

# 국가 과학데이터센터 구축의 경제적 파급효과 분석

## An Economic Ripple Effect Analysis of National Scientific Data Center Construction

박성욱\* · 한선화\*\*

Sung-Uk Park · Sun-Hwa Hahn

### 차 례

1. 서론	4. 경제적 파급효과 분석
2. 과학데이터 정의	5. 결론
3. 국내외 과학데이터 현황	· 참고문헌

### 초 록

현대 과학기술연구개발에서 데이터의 수집, 큐레이션, 분석, 시각화의 효율화는 과학기술발전의 중요한 요소이다. 데이터 집중연구에서 이러한 과학데이터의 가치는 절대적으로 필요하다고 할 수 있다. 과학데이터의 생산량은 기하급수적으로 증가했지만 데이터를 수집하고 배포할 수 있는 수단을 마련하는 측면에서 보면, 현재 분야별로 과학데이터의 개별적인 활용 체제만 운영되고 있어 걸음마 단계이다. 이에 본 논문에서는 국가차원의 국가 과학데이터센터 구축의 경제적 파급효과를 한국은행(2009)의 산업연관분석을 이용하여 생산유발효과, 부가가치유발효과, 취업자유발효과 및 전·후방연쇄효과로 나누어 분석하여 국가 과학데이터센터 구축에 대한 경제적 타당성을 검증하고자 한다.

### 키 워 드

과학데이터, 데이터 센터, 생산유발효과, 부가가치유발효과, 취업자유발효과, 전·후방연쇄효과

\* 한국과학기술정보연구원 정책연구실 선임연구원  
(Senior Researcher, Dept. of Policy Research, Korea Institute of Science and Technology Information, supark@kisti.re.kr)

\*\* 한국과학기술정보연구원 정책연구실 책임연구원  
(Principal Researcher, Dept. of Policy Research, Korea Institute of Science and Technology Information, shhahn@kisti.re.kr)

• 논문접수일자: 2011년 3월 22일  
• 최종심사(수정)일자: 2011년 7월 5일  
• 게재확정일자: 2011년 7월 7일

## ABSTRACT

In the modern scientific R&D, the efficient acquisition, curation, analysis and visualization are core elements of the science development. The value of scientific data is very important in data intensive research. An output of scientific data is drastically increasing. However we have only each individual system of scientific data in now. Therefore We feel a lack of efficiency of scientific data. In this paper, We analyze an economic ripple effects in terms of production inducement effect, added value inducement effect, labor inducement effect and forward-backward linkage effect of national scientific data center construction using an input-output analysis of the bank of Korea(2009). We also examine an economic propriety of national scientific data center construction.

## KEYWORDS

Scientific Data, Data Center, Production Inducement Effect, Added Value Inducement Effect, Labor Inducement Effect, Forward-Backward Linkage Effect

## 1. 서 론

### 1.1 연구배경 및 목적

아날로그에서 디지털로 패러다임이 변화면서 우리가 체감하는 큰 변화중의 하나는 일상의 모든 활동이 정보시스템을 기반으로 움직이고 그 흔적의 대부분이 데이터로 기록돼 DB에 쌓인다는 사실이다. 또 아날로그로 존재하는 정보나 지식은 별도의 디지털화 과정을 거쳐 DB로 축적해 활동이 용이하도록 하고 있다.

이러한 변화가 있는 지 10여년이 지난 지금, 이제는 범람을 우려할 만큼 우리의 주위에 데이터가 넘쳐나고 있다. 그러나 '데이터의 범람'의 한편에는 생산된 데이터가 체계적으로 축적되거나 관리되지 못해 데이터의 유실이 일

어나는 DB 구축의 사각지대가 여전히 존재하며 이들 중 하나가 바로 '과학데이터'라고 김선영(2010)은 정의하고 있다.

학술논문은 과학데이터를 분석하고 재해석해 새로운 의미를 제시했을 때에 비로소 가치가 있다면, 과학데이터는 존재 그 자체로 의미가 있다. 이는 동일한 연구에서도 전혀 다른 데이터가 생산되기도 하고 데이터에 따라서는 재연이 어렵거나 생산에 천문학적 비용이 소요되기도 하기 때문이다.

이렇듯 과학데이터에 대한 중요성과 이의 효과적 활용은 오래전부터 연구자들에게 중요한 관심사이다. 이에 관한 대표적인 국제기구로 CODATA(Committee on Data for Science and Technology)가 있다. CODATA는 유용하고 신뢰할 수 있고 손쉽게 입수할 수 있는

과학기술 데이터에 대한 세계 각계의 요구를 인식하고 국제학술연합(ICSU: International Council for Science Union)이 1966년에 전 세계에 걸쳐 과학기술분야 데이터의 평가, 축적, 보급을 진흥하고 이 분야에서 국제 협력을 촉진하기 위해 설립했다. 최근 정보통신기술의 발달과 더불어 데이터 집중연구시대에 돌입하면서 미국, 독일, 영국 등은 과학데이터를 효과적으로 활용하기 위한 정책을 추진하고 있다.

KISTI(2010)에 의하면, 창의적 연구개발 성과는 기초 데이터의 효과적 활용으로 창출 가능하며 Raw Data를 지식의 형태로 해석할 수 있도록 지원하는 새로운 소프트웨어가 다양하게 개발되고 있다. 또한 우리나라를 포함한 세계 각국은 연구개발에 많은 투자를 하고 있으며, 그 결과 디지털 정보의 양이 계속 확대되고 있으나 많은 데이터가 사장되어 사용되고 있지 못하며 소중한 지식자원의 효율과 효과에 대한 심각한 의문이 제기되고 있다. 이에 과학기술분야 연구개발의 창의성과 효율화를 위해 국내외에 숨어 있는 과학데이터를 수집, 관리, 제공하는 시스템의 구축이 필요한 것이다.

