

과학기술분야 출연연구기관 연구데이터 관리 및 공유 사례 분석 연구

A Study on the analysis of Research Data Management and Sharing of Science & Technology Government-funded Research Institutes

박 미 영 (Miyoung Park)*

안 인 자 (Inja Ahn)**

남 승 주 (Seungjoo Nam)***

초 록

본 연구에서는 오픈사이언스 정책의 일환으로 학문분야별 연구데이터 공유·활용에 관한 인식을 비교분석하였다. 이를 기반으로 과학기술분야 정부출연연구기관 27개 기관 데이터 업무담당자를 대상으로 반구조화된 질문을 통하여 심층인터뷰를 실시하였다. 그 중 데이터 관리 면에서 우수기관 9개 기관을 선정하여 연구데이터 수집 및 관리현황 사례를 구체적으로 제시하였다. 결과 현재 출연연구기관의 연구데이터 수집·관리는 전반적으로 시범사업 단계이며, 데이터의 수집 구축 수준 또한 기관별로 상이하다. 기관별로 살펴보면 수집을 시작하는 단계(KIOM), 수집·관리가 고도화되고 있는 단계(KIST), 공유·활용을 시작하는 단계(KRIBB, KRICT) 등으로 구분된다.

ABSTRACT

As a part of the open science policy, this study compared the perception of research data sharing and utilization by academic field. Based on this, in - depth interviews were conducted with semistructured questions to the data task managers of 27 government - funded research institutes in science and technology. Among them, nine excellent organizations were selected from the viewpoint of data management and cases of research data collection and management were specifically presented. The State of the collection and management of research data by the participating research institutes is generally a pilot project stage, and the level of collection and establishment of data also differs by institution. In terms of institutions, they are divided into three levels: the level of collection and establishment of data(KIOM), the advanced level of it (KIST), And level of steps to start sharing (KRIBB, KRICT).

키워드: 연구데이터, 과학데이터, 데이터관리, 오픈사이언스, 데이터공유, 데이터재사용, 한국과학기술연구원, 한국과학기술정보연구원, 한국생명공학연구원, 한국지질자원연구원, 한국표준과학연구원, 한국한의학연구원, 한국화학연구원

Research Data, Scientific Data, Data Management, Open Science, Data Sharing, Data Reuse, KIST, KISTI, KRIBB, KIGAM, KRISS, KIOM, KRICT

* KSPP 수석연구원(mypark@kspp.re.kr) (제1저자)

** 동원대학교 문헌정보데이터관리과 교수(ijahn@tw.ac.kr) (교신저자)

*** KSPP 선임연구원(sjnam@kspp.re.kr) (공동저자)

논문접수일자 : 2018년 11월 28일 논문심사일자 : 2018년 12월 26일 게재확정일자 : 2018년 12월 27일
한국비블리아학회지, 29(4): 319-344, 2018. [http://dx.doi.org/10.14699/kbiblia.2018.29.4.319]

1. 서론

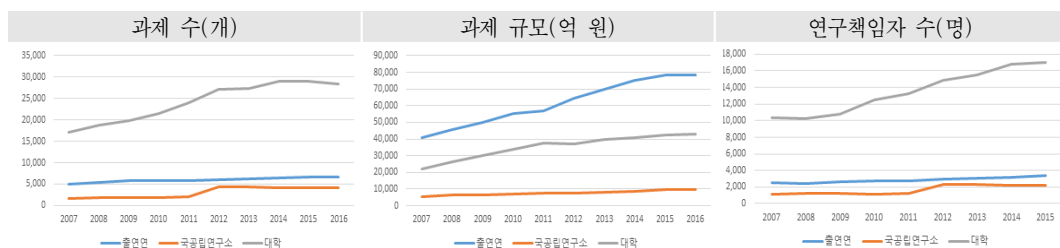
2000년을 전후로 하여 글로벌 차원에서 오픈 사이언스정책이 가시화되었다. 오픈사이언스는 디지털 기술을 통해 연구의 전 과정을 보다 개방적으로 전환하려는 일련의 움직임이다(OECD 2016). 2002년 부다페스트 오픈사이언스 선언을 시작으로 2010년 이후 미주, 유럽을 중심으로 국가차원에서 오픈사이언스정책이 수립, 이행되어가고 있다. 오픈사이언스의 주요특징은 공개연구, 오픈 액세스, 연구데이터공유, 공개적 동료평가, 투명한 연구 성과 평가이다(김순, 이보람, 김환민, 김혜선 2017). 특히 연구결과를 디지털로 공개함으로써 연구성과나 과정에 대한 후속 검증은 물론 추가 연구와 새로운 연구 방법 개발 등에 활용할 수 있기 때문이다.

이러한 운동은 과학 정보를 공유하는데 있어서 제약이 되는 규약과 장애물을 없애는데 초점을 맞추고 있다. 왜냐하면 오픈사이언스정책은 기본적으로 학문의 자유를 위한 연구 성과에의 접근성 증대, 공공 연구 성과의 사회경제적 편익 제고와 같은 사회적 요청에 부응하고자 마련된 사회적·정책적 산물이기 때문이다. 이러한

측면에서 외국의 적극적 오픈 사이언스 정책에 비하여 한국은 오픈 사이언스 특성 중 연구데이터 관리나 협력 연구에 있어서 특히 미약한 것으로 조사되고 있다.

국내 공적 자금이 투자된 연구데이터 주생산기관을 살펴보면 2016년 기준 정부 연구개발사업에 참여한 연구수행 주체로 정부출연기관이 41.2%, 대학이 22.5%로서 전체 정부투자자의 60%를 상회할 정도로 큰 비중을 차지하는 것을 알 수 있다.¹⁾ 이전 10년간인 2007년부터 2016년까지 연구수행주체별 국가 R&D 수행현황을 살펴봐도 과제규모 등 양적 측면에서 출연연구기관(누적수치 61조원)이 국가 R&D 연구데이터의 주요 생산자 및 이용자로 파악된다(〈그림 1〉 참조). 따라서 출연기관의 연구데이터 공유 및 관리 현황을 파악하는 것은 글로벌 차원의 오픈사이언스 정책으로 나아가는 첫걸음일 것이다.

최근 국내에서는 과학기술정보통신부를 중심으로 국가연구데이터센터 설립과 통합플랫폼 구축, 연구자용 데이터관리를 위한 클라우드 인프라 구축, 나아가 연구데이터 관리제도를 도입하고 연구자의 권리를 보장하기 위해 관련 법령



〈그림 1〉 10년간 주요 연구수행주체별 국가 R&D 수행 현황

1) 국가과학기술지식정보서비스(NTIS, www.ntis.go.kr)

정비를 위한 계획을 가지고 있다(The Science Times 2018.10.14). 이러한 제도와 법령, 플랫폼 구축이 실용적이고 이용자 관점에서 편의성을 가지기 위해서는 관련기관의 인식과 현황을 잘 파악하는 것이 선제이다.

따라서 본 연구는 연구데이터를 주로 생산하고 있는 과학기술분야 정부출연연구기관과 인문사회과학 분야의 데이터 공유에 대한 이용자 인식차이를 도출하고, 국내 과학기술정보통신부 산하 25개 출연기관을 중심으로 연구데이터와 공유 및 관리하는 사례 분석을 목적으로 한다. 사례분석 결과 우수한 기관을 선정하여 모델로 제시하고자 한다. 연구데이터에 관한 학문분야별 연구자들의 인식차이와 출연기관의 현황분석, 우수기관 모델화 결과는 향후 국가 및 기관 차원에서의 연구데이터 플랫폼개발과 메타데이터 개발에 반영할 수 있는 기초자료가 될 것이며, 오픈사이언스로 나가기 위한 로드맵 개발에 기초자료가 될 것이다.

2. 학문분야별 데이터공유 인식 차이분석

본 연구와 관련된 선행연구를 크게 네 부분으로 구분하여 고찰할 수 있다. 첫째, 법, 제도, 정책 연구, 둘째, 국내 연구자들의 데이터 관리 및 공유에 관한 인식연구, 셋째, 연구데이터 관리현황 연구, 넷째, 메타데이터에 관한 연구로 구분할 수 있다. 이에 대해서 구체적으로 살펴보면 <표 1>과 같다.

본 연구는 데이터 공유에 대한 이용자 인식과 이를 기반으로 한 국내 과학기술정보통신부 산하 출연기관 중심의 연구데이터와 공유 및 관리하는 사례 분석을 하고자 한다. 이를 위한 선제 연구로서 연구자 인식조사, 연구데이터 관리 분야의 선행연구를 중심으로 분석하였다. 인문사회학 분야의 연구데이터 관리 및 현황에 관한 연구는 국내의 기초학문연구센터의 시스템과 현황을 분석한 것이 거의 대부분이다. 자연과학분야의 경우는 출연연구기관에 관한 선행연구가 대부분이다. 하

<표 1> 선행연구

영역	논문 제목
법 제도 정책	- 윤종민, 김규빈 (2013) 과학데이터에 관한 입법례와 관리정책 그리고 대응방안 - 김지현 (2014) 국외 정부연구비지원기관의 연구데이터 관리정책 분석 - 김규빈 (2015) 과학데이터관리에 관한 해외 법제의 비교법적 고찰
연구자 인식조사	- 김지현 (2015) 데이터 관리와 공유에 대한 대학 연구자들의 인식에 관한 연구
연구데이터 관리	- 강주연 (2017) 생명공학분야 연구데이터 관리 방안 연구 - 김지현 (2012) 대학 내 연구자들의 연구데이터 관리에 관한 연구 - 김지현 (2011) 자연과학분야 대학실험실에서의 연구노트 작성 및 관리에 관한 연구 - 박미영, 안인자 (2017) 바이오뱅크 분야의 연구데이터 관리 및 서비스 동향에 관한 연구 - 최명석, 이승복, 이상환 (2017) 국내 과학기술분야 연구기관의 과학데이터 관리 현황
레포지터리	- 김문정 (2015) 과학기술분야 연구자의 레포지터리 활성화를 위한 연구데이터 공유에 대한 연구 - 김지현 (2016) 연구데이터 레포지터리의 데이터 접근 및 이용 통제 정책 요소에 관한 연구
연구정보 수집	- 김은정, 남태우 (2012) 연구데이터 수집에 영향을 미치는 요인 분석
메타데이터	- 김선태 (2012) 해양관측분야 메타데이터 표준요소 선정 및 설계에 관한 연구 - 김선태, 한선화, 이태영, 김용 (2010) 과학데이터 보존 및 활용모델에 관한 연구

지만 출연기관의 공유현황과 인식에 관한 연구는 사례적 제시, 시스템의 모형화가 아니라 통계적 제시이다(김은정 2012; 김문정, 김성희 2015; 최명석 2017). 본 연구는 실제 관리와 공유가 우수한 기관을 대상으로 담당자의 심층 인터뷰를 한 것으로 실제 사례에 대한 제시와 과제를 살펴볼 수 있으며, 홈페이지를 통하여 연구데이터 관리 현황을 살펴볼 수 있다. 또한 담당자를 대상으로 반구조화된 인터뷰 결과를 바탕으로 논문을 작성하였다. 따라서 본 논문의 의의는 실제적 사례를 통한 모형화에 있다.

