

과학기술분야 연구기록의 평가에 관한 연구

이 미 영*

1. 서론
2. 연구기록의 범위와 특징
3. 연구기록의 평가
 - 1) 평가의 주체
 - 2) 평가기준
 - 3) 평가방식
4. 결론 및 제언

* 지질연구원, 한국지질자원연구원 기록물관리전문요원 및 사서, 주요 논저 : 「재해 지역의 기록화 사례 연구 : 한신·아와지 대지진 기록관리 사례를 중심으로」 『기록학연구』 제21호, 2009; 「지질분야 연구기관의 연구기록 관리 사례」 『제5회 전국기록인대회 발표자료집』, 2013.

▪투고일 : 2014년 6월 26일 ▪최초심사일 : 2014년 7월 1일 ▪게재확정일 : 2014년 7월 21일.

[국문초록]

연구데이터의 양적 증가와 함께 막대한 보존비용, 공유 확대 등의 이슈는 이제 조직에게 그들이 무엇을 수집할지 우선순위를 정하고 가치있는 기록만을 보존할 것을 요구할 것이다. 따라서 생산된 직접적인 목적을 넘어서는 계속적 가치를 확인하는 기록의 평가는 현 시기에 더욱 중요하다. 그러나, 연구기록 생산기관으로서 대학과 정부출연연구소와 같은 공공기관의 비중이 높다 보니, 연구기록관리 전반에 대한 고민이 ‘공공기록’과 ‘행정기록’의 틀을 넘어서지 못하고 있다. 따라서 본 연구에서는 연구기록이 행정기록과는 다른 관점으로 관리되어야 한다는 주장의 배경과 생산자 측면, 연구 행위의 측면, 기록의 측면에서 과학기술분야 연구기록의 특성을 살펴보고, 이를 바탕으로 연구기록 평가의 주제, 기준, 방식에 있어서의 이슈와 고려사항을 제시하였다.

주제어 : 연구기록, 연구기록물, 연구데이터, 과학데이터, 과학연구데이터, 실험실데이터, 평가, 기록 평가, 연구기록평가, 평가기준, 평가방식, 연구기록관리

1. 서론

디지털환경을 기반으로 데이터의 폭발적 증가를 경험하는 오늘날을 과학의 제4패러다임이라고 일컫는다. 그리고 제4패러다임은 ‘데이터 중심 과학’을 특징으로 한다.¹⁾ EMC가 시장조사 기관인 IDC에 의뢰해서 만든 보고서에 따르면 2012년에 만들어진 전 세계 데이터의 양은 2.8제

1) Tony Hey, Stewart Tansley, Kristin Tolle, The fourth paradigm, Microsoft Research: Washington, 2009, p.xviii. <http://research.microsoft.com/en-us/collaboration/fourth-paradigm/4th_paradigm_book_complete_lr.pdf>

타바이트(ZB)에 이르며, 2020년에는 약 40ZB까지 커질 것이라고 예측했다. 전 세계 해변에 있는 모래알 수가 7억 50만조개라고 추정하는데, 적어도 2020년이 되면 전 세계 해변에 있는 모래알의 수보다 약 57배나 더 많은 데이터가 시중에 유통될 전망이다.²⁾ 연구분야, 특히 과학기술분야도 예외는 아니어서 실험, 관측, 모니터링 등의 연구활동을 통해 정기적으로 상당한 양의 데이터세트를 생산하거나 활용, 분석하고 있다.

그러나, 이러한 상황에도 불구하고 인식부족, 예산부족, 전문인력부족 등의 이유로 중요한 연구데이터가 사라지고 있다는 지적이 더 많은 듯하다. 즉 생산의 패러다임 변화는 명확하나, 관리의 패러다임 변화는 그것을 못 쫓고 있는 것이다.