현재 대부분의 과학데이터는 소프트웨어나 첨단장비를 통해 관리되기 쉬운 디지털 형태로 생산된다. 그러나 아직까지 과학데이터의 공유나 유통의 중요성에 대한 연구자들의 인식도 부족하고 이를 지원할 수 있는 국가차원의 관리 체계도 결핍한 단계이다 보니 어디에 어떤 데이터가 있는지 소재파악도 어렵고 생산된 데이터는 일부 연구자만이 독점적으로

사용하다가 연구가 끝나면 휴면자원이 되는 경우가 많이 발생하고 있다.

2011년 정부의 국가 R&D 투자금액은 14.9조원으로 이 막대한 예산을 통해 생산되는 과학데이터의 양은 어느 정도인지, 또 이를 체계적으로 축적하지 않아 재사용되지 못해 발생하는 손실비용은 도대체 얼마나 되는지 기늙조차 어려운 실정이다. 과학데이터의 경제적인 효과 분석과 관련된 선행 논문이 존재하지 않아 본 논문에서는 우선 기초기술연구회 소관기관을 대상으로 국가 과학데이터센터 구축을 통해 국가 과학데이터센터 구축이 우리나라 경제 각 부문에 미치는 경제적인 효과를 한국은행(2009)의 산업연관분석을 이용하여 생산유발효과, 부가가치유발효과, 취업자유발효과로 나누어 경제적 파급효과를 분석한다. 또한, 감응도계수와 영향력계수를 통한 전·후방연쇄효과를 살펴봄으로써 국가 과학데이터센터 구축의 경제적 타당성을 검증한다. 이를 바탕으로 과학데이터 관련 발전 방향을 제공할 수 있다는 점에서 본 논문은 의미가 있다.

## 1.2 연구내용 및 방법

과학데이터에 대한 관심이 최근 급격히 확산되고 있는 추세이며 특히, 중국과 호주는 국가 주도적으로 관련 사업을 추진하고 있다. 특히 중국은 2003년 국가과학기술계획항목 과학데이터 제출 임시법 제정을 위한 초안을 작성하였으며, 연차적으로 자원환경분야 과학데이

터 제출 임시법을 비롯하여 과학기술 분야별 관련법과 시행규칙을 제정하여 운영하고 있다. 호주정부는 2007년 연구책임수행법을 공포하였고, 미국, 영국 등도 과학데이터를 위한 정책을 추진하고 있다.

현재 국내에서는 과학데이터와 관련해서 분야별로 과학데이터를 개별적으로 활용하는 체계가 구축되어 국가차원에서의 수집 및 관리가 되지 않고 있다. 이에 미국, 호주, 영국 등에서 진행되고 있는 과학데이터의 보존과 활용 사례를 간략히 설명하고 국가 과학데이터센터 구축의 효과에 대해 설명하고자 한다. 또한 국가 과학데이터센터 구축이 경제 각 부문에 미치는 효과를 한국은행(2009)의 산업연관분석을 이용하여 생산유발효과, 부가가치유발효과, 취업자유발효과 및 전·후방 연쇄효과로 나누어 분석한다. 이러한 경제적 파급효과를 국민경제적 차원에서 측정하여 향후 관련 연구사업에 대한 예산 배분 등 미래의 투자 계획 및 정책수립 등의 기초자료로 활용하고자 한다.

국가 과학데이터센터 구축의 경제적 파급효과를 분석하기 위해 한국은행(2009)의 투입·산출표인 산업연관표에 의거하여 분류함으로써 국가 과학데이터센터 구축이 국내 산업부문 중 어느 부문에 속해 있는지 파악할 것이다. 그리고 국가 과학데이터센터 구축에 대한 파급효과를 이론적 차원에서 생산유발효과, 부가가치유발효과, 취업자유발효과 및 전·후방 연쇄효과를 나누어 분석하고 국가 과학데이터센터 구축에 대한 각각의 파급효과를 한국은행의

투입-산출모형에 입각하여 분석할 것이다.

본 논문의 제2장에서는 과학데이터의 정의에 대해 개략적으로 설명한다. 제3장에서는 국내외 과학데이터 현황을 살펴보고, 제4장에서는 경제적 파급효과 분석 방법론에 대해 설명하고 본 연구에서 사용하고 있는 산업연관분석모형에 대해 정리한다. 제5장에서는 본 논문의 결론 및 연구의 한계에 대해 정리해 본다.

## 2. 과학데이터 정의

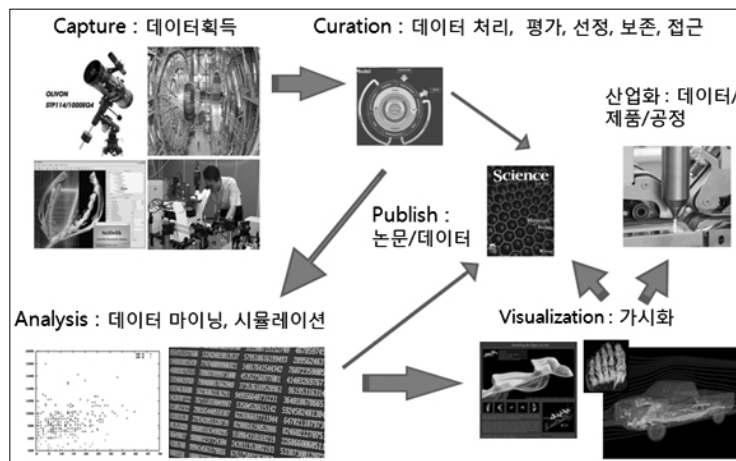
Cheng(2006)에 의하면 과학데이터는 과학기술활동의 결과 얻어진 데이터로 정의하고 있다. 즉, 망원경, 전자현미경, 인공위성 등을 통한 관측, 센서 등을 통한 감시, 설문조사, 기술 및 시장조사, 기술가치평가 등을 통한 조사, 가속기, 화학 및 바이오 실험장비 등을 통한 실험, 분석도구 등을 통한 연구분석, 슈퍼컴퓨터 등을 통한 계산 데이터를 모두 포함한 개념으로 정의하고 있다. 또한, OECD(2007)에 의하면 데이터를 과학연구 수행을 위한 주요한 원천으로 사용하는 숫자, 문자정보, 이미지 및 소리 등의 사실적인 기록으로 정의하고 있다.