2.1 인문사회과학분야 과학자들의 연구데이터 공유 인식

학문분야별로 연구자들이 연구데이터 공유와 관리에 관한 태도, 요구사항, 데이터 관리를 위한 인프라 및 정책이 많은 차이를 보인다는 연구 결과는 다수의 선행 논문에서 언급되고 있다(Swan and Sheridan 2008; Beagrie 2009; 김지현 2012). 본 연구에서는 이를 바탕으로 학문분야별 연구데이터 공유 인식 차이를 정리해보고자 한다.

반면 인문사회 연구데이터란 관련 후속 연구와 교육에 지속적으로 활용될 수 있으며 학술적 보존의 가치가 있는 양적 및 질적 자료로 연구적, 교육적, 그리고 증거적 활용 가치가 있다(신영란, 정연경 2012). 주제적으로는 사회, 정치, 경제 관련 정량적 데이터와 디지털형태의 정성적 데이터, 실험데이터, 이미지와 동영상, 사운드 데이터베이스로 제시되고 있으며, 세미나, 워크샵, 공청회, 토론회, 심포지엄 등에 관련된 자료들이 포함된다. 그 중 사회과학 분야에서도

인구통계 데이터, 체계적인 표본을 대상으로 한 다양한 설문 데이터의 축적이 많은 경우는 공유 및 활용도가 높은 편이다.

인문학 분야는 주로 독자적 작업을 수행하며, 활용자료에 대한 개인적 해석이 연구결과를 도출하는데 있어 중요한 역할을 한다. 또한 타 분야에 비해 다양한 학문분야로 구성되어 있으며, 연구 방법론 또한 고전을 비롯한 텍스트 자료의 분석, 실험, 설문 기반 연구, 현장 관찰, 창작 등으로 매우 다양하다. 인문학 연구에서는 주로 단행본과 학술자료를 이용하며, 생산되는 자료는 텍스트 자료, 원고, 녹음/동영상, 이미지, 낱장자료, 기사자료 등으로 매우 다양하다. 따라서 타 분야보다 학문분야별 데이터의 수집, 관리 및 공유를 위한 제도적, 시스템적 기반이 미약한 실정이다. 더욱이 교수와 연구진들은 그들의 데이터를 공유하고자 하지만 실제 이를 효과적으로 실행하기 위한 기술이 부족하다.

국내에서는 한국연구재단이 지원한 인문사회과학분야의 연구과제에서 생산되는 연구성과물 및 데이터를 관리하기 위해 2006년부터 기초학문자료센터(Korea Research Memory, KRM)를 운영하고 있다. 특히 학문의 특성상 비디지털 데이터가 생산되는 인문사회학 분야는 데이터베이스 작업이나 시뮬레이션 기술 등의 작업에 대한 비용이 높다. 반면 정량적 연구분야에서는 데이터 시각화와 시뮬레이션을 위한 도구와 과정이 잘 정립되어 데이터 재사용이 상대적으로 원활한 것으로 조사되고 있다(심원식, 안혜연, 변제연 2015).

인문사회학 분야에 관하여 데이터 공유 및 관리에 관하여 분석한 학자들의 동기요인, 방해요인, 분야의 특성에 대한 분석을 살펴보면 <표 2>와 같다.

〈표 2〉 인문사회과학분야 연구자 연구데이터 공유활용 인식

연구자	American Council of Learned Societies (2006)	Perry (2008)	김현희 (2006)	신영란, 정연경 (2012)	심원식, 안혜연, 변제연 (2015)
연구대상분야	인문사회과학	인문사회과학분야	경제, 사회분야	인문사회과학	인문학
동기 요인	• 제도, 시스템 기반 필요	• 평판효과, 상호호환성 지식의 효력 • 타인도움의 즐거움	• 기관의지, 인식 강화	• 협력체계 필요 • 메타데이터 보완	• 보상체계, 관행, 절차 필요
방해 요인	• 연구데이터 규정, 분류어려움	• 공유의식 부족 • 추진주체의 의지부족 • 공식공유 시스템미비	• 기술, 인프라부족 • 인식부족 • 정책, 설득 부족	• 데이터부족 • 시스템부족 • 정책정비부족	• 정형화어렵고(기술) • 관리싸이클, 절차 부재
분야의 특징	• 비디지털데이터의 경우 DB작업과 메타데이터 작업 비용이 과함 • DMP 관련 제도, 시스템 기반 미약 • '데이터와 데이터의 근원(data and sources of data)'간 경계 모호 • 아카이브 자체가 의미 있음				

2.2 자연과학분야 과학자들의 연구데이터 공유 인식

과학데이터는 연구를 수행하는 과정에서 생산되는 모든 데이터를 포함하며 이러한 데이터는 통계수치, 수식, 이미지지도, 도표, 문서 등을 포함한다. 데이터로서 학술과정에서 생산되는 연구데이터 또는 과학데이터를 의미한다. 일반적으로 과학데이터는 관측, 감시, 조사, 실험, 연구 분석, 계산 등 일련의 과학기술 활동의 결과로서 수치, 공간, 도표, 문서 등의 형태를 취한다. 따라서 그 형식이 복잡하고 널리 배포되며 범위가 넓은 것이 큰 특징이다. 이공계열은 대부분이 실험데이터, 계산데이터, 관찰데이터, 데이터의 기록화는 적은 형편이다. 제공하는 연구자보다 이용하는 연구자는 세 배에 이른다(김지현 2012).

자연과학 또는 공학 분야에 있어서 생산되는 데이터는 연구자들의 많은 시간, 노력, 인력 또한 수많은 예산이 투입되는 경우가 빈번하다. 국가적 차원에 있어서 국가 예산 및 자원의 효율적인 활용과 재사용성에 있어서 연구를 통하

여 생산된 모든 과학데이터는 매우 중요한 국가 자원이다. 특히 해양 환경, 천체물리학, 생명공학 등의 분야는 데이터를 반복해서 수집하기 어렵고, 대규모 장비 및 설비투자를 통해서 수행되거나 데이터 공유의 효과가 크고 직접적이라는 것이다(심원식 외 2015).

국내에서는 과학기술정보통신부를 중심으로 국가연구데이터센터 설립과 시스템과 제도 정비 위한 계획을 가지고 있다. 이러한 시점에 자연과학 분야에 관하여 데이터 공유 및 관리에 관하여 분석한 학자들의 공유 동기요인, 방해요인에 대한 분석은 시스템 및 제도 도입에 중요한 데이터가 될 것이다.

선행연구를 통하여 연구데이터 공유현황과 인식전반에 관하여 비교, 분석 결과 인문과학, 사회과학과 자연과학 분야의 여러 면에서 상이한 결과를 발견할 수 있다. 내용은 다음과 같다.

첫째, 데이터의 유형 측면에서 인문과학분야의 데이터는 데이터 소스와 결과 데이터와의 경계 구분이 모호하며, 사회과학 분야의 데이터는 데이터의 규정과 분류의 어려움, 데이터의 양적

〈표 3〉 자연과학기술분야 연구자 연구데이터 공유활용 인식

연구자	김은정 (2012)	김문정, 김성희 (2015)	Williams (2013)	박미영, 안인자 (2017)	최명석 외 (2017)
대상	과학기술분야	과학기술분야	작물학	생명과학	과학기술분야
동기요인	과제완료의 안정감, 의무감, 만족감, 실적 인정, 금전적 보상, 좋은평판	- 보상체계 - 인지성	- 타데이터 공유 기대 - 다른 연구자에게 기여 - 연구비지원기관의 요구사항	- 시스템 구축필요, 제도기반 필요 - 가이드라인 필요	- 제도 개선, 데이터 관리·활용 지원체계 - 공감대 형성
방해 요인	- 표절, 아이디어 도용 - 논문 출판의 주도권 상실	아이디어 도용, 표절, 논문 출판의 주도권 문제	- 업무가중 - 데이터 표준화어려움	연구 성과 인정없음	- 성과인정없음 - 추가적 업무관리, 품질 및 신뢰성 책임 이슈 - 관리·공유가 부담
분야별 특징	정보인프라요인(관리시스템, 메타데이터, 데이터표준화, 관리 및 교육프로그램)				

이 부족 현상이 보인다. 반면 자연과학분야 데이터는 통계수치, 수식, 이미지지도, 도표, 문서 등을 포함하며, 양적으로 방대하고, 구조화하기가 용이한 특성을 지닌다. 자원의 가치 측면에서 자연과학 분야의 데이터는 연구자들의 많은 시간, 노력, 인력이 투입되는 경우가 빈번하고, 특히 국가적 차원에서 예산 및 자원의 투입에 대한 효율적인 활용과 재사용이 요구된다.

둘째, 인식 측면에서 보상, 관행, 절차 등을 통하여 인문학분야 과학자들과 소통이 요구된다. 사회과학분야의 경우 기관의 의지와 인식이 우선적으로 요구되며, 평판 효과 나아가 타인에게 도움이 되는 자신의 가치 상충의 만족에 대한 인문적 공감과정이 요구된다. 특히 인문사회과학 모든 분야에서 공유인식 부족과 주지주체의 의지 부족을 중요한 방해요인으로 제시하고 있다. 자연과학 분야의 경우 제도, 시스템, 가이드라인 등의 환경 구축이 갖추어 질 경우 기대되는 과제완료의 안정감, 의무감, 만족감, 실적인정 등의 인식공유, 기관의 요구사항 충족에 대한 기대가 동기부여로 제시되고 있다. 특히 자연과학 분야 공유인식의 방해요인은 아이디어 도용, 표절, 논문 출판의 주도권

문제, 업무가중의 불편함 등으로 연구자들에게 추가적으로 가중되는 구체적 불편함이 제시되고 있다.

셋째, 제도와 시스템화 측면에서 모든 분야에서 아직은 초기단계이지만 특히 인문학분야는 정형화하여 기술을 적용하기가 어렵고, 관리 절차가 부재한 상황이다. 인문사회과학분야도 제도, 시스템 기반이 요구되며, 특히 협력체계와 메타데이터의 보완이 요구된다. 반면 자연과학 분야 데이터는 정보인프라요인이 비교적 명확하고, 공유, 활용에 대한 공감대가 일부 형성된 바 과학기술정보통신부를 중심으로 국가연구데이터센터 설립과 시스템과 제도 정비를 위한 계획이 실행 단계 있다.