생물학연구정보센터 BRIC이 899명의 응답의견을 토대로 발표한 조사 결과에 따르면³⁾, 연구실에서 생산되는 실험데이터를 관리자를 정해 관리한다는 응답은 4%에 불과한 반면, 각자 알아서 개별적으로 관리한다는 응답은 61%에 달했고, 연구실에서 생산되는 모든 실험데이터가 연구실 차원에서 잘 관리되는가라는 질문에, 관리되지 않고 있다가 21%를 차지했다. 과학분야 대표 저널인 『Science』도 “잘 만들어진 데이터 아카이브를 가지고 있는 연구분야 조차도 데이터 양과 관리에 있어 새로운 도전에 직면하고 있다”고 지적하면서, 1,700명이 참여한 데이터 활용가능성에 대한 조사에서 응답자의 절반(50.2%)이 연구를 위해 생산된 대부분의 데이터는 오직 연구실에 저장된다고 응답했다.⁴⁾

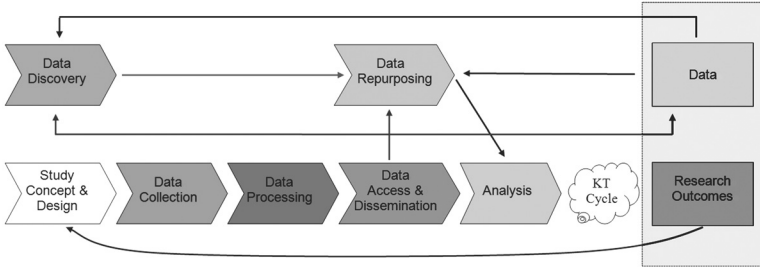
Charles Humphrey는 <그림 1>에서 연구로부터 지식이 생성되는 라이프 사이클의 각 단계를 제시하였는데, 각 상자간의 간격은 각 단계들 간의 전환(transition)을 상징하며 연구단계가 전환될 때마다 데이터의 손실이 발생함으로 각 단계에서 생성된 디지털 객체를 누가 책임지고 있는가를 확인하는 것이 중요하다고 하였다.⁵⁾

2) EMC, 「2020년 전 세계 데이터 양은 40제타 바이트가 될 것」 [cited 2014.6.20.] <<http://hallet.tistory.com/659>>.

3) BRIC, 「연구실 실험데이터 관리 실태조사」, 『SciON Survey Report』 No.192, 2009, 4쪽.

4) Science staff, “Challenges and opportunities” Science Vol.331, No.6018, 2011, p.692.

〈그림 1〉 연구지식 생산의 라이프 사이클 모델



데이터관리를 지원하기 위해 진행된 한 연구도 궁극적으로 데이터를 관리할 담당자와 장소를 요구하지 않는 “마법의 탄환(magic bullet)”은 없음을 확인했다.⁵⁾ 가령 3년간의 연구프로젝트에서 수집된 데이터는 프로젝트 기간 동안만이 아니라, 프로젝트 이후 십년 또는 그 이상 다른 연구자나 공동연구자들에게 유용하다. 그러나, 처음 3년 동안 관리 비용은 보장되지만, 지속적인 지원은 이뤄지지 않고 있는 것이 현실이다.

이렇듯 연구기록의 생산단계부터 일련의 기록관리 프로세스에는 해결해야 할 이슈들이 산재해 있으나 연구기록과 관련한 국내외 연구에서 연구자들의 생산 프로세스, 기록 유형, 관리방식 등에 관한 심도깊은 분석은 많이 진행되지 않고 있다.

연구과정에서 기록의 생산과 유지는 필수적이다. 기록은 연구가 실행되었음을 증명하고 연구의 증거로써 신뢰성을 높일 수 있으며 연구 부정행위 혐의로부터 연구자와 연구기관을 보호할 수 있고, 개인과 기관의 지적 재산권을 보호함은 물론 법률과 규정 및 다른 요건을 준수했음을 증명할 수 있다. 그리고 많이 인용되거나 활용되는 경우에는 연구

5) Charles Humphrey, “e-Science and the Life Cycle of Research”. [cited 2014,6,21.] <http://datalib.library.ualberta.ca/~humphrey/lifecycle-science060308.doc>

6) Francine Berman and Vint Cerf, “Who will pay for public access to research data?” Science Vol.341, 2013, p.616.

자의 권위를 높이는 역할도 할 수 있다.