과학데이터는 처음에 데이터 획득이라는 수집단계에서 생성되어 데이터 처리, 평가, 선정, 보존, 접근이라는 큐레이션 단계와 데이터 마이닝이나 시뮬레이션을 통한 분석 단계를 거쳐 가치화 단계에 이르면 논문으로 출판되거나 제품으로 출시되는 산업화 과정을 거쳐 소멸된다.

<그림 1>은 이러한 과정을 설명하고 있다.

과학데이터의 종류는 숫자로 구성된 표, 비트로 구성된 배열, 종이에 인쇄된 문자, 음성 녹음 및 월석표본 등 다양하게 분류된다. 전문가나 사용자의 데이터 해석을 거친 후에 데이터의 가치가 나타나며 데이터를 재사용할 수 있는 가능성이 높아지면서 많은 연구자가 데

이터의 본질적인 가치에 대해 인정하고 있다. 공개된 데이터 컬렉션을 대개 연구데이터, 자원 혹은 커뮤니티 데이터, 참조데이터 세 가지 형태로 구분하면 데이터 컬렉션의 수는 점점 증가하고 있다. 공개 데이터의 표준적 인용방법이 없는 실정이나 최근 표준화를 위한 다양한 접근이 진행 중에 있다(<표 1> 참조).



<그림 1> 과학데이터의 생애(류범중 2010)

<표 1> 과학데이터 컬렉션 구분(이상환, 심원식 2009)

구분	내용
연구데이터 컬렉션 (Research Data Collections)	- 단일 혹은 복수의 연구 프로젝트에서 생성되는 데이터에서 파생 - 원시 데이터의 가공 및 큐레이션 기능이 거의 없음 - 이용자가 소수이거나 재정 지원 부족 등으로 인해 연구자가 프로젝트 종료 이후에 데이터를 유지할 의사가 대부분 없음
자원 혹은 커뮤니티 데이터 컬렉션 (Resource or Community Data Collections)	- 단일 학문 집단이나 세부 연구 주제를 대상으로 구축 - 대개 학문 집단 수준의 표준에 부합하며 해당 표준의 선택이나 구축에 중요한 역할을 했을 가능성이 있음 - 연구지원기관은 데이터 컬렉션의 유지를 지원하고 있으나 정책 변화의 가능성 때문에 지원 기간은 불투명함
참조데이터 컬렉션 (Reference Data Collections)	- 다양한 분야와 기관에 있는 연구자, 학생, 교사를 포함한 대단위의 이용자 커뮤니티를 대상으로 구축 - 장기적인 재정 지원, 표준 및 관례에 대한 부합, 그리고 전문적인 관리가 참조데이터 컬렉션에 구축에 필수적임

데이터를 원활하게 수집, 접근하려면 연구자들이 데이터를 적극적으로 공유할 필요가 있으며 과학데이터를 공유하는 목적은 크게 5가지로 요약할 수 있다. 첫째, 데이터의 보존과 재사용을 위해, 둘째, 반복 실험 및 결과의 검증을 위해, 셋째, 반복할 수 없거나 반복하는데 많은 비용이 드는 실험 및 관찰의 경우, 넷째, 교육적 목적을 위해서, 마지막으로, 이외에도 비교연구나 여러 실험에서 생성된 데이터를 결합하는 메타연구, 모델링을 위해 과학데이터를 공유하며 국내에 산재되어 있는 과학데이터를 국가 과학데이터센터를 구축함으로써 공유 및 활용하고자 한다.

### 3. 국내외 과학데이터 현황

엄청난 공공재원으로 수행된 연구과제의 성과물인 학술지 논문, 연구보고서 등은 다양한 경로를 통해 배포, 접근하는데 반해 과학데이터의 경우는 체계적으로 수집하고 배포하는 경로가 상대적으로 미흡한 실정이다. 이러한 이유로는 학술논문 발표는 연구자의 임용, 승진이나 명성에 직·간접적으로 긍정적인 영향을 미치지만 데이터 공개는 아직까지 구체적인 혜택이 없다. 또한 정부재원으로 수행된 연구의 경우 연구지원기관에 데이터를 기탁하는 것을 연구수행 조건으로 제시하고 있으나 데이터를 기탁할 수 있는 구체적인 절차나 체계가 마련되어 있지 않다. 과학데이터의 생산량

은 매우 방대하지만 데이터의 가치, 재사용, 보존에 대한 기준이 없다. 특히 과학데이터의 경우 데이터를 생성하는 연구 환경, 절차, 기기, 세부적인 조건 등이 데이터 속성 및 내용에 큰 영향을 주고 있다고 이상환, 심원식(2009)은 설명하고 있다.