분석 내용으로 볼 때 우선적으로는 자연과학 분야의 데이터가 단순하고 정형화하기 용이하고, 인식 면에서도 구체화 되어 있으므로 국가출연연구기관을 중심으로 데이터 공유 및 활용을 위한 프로세스를 시도하는 것이 적합하다. 2차적으로 자연과학분야의 기술, 시스템의 구조화 및 경험을 바탕으로 이를 인문과학분야에 적용하기 위한 융합적 연구, 인문과학 특성화, 시스템 고도화 등의 과정이 필요하다.

3. 연구설계 및 방법론

수기관을 선정하여 모델을 제시하였다.

3.1 연구설계

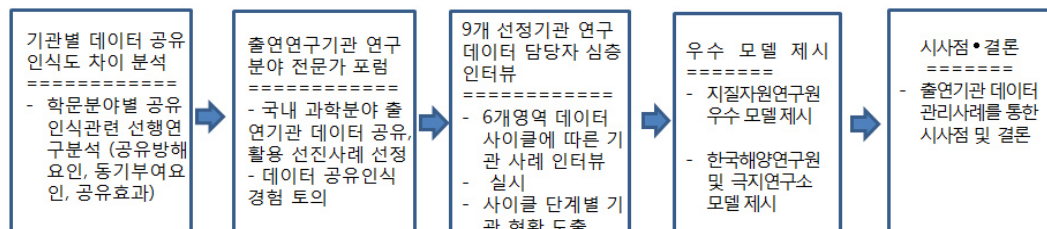
본 연구에서는 국내 과학기술분야 정부 출연 연구기관 27개를 대상으로 연구데이터 관리현황 포럼과 심층 인터뷰를 개최하였으며, 각 기관의 2017년 수행과제에 관하여 관리 및 공유 현황조사를 실시하였다. 이중 25개 정부출연연구기관은 과학기술정보통신부 국가과학기술연구회 소속이고, 2개 기관인 한국해양과학기술원과 극지연구소는 2012년 이후 과학기술정보통신부에서 해양수산부 산하로 이관된 기관이다. 이관된 기관임에도 성격적으로 유사성이 있으며, 연구데이터 관리 면에서 매우 우수한 기관이기 때문에 타 기관의 모델적 성격으로 포함되었다.

세부내용은 <그림 2>와 같다. 첫째, 연구데이터 공유, 활용 체계에 대한 이해도증진 및 공감대 형성을 위하여 연구기관 전문가 32명을 대상으로 전문가 포럼을 개최하였다. 본 포럼에서는 공유 경험을 나누고, 국내기관의 연구데이터 관리 체계 구축현황 및 우수사례를 제시하였다. 둘째, 조직차원에서 연구데이터 관리를 진행 중인 9개 기관을 대상으로 연구데이터 단계별 현황에 관하여 인터뷰를 실시하였다. 셋째, 3개 우

3.2 연구질문

연구데이터 관리에 관한 통합관리체계는 제도, 인력, 조직 등의 인프라, 물리적 리포지토리, 디지털 리포지터의 구성과 활용으로 모듈화하는 것이 통상적 방법이다. 이를 기반으로 반구조화된 인터뷰(semi structured interview) 설문이 구성되었고, 포럼과 인터뷰를 통하여 질의되었다(한국지질자원연구원 2015).

- 1) 제도: 기관내 연구데이터 수집관리를 위해 제도 방안이 마련되어 있는가? 연구데이터관리계획(DMP)수립되어 있는가? 내부적으로 관련 규정이 마련되어 있는가?
- 2) 인력 및 조직: 연구데이터 수집관리를 위한 조직, 인력 등 추진체계가 구축되어 있는가?
- 3) 수집 및 저장: 데이터의 저장, 시스템, 메타데이터 등 관련 가이드라인이 있는가? 어느 정도 기간의 데이터가 축적되어 있는가? 데이터의 타입과 특징은 무엇인가?
- 4) 관리 및 보존: 공유 활용을 위한 기관내 레포지토리, 플랫폼이 있는가? 공유한 사



<그림 2> 연구내용과 방법

례가 있는가? 혹시 국외에서 요청한 적이 있는가?

- 5) 공유 및 활용: 기관 내에서 추진하고 있는 수집관리 혹은 공유 활용을 위한 사업이 있는가? 예산이 확보되어 있는가? 기관 경영진의 관심은 어느 정도인가? 연구자들의 공유 활용에 대한 인식은 적극적인가? 내부 규정으로 보상체계가 마련되어 있는가?

4. 연구데이터 수집 및 관리 현황분석

4.1 분석의 대상

본 연구에서는 국가연구개발사업 수행을 통해 연구데이터를 주로 생산하고 있는 과학기술 분야 정부출연연구기관을 분석 대상으로 삼고 해당 기관의 연구데이터 수집·관리 현황을 살

펴보았다. 분석 대상기관인 정부출연연구기관 구기관은 '과학기술분야 정부출연연구기관구기관 등의 설립·운영 및 육성에 관한 법률'에 근거하여 설립된 특수법인으로서 부설연구기관들을 포함하여 모두 과학기술정보통신부, 해양수산부 산하 공공기관(기타공공기관)으로 지정되어 있으며, 국가과학기술연구회(한국해양과학기술원, 극지연구소 제외)의 지원을 받는다.

본 연구에서는 과학기술정보통신부, 해양수산부 소관 27개 연구기관 중 자체적으로 연구데이터 수집·활용 체계를 구축하고자 관련 사업을 추진 중에 있는 9개 기관을 대상으로 2018년 5월 30일~6월 10일 연구데이터 담당자 심층인터뷰를 통해 분석을 진행하였으며, 기관 목록은 <표 4>와 같다.

4.2 연구데이터 수집·관리 근거

국가연구개발사업의 관리 등에 관한 규정(이

<표 4> 과기정통부, 해양수산부 산하 정부출연연구기관

	연구소명	분야		연구소명	분야
1	국가보안기술연구소	정보보안	15	한국에너지기술연구원	에너지
2	국가핵융합연구소	핵융합	16	한국원자력연구원	원자력
3	녹색기술센터	기후기술	17	한국전기연구원	전기
4	세계김치연구소	식품(김치)	18	한국전자통신연구원	전자통신
5	안전성평가연구소	생명과학(비임상)	19	한국지질자원연구원	지질
6	재료연구소	소재기술	20	한국천문연구원	천문
7	한국건설기술연구원	건설기술	21	한국철도기술연구원	철도
8	한국과학기술연구원	기초연구전분야	22	한국표준과학연구원	국가 측정표준
9	한국과학기술정보연구원	과학기술정보	23	한국한의학연구원	한의학
10	한국기계연구원	기계	24	한국항공우주연구원	항공우주
11	한국기초과학지원연구원	기초연구전분야	25	한국화학연구원	화학
12	한국생명공학연구원	생명분야	26	한국해양과학기술원	해양
13	한국생산기술연구원	제조기술	27	극지연구소	극지환경
14	한국식품연구원	식품	* 분석 대상기관은 굵은 테두리, 음영 표기		

하 공동관리규정) 제25조에서는 논문, 특허, 보고서 원문, 연구시설·장비, 기술요약정보, 생명자원, 소프트웨어, 화합물, 신물질 등 공식 R&D 성과물을 지정하였다. 이를 효율적으로 관리하고 유통하기 위해 과학기술정보통신부장관이 지정한 전담기관에 분야별로 등록하거나 기탁하도록 규정하고 있다(〈표 5〉 참조). 예를 들면 논문과 보고서 원문은 한국과학기술정보연구원, 특허는 한국지식재산전략원, 기술요약정보는 한국산업기술진흥원 등에서 전담하고 있다.

국가연구개발사업의 연구성과의 등록대상과 기탁대상, 범위는 〈표 6〉과 같다. 실험·관측·분석 등 연구개발과정에서 산출된 ‘연구데이터’는 연구결과로서 가치가 높음에도 불구하고 현재 공동관리규정 상 공식 R&D 성과물에 포함되지 않아 국가차원에서 수집하고 관리할 수 있는 제도적 기반이 마련되어 있지 않은 실정이다.

공적자금을 투입한 국가연구개발사업에서 얻은 연구데이터를 관리하고 공유하기 위해서 또 하나 뒷받침되어야 할 제도적 기반은 데이터관리계획(Data Management Plan, 이하 DMP)

〈표 5〉 국가 R&D 연구성과물 관리전담기관 현황

연구성과물	관리전담기관	연구성과물	관리전담기관
논문	한국과학기술정보연구원	연구시설·장비	한국기초과학지원연구원
특허	한국지식재산전략원	화합물	한국화학연구원
보고서원문	한국과학기술정보연구원	생명 자원	한국생명공학연구원
기술요약정보	한국산업기술진흥원		한국생명공학연구원
소프트웨어	한국저작권위원회		농업유전자원센터
	정보통신산업진흥원	신물질정보	농림수산식품교육문화정보원

〈표 6〉 국가연구개발사업 연구성과의 등록 및 기탁 범위

등록대상		기탁대상	
논문	국내외 학술단체에서 발간하는 학술지 및 학술대회지에 수록된 학술논문	생물자원 (생명자원)	미생물자원(세균, 곰팡이, 바이러스 등), 동물자원(사람·동물세포, 수정란 등), 식물자원(식물세포, 종자 등), 유전체자원(DNA, RNA, 플라스미드 등) 및 관련 정보
특허	국내외에 출원 또는 등록된 특허정보		
보고서 원문	연구개발 종료 시 제출하는 최종보고서 및 연차보고서(전자원문 포함)		
연구시설·장비	국가연구개발사업을 통하여 취득한 3천만원 이상 연구·시설장비 또는 3천만원 미만이라도 공동활용이 가능한 연구·시설장비	화합물	합성 또는 천연물에서 추출한 유기화합물 및 관련 정보
기술요약정보	기초·응용·개발단계 등의 최종보고 및 연차보고가 완료된 결과물의 기술정보를 요약하여 공유·활용(기술이전·사업화 등)할 수 있도록 작성된 기록정보		
생명정보 (생명자원)	유전체정보(서열, 발현정보 등), 단백질정보(서열, 구조, 상호작용 등), 발현체정보(유전자(DNA)칩, 단백질칩 등) 및 그 밖의 관련 정보		
소프트웨어	창작된 소프트웨어 및 등록에 필요한 관련 정보		

이다. 데이터관리계획은 연구개발 과정에서 생성된 데이터에 대한 관리 및 계획을 담은 문서로서 연구수행과정에서 누가, 어떤 데이터를 언제, 어떤 방법을 통해 어떤 형태로 취득할 것인지를 사전에 계획함으로써 연구수행의 효율성을 높이고, 연구과정에서 생산되는 데이터를 체계적으로 관리하기 위한 목적으로 작성한다.

