학데이터 제출의무는 연구수행 책임자가 부담하게 되며, 적법절차에 따라 데이터 목차와 데이터 전자문서를 포함하여 국가데이터공유센터에 제출하여야 한다. 제출의무자는 크게 제출의무, 보관의무, 책임의무, 설명의무, 기밀보장의무, 심사받을 의무 등을 부담한다(제15조-제16조).

(5) 과학데이터의 표준화

국가과학데이터공유센터는 과학데이터 ‘기술표준’과 ‘기술규범’을 제정하고 발표해야 한다. 그러나 과학데이터의 종류가 매우 다양하기 때문에 과학데이터 관련 부서별, 영역별로 표준규범을 제정해야 함을 규정하고 있다(제19조).

(6) 과학데이터의 공유

제출된 과학데이터는 공유에 앞서서 국가과학데이터공유센터에 공시해야 한다(제24조-제25조). 이는 표절, 모방 등 과학데이터의 지적재산권 분쟁여지를 줄이고 원활한 과학데이터의 공유를 통하여 사회이익과 경제이익에 기여하기 위함이다. 국가과학데이터공유센터는 과학데이터 편집권, 발표, 데이터품질 감독권, 데이터가공처리권을 가지며, 과학데이터서비스 이용요금을 부과할 수 있다(제28조-제29조).

(7) 과학데이터의 이용

과학데이터의 합리적 이용을 담보하기 위하여 과학데이터의 등급과 종류를 분류하여 그 등급과 종류에 맞게 사용할 수 있도록 규정하고 있다(제31조). 과학데이터 이용자는 국가과학데이터 공유 사이트를 방문, 필요한 과학데이터 기본정보를 검색하고 무료로 온라인 서비스를 받을 권리를 지닌다(제32조, 제35조).

(8) 과학데이터 공유 관리 및 분쟁해결절차

국가과학데이터공유센터는 과학데이터의 공유가 안전하고 효율적으로 이루어질 수 있도록 과학데이터 품질감독과 비밀유지가 필요한 과학데이터는 필요한 조치를 취해야 한다(제39조, 제41조). 그리고 국무원 과학기술 행정주관부서는 과학데이터공유 분쟁해결기구를 설립하고, 개인 및 단체가 제출한 과학데이터 관련 분쟁을 해결하도록 한다. 분쟁해결기구의 결정은 법적 효력을 지니며, 중재결정에 불복한 경우 인민법원에 소를 제기할 수 있다(제42조, 45조).

V. 결론 및 대응방안

연구의 효율성 개선과 공유, 검증 및 재이용 가능한 과학데이터를 생산할 수 있는 기반은 과학데이터관리체제를 어떻게 운용하느냐에 달려있다. 과학데이터 접근에 대한 장벽은 혁신, 성장, 그리고 연구를 통해 발생할 수 있는 다양한 혜택을 제한하는 요인이 되기도 한다. 반면 데이터의 대규모 증가는 데이터에 대한 접근, 수집, 이용에 있어 그 배제성이 없는 이유로 “무임승차”(free rider)할 유인을 확산시킨다. 이러한 유인이 부정적 외부효과만을 준다고는 말할 수 없다. 그러나 연구자 또는 연구기관에게는 연구의 효율성을 저하시키는 요인이 되기도 한다. 이를 해결할 수 있는 방안은 국가차원의 과학데이터관리체제의 구축과 이를 위한 법적 근거를 마련하는 것이다. 국가 별로 다소 차이는 있으나 호주, 미국, 중국은 국가차원에서 과학데이터관리를 위해 중장기적

계획을 수립하고 예산을 확보하여 단계적으로 과학데이터관리사업을 추진하고 있으며, 이를 위한 법제도적 정비도 지속적으로 이루어지고 있다.