이에 과학데이터 공유와 관련하여 영국에서는 과학데이터에 대한 영구적 접근을 위해 디지털 큐레이션 센터를 설립하여 관련 도구와 정책을 검토하고 있으며 미국에서는 연구와 교육 목적을 위해 국가적, 국제적 요구조건을 다루고 있는 정책문서를 발간하고 있다. 이들 나라에서는 어떤 경우에는 연구자들이 연구비 지원 계약 조건에 의해 데이터를 공유하도록 의무화 하고 있다. 국내의 경우에는 한국연구재단의 인문사회분야 학술연구지원사업 처리규정(교육과학기술부 훈령 제111호) 제 34조에서 연구성과물의 데이터베이스화와 공유 촉진을 언급하고 있다.

과학데이터 전반에 관한 주요 국가의 동향을 정리해 보면 <표 2>와 같다.

국내에서는 우선 한국과학기술정보연구원에서 생물다양성 데이터, 화합물, 유전자, 단백질, 무기결정구조, 플라즈마 물성, 디지털 코리언, 전통한의학, 식품, 천연물질 DB 등 과학데이터와 관련하여 수집 및 제공에 관한 서비스를 하고 있다. 또한 대용량 실험데이터를 구축하여 제공하고 있으며 가상세포, 생태 데이터 축적 및 시뮬레이션, 데이터 가시화 등 과학데이터 관련 과제를 자체적으로 현재 수행하고 있다.

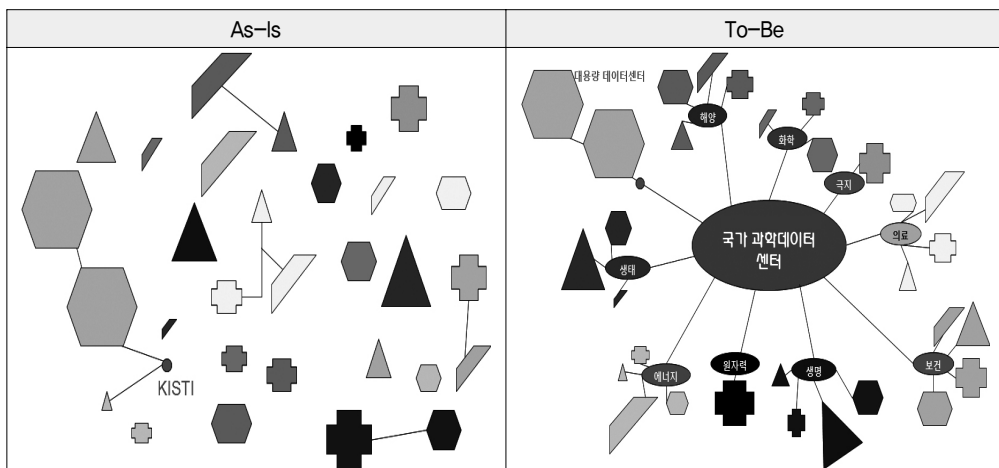
〈표 2〉 주요 국가 과학데이터 현황(류범중 2010)

구분	내용
미국	- NSTC 산하 디지털 과학데이터 보존, 접근, 상호운용을 위한 위원회 신설 제안 - 데이터관련 정책수립의 필요성과 공표 제안
호주	- 국가차원에서 과학데이터 수집, 관리, 서비스 - Australian National Data Service - Australian Research Collaboration Service
유럽연합	- DRIVER 프로젝트 진행 - PARADE에서 유럽의 데이터 인프라 구축을 위한 전략보고서 발행 - 디지털 연구성과물 보존관련 설문조사 프로젝트
중국	- 중국과학기술부(MOST)-중국과학원(CAS)에서 National Scientific Data Sharing Program를 2002년부터 진행 - 40개 이상의 과학데이터센터, 300개 마스터 데이터베이스, 1개 포털을 2010년 구축

또한, 한국표준과학연구원 국가참조표준센터에서는 국가적으로 개발능력이 있고 필요한 분야에서 총 17개 데이터 센터를 지정하여 협력하고 있다.

또한 주체별로 명확한 경우 단일 주체의 경우에는 생물정보 데이터, 대기과학 데이터, 다수주체인 경우에는 인체데이터, 의료데이터로 구분되고 주체가 불명확한 경우에는 생태계 데

이터, 뇌과학 데이터로 구분된다. 그리고 데이터 구축 정도에 따라 데이터가 발생하나 미수집 되는 초기상태인 의료 데이터와 일부 데이터를 수집하는 극지 및 해양데이터, 일부 데이터를 수집하여 활용하는 인체 데이터 등으로 구분할 수 있다. 이렇듯 국내에서는 각 분야별로 해당 분야의 과학데이터가 산재되어 있는 실정이다. 이에 국가 과학데이터는 〈그림 2〉의



〈그림 2〉 국가 과학데이터 향후 방향(한선화 2011)

〈표 3〉 국가 과학데이터센터 체제 효과(김중권 2011)

분야별 과학데이터센터	국가 과학데이터센터
<ul style="list-style-type: none"> <li>- 분야별 과학데이터센터만 존재</li> <li>- 기관에 따라 개별적 추진</li> <li>- 분야에 따라 불균형 발전</li> <li>- 분야별 활용체제 구축</li> <li>- 과학데이터의 수집부터 출판까지 단독추진</li> <li>- 개별기관의 자원만 활용</li> <li>- 데이터 보존 단일기관 의존</li> <li>- 기존의 데이터 분석도구, 가시화 도구활용</li> <li>- 분야별 데이터만 활용</li> <li>- 범지구적 문제 단일 대처</li> <li>- 국가 현안문제 개별적 대처</li> <li>- 데이터 축적량 증가에 따라 활용효과 상승</li> <li>- 예산 확보 난이</li> </ul>	<ul style="list-style-type: none"> <li>- 분야별 과학데이터 센터 + 국가 과학데이터 센터</li> <li>- 합목적에 따라 공동 추진</li> <li>- 각 분야의 균형 발전 가능</li> <li>- 국가적 공동 활용체제 구축</li> <li>- 과학데이터의 수집부터 출판까지 효율적 추진 가능</li> <li>- 대형장비 활용가능: 슈퍼컴, 대용량 디스크, 가시화장비, 국내외 초고속 전산망</li> <li>- 데이터 보존 공동 대처 가능</li> <li>- 고성능 데이터 분석도구, 가시화 도구 활용 가능</li> <li>- 이종 데이터 융합을 통한 시너지</li> <li>- 범지구적 문제 해결 공동 대처</li> <li>- 국가 현안문제 공동 대처</li> <li>- 데이터 축적량 증가에 따라 활용효과 극대화</li> <li>- 예산확보 용이</li> </ul>