미국, 영국, 호주 등 데이터관리 선진국은 공적 자금지원의 연구개발사업에 대해서는 DMP 제출을 의무화하고 있으며, 사례로 제시한 <표 7>은 미국연구재단(NSF)의 DMP 주요 요구사항에 관한 내용이다. 반면, 국내는 공동관리규정 상 데이터관리계획에 관한 조문이 마련되어 있지 않다. 한국지질자원연구원, 극지연구소 등 개별 기관차원에서 소관 분야 연구데이터를 수집하여 활용하고자 자체적으로 DMP를 마련한 기관이 있으나 해당 데이터의 공유 및 활용은 제한적이다.

최근 정부에서는 ‘연구데이터 공유·활용 전략’ 발표(‘18.01)를 통해 연구데이터 관련 주요 핵심 과제 및 개략적인 추진 내용을 도출하였다. 여기에는 공동관리규정 조문 개정안이 포함되어 현재 ‘국가 R&D사업 공동관리규정’에 공식 연구성과물로 연구데이터를 포함하기 위해 연구데이터 정의를 담고, 연구제안자로부터 ‘데이터관리계획(DMP)’을 접수하는 내용을 추가하

는 개정 작업이 진행 중이다. 이 밖에 ‘연구관리 표준매뉴얼’에 연구데이터 관련 연구자 권리 개정도 함께 추진 중에 있다. 공동관리규정 개정안은 2018년 하반기 국가과학기술심의회 운영위원회를 거쳐 개정 작업을 진행 할 예정이다.

4.3 출연연구기관 연구데이터 수집·관리 사례 분석

4.3.1 연구데이터 수집·관리 사례 진단 분석의 틀

본 연구에서는 연구데이터를 생산해서 활용하기까지 연구데이터 사이클(규정 및 제도, 인력 및 조직, 수집 및 저장, 관리 및 보존, 공유 및 활용(인프라))에 따라 과학기술정보통신부 산하 7개 출연연구기관을 대상으로 연구데이터 수집 및 관리를 위한 체계를 갖추었는지 살펴보고자 한다. 연구데이터 수집·관리 사례 진단 분석의 틀은 <표 8>과 같다.

4.3.2 출연연구기관 연구데이터 수집 및 관리 현황

- 1) 한국과학기술연구원(Korea Institute of Science and Technology, KIST)
한국과학기술연구원(이하 KIST)은 기관고유

<표 7> 미국 연구재단(NSF)의 DMP 주요 요구사항

데이터 유형	데이터, 샘플, 물리적 수집물, 소프트웨어, 커리큘럼 자료, 연구과정에서 생산되는 기타 자료에 대한 유형
데이터 표준	데이터, 메타데이터 포맷 및 내용에 사용되는 표준 (기존 표준이 부재하거나 부적절한 경우, 해결방법을 문서화하여 제출)
공유·개방 정책	프라이버시, 기밀, 보안, 지적재산권, 기타 권리·요구사항 보호 관련 사항을 포함한 데이터 공유·개방 정책
재사용 정책	재사용, 재배포, 파생연구를 통해 생성된 데이터를 위한 정책·조항
데이터 아카이빙 계획	데이터, 샘플, 기타 연구 결과물에 대한 아카이빙 및 접근(공유)을 위한 계획

〈표 8〉 본 연구의 분석 틀

구 분	분석의 핵심 주제
규정 및 제도	- 기관차원에서 연구데이터를 체계적이면서 효율적으로 수집하고, 관리·보존·활용하기 위한 ‘연구데이터관리 규정(안)’ 마련 여부 - 연구계획서 제출 시 DMP 제출을 요구하는지 여부
전담인력 및 조직	조직 내 연구데이터 수집·관리를 전담할 인력 및 조직 유무
수집 및 저장*	- 연구자의 데이터 업로드를 통한 수집 기능 - 표준 프로토콜을 통한 데이터 자동 수집 기능
관리 및 보존*	- OAIS 표준에 의한 데이터의 체계적인 관리 - 데이터 라이프사이클** 기반의 보존
공유 및 활용* (데이터 출판/서비스)	- 연구자의 데이터에 글로벌 식별자(DOI) 부여 - 글로벌 데이터 핸들러 서버에 데이터 등록 - 조직별, 컬렉션별, 저장별, 주제분류별 검색 가능 - 데이터와 학술문헌과의 연계서비스

* 해당 기능이 실현되기 위해서는 물리적/디지털 리포지터리 시스템 구축 필요함

** 데이터 생산-처리-분석-출판-아카이브-재활용

〈표 9〉 KIST 연구데이터 인프라 서비스

서비스 구분	주요 기능
KIST R&D Platform (KiRI Note) 구축 및 서비스	- 연구자의 통상적 연구활동이 DB화 될 수 있는 플랫폼 구축 - 특성분석센터 데이터 통합관리 시스템 업그레이드 * 실험·측정 데이터의 자동수집·저장시스템, 연구노트 생성기 등 - 주제별 계산과학 플랫폼과의 연계 프로토콜 구축 - AI 기반 실시간 연구활동 지원 서비스 구현
KIST R&D Data Bank 구축 및 서비스	- 지속적 경쟁력을 가진 소재(축매) 데이터 관리 체계 확립 및 데이터 축적 - KIST 내외부의 정형·비정형 데이터의 통합, 저장 및 관리 시스템 확립 - 데이터 기반 실시간 연구활동 지원 서비스 개발 - 이질적인 과학기술 데이터의 체계적 관리를 위한 메타데이터/마스터 데이터 관리
KIST R&D Informatics Platform (KiRI) 구축 및 서비스	- KIST 자체 데이터를 포함한 글로벌 스케일의 데이터를 통합 검색 및 활용을 위한 플랫폼 개발 - 글로벌 과학기술 데이터 연계를 위한 분석대상 데이터 통합관리 시스템 구축 - 데이터 가시화 및 분석이 가능한 KIST R&D Data Bank 서비스 포털 구축 - 머신러닝 기반 축매 설계 시스템 개발 및 서비스

* 출처: 서덕록, KIST 데이터기반 R&D환경 구축사업 설명자료(2018.5.28.)

사업으로 '17년 10억원, '18년 8억원의 예산을 투입하여 데이터 기반 R&D 환경 구축 사업을 진행 중이다. KIST 내 모든 연구 분야에 적용될 수 있는 범용의 인프라 구축을 목표로 하되, 인프라의 실효성 입증에 위한 적용 분야로서 소재(축매) 분야를 선정하여 〈표 9〉와 같이 3종의 서

비스를 구축 중에 있다.

① 제도적 측면에서 KIST는 기관 자체적으로 연구데이터 관리규정을 마련하고 있지 않으나, 데이터 기반 R&D환경 구축 사업을 통해 연구데이터 수집·관리·공유 인프라

라를 구축하여 '18년 말 베타서비스를 제공할 예정이다.

- ② 연구데이터 수집·저장 현황에 있어, '18년 말 서비스 구축 완료시점에 약 15만 건의 데이터(축매분야 논문데이터 10만건, 측정데이터 5만 건)를 축적할 예정이다.
- ③ 연구데이터 관리 및 보존에 있어서는 현재 인프라 구축 중으로서 연구데이터 공유·활용이 되고 있지는 않으나, 시스템 개발 후 KIST 내부에 우선 공개 예정이며 단계적으로 확대할 계획이다. 다만, 측정데이터 및 논문 원문은 라이선스 상의 문제로 공개가 불가하며, 자체 측정데이터는 일정 기간의 엠바고 기간 후 GPL, CCL등의 공공라이선스(public license)를 활용한 공개를 검토 중에 있다.

2) 한국과학기술정보연구원(Korea Institute of Science and Technology Information, KISTI)

한국과학기술정보연구원(이하 KISTI)은 기초연구실험데이터글로벌허브구축사업(과학기술정보통신부 지원)으로 '15년 28억원, '16년 25억원, '17년 21억원의 예산을 투입하여 첨단 연구 장비에서 생산되는 대용량 데이터를 확보하고, 실험데이터 분석을 위한 컴퓨팅 인프라를 구축 중에 있다.

- ① 제도적 측면에서 KISTI는 기관 자체적으로 연구데이터 관리규정을 마련하고 있지 않으나,
- ② 조직/인력측면에서 기관 내 연구데이터허브센터, 연구데이터플랫폼센터 등의 전담

조직을 두어 연구데이터 관련 사업을 준비 중에 있다.

- ③ 급증하는 데이터 분석 수요 대응을 위한 데이터 저장·공유·분석 인프라를 구축하고자 컴퓨팅 인프라와 저장인프라를 확장하고 있다. 일례로, KISTI는 컴퓨팅 인프라를 '15년 CPI 560코어에서 '17년 1,260코어로 2배 이상 확장하였으며, 저장인프라 또한 1.5PB의 CERN 가속기 원천데이터 확보 및 테입시스템의 활용이 90%를 넘어, 추가적으로 1.5PB 테입(Tape) 저장인프라를 확장하여 3PB 규모의 원천데이터 확보가 가능해졌다.
- ④ 연구데이터 공유·활용 측면에서 ALICE, CMS 등 연구데이터 공유·분석 환경을 제공하여 국내외 연구자가 사용 중에 있다.

3) 한국생명공학연구원(Korea Research Institute of Bioscience and Bioechnology, KRIBB)

한국생명공학연구원(이하 KRIBB)은 생명연구자원 통합시스템 활용기반 구축사업(과학기술부 지원)으로 '10년부터 '19년까지 9년간 약 270억원의 예산을 투입하여 국가생명연구자원 통합시스템(KOBIS)을 구축하고 생명연구 분야 연구데이터 수집·활용을 지원 중에 있다.

- ① 제도적 측면에서 공동관리규정 상 생명정보는 연구성과물로 지정되어 KRIBB는 생명연구자원법 제11조에 따라 전담조직(국가생명연구자원정보센터)을 설립하고 생명정보의 통합적 관리와 유통을 제공하고 있다. 반면 관리계획, 문서화면에서 단계 시행이 비미하다(박미영 외 2017).