앞서 지적한 연구데이터의 양적 증가와 함께 막대한 보존비용, 공유 확대 등의 이슈는 이제 조직에게 그들이 무엇을 수집할지 우선순위를 정하고 가치있는 기록만을 보존할 것을 요구할 것이다. 따라서 생산된 직접적인 목적을 넘어서는 계속적 가치를 확인하는 기록의 평가는 현 시기에 더욱 중요하다. 더군다나 연구기록관리에 관한 별도의 법률없이 『공공기록물 관리에 관한 법률』 및 관련 규정만이 존재하는 상황에서 연구기록 생산기관으로서 대학과 정부출연연구소와 같은 공공기관의 비중이 높다보니, 연구기록관리 전반에 대한 고민이 ‘공공기록’과 ‘행정기록’의 틀을 넘어서지 못하고 있다는 점도 간과해서는 안 될 이슈이다.

따라서 본 연구에서는 연구기록이 행정기록과는 다른 관점으로 관리되어야 한다는 주장의 배경을 살펴보고, 특히 과학기술분야 연구기록의 특성을 고려한 평가의 방향을 고민해 보고자 한다. 이를 위해 문헌 및 사례분석 방법을 사용하였다. 연구기록관리와 관련한 일련의 역사적 발전 배경과 이슈는 기록학 분야 및 과학분야 문헌을 통해 분석하였고, 연구기록 평가와 관련해서는 해외 대학 및 연구기관의 사례를 중심으로 살펴보았다.

2. 연구기록⁸⁾의 범위와 특징

우선, 연구기록의 평가를 얘기하기에 앞서 연구기록의 정의와 범위,

7) 한국기록학회, 『기록학용어 사전』, 역사비평사, 2008, 265쪽.

8) 연구기록은 research records, records of science, 연구기록물로, 연구데이터는 scientific data, research data, 과학데이터, 연구데이터, 과학연구데이터, 실험실데이터 등으로 연구마다 표기가 다르나, 본 논문에서는 연구기록과 연구데이터를 사용하고 해외 논문의 인용인 경우 본래의 용어를 병기하였다.

특징에 관해 살펴보고자 한다.

정보학분야 연구자인 Kalpana Shankar의 1999년 논문에는 연구데이터와 연구기록 간의 흥미로운 시각이 존재한다.⁹⁾ 그 내용은 실험이 수행된 조건에 관계없이 컴퓨터 또는 다른 기기장치로부터 출력 형태의 원시 데이터(raw data)는 일반적으로 활동의 기록으로써 유용하지 않으나, 이 데이터가 실험노트에 포함되어 있다면 그것은 연구 수행 기록의 일부를 구성한다는 것이다. 이는 연구기록 생산의 특수성과 다양한 기록 유형을 이해하기보다 관리자의 입장에서의 방법적인 접근이라 유추해볼 수 있다. Kalpana Shankar 스스로도 1990년대의 문제는 연구기관, 특히 학문기관에 있어 디지털 기록의 역할은 초창기이며 완전히 이해되지도 않는다고 인정하였다. 더욱이 연구활동의 디지털 기록은 아키비스트에게는 문제로 남아 있으며 보다 큰 몇몇 연구 아카이브는 전자기록을 다루길 꺼리한다고 밝히면서 아키비스트는 종이기록이 점차 연구자들의 연구수행 방법 중 오직 일부분이 되어 가고 있음을 명심할 필요가 있다고 지적했다.

전자기록이 기록관리자에게 결코 쉬운 대상이었던 1990년대의 상황은 20년이 지난 지금도 크게 달라지지 않은 듯하나, 데이터 중심의 연구환경에서 다양한 형태의 디지털 연구기록을 관리해야 한다는 것은 더 이상 피할 수 없는 현실이 되었다.