As-Is인 각 데이터들이 결합되지 못하고 각 분야별로 산재되어 있으나 과학데이터의 중요성 및 필요성으로 본 논문에서는 기초기술연구회 소관기관을 중심으로 국가 차원의 국가 과학데이터센터를 구축하여 To-Be 모형으로 국가 과학데이터센터를 구축하여 분야별 과학데이터를 공유 및 활용 체계를 갖춰 나가야 한다고 제안한다.

국가 과학데이터센터 구축은 국가 과학데이터 공유 및 활용체제를 구축하여 분야별 과학데이터의 개별적 활용 체제와 서로 윈-윈 효과를 〈표 3〉과 같이 가져올 것으로 판단된다.

#### 4. 경제적 파급효과 분석<sup>1)</sup>

##### 4.1 산업연관분석의 의의 및 구조

한국은행(2009)에 의하면 한 국가 경제에서 각 산업들은 생산활동을 위해 상호 간에 재화와 서비스를 구입하고 판매하는 과정을 통해 직접 또는 간접적으로 서로 관계를 맺게 되는데 산업연관표는 일정기간(보통 1년) 동안의 이러한 산업간 거래관계를 일정한 원칙에 따라 행렬식으로 기록한 통계표이다. 산업연관분석 또는 투입산출분석은 산업연관표를 바탕으로 하여 산업간 상호연관관계를 수량적으로 분석한 방법인 것이다. 산업연관분석은 최종수요가 유발하는 생산, 고용, 소득 등 각종의 파

1) 박성욱, 한선화, 성원경(2010) 재구성.



〈표 4〉 산업연관표 기본 구조(한국은행 2009)

내생부문	중 간 수 요						최종수요	수입(공제)	총산출액	
	1	2	·	·	·	n				
중 간 투 입	1	X <sub>11</sub>	X <sub>12</sub>	·	·	·	X <sub>1n</sub>	Y <sub>1</sub>	M <sub>1</sub>	X <sub>1</sub>
	2	X <sub>21</sub>	X <sub>22</sub>	·	·	·	X <sub>2n</sub>	Y <sub>2</sub>	M <sub>2</sub>	X <sub>2</sub>
	·	·	·	·	·	·	·	·	·	·
	·	·	·	·	·	·	·	·	·	·
	·	·	·	·	·	·	·	·	·	·
	n	X <sub>n1</sub>	X <sub>n2</sub>	·	·	·	X <sub>nn</sub>	Y <sub>n</sub>	M <sub>n</sub>	X <sub>n</sub>
부가가치	V <sub>1</sub>	V <sub>2</sub>	·	·	·	V <sub>n</sub>	외생부문			
총투입액	X <sub>1</sub>	X <sub>2</sub>	·	·	·	X <sub>n</sub>				

- 주) 1. 총투입액 = 중간투입+부가가치 (투입구조)  
 총산출액 = 중간수요+최종수요-수입 (배분구조)  
 총투입액 = 총산출액
2. 세로방향: 각 산업부문이 해당상품 생산을 위하여 지출한 생산비용의 구성, 즉 투입구조를 설명함  
 가로방향: 각 산업부문의 생산물이 어떤 부문에 중간 수요 또는 최종수요 형태로 얼마나 판매되었는가를 하는 배분구조를 의미

급효과를 산업부문별로 구분하여 분석할 수 있기 때문에 경제정책의 수립, 정책효과의 측정 등에 활용되고 있다. 다만 한국은행 산업연관표를 통한 생산파급효과 등은 국가 과학데이터 구축에 대한 R&D 투자 등의 직접적인 효과를 고려하지 못하는 점에서 과소평가의 우려가 있지만 가격의 변화를 고려하지 못하는 정태분석이라는 점에서 효과를 과대평가할 우려가 상존하고 있어서 과소 또는 과대평가의 상대적 크기에 대해 단정적인 평가를 내리기는 어려운 점이 존재하고 있다.

한국은행의 산업연관표를 이용하여 국가 과학데이터센터 구축에 대한 생산, 부가가치, 취업 등의 파급효과를 추정하기 위해 산업연관표의 통합소분류에서 '연구기관'에 속하는 행과 열을 모두 '0'으로 대체하여 만들고 파급효

과 추정을 위한 각종 계수도 이를 통해 산출하여 유발효과를 얻을 수 있다.

#### 4.2 각종 계수의 산출 및 산업연관효과

한국은행(2009)과 신용보증기금(2009)에 의하면 각 산업부문이 해당 부문의 재화나 서비스 생산에 사용하기 위하여 다른 부문으로부터 구입한 원재료 등의 중간 투입액을 총투입액으로 나누어 산출한 계수를 투입계수라 하고 투입계수를 산업연관표의 내생부문과 같은 모양으로 배열한 행렬이 투입계수표라 설명하고 있다. 국가 과학데이터 구축에 대한 경제적 파급효과를 산출하기 위해서는 수입된 재화 또는 서비스로 인한 산출효과를 배제해야 하므로 수입거래 금액을 제외한 국산 투입계수

는 식(1)과 같다.

ij 산업 간의 국산투입계수

$$a_{ij} = \frac{X_{ij} - M_{ij}}{X_j} \text{ ----- 식(1)}$$

생산유발계수는 식(1)을 이용하여 국가 과학데이터센터 구축에 대한 ‘연구기관’ 부문을 외생변수화 한 후에 식(2)를 사용하여 생산유발계수를 산출한다.