〈표 10〉 연구데이터 분야별 국내외 참여연구자 규모

실험명	참여연구자 규모
ALICE 실험(CERN)	• (국외) 37개국 151개 기관, 1,550명 참여 • (국내) 한국 ALICE 연구그룹 - 국내 6개 대학, 40여명 참여
CMS 실험(CERN)	• (국외) 40개국 172개 기관, 약 3,000명 참여 • (국내) 한국 CMS 연구그룹 - 국내 8개 대학, 80여명 참여
Belle II (일본 KEK)	• (국외) 20개국 67개 기관, 600여명 참여 • (국내) 한국 Belle 연구그룹 - 국내 9개 대학 40여명 참여
중력파(LOGI/KAGRA)	• (국외) 16개국, 60개 기관, 1,000여명 • (국내) 한국중력파연구협력단 및 해외 공동연구자 - 10개 대학, 4개 연구소, 40여명 참여
RENO	• (국내) 한국중성미자연구센터 - 국내 5개 대학 50여명 참여
유전체	• (국내) 유전체 연구그룹 - 국내 40여명 참여
구조생물학(KBSI 전자현미경)	• (국내) 한국기초과학지원연구원(KBSI), KAIST 등 15개 연구그룹 50여명

* 출처: 노서영, 2018년도 기초연구실험데이터글로벌허브구축사업 설명자료(2018.1.29.)

② 연구데이터 수집 측면에서 KOBIC과 정보연계된 생명정보 연구성과물 등록시스템인 Biodata의 경우, '18년 4월 기준 총 1,198,937건(23TB)의 연구성과물이 등록되어 있으며, 다부처 유전체 정보연계 시스템인 Genome-InfraNET에도 '18년 4월 기준 14,568건(82TB)의 정보가 연계되어 있으나 해당 데이터에 대한 공유·활용 실적은 저조한 실정이다.

③ 공유·활용 측면에서 KRIBB은 '18년 1월

기준 CPU 3,100 코어, 저장 공간 10PB 등 컴퓨팅 저장 인프라를 구축하고 구축된 KOBIC 시스템 활용으로 5개 부처 유전체 정보 및 생명연구자원정보의 수집·관리 서비스를 지원 중에 있다. '18년 4월 기준 생명연구자원정보 연계의 누적건수는 4백만건을 돌파하였으나 생명정보 분석·활용 서비스 지원은 27건에 불과해 생명연구자원정보 공유·활용 수준은 수집·구축된 데이터 대비 아직 미비한 실정이다.

〈표 11〉 생명연구자원통합시스템의 생명연구자원 수집·관리 현황

(단위: 건)

지표[지표유형]	2015		2016		2017		2018.04	
	목표	실적	목표	실적	목표	실적	목표	실적
생명연구자원정보 연계(누적)	-	3,166,417	3,260,000	2,336,451	3,360,000	3,199,049	3,406,667	4,837,155
생명정보 분석·활용 서비스 지원	-	50	55	55	60	63	21.67	27

* 출처: 진태은, 국가생명연구자원 통합시스템 구축 설명자료(2018.5.25.)

4) 한국지질자원연구원(Korea Institute of Geoscience and Mineral resources, KIGAM)

한국지질자원연구원(이하 KIGAM)은 기관 고유사업으로 '15년 2.7억원, '16년 3.6억원, '17년 2.1억원의 예산을 투입하여 KIGAM 지질자원정보 통합관리체계 구축사업을 진행하였다.

- ① 사업수행의 결과, 제도적 측면에서 KIGAM은 기관 자체적으로 DMP양식을 포함한 연구데이터 관리규정을 마련하였으며,
- ② 조직·인력 차원에서 기관 내 지질자원데이터센터를 구축하여 연구데이터 대응 전담 인력과 조직을 구축하였다.
- ③ 연구데이터 수집·관리를 위해 연구데이터 분류코드를 작성하고 메타데이터 항목을 선정하며 데이터 유형별 XML 스키마를 설계하고 구현하는 등 수집·저장을 위한 표준 프로토콜을 구축하였다.
- ④ 이후 2016년 연구데이터 공유·활용을 위해 웹기반의 디지털 지질자원 연구데이터 리포지토리 시스템인 Geoscience Data Repository System(GDR)을 개발하여 현재 1차 파일럿 시스템을 서비스하고 있다.

5) 한국표준과학연구원(Korea Research Institute of Standards and Science, KRISS)

한국표준과학연구원(이하 KRISS)은 기관의 연구결과인 측정데이터의 수집, 구축을 통한 고객 맞춤형 보급 확산을 목표로 KRISS 데이터센터(KDC) 지정과 KDC 플랫폼 구축사업을 추진 중에 있다.

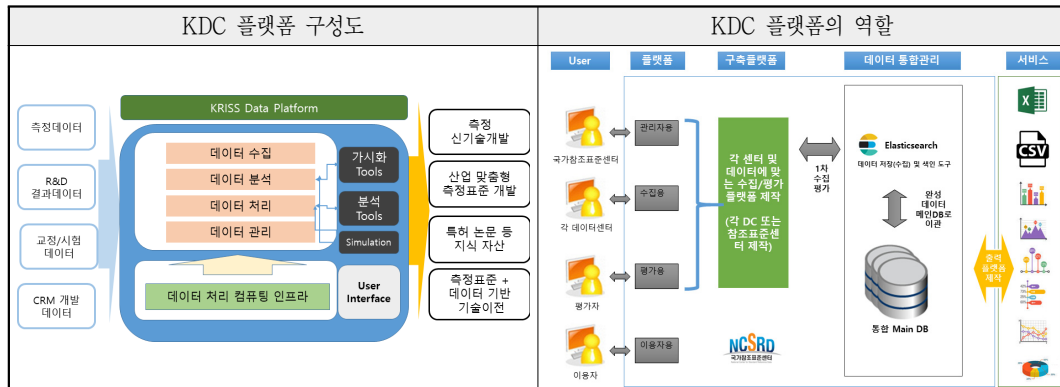
① 제도적 측면에서 KRISS는 기관 자체적으로 연구데이터 관리규정을 마련하고 있지는 않으며, KDC 플랫폼 구축사업에 따른 사업담당 인력은 존재하나

② 기관 차원의 연구데이터 전담조직·인력이 구성되지는 않은 실정이다. 현재 KRISS의 연구데이터 공유·활용 체계는 기획단계로 KDC 플랫폼의 구성 및 역할까지 도출된 상태이며, '20년까지 신뢰성 있는 측정데이터의 생산능력을 갖춘 연구센터에 KDC를 5개까지 지정하는 것을 목표로 시범사업(금속 박막 복소굴절률 측정 데이터를 시범사업으로 운영 중)을 운영 중에 있다.

③ KDC 플랫폼의 주요 기능은 데이터 구축, 분석, 처리, 관리기능이다. 데이터 수집 관리 측면에서는 데이터 구축기능으로 기존 또는 신규 데이터를 수집 및 구축하고 외부 사이트의 데이터와 연계가 가능하다(〈그림 3〉 참조). 구축된 데이터는 분석기능을 통해 활용성을 제고할 수 있으며, 분석 소프트웨어를 통한 데이터 분석 처리, 데이터 처리결과 발간(publication), 측정표준 등을 위한 2차 데이터 가공 등 데이터 처리도 할 수 있다. 마지막으로 통계, 데이터 업데이트, 참조표준화, DOI 부여 등 관리기능을 통해 데이터 구축의 편의성을 제공받고 데이터 양적 확대를 제고할 수 있다.

6) 한국한의학연구원(Korea Institute of Oriental Medicine, KIOM)

한국한의학연구원(이하 KIOM)은 연구개발적립금사업으로 한의학 연구데이터 통합관리 체계를 구축하고자 사전준비단계에 있다.



* 출처: 채균식, 참조표준사업과 KRIS data activities 설명자료(2018.5.31.)

〈그림 3〉 KRIS의 KDC플랫폼 개념도 및 역할

- ① 제도적 측면에서 연구데이터 공유·활용 전략 수립을 위해 TF팀을 구성하고 연구원 내 연구데이터 현황 자료 설문조사('18.4)를 바탕으로 연구분야별 특성을 반영한 연구데이터 분류체계 및 DMP를 개발하고 DMP보급을 위해 '18년 말 기관 내 시범운영할 예정이다.
- ② 현재 KIOM의 연구데이터 공유·활용 체계는 기획단계로 연구데이터 수집·관리·공유·활용은 이루어지고 있지 않으나, 내부적으로 연구데이터 관리체계 및 운영지침을 개발하고 DMP를 시범적용하고 있는 사실은 고무적이다. 이를 바탕으로 '19년 KISTI의 연구데이터 리포지토리 시스템을 활용하여 연구원 환경에 맞도록 한의학 연구데이터 공유·활용 플랫폼을 구축할 예정이다.

7) 한국화학연구원(Korea Research Institute of Chemical Technology, KRICT)

한국화학연구원(이하 KRICT)은 내부 주요 사업 및 과기정통부 사업(나노소재원천기술개발사업, 인공지능신약개발구축사업)을 통해 열전소재 및 신약분야 플랫폼을 구축하여 연구데이터를 수집하고 활용하려는 사업을 진행 중에 있다.

발사업, 인공지능신약개발구축사업)을 통해 열전소재 및 신약분야 플랫폼을 구축하여 연구데이터를 수집하고 활용하려는 사업을 진행 중에 있다.

- ① 제도적 측면에서 KRICT는 자체적으로 관리규정 및 DMP를 개발하고 있지 않지만,
- ② 조직·인력 측면에서 화학데이터기반연구센터를 통해 연구데이터 수집 및 관리 플랫폼 개발, 데이터 기반 소재 물성 예측 플랫폼 개발, 화학시뮬레이션 기반을 활용한 연구 수행 등 전담조직 및 인력을 구축한 상태이다.
- ③ KRICT는 '17년부터 '22년까지 5년간 연간 10억원을 투입하여 고효율 열전 신소재 예측 및 개발을 위한 열전소재 데이터 플랫폼을 구축 중에 있다. 이를 위해 KRICT는 우선 연구데이터 수집·관리측면에서 열전소재 DB 및 실험데이터를 확보하고 계산데이터를 생성하여 기계학습 기반 열전 특성 예측 프로토콜을 개발 중이다. 이후 열전소재 데이터 수집의 표준화 기술을 확

보하고 웹기반 열전소재 데이터 플랫폼을 공개하여 기계학습 기반 열전소재 성능 예측 서비스를 제공할 예정이다.

- ④ KRICT는 또한 내부 주요사업을 통해 연 3억원을 투입하여 신약분야 데이터 플랫폼 구축사업을 진행 중이며, 연구데이터 공유·활용측면에서 기구축된 DB에 대한 공유 사업을 '18년부터 진행하고 있다. 현재 화합물 연구데이터 관련 130개 타겟에 대한 100만건의 데이터 공유가 가능하며, 화학연은 본 사업을 통해 향후 2년간 국내외 개방형/폐쇄형 데이터를 1,000만 건(국내외 개방형 화합물 라이브러리) 이상 구축할 예정이다. 이를 통해 질병, 타겟 및 유효물질을 선정, 인공지능 예측 모델을 통해 유효물질 최적화 등 인공지능 신

약개발 플랫폼 구축을 통해 신약개발 단계별 예측모델을 개발하고 시각화할 예정이다.