원시 데이터의 수집은 대부분 컴퓨터로만 수행되고 분석결과는 소프트웨어 패키지로 쉽게 처리되며 실험실 또는 필드 노트는 전자적 형태로 관리되기도 한다. 이런 연구환경에서 기록과 데이터는 사전적 개념을 넘어 중첩되기도 하고 구분되기도 한다. 따라서 연구기록의 개념은 국제적으로 국가적으로 기관들마다 매우 다양하다. 연구기록관리에 대

9) Kalpana Shankar, "Understanding the record-keeping practices of scientists" Technical leaflet series, No.10, Mid-Atlantic Regional Archives Conference: Carlisle, 1999. [cited 2014.6.24.] <http://www.marac.info/assets/documents/marac_technical_leaflet_10.pdf>

한 역사적, 법률적 배경과 전통이 다르기 때문에 기록을 관리하는 방법에 영향을 미치는 특정 부분의 내부적 규칙까지도 서로 다른 이해가 있었고, 기록관리 실무를 하면서 이는 다양하게 적용될 수밖에 없었기 때문이다.

〈표 1〉 연구기록의 정의¹⁰⁾

국가 or 기관	연구기록의 정의
오스트리아/ 벨기에	연구 프로젝트 중에 생산된 기록
하이파(Haifa) 대학 (이스라엘)	특정 연구자와 관련된 기록은 물론, 연구 프로젝트 및 전체 연구활동(research activity) 중에 생산된 기록
미시건 대학	연구자의 개인 문서(private paper)도 포함
일리노이 대학	특정 연구자 및 개인 문서(private paper)를 포함하는 모든 기록
브라질	연구 프로젝트 중에 생산된 기록뿐 아니라 개인 문서를 포함한 특정 연구자의 기록
폴란드, 체코, 아이슬란드, 스웨덴	전체 연구활동 중에 생산된 기록
호주	연구활동 및 문화적 배경의 결과이자 조직의 기록까지도 포함하는 것으로 연구기록의 의미를 확대
캐나다	연구활동과 관련된 모든 기록

〈표 1〉의 정의를 살펴보면, 연구프로젝트 중에 생산된 기록으로 보는 시각에서부터 연구자 개인의 기록을 포함하여 전체 연구활동 중에 생산된 기록으로 보는 시각까지 그 스펙트럼이 다양하다. 연구기록과 관련한 국내 연구에서도 ‘연구활동과 관련된 모든 기록’이라는 큰 범주로 대부분 수렴되고 있다.

기록에 대한 기록관리자와 생산자들의 이해가 차이를 확인했던 InterPARES²¹¹⁾는 연구기록의 범위를 ‘연구과정과 그 결과에 대한 정확한

10) Renata Arovelius, “Archives of science: An international perspective and comparison on best practices for handling of scientific records”, 2003, p.2. [cited 2014.6.24.] <<http://www.library.illinois.edu/ica-suv/SciIntPers.pdf>>

11) Luciana Duranti and Randy Preston, International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic

기록이 생산되도록 확실하게 보장하는 것으로 넓게 정하고, “특정한 주제(topic)나 분야와 관련된 데이터세트에 모아진 미가공 형태의 데이터, 숫자, 그래프 또는 텍스트일 수 있으며 화학의 복잡한 구조, 결정학 데이터, 유전자 배열 등의 논리구조를 포함한다”고 정의한 뒤, 무엇보다 연구기록에 관한 정의에서 강조하고 싶은 것은 연구목적에 의해 모아지고 이용된 데이터와 이들의 다양한 집합에 있다고 정리하였다. InterPARES2의 중요한 성과는 대부분의 연구자들이 기록학의 시각으로 기록을 보는 것이 아니라 연구기록과 연구데이터를 동일시하는 경향이 있으며, 때론 연구데이터기록이란 용어를 사용하고 있음을 확인하였다는 점이다.

스웨덴은 기록, 데이터, 자료로 연구기록을 구분하고 있는데 ‘데이터’는 연구과정에서 관측, 실험 등의 과정과 결과를 기록한 것이고, 자료는 연구단계별 결과물로 보고 있다. 이는 해외 연구기관들처럼 아카이브, 도서관, 데이터센터로 관리주체를 구분하고 그에 따라 관리대상을 특정할 때 좋은 참고가 될 수 있을 듯하다.