생산유발계수

$$= A_s^d (I - A^d)^{-1} \text{ ----- 식(2)}$$

- $A_s^d$  : 국가 과학데이터센터 구축의 국산투입계수 행벡터
- $I$  : 1로 이루어진 대각 행렬(diagonal matrix)
- $A^d$  : 국산투입계수( $a_{ij}$ ) 행렬

부가가치계수는 총 산출에서 부가가치가 차지하는 비중을 나타내며, 산업연관표에서 각 산업의 부가가치 합계를 총 산출로 나누어 식(3)처럼 구한다.

i 산업의 부가가치계수

$$v_i = \frac{V_i}{X_i} \text{ ----- 식(3)}$$

노동계수란 일정기간 동안 생산활동에 투입된 노동량을 총산출액으로 나눈 계수로 한 단위의 생산에 직접 필요한 노동량을 의미하며, 노동량에 자영업주 및 무급가족종사자를 포함하느냐의 여부에 따라 취업자계수와 고용계수

로 구분한다.

$$\text{취업계수 } l_w = \frac{L_w}{X} ,$$

$$\text{고용계수 } l_e = \frac{L_e}{X} \text{ ----- 식(4)}$$

- $L_w$  : 취업자수,  $L_e$  : 피용자수,
- $X$  : 총산출액

노동유발계수는 어느 산업의 생산물을 한 단위 생산하는데 직접 필요한 노동량 뿐 아니라 생산과급과정에서 간접적으로 필요한 노동량까지 포함하고 있으며 식(5)로 나타내며, 본 논문에서는 취업자유발계수를 이용하여 취업자유발효과를 분석한다.

취업자유발계수

$$= \hat{\Gamma}_w (I - A^d)^{-1},$$

고용유발계수

$$= \hat{\Gamma}_e (I - A^d)^{-1} \text{ ----- 식(5)}$$

- $\hat{\Gamma}$  : 노동계수의 대각 행렬,  $I$  : 단위 행렬,
- $A^d$  : 국산투입계수 행렬

또한 감응도계수는 전 부문의 최종수요를 모두 한 단위씩 증가시키기 위해 I번째 산업이 생산해야 할 단위의 산업 평균치에 대한 비율로 계산된다. 영향력계수는 전 산업 평균 생산유발계수에 대한 산업별 생산유발계수의 비율을 의미한다.

- 감응도계수 = 생산유발계수의 행합 / 생산유발계수의 총합

- 영향력계수 = 생산유발계수의 열합 / 생산유발계수의 총합

산업연관분석을 적용한 국내 기존연구로는 박성욱 등(2010)은 시맨틱 기술에 대한 경제적 파급효과를 분석하였고, 박성욱(2008)은 국가 지식정보시스템 개발에 대해 경제적 파급효과를 분석하였다. 박성욱 등(2008)은 정보보호 산업의 경제적 파급효과를 분석하였다. 김도환(2007)은 정보통신산업에서의 경제적 기여도를 파악하고 정보통신산업의 특정 세부산업이 국민경제에 미치는 영향력을 분석하고 있다. 또한, 김수현 등(2006)은 IT기기 수출이 국내 산업에 미치는 경제적 파급효과를 분석하고 있는 등 최근에 관련 문헌상에 산업연관분석이 경제적 파급효과의 분석도구로 사용되고 있다. 산업연관분석은 경제부문 간의 재화와 서비스의 흐름이 비교적 안정적이라는 점을 활용하여 경제체계의 모습을 보다 자세하게

통계적으로 분석함으로써 경제현상에 대한 설명을 보다 구체적으로 해주는 역할을 하고 있다. 특히 산업연관분석은 한 나라의 경제정책 수립 및 효과분석과 관련된 분야에 많이 이용되고 있어 본 논문에서 분석하고자 하는 국가 과학데이터센터 구축에 경제적 파급효과를 분석하는데 적절하며 타당하다고 볼 수 있다.

우선, 산업연관분석을 적용하기 위해서는 본 논문에의 국가 과학데이터센터 구축은 통합소분류의 '연구기관'에 적용시켰다(〈표 5〉 참조).

본 논문에서는 식(1)~식(5)를 통하여 생산유발계수, 부가가치유발계수, 수입유발계수, 취업자유발계수를 산출하여 〈표 6〉에 정리를 하였으며 각각의 유발계수는 국가 과학데이터센터 구축에 대한 최종수요가 한 단위 증가하는 경우 전 산업에서 직·간접적으로 일어나는 유발효과를 설명하고 있다.

〈표 6〉을 살펴보면 생산유발계수는 1.487496으로 전체 평균인 1.940보다 작지만 부가가치유

〈표 5〉 국가 과학데이터센터 구축의 산업연관표 해당부문(한국은행 2009)

통합대분류(28부문)	통합중분류(78부문)	통합소분류(168부문)
26. 교육 및 보건	66. 연구기관	148. 연구기관

〈표 6〉 국가 과학데이터센터 구축에 대한 유발계수(한국은행 2009)

통합소분류	생산유발계수	부가가치유발계수	취업자유발계수(명/10억원 당)	감응도계수	영향력계수
연구기관(148)	1.487496	0.924146	22.7689	0.624208	0.767042
전체 산업 평균	1.940	0.722	10.4	'1'과 비교	

발계수는 전체 평균 0.722보다 큰 0.924146이고, 취업자유발계수도 전체 평균 10.4명보다 훨씬 큰 22.8명으로 부가가치유발과 취업자유발효과가 전체 평균보다 훨씬 큼을 알 수 있다.