8) 종합

국내 국가과학기술연구회 소관 25개 정부출연연구기관 중 자체적으로 연구데이터 수집·활용 체계를 구축하고자 적극적으로 사업을 추진 중인 9개 연구기관을 살펴보았다(〈표 13〉 참조). 구축의 수준은 상이하였으나 전반적으로 최근 2~3년을 기준으로 과학계에 오픈사이언스 기조에 따른 데이터기반 연구개발이 중요해짐에 따라 여러 출연연구기관에서 연구데이터 공유정책을 수립하고, 수집·관리·공유를 위한 플랫폼 구축 사업이 활발해지고 있다.

〈표 12〉 KRICT 화합물 연구데이터 현황('18.05.기준)

활용자	타겟수	실험데이터(건)
전체	760	4,672,450
KRICT	267	2,606,714
공개 가능	130	1,026,910

* 출처: 이선경, 빅데이터·인공지능기반 신약개발 플랫폼 구축 설명자료(2018.5.29.)

〈표 13〉 연구데이터 수집·관리 중인 과학기술분야 출연연구기관('18.11.기준)

연구소명	분야	연구데이터 관리 단계				
		규정	→ 인력/조직	→ 수집	→ 관리	→ 활용
한국과학기술연구원(KIST)	기초연구 전분야	×	△	○	○	×
한국과학기술정보연구원(KISTI)	과학기술정보	×	○	○	○	○
한국생명공학연구원(KRIBB)	생명분야	×	○	○	○	○
한국지질자원연구원(KIGAM)	지질	○	○	○	○	○
한국표준과학연구원(KRIS)	국가 측정표준	×	△	×	×	×
한국한의학연구원(KIOM)	한의학	○	○	×	×	×
한국화학연구원(KRICT)	화학	×	○	○	○	○
한국해양과학기술연구원	해양	△	○	○	○	○
극지연구소	극지	○	○	○	○	○

국가과학기술위원회에서도 출연연구기관 연구데이터 공유 플랫폼 정책의 일환으로 연구데이터 협의체를 구성하였고 '18년 말 관련 가이드라인을 제정할 예정이다. 또한 KISTI주도로 데이터 리포지토리 시스템을 설계 중이며 '18년 말 프로토타입이 배포 예정인 바, 각 출연연구기관기관은 이를 적용하여 해당 분야에 특화된 데이터 리포지토리를 구축할 수 있을 것이다.

4.4 연구데이터 수집·관리 우수 모델

연구데이터 수집, 관리 모델선정에 있어서 <표 13>을 통하여 볼 때 한국지질자원연구원, 한국해양과학기술원, 극지연구소 등이 5단계 연구데이터 사이클에서 가장 선도적 역할을 하고 있다. 이는 과학기술분야 연구데이터 생산주체 및 활용주체를 고려하여 우수사례 대상을 선정하였으며 5단계 연구데이터 사이클 모든 분야에서 선도적 역할을 하고 있으며 3개 기관의 현황을 제시하면 다음과 같다.

4.4.1 한국지질자원연구원(Korea Institute of Geoscience and Mineral resources, KIGAM)

지질자원연구원(이하 KIGAM)은 출연연구기관 가운데 선도적으로 개별기관 차원에서 연구데이터 통합관리체계를 구축하여 운영을 개시한 모델로 볼 수 있다. 모범사례로 꼽히는 이유는 첫째, 연구데이터 수집·관리·공유 체계 구축을 위한 제도, 교육, 조직 등 소프트웨어 마련하고, 둘째, 육상과 해저에서 생산되는 시료를 체계적으로 수집, 저장, 보존, 활용할 수 있도록 물리적/디지털 리포지토리 구축 등 하드웨어

인프라를 마련했다는 점이다. KIGAM은 연구데이터 통합체계 구축을 위해 우선 규정, 조직, 교육안을 마련하였다. 해당 분야 데이터 특성에 맞게 연구데이터 관리규정 및 데이터관리계획(DMP)을 마련하고 '18년 말 시범적으로 시행할 예정이다. 데이터에 관한 제반 사항을 심의, 의결할 수 있는 데이터관리위원회를 구성하여 운영하고 있으며 데이터 생산자가 데이터 관리·공유 활동에 능동적으로 참여할 수 있도록 연구실적 평가제도를 개선하여 데이터관리계획 이행실적을 과제평가에 반영하도록 하였다. 또한, 안정적이면서 지속적인 데이터 관리가 가능하도록 전문인력을 확보하고 전담조직(지질자원데이터센터)을 운영하고 있다. 지질자원데이터센터는 데이터 생산자와 관리자의 데이터 관리 능력을 배양하고, 고품질의 데이터 생산이 가능하도록 국제지질자원교육센터와 연계하여 데이터 품질관리(QA/QC)에 관한 교육프로그램을 개발하여 운영하고 있다.

- ① KIGAM은 하드웨어 인프라 측면에서 시료관리·연구시설을 갖춘 물리적 리포지토리를 구축하고 연구데이터를 수집, 등록, 관리, 공유, 검색할 수 있는 오픈소스 기반 디지털 리포지토리 시스템을 개발하였다.
- ② 구축된 디지털 리포지토리 시스템이 지질자원연구와 교육분야에서 보다 활용될 수 있도록 On-line KIGAM Data Journal 발간 시스템을 통해 개인이 보관 중인 과거 데이터를 공개하여 재사용하는 등 연구데이터 공유를 위한 다양한 방안을 모색하고 있다.
- ③ 또한 국내 지질자원 데이터 리포지토리가 국제적 수준으로 인정받을 수 있도록 Thomson Reuters사의 Web of Science에서 제공하

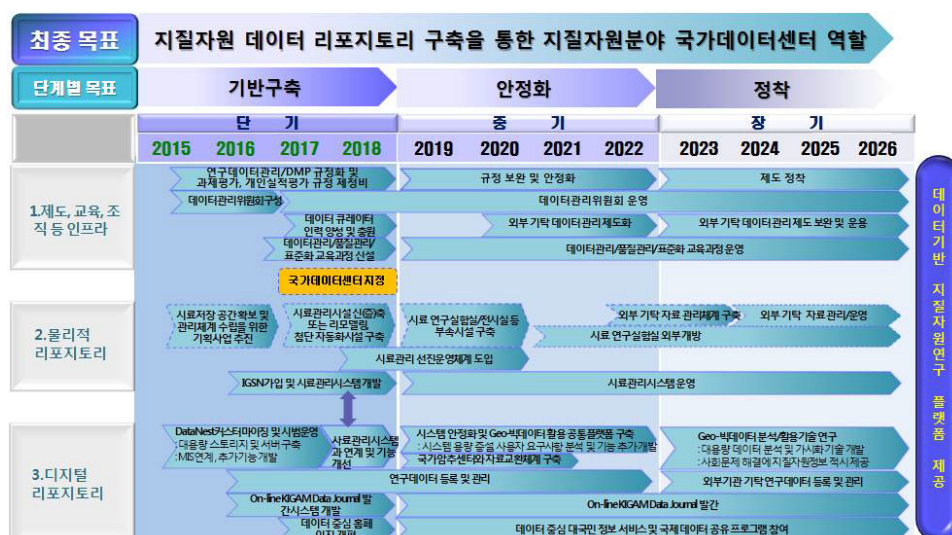
는 DCI(Data Citation Index)의 인증 리포지토리 등록을 준비 중이며, 지질자원분야 국제 데이터 공유 프로그램인 OneGeology, CCOP-GSJ, GeoInformation Sharing Infrastructure for East and Southeast Asia (GSi) 프로그램에 적극적으로 참여하는 등 지질자원분야 국가데이터센터로서 핵심 역할을 수행하고 있다.

4.4.2 한국해양과학기술원(Korea Institute of Ocean Science and Technology, KIOST)

한국해양과학기술원(이하 KIOST)은 해양조사법 제21조, 제22조 및 국제해양자료정보교환 체계(International Oceanographic Data and Information Exchange, IODE) 관련 정부간해양학위원회(Intergovernmental Oceanographic Commission, IOC) 회담 800(1980.10.02.)에 따라 기관 내 해양조사자료 관리체계를 구축하였

다. KIOST는 해양자료 관리업무의 일관성 유지를 위해 업무 전반에 대한 매뉴얼을 작성하고 해당 매뉴얼에 따라 주요사업 및 국가연구사업 생산 자료에 대한 기관 차원의 자료관리업무를 수행하고 있다. 이 외에도 매년 해양조사자료목록, 해양조사계획(NOP), 항해요약보고서(CSR)을 작성하여 국립해양조사원과 국립수산물과학원에 제공하고 있다.

① KIOST는 데이터 수집·관리를 위한 인프라로서 해양자료 통합 DB와 KIOST 해양자료 포털을 구축·운영하고 있다. 해양물리, 해양화학, 해양생물, 해양지질, 지구물리 등 해양분야 통합 자료 관리를 위한 Oracle 기반의 해양자료 DB를 구축하였고, 기관 내 해양자료와 분석정보 제공을 위한 KIOST 해양자료포털을 운영하며 지속적으로 정보를 업데이트하고 있다.



* 출처: 한국지질자원연구원, KIGAM 지질자원정보 통합관리체계 구축(2015.4.)

〈그림 4〉 지질자원연 지질자원정보 통합관리체계 중장기 로드맵

〈표 14〉 해양과학조사자료 관리를 위한 제출자료

구분	내용	근거
조사자료목록(메타)	조사자료 항목 등 메타정보	해양과학조사법 - 국립해양조사원에 제출
해양과학조사계획서	당해연도 해양조사계획	IOC 회람 800
해양과학조사보고	연구선을 이용한 해양조사 간략보고	- 국립수산물과학원 해양자료센터에 제출
조사자료	현장관측 데이터 및 메타정보	KIOST 자료관리규정

* 출처: 최현우, KIOST 해양자료실 소개 설명자료(2018.5.30.)

- ② 해양과학데이터 허브역할을 수행할 전담 조직으로 원내 해양과학데이터센터를 운영 중이며, 총 10인의 인력이 전담인력으로 종사 중이다. 해양과학데이터센터의 주요 기능은 해양과학데이터 기술 개발 및 지원, 해양공간정보 활용 기술 개발, 해양 정보시스템 구축 및 운영이다.
- ③ 이를 위해 해양데이터 수집·처리 및 관리 뿐만 아니라 데이터의 품질관리와 빅데이터 분석기술도 지원하고 있으며 상시해양관측자료 체계(KUMOS)²⁾를 구축하고 운영하는 등 전담조직으로서 해양데이터과학 기반 가치창출을 도모하고 있다.