호주는 국가차원에서 과학데이터에 관한 관리와 공유체제가 구축되어 운용되고 있다. 2009년부터 운영되고 있는 ANDS는 국가적인 연구협력 체제하에서 과학데이터의 관리, 공유정책, 서비스개발을 추진하고 있으며, ARCS는 협업서비스 및 워크플로우를 담당하고 있다. ANDS는 네 개의 통합된 상호관련 서비스전송프로그램(service delivery programs)을 기반으로 사용자가 온라인접속을 통하여 ARDC 메타데이터에 접속하고, ARDC에서 제공하는 각종 서비스 유틸리티를 이용하여 연구기관에 저장된 실제과학데이터를 확보한다. 호주정부는 이러한 과학데이터공유와 관리체계를 안정적이고 효율적으로 운용하기 위하여 호주책임연구수행법(ACRCR)을 제정·시행하고 있다.

미국 NSF의 과제신청 및 선정절차지침(PAPPG)은 미국의 연방법과 NSF의 과학데이터 관련정책을 토대로 마련되었으며, 2011년부터는 NSF의 연구성과물에 대한 공유·관리정책에 따라 모든 과제신청에는 과학데이터관리계획과 연구 성과물의 공유계획을 작성·제출하도록 의무화하고 있다. 과제지원을 하는 연구기관 또는 연구자는 과학데이터관리계획을 위하여 ① 데이터, 표본, 물리적 수집, 소프트웨어, 교육과정자료, 기타 과제수행 중 생산되는 자료의 유형, ② 데이터, 메타데이터 포맷 및 내용에 사용되는 기준, ③ 개인정보, 비밀정보, 보안, 지적재산권, 또는 다른 권리나 요건의 적절한 보호규정을 포함한 데이터접근 및 공유정책, ④ 과학데이터의 재활용, 재보급(re-distribution), 파생물(derivatives) 생산에 관한 정책 및 규정, ⑤ 데이터, 표본, 기타 연구 성과물의 보관 및 접근유지계획 등을 구체적으로 기술해야 한다.

1980년대 이후 중국 정부는 과학기술연구발전에 관한 총체적인 국가과학기술계획을 수립하면서 이를 지원하기 위하여 과학기술법제 마련에 중점을 두기 시작했다. 과학데이터와 관련된 정책이 수립·추진된 시기는 2001년부터이다. 과학데이터공유공정은 국가과학기술의 기본요건과 기반구조 구축체제로 2002년부터 중국과학원에서 운용하고 있다. 과학데이터공유공정의 운용은 2020년까지 과학기술의 진보와 혁신, 경제성장, 사회발전 및 국가안보를 위한 통합계획, 정책 및 관련법 정비 등을 통하여 과학데이터를 체계적·통합적으로 관리하고 효율적으로 활용하는데 목적을 둔다. 과학데이터공유공정은 데이터베이스 구축, 과학데이터센터 구축, 인터넷포털사이트 구축이라는 3가지 주요 부분으로 구성되어 있다.

중국은 지난 십여 년간 과학데이터 공유정책과 입법논의가 진행 중이다. 아직까지는 과학데이터관리·공유에 관한 통일된 법제는 마련되지 않고 있으나, 2004년부터 2006년까지는 “과학데이터공유조례”와 “과학데이터공유법”의 입법근거를 마련하였으며, 2006년부터 2012년까지는 과학기술부에서 제정한 과학데이터 공유관련 부위규장(部委規章)을 토대로 국무원은 “과학데이터공유조례” 입법을 추진하여 과학데이터공유 관련법의 초안을 준비하였으며, 2012년부터 2020년까지는 현재까지 마련된 입법정책 및 법안을 지속적으로 정비·개선하여 “중화인민공화국 과학데이터공유법(안)”을 확립하고자 노력하고 있다.