또한, 국민경제는 특정산업군에서 생산해낸 산출물을 타 산업에서 이용하고 또한 타 산업에서 생산해낸 산출물이 여타산업에서 이용되는 파급효과, 혹은 가치사슬에 의하여 생산활동이 이루어진다. 본 논문의 가치사슬 상에서 국가 과학데이터센터 구축이 앞에 있는 산업에 영향을 미치는 효과를 전방연쇄효과라고 하며 가치사슬 상에서 뒤에 있는 산업에 영향을 미치는 효과를 후방연쇄효과라고 한다. 국가 과학데이터센터 구축에 대한 감응도계수가 0.624208로 1보다 작으므로 전방연쇄효과가 적고 영향력계수도 0.767042이므로 1보다 작으므로 후방연쇄효과가 적다고 볼 수 있다.

여기서 국가 과학데이터센터 구축의 경제적 파급효과를 산출하기 위해서는 국가 과학데이터센터로부터 생성되는 과학데이터가 어떤 투입변수에 가장 영향을 받는지 판단해야 한다.

과학데이터의 정의에서 과학데이터의 생성은 국가 R&D 전반에 걸쳐 나타나는 모든 생산물로 파악될 수 있으며 이는 정부의 국가 R&D 예산의 투입에 따른 결과지표로 해석될 수 있다. 하지만 본 논문에서는 우선 기초기술연구회 소관기관을 대상으로 먼저 국가 과학데이터센터를 구축하고자 하기 때문에 기초기술연구회 소관기관의 예산을 투입변수로 사용하여 국가 과학데이터센터 구축의 경제적 파급효과를 산출할 수 있다.

〈표 7〉을 통해 2011년 정부 R&D 투자는 14.9조원으로 2010년에 비해 8.7% 증가한 수치이며, 최근 4년간 연평균 정부 R&D 투자 증가율은 10.3%로 높은 수준이다. 하지만 선진국에 비하면 아직도 낮은 수준임을 〈표 8〉을 통해 알 수 있다.

기초기술연구회 소속 출연(연)은 전체 13개 기관으로 2011년 예산은 1.9조원임을 〈표 9〉를 통해 확인할 수 있다.

이제 〈표 6〉에서 구한 국가 과학데이터센터 구축의 유발계수를 바탕으로 경제적 파급효과

〈표 7〉 국가 R&D 연도별 예산 현황(교육과학기술부 2011)

(단위: 억원/%)

구분	2008년	2009년	2010년	2011년	연평균성장률
합계	110,781	123,437	137,014	148,902	10.3

〈표 8〉 국가별 정부재원 R&D 예산 비율(교육과학기술부 2011)

구분	한국 (2011년 추정)	프랑스 (2008년)	영국 (2008년)	미국 (2008년)	독일 (2008년)	일본 (2008년)
정부재원비중	28.3%	41.5%	36.8%	32.7%	28.1%	21.4%

〈표 9〉 2011년 기초기술연구회 소관기관 예산(기초기술연구회 2011)

(단위: 억원)

KIST	KBSI	핵융합(연)	수리(연)	천문(연)	생명(연)	한의학(연)
2,543	758	1,717	131	504	1,417	430
KISTI	표준(연)	해양(연)	극지(연)	항우(연)	원자력(연)	전체
1,216	1,185	1,554	646	2,980	3,654	18,735

〈표 10〉 국가 과학데이터센터 구축의 경제적 파급효과

(단위: 억원/명)

구분	생산유발효과	부가가치유발효과	취업자유발효과
2011년	27,868	25,754	63,453

를 계산하기 위해 우선 생산유발효과는 국가 과학데이터센터 구축을 위한 산업연관표를 통해 산출한 생산유발계수를 곱하여 추정하고, 부가가치유발효과는 국가 과학데이터센터 구축을 통해 얻을 수 있는 국민경제적인 순가치이며 생산유발효과 중 부가가치 귀속부분을 부가가치계수를 통해 추정할 수 있다. 국가 과학데이터센터 구축으로 관련 시장이 증대함에 따라 취업자의 증가로 이어지는 관점에서 취업자유발효과를 추정할 수 있으며 이는 생산유발효과에 취업자유발계수를 곱하여 산출할 수 있다.

〈표 10〉은 기초기술연구회 소관기관을 중심으로 국가 과학데이터센터를 구축하였을 경우 유발되는 경제적 파급효과를 설명하고 있다. 먼저 국가 과학데이터센터 구축의 생산유발효과는 2.8조원, 부가가치유발효과는 2.6조원, 취업자유발효과는 63,453명으로 분야별로 과학데이터의 개별적 활용 체제를 구축하였을

경우보다 국가 과학데이터센터가 구축되면 1.5배의 생산유발효과가 유발될 것으로 예상된다.