〈표 2〉 스웨덴의 사례-연구기록(records of science)의 4가지 주요 그룹¹²⁾

구분	유형
행정기록 (Administrative records)	프로젝트 계획과 기술(descriptions), 자금운용, 계약, 후원기관과의 서신 등
원시 데이터 (Raw data)	연구 진행을 위해 사용된 모든 정보 - 서베이, 실험실 및 필드 노트, X-ray 사진, 토양 및 혈액 샘플 또는 코어 시료
분석 데이터 (Analyzed data)	보고서 초안, 발췌 부분, 데이터 처리과정의 일부로써 계산 기록 또는 전자 기록과 같은 “작업용(working)” 데이터
모든 단계별 결과에 관한 보고 자료 (All material)	중간보고서, 최종보고서, 출판물, 기사

Records, Associazione Nazionale Archivistica Italiana: Padova, 2008. [cited 2014,6,15.]
 〈http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_complete.pdf〉

12) R. Arovelius et al., Management and preservation of scientific records and data, International Council on Archives: Paris, 2010, pp.11-12.

영국의 JISC¹³⁾는 연구 프로세스와 관련된 연구기록을 4가지 범주로 구분하고 있다. 첫 번째는, 연구 프로세스를 기록화(documenting)한 것으로 연구 프로토콜, 어플리케이션이 이에 속하며, 두 번째는 연구 성과 또는 결과물을 기록한 것으로 대표적 유형은 기술보고서, 단행본이다. 세 번째는 연구프로세스/프로젝트의 관리를 기록한 것으로 예산 신청서, 계약서, 인보이스, 직원 근태기록 등이다. 네 번째는 ‘가공되지 않거나(raw) ‘분석된(analyzed)’ 형태, 양쪽 모두의 연구데이터로, 노트, 질문지, 오디오/비디오 기록, 사진, 계기 판독, DB, 샘플이 이에 속한다. JISC의 2~4번째 특징은 스웨덴과 유사하며 연구 프로세스를 기록하는 프로토콜이 추가된 점이 특징적이다.

미국 국립지질조사소(United States Geological Survey, 이하 USGS)¹⁴⁾는 연구기록(research records)을 “연구데이터는 물론, 프로젝트 개발, 시연, 배포, 평가, 테스트 및 관련 업무를 포함한 연구의 모든 측면을 나타내는 것으로, 연구자들(연구 책임자, 섹션별 책임자, 연구자, 기술자, 엔지니어, 컴퓨터 전문가 등)이 연구기록을 생산한다”고 정의하고 있다. 또한 연구기록은 “연구 프로젝트의 계획, 역사, 결과, 그리고 성과를 기록화해야 하며, 이 부분에서 기록과 데이터가 구분되기도 하고 중첩되기도 하는데, 둘 다 연구의 과정과 결과를 증명하는 것이지만 구분하자면 데이터는 별도의 특정한 시스템으로부터 생산된 것을 의미한다고 볼 수 있다”고 정리하였다.

연구기록에 관한 많은 정의들이 생산주체를 명확하게 드러내지 않는 데 반해, USGS의 정의에서는 연구기록의 생산주체를 직접적으로 연구

13) ISC/JISC infoNet, HEI records management: Guidance on managing research records, 2007, pp.3-5. [cited 2014.6.14.] <<http://tools.jiscinfonet.ac.uk/downloads/bcs-rs/managing-research-records.pdf>>

14) USGS, U.S. Geological Survey Manual: Geology discipline research records schedule 432-1-S5, 2009. [cited 2014.6.22.] <<http://www.usgs.gov/usgs-manual/schedule/432-1-s5/gd.html#sked>>

를 수행하는 연구그룹으로 지시한 것이 눈에 띈다. 앞서 지적했듯이 연구기록이 ‘연구활동과 관련된 모든 기록’이라는 큰 범주로 정의되다 보니, 이와 관련된 다양한 주체들이 연구기록 생산자의 범주에 포함되는 것으로 이해될 수 있었으나, USGS에서 연구기록은 바로 ‘연구자’들이 생산한 기록임을 명확히 하고 있다. 또한 USGS는 연구기록의 처분일정표(records schedule)를 연방기록(일반 행정기록)의 그것과는 별도로 구분하고 있는데, 이 처분일정표의 기록은 물론 연구자들이 연구과정에서 생산한 기록을 의미한다.