4.4.3 극지연구소(Korea Polar Research Institute, KOPRI)

KIOST의 부설연구기관인 극지연구소는 남극 조약협약당사국(Antarctic Treaty Consultative Party, ATCP) 협약에 따라 '11년부터 극지활동을 통해 수집되는 연구데이터의 수집 및 관

리·공유를 의무적으로 시행하고 있다.

- ① 이에 따라 연구소 내 사업기획 시 DIF 포맷의 데이터관리계획서(DMP)를 포함한 연구계획서를 제출하고 사업 수행 후 DMP 대비 데이터 등록 이행률을 평가에 반영하고 있다.
- ② 극지연구소는 극지데이터센터(KPDC)를 구축·운영 중이며, 이는 전담조직이라기보다 데이터 플랫폼의 개념이다. 극지데이터센터가 수집 대상인 연구데이터는 지상·고층 기상, 전파레이더, 중력계, 지진계, GPS, 해수, 지형, 탄성과 등의 디지털 데이터와 빙하, 퇴적물, 운석, 암석, 화석, 해수, 동물 등 아날로그(실물) 데이터이다. 메타데이터는 DIF(Directory Interchange Format) 표준³⁾을 준수하고, 〈그림 5〉와 같이 3가지 등급으로 구성하여 수집·공유 중에 있다.
- ③ 극지연구소는 '12년부터 시스템화하여 극

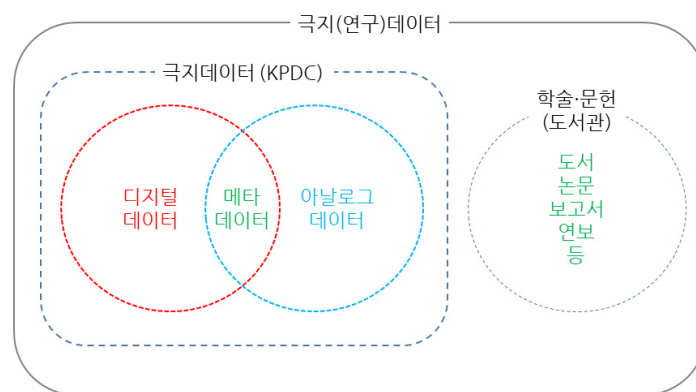
2) KUMOS(KIOST Underway Meteorological & Oceanographic System)는 내외부 서비스를 모두 지원하여, 누구나 해양관측자료에 대한 접근이 가능(내부: kumos.kioست.ac.kr/ 외부: rkumos.net)

3) DIF(Directory Interchange Format)는 메타데이터 표준으로 1987년 Earth Science and Applications Data Systems Workshop (ESADS)에서 개발되어 1991년 미국 NASA GCMD(Global Change Master Directory)에서 채택하였으며, ISO 19115에서 메타데이터 표준으로 준용하였다.

지연구데이터를 관리 중에 있으며, 현재 연평균 5회 정도(최대 10회) 연구데이터를 공유 중이다. 극지연구데이터 이용자는 KPDC 홈페이지에서 메타데이터를 검색하여 공유 요청이 가능하며, 메타데이터 요소는 <표 15>와 같다. 최근 시스템을 고도화하여 데이터 활용성을 강화하고자 지리정보시스템(GIS) 및 데이터 가시화 기법을 적용한 실시간 제공 가능 서비스(아

라운호 항해정보, 기상관측자료 등)를 개발하고 있다.

- ④ <그림 6>과 같이 극지연구소는 극지연구데이터 수집부터 활용까지 체계적인 시스템 구축을 위해 극지데이터관리위원회를 운영하여 극지자료 관련 정책을 입안하고 마련하고 있으며, 그에 따라 현업부서와 극지데이터센터 간 협업구조로 연구데이터를 수집·관리·활용하고 있다.



* 출처: 주동찬, 한국극지데이터센터 소개 설명자료(2018.5.29.)

<그림 5> 극지데이터센터 수집 대상인 극지연구데이터

<표 15> 극지데이터 메타데이터의 등급별 내용

Required	Highly Recommended	Recommended
• Entry ID • Entry Title • Science Keywords • ISO Topic Category • Organization • Summary • Related URL • Platform/Instrument • Temporal Coverage • Spatial Coverage • Project • Metadata Dates	• Dataset Citation • Personnel • Location • Data Resolution • Quality • Access Constraints • Use Constraints • Distribution Information • Dataset Progress • Dataset Language	• Originating Center • Multimedia Sample • Metadata Association • IDN Node • DIF Revision History • Version Description • Additional Attributes • Product Level Id • Collection Data Type • Extended Metadata • Ancillary Keyword • Publication/Reference

* 출처: 주동찬, 한국극지데이터센터 소개 설명자료(2018.5.29.)



* 출처: 주동찬, 한국극지데이터센터 소개 설명자료(2018.5.29.)

〈그림 6〉 극지연구소 연구데이터 수집·관리·활용 정책

5. 시사점 및 결론

출연연구기관 연구데이터 수집·관리 현황에서 살펴본 것과 같이 현재 출연연구기관의 연구데이터 수집·관리는 전반적으로 시범사업의 단계이며, 데이터의 수집 구축수준 또한 기관별로 상이하여, 수집을 시작하는 단계(KIOM), 수집·관리가 고도화되고 있는 단계(KIST), 공유·활용을 시작하는 단계(KRIBB, KRICT) 등으로 구분된다.

현황분석 및 우수사례를 통해 출연연구기관 연구데이터 수집·관리체계 구축에 적용될 수 있는 시사점은 다음과 같다.

첫째, 연구데이터 수집·관리를 위해서는 데이터 관리규정, 데이터관리계획서(DMP) 마련 등 기관차원에서의 제도화 구축이 선결적으로

필요하다. 이를 위해서 데이터관리위원회, 전담 인력·조직 등 거버넌스 차원에서의 구조적 변화가 요구되며, 기관 내에서 데이터관리의 필요성에 대한 인식 확산과 관련 예산 확보를 위한 지속적인 홍보 등의 노력이 필요하다. 특히, 출연연구기관 원장이 3~4년 주기로 바뀌는 점을 고려할 때, 기관 전체차원의 공감대 형성이 매우 중요하며 이는 국가과학기술위원회 차원의 정책적 지원과 협조가 요구된다.

둘째, 연구데이터 관리 업무를 수행할 인력은 해당 연구 분야에 대한 지식과 정보관리 및 데이터 기술에 대한 지식을 동시에 갖추어야 하므로, 두 분야에 대한 전문지식을 갖춘 인력의 양성에 힘써야 한다. 특히, 출연연구기관 대상 지속적인 데이터 관리 교육을 통해 데이터 품질관리 능력을 배양시키는 것이 필요하다.

셋째, 연구데이터를 지속적으로 공유하고 활용하기 위해서는 연구자의 적극적인 참여가 전제되는 바, 연구자 참여 독려를 위한 보상체계 등 유인책 마련이 필요하다. 가령, 데이터 인용 지수를 반영한 연구실적 평가제도 개선 등 적절한 인센티브 부여 방안을 고려할 수 있을 것이다.

넷째, 국가연구개발사업 등 공적자금을 통해 얻어진 데이터는 개인소유물이 아닌 국가자산으로 인식하여 가능한 공개함으로써 많은 분야에서 새로운 가치를 창조할 수 있도록 연구자 인식 개선이 필요하다. 여러 출연연구기관에서 조사된 연구데이터 인식조사 결과('14, KIGAM, '18 KIOM 등)에 따르면 연구데이터를 개인소유물로 인식하여 데이터 공개·공유가 잘 되지 않고 있으며 재사용성이 낮다. 이를 위해 지속적인 데이터 관리 교육을 통해 데이터 수집 관리·공유 활용의 중요성을 인식시키는 것이 필요하다.

마지막으로, 국제적 수준의 데이터 활용을 제고하지 않으면 연구데이터는 수집되더라도 활용될 수 없다는 점을 고려하여 국내 연구데이터의 국제화를 위한 질적 제고를 고려해야 한다. 이를 위해 국제표준에 부합하는 데이터 품질관리 및 표준화가 요구되며, 분야별 국제 데이터 공유 프로그램 참여를 통한 국제적 수준의 데이터 공유·활용 노력이 필요하다.

오픈사이언스정책은 기본적으로 학문의 자유

를 위한 연구 성과의 사회경제적 가치 증대를 위해 공적 자금이 투입된 연구과제의 연구 결과물을 연구자간은 물론 국민에게도 공개·공유하자는 것이다. 즉 데이터를 같이 나눠서 쓰고, 모두가 공유를 하게 되면 사회적 비용이 훨씬 더 감소할 수 있을 것이라 보는 것이다.

오픈사이언스의 구성 요소 중 하나인 연구데이터 공유·활용 체계를 위해서는 오픈데이터는 연구자의 연구결과물에 대한 재현성 확보나 재사용성을 제고해 연구의 투명성을 확보하거나 새로운 과학적 발견을 가능케 한다. 또 기후변화, 재난 등 많은 사회현안 문제를 해결하거나 융복합 연구를 수행하기 위해서도 기본적으로 연구데이터를 공유·활용해야 한다. 이를 위해 국가연구데이터센터와 과학기술 분야별 전문센터 체계를 구축하려 하는 것이다. 국가연구데이터센터는 관리체계 총괄로서 연구데이터를 원스탑으로 검색하고 활용할 수 있고, 분야별 전문센터는 분야별 데이터 수집관리와 활용촉진을 담당하게 된다. KISTI만의 축적되고 상호연계된 인프라를 통해 가깝게는 25개 출연연구기관, 대학, 연구소, 중소기업, 해외 유관그룹 등이 함께 상호작용하며, R&D 혁신과 문제해결, 혁신성장과 일자리 창출로 이어지는 데이터 생태계 중심기관으로 성장시켜 나갈 것으로 기대된다.

참 고 문 헌

- 강주연. 2017. 『생명공학분야 연구데이터 관리방안 연구』. 석사학위논문. 전북대학교대학원, 문헌정보학과.
- 김규빈. 2015. 과학데이터관리에 관한 해외법제의 비교법적 고찰. 『과학기술과 법』, 6(1): 1-38.