## 5. 결론

과학관련 연구의 주요 산출물인 과학데이터는 수많은 과학자 등의 활용을 통해 그 가치가 점차 증대되고 있다. 하지만 과학데이터의 체계적인 수집이나 배포 경로가 미흡하고 또한 방대한 생산에도 불구하고 데이터의 활용, 재사용, 보존에 대한 기준이 미비한 시점에 있다. 본 논문은 현재 분야별로 과학데이터의 개별적 활용체계가 구축되어 활용되고 있는데 국가 과학데이터의 공유 및 활용 체제에 대한 원-원 효과를 위해 국가 과학데이터센터 구축을 제안하고 있다. 이에 우선 기초기술연구회 소관기관의 과학데이터부터 구축하여 활용 성과

를 파악하기 위해 국가 과학데이터센터 구축의 경제적 파급효과를 한국은행(2009)의 산업연관표를 이용하여 분석했다. 이를 통해 경제적 타당성을 파악하고자 했는데 분석결과 분야별 과학데이터를 개별적으로 활용하는 체제보다 국가 과학데이터센터를 구축하였을 경우 생산유발효과에서 1.5배의 경제적 효과가 발생함을 분석결과를 통해 알 수 있었다. 또한, 부가가치유발효과와 취업자유발효과가 전체 산업 평균보다 훨씬 높음을 분석결과를 통해 알 수 있었다. 즉, 과학데이터는 데이터 획득, 수집, 큐레이션, 분석, 가시화 단계 등의 생애주기를 거치면서 많은 비용과 시간이 소요되거나 데이터 관리주체가 산재되어 있어 상호협력체제가 이루어지지 않아 비효율성을 보이고 있지만 국가 과학데이터센터를 구축함으로써 그 효율 및 효과를 극대화 할 수 있다.

하지만, 본 논문에서 활용된 한국은행의 산업연관표를 통한 경제적 파급효과 분석은 국가 과학데이터센터 구축에 대한 R&D 투자 등의 직접적인 효과를 고려하지 못하는 점에서 과소평가의 우려가 있지만 가격의 변화를 고려하지 못하는 정태분석이라는 점에서 효과를 과대평가할 우려가 상존하고 있어서 과소 또는 과대평가의 상대적 크기에 대해 단정적인 평가를 내리기는 어려운 점이 존재하고 있다.

마지막으로, 본 논문을 통해 과학데이터는 지속적으로 관심을 갖고 투자가 필요한 분야이며 성공적으로 국가 과학데이터센터를 구축하였을 경우 무한한 부가가치를 창출해 낼 수

있을 거라 생각한다. 따라서 이 연구결과를 활용하여 과학데이터에 대한 연구개발의 필요성과 중요성을 널리 알리고 국가의 핵심 과제로 부각이 되어 국내에서도 과학데이터를 효과적으로 활용하기 위한 정책이 수립되었으면 한다. 향후 국가 과학데이터센터를 통해 지구적 문제 공동해결을 위한 연구 인프라를 제공하고 국가적 아젠다의 효율적 수행을 통해 물리, 화학, 생의학 분야에 노벨상 수상자를 배출하여 국가적 위상이 높아졌으면 한다.

## 참고문헌

- 기초기술연구회. 2011. 『소관기관 예산(출연금 및 자체수입)』. [인용 2011. 3. 21].  
 <<http://www.krcf.re.kr/site/krcf/sub.do?Key=1999>>.
- 김도환. 2007. 산업연관분석에 의한 정보통신산업의 경제적 파급효과. 『한국경영과학회지』, 32(3): 81-96.
- 김수현, 강희일, 홍승표, 정해식. 2006. IT 기기 수출의 경제적 파급효과. 『기술혁신학회지』, 9(2): 279-303.
- 김선영. 2010. 『DB 구축의 사각지대, 과학데이터』. [인용 2011. 3. 2].  
 <<http://blog.naver.com/dbstory2009?Rdirect=Log&logNo=120107578549>>.
- 김중권. 2011. 국가 과학데이터 공유·활용체제 원-원 효과. 『과학데이터 자문위원회 준

- 비회의』, 2011년 2월 25일. [대전: 한국 과학기술정보연구원].
- 류범중. 2010. 과학데이터(Scientific Data) 활용방안 고찰. 『데이터베이스 그랜드 컨퍼런스 2010』, 2010년 10월 19일. [서울: 코엑스 컨퍼런스 그랜드 볼룸].
- 박성욱. 2008. 국가 지식정보시스템 개발의 경제적 효과분석 - 한국과학기술정보원(KISTI)의 연구활동을 중심으로. 『정보관리연구』, 39(1): 73-94.
- 박성욱, 한선화, 성원경. 2010. 시맨틱 기술의 경제적 파급효과 분석. 『정보관리연구』, 41(3): 175-190.
- 박성욱, 이상호. 2008. 한국의 정보보호산업과 경제적 파급효과 - 산업연관분석을 이용하여 -. 『산업혁신연구』, 24(2): 1-22.
- 신용보증기금 성과평가부. 2009. 『2008년도 신용보증 성과 및 국민경제기여도 분석』. 서울: 신용보증기금.
- 이상호. 2010. 『과학데이터의 아카이빙 및 관련 자료』. [인용 2011. 3. 21]. <<http://www.slideshare.net/glorykim/2010-0603>>.
- 이상환, 심원식. 2009. 과학데이터의 공유와 활용. 『지식리포트』, 5: 1-36.
- 한국. 교육과학기술부. 2011. 주간보도자료(2011년 투자현황). [인용 2011. 3. 2]. [서울: 교육과학기술부].
- 한국과학기술정보연구원. 2010. 『과학데이터의 체계적 보존/활용을 위한 큐레이션 정책 연구 및 시범 구축』. 대전: 한국과학기술정보연구원.
- 한국은행. 2009. 『2007년 산업연관표』. 서울: 한국은행.
- 한국은행. 2009. 『산업연관분석 해설』. 서울: 한국은행.
- 한선화. 2011. 국가 과학데이터 공유·활용체제 구축. 『과학데이터 TFT 제3차 자문회의』, 2011년 2월 28일. [대전: 스파피아호텔].
- Cheng, Jinpei. 2006. *Strategies for Preservation of and Open Access to Scientific Data in China: Summary of a Workshop*, Washington, DC: National Academies Press.
- OECD. 2007. OECD Principles and guidelines for access to research data from public funding. [cited 2011. 3. 21]. <<http://www.oecd.org/dataoecd/9/61/38500813.pdf>>.