우리나라는 아직까지 국가차원의 과학데이터관리·공유에 관한 정책수립 및 입법계획이 미흡한 수준이다. 다만 “과학기술기본법”에 따른 “국가연구개발사업의 관리 등에 관한 규정”의 성과물 관리제도에서 과학데이터와 관련된 부분을 포함하고 있을 뿐이다. 그리고 2010년 “생물다양성 협약”의 이행을 위한 ABS 나고야의정서(Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization) 체결로 국내이행을 위해 제·개정된 “생명연구자원의 확보·관리 및 활용에 관한 법률”, “해양생명자원의 확보·관리 및 이용 등에 관한 법률”, “농수산생명자원의 보존·관리 및 이용에 관한 법률” 등에 근거한 기탁·등록제도 등에 의해서 과학데이터의 공유 및 관리가 부분적으로 이루어지고 있다.

현재 학계뿐만 아니라, 여러 분야의 국가연구기관에서 활발하게 논의되고 있는 과학데이터의 공

유·관리 문제를 효과적으로 해결하기 위해서는 과학기술정책을 포함하여 국가과학데이터의 생성, 접근, 보존, 공유, 관리, 평가, 활용 등 일련의 과정을 충분히 이해하고 정책을 추진하여야 할 것이다. 먼저 과학데이터라는 이슈의 핵심을 간파하여 문제의식을 갖고, 국가차원의 해결방안을 마련해야 한다. 그 다음 단계로 그 문제에 대한 정책결정과정과 그 결과물로서의 입법적 대응을 모색하여야 할 것이다. 호주와 미국은 정책관련성에 있어서 우리나라의 중장기적 과학데이터의 공유·관리정책수립 및 수행에 많은 시사점을 주고 있다. 반면 입법적 대응이라고 할 수 있는 절차와 기준정립의 문제에 있어서는 호주와 중국의 입법례가 좋은 사례라고 볼 수 있다. 호주는 과학데이터와 관련하여 연구자, 연구기관의 책임과 의무를 포괄적으로 규정하고 있고, 중국은 과학데이터의 관리·공유에 적용될 수 있는 구체적인 사항들을 규정하고 있기 때문이다.

참고문헌

- 한국학술진흥재단(2007), 연구윤리 확립을 위한 정부 정책 방안 연구, 서울: 33.
- ANDS(2011), “ANDS Guide: What is Reserach Data?”, ANDS, <http://ands.org.au/guides/what-is-research-data.pdf> (29. August 2012).
- ANDS Technical Working Group(2007), “Towards the Australian Data Commons: A proposal for an Australian National Data Service”, DEST, 11.
- Clifford Lynch(2009), “Jim Gray’s Fourth Paradigm and the Construction of the Scientific Record”, in Tony Hey et al. (eds.), *The Fourth Paradigm*, Microsoft Research, Washington, 177.
- National Science Board(NSB)(2005), “Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century”, NSF, 37 ff.
- National Science Foundation(NSF)(2007), “Cyberinfrastructure Vision for 21st Century Discovery”, NSF Cyberinfrastructure Council, 2.
- NSF(2011), Data Management for NSF Engineering Directorate Proposals and Awards, http://nsf.gov/eng/general/ENG_DMP_Policy.pdf (1. August 2012).
- NSF(2011), “Digital Research Data Sharing and Management”: 1, <http://www.nsf.gov/nsb/publications/2011/nsb01211.pdf> (8. September 2012).
- NSF(2010), “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.
- NSF(2010), “Proposal and Award Policies and Procedures Guide, Part II – Award & Administration Guide(AAG)”, Chapter VI, D.4, NSF 11-1, VI-8 f.
- NHMRC/ARC(2007), Australian Code for the Responsible Conduct of Research, Australian Government, <http://www.nhmrc.gov.au/guidelines/publications/r39> (18. August 2012).
- Peter Fox/James Hendler(2011), “Changing the Equation on Scientific Data Visualization”, *Science*(331), (2011.2.11.), 706, <http://www.sciencemag.org/content/331/6018/705.full.pdf> (7. August 2012).
- Rhys Francis(2008), “eResearch: from research by the knowledge poor to research by the knowledge rich”, *International Association of Technological University Libraries 29th Annual Conference: Digital Discovery: Strategies & Solutions Auckland New Zealand*, 21-24 April 2008, Australian eResearch Infrastructure Council, Department of Innovation, Industry Science and Research, 3