- 김문정, 김성희. 2015. 과학기술분야 연구자의 연구데이터 공유의 영향요인에 대한 연구. 『한국문헌정보학회지』, 49(2): 313-334.
- 김선태 외. 2010. 과학데이터 보존 및 활용모델에 관한 연구. 『한국비블리아학회지』, 21(4): 81-93.
- 김선태. 2012. 『해양관측분야 메타데이터 표준요소선정 및 설계에 관한 연구』. 박사학위논문. 전북대학교대학원, 문헌정보학과.
- 김순, 이보람, 김환민, 김혜선. 2017. 오픈 사이언스 시대를 위한 과학기술 연구지원 서비스 동향분석. 『정보관리학회지』, 34(3): 229-249.
- 김은정, 남태우. 2012. 연구데이터 수집에 영향을 미치는 요인 분석. 『정보관리학회지』, 29(2): 27-44.
- 김지현. 2011. 자연과학분야 대학실험실에서의 연구노트 작성 및 관리에 관한 연구. 『한국기록관리학회지』, 11(1): 139-159.
- 김지현. 2012. 대학 내 연구자들의 연구데이터 관리에 관한 연구. 『한국도서관·정보학회지』, 43(3): 433-455.
- 김지현. 2013. 국외 정부연구비지원기관의 연구데이터 관리정책 분석. 『한국문헌정보학회지』, 47(3): 251-274.
- 김지현. 2015. 데이터 관리와 공유에 대한 대학 연구자들의 인식에 관한 연구. 『한국도서관·정보학회지』, 47(3): 413-436.
- 김지현. 2016. 연구데이터 레포지터리의 데이터 접근 및 이용 통제 정책 요소에 관한 연구. 『한국도서관·정보학회지』, 47(3): 213-239.
- 김현희, 정경희, 김영석, 안태경. 2006. 경제사회 분야 정부 출연 기관의 오픈 액세스 기반 지식 공유 활성화 연구: IKIS를 중심으로. 『한국도서관·정보학회지』, 37(2): 145-171.
- 노서영. 2018. 『2018년도 기초연구 실험데이터 글로벌 허브 구축사업』. 서울: KIST.
- 박미영, 안인자, 김준모. 2017. 생명공학분야의 연구데이터 공유 사례에 관한 연구. 『한국비블리아학회지』, 29(1): 393-416.
- 서덕록. 2018. 『KIST 데이터기반 R&D환경 구축사업』. 서울: KIST.
- 신영란, 정연경. 2012. 국내 인문사회 연구데이터 아카이브의 개선방안에 관한 연구. 『한국기록관리학회지』, 12(3): 93-115.
- 심원식, 안혜연, 변제연. 2015. 인문학 분야 연구데이터의 수집 및 활용성 증진을 위한 전략 연구: 기초학문자료센터를 중심으로. 『한국문헌정보학회지』, 49(3): 155-183.
- 유소영. 2015. Cyberscholarship 환경에서의 인문학 분야 연구데이터 사용 패턴 및 인식에 관한 연구. 『한국문헌정보학회 학술발표논문집』, 35: 9-28.
- 윤종민, 김규빈. 2013. 과학데이터에 대한 입법례와 관리정책 그리고 대응방안: 호주, 미국, 중국을 중심으로. 『기술혁신학회지』, 16(1): 63-100.
- 이선경. 2018. 『빅데이터·인공지능기반 신약개발 플랫폼 구축』. 대전: KRICT.

- 이성호 외. 2013. 『과학데이터의 관리와 공유를 위한 제언』. 대전: 한국과학기술정보연구원.
- 주동찬. 2018. 『한국극지데이터센터 소개』. 인천: 극지연구소.
- 진태은. 2018. 『국가생명연구자원 통합시스템 구축』. 대전: KRIBB.
- 채균식. 2018. 『참조표준사업과 KRISS data activities』. 대전: KRISS.
- 최명석, 이승복, 이상환. 2017. 국내 과학기술분야 연구기관의 과학데이터 관리 현황. 『한국콘텐츠학회논문지』, 17(12): 117-126.
- 최현우. 2018. 『KIOST 해양자료실 소개』. 부산: KIOST.
- 한국지질자원연구원. 2018. 『KIGAM 지질자원정보 통합관리체계 구축』. 대전: 한국지질자원연구원.
- 한선화 외. 2011. 국가과학데이터 공유·융합 체제 구축에 관한 연구. 한국과학기술정보연구원.
- 한종규 외. 2015. 『KIGAM 지질자원정보 통합관리체계 구축』. 대전: 한국지질자원연구원.
- 한종규 외. 2016. 『KIGAM 디지털 데이터 리포지토리 시스템 인프라 구축』. 대전: 한국지질자원연구원.
- American Council on Learned Societies. 2006. *Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyber Infrastructure for the Humanities and Social Sciences*. New York: American Council on Learned Societies.
- OECD. 2016. *Open and Inclusive Collaboration in Science: A Framework, Room Document of the Global Science Forum (2016.10.13.-10.14., Paris)*, DSTI/STP/MS(2016)10.
- Swan, A. et al. 2005. "Delivery, management and access model for E-prints and open access journals within further and higher education." *Learned Publishing*, 18(1): 25-40.
- Williams, S. C. 2013. "Data sharing interviews with crop sciences faculty: why they share data and how the library can help." *Issues in Science and Technology Librarianship*, 72. [online]. [cited 2018.9.10].
 <http://www.isrl.org/13-spring/refereed2.html?utm_source=dlvr.it&utm_medium=twitter>.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Chae, Kyun-sik. 2018. *Reference standard business and KRISS data activities*. Daejeon: KRISS.
- Choi, Hyun Woo. 2018. *Introduction of KIOST maritime data center*. Busan: KIOST.
- Choi, Myung-Seok, Seung-Bock Lee, and Sanghwan Lee. 2017. "Research Data Management of Science and Technology Research Institutes in Korea." *Journal of the Korea Contents Association*, 17(12): 117-126.
- Gang, Ju-yeon. 2017. *A Study on Analysis of the Current Status and the Method to Improve*

- Biotechnology Research Data Management*. Master thesis, Dept. of Library and information Science, The Graduate School Chonbuk National University.
- Hahn, Sun-Hwa et al. 2011. *A Study on the National Scientific Data Sharing and Convergence System*. Daejeon: Korea Institute of Science and Technology Information.
- Han, Jongkyu et al. 2015. *KIGAM Establishment of integrated management system of geological resource information*. Daejeon: Korea Institute of Geoscience and Mineral Resources.
- Han, Jongkyu et al. 2016. *KIGAM Building a Digital Data Repository System Infrastructure*. Daejeon: Korea Institute of Geoscience and Mineral Resources.
- Jin, Tae Eun. 2018. *Establishment of Integrated System for National Life Research Resources*. Daejeon: KRIBB.
- Joo, Dong Chan. 2018. *Introduction of Korea Polar Data Center*. Incheon: KOPRI.
- Kim, Eun-Jeong and Tae-Woo Nam. 2012. "Factor Analysis of Effects on Research Data Collection." *Journal of the Korea Society for Information Management*, 29(2): 27-44.
- Kim, Hyun-Hee, Kyoung-Hee Joun, Young-Seok Kim, and Tae-Kyoung Ahn. 2006. "Promoting Resource Sharing among Government-Sponsored Economic and Social Research Institutes Focusing on IKIS." *Journal of Korean Library and Information Science Society*, 37(2): 145-171.
- Kim, Ji-hyun. 2011. "A study on the creation and maintenance of laboratory notebooks in scientific laboratories of a university." *Journal of Korean Society of Archives and Records Management*, 11(1): 139-159.
- Kim, Ji-hyun. 2012. "A Study on University Researchers' Data Management Practices." *Journal of Korean Library and information Science Society*, 43(3): 433-455.
- Kim, Ji-hyun. 2013. "An Analysis of Data Management Policies of Governmental Funding Agencies in the U.S., the U.K., Canada and Australia." *Journal of Korean Library and Information Science Society*, 47(3): 251-274.
- Kim, Ji-hyun. 2015. "A Study on the Perceptions of University Researchers on Data Management and Sharing." *Korean Society for Library and Information Science*, 47(3): 413-436.
- Kim, Ji-hyun. 2016. "A Study on Policy Components of Data Access and Use Controls in Research Data Repositories." *Korean Library and Information Science Society*, 47(3): 213-239.
- Kim, kyu bin. 2015. "A Comparative Legal Analysis of Research Data Management." *Science, Technology and Law*, 6(1): 1-38.
- Kim, Moonjeong and Seonghee Kim. 2015. "A Study on the Factors Affecting Sharing of Research Data of Science and Technology Researchers." *Journal of the Korean Society*

- for Library and Information Science*, 49(2): 313-334.
- Kim, Soon. Boram Lee, Hwanmin Kim, and Hyesun Kim. 2017. "HyScience and Technology Research Support Service Trends for Open Science Era." *Journal of Korean Library and Information Science Society*, 34(3): 229-249.
- Kim, Sun-Tae et al. 2010. "A Study on a Model for Using and Preserving Scientific Data." *Journal of the Korean Biblia Society for Library and Information Science*, 21(4): 81-93.
- Kim, Sun-Tae. 2012. *A Study on Extraction and Design of Standardized Elements on Metadata for Ocean Observational Dat*. Ph. D. diss. Dept. of Library and information Science, The Graduate School Chonbuk National University.
- Korea Institute of Science and Technology Information. 2018. *Document of Establishment of integrated management system of KIGAM geological resource information*. Daejeon: Korea Institute of Geoscience and Mineral Resources.
- Lee, Seong Ho et al. 2013. *Suggestions for the management and sharing of scientific data*. Daejeon: Korea Institute of Geoscience and Mineral Resources.
- Lee, Sun-kyung. 2018. *Big Data · Building a new drug development platform based on artificial intelligence*. Daejeon: KRICT.
- Noh, Seo Young. 2018. *Global hub project for basic research data for 2018*. Seoul: KIST.
- Park, Miyoung, Inja Ahn, and Junmo Kim. 2018. "A Study on Use Case of Research Data Sharing in Biotechnology." *Korean Biblia Society for Library and Information Science*, 29(1): 393-416.
- Seo, Duk Rok. 2018. *KIST Data-Based R&D Environment Construction Project*. Seoul: KIST.
- Shim, Wonsik, Hye-yeon Ahn, and Jeayeon Byun. 2015. "Strategies for Improving the Collection and Use of Research Data in the Humanities." *Journal of Korean Library and Information Science Society*, 49(3): 155-183.
- Shin, Young-Ran and Yeon-Kyoung Chung. 2012. "A study on the improvement plans of the humanities and social sciences research data archives in Korea." *Journal of Korean Society of Archives and Records Management*, 12(3): 93-115.
- Yoon, Chong-min and Kyubin Kim. 2013. "Legislation Cases, Management Policies and Countermeasures on Scientific Data: Focusing Australia, the United States and China." *Journal of Korean technology innovation society*, 16(1): 63-100.
- Yu, So-Young. 2015. "Cyberscholarship Hwangyeongeseoui Inmunhak Bunnya Yeongudeiteo Sayong Paeteon Mit Insigne Gwanhan Yeon-gu." *Korean Society for Library and Information Science Occasional Papers Series*, 35: 9-28.