일반적으로 연구 프로세스에는 계획, 실험, 분석, 외연화(배포, 공유)가 포함되는데 그 과정에서 연구수행을 둘러싼 다양한 행정기록이 포함될 수는 있으나, 우리는 이를 연구과정을 그대로 지시하고 기록한 연구기록과는 좀 더 명확하게 구분할 필요가 있다.

이미 연구의 중심 토대가 된 연구데이터의 정의에 대해서도 좀 더 살펴보면, Nil Beagrie(2009)는 KRDS2 연구에서 연구데이터(research data)를 “연구자들(academic researchers)이 연구를 수행하기 위해 사용하거나 연구의 증거기록으로써 제공할 수 있는, 어떤 학문 또는 정보자료로부터의 구조화된 디지털 데이터의 집합(collections of structured digital data)”으로 정의하고 있다.¹⁵⁾

I2S2 프로젝트(Infrastructure for integration in structural science)¹⁶⁾에서는 “연구 실험으로부터 생산되거나 수집된 가공되지 않은 이미지와 숫자로 된 데이터세트뿐만 아니라, 그러한 데이터들과 연결된 보다 광범위한 정보까지도 고려해야 한다. 연구데이터에는 연구 및 실험 제안서

15) Neil Beagrie, Brian Lavoie, Matthew Woollard, Keeping research data safe 2: KRDS2 data survey-selection criteria, 2009, p.2. [cited 2014.6.15.] <http://www.data-archive.ac.uk/media/1687/KRDS2_finalreport.pdf>

16) Patel, M, I2S2 idealised scientific research activity lifecycle model, 2011, pp.1-2. [cited 2014.6.15.] <http://opus.bath.ac.uk/35186/1/I2S2_ResearchActivityLifecycleModel_110407.pdf>

(proposal), 같은 분야 전문가들의 검토(peer-review) 결과, 실험 노트, 실험 장비 구성 환경과 측정결과, 프로세싱 소프트웨어와 관련된 제어변수, 위키(wikis), 블로그, 맥락과 출처 등에 대한 메타데이터, 그리고 연구데이터(의미론적)에 대한 해석과 이해를 위해 필요한 그밖에 다른 도큐멘테이션뿐만 아니라, 실험에 수반되는 행정 및 보안데이터도 포함한다”고 정의하고 있어 연구의 과정과 결과를 입증할 수 있는 모든 정보를 광범위하게 포괄하고 있다.

호주의 ANDS가이드¹⁷⁾는 연구데이터(research data)에 대한 다양한 정의들이 인용되고 있는데 ANDS는 연구기관마다 이미 다른 형태의 데이터를 관리하고 있음을 지적하며 무엇이 연구데이터가 아닌지의 측면에서 고려하였다. 즉, 행정데이터(administrative data)는 급여대상자 명단, 연구평가 등의 기록으로 이뤄져 있고 몇몇 행정데이터는 연구프로젝트와 관련하여 연구데이터로써 취급될 필요도 있으나 거의 대부분은 데이터 관리 정책, 절차, 전략 측면에서 기관내에서 별개로 취급된다. 연구출판물(research publications)은 데이터로 간주될 수 있으나 대부분의 경우에 이것은 외부 기관이나 출판업자 등에 의해 잘 관리되고 있으며 심지어 기관 내에 소장된 경우에도 이것은 다른 연구데이터와는 별개로 관리되는 경향이 있는 것으로 보았다. 이와 함께 연구데이터는 입증과 재활용을 위해 무엇이 필요한가에 초점을 맞추는 것이 유용하다는 지적도 있었다. 이 같은 ANDS의 연구데이터 구분과 정의 역시 앞선 사례와 유사하다고 볼 수 있다.

17) ANDS, ANDS Guide: What is research data?, 2011, [cited 2014.6.15.] <<http://ands.org.au/guides/what-is-research-data.html>>

〈표 3〉 호주 대학의 연구데이터(Research Data)의 정의

기관	정의
Queensland Univ.	연구데이터는 주장, 이론, 테스트나 가설, 또는 다른 연구 성과의 기반이 되는 사실, 관측, 이미지, 컴퓨터 프로그램의 결과, 기록, 측정 또는 경험 형태의 데이터를 의미. 데이터는 숫자로 되어 있거나, 수치적, 서술적, 시각적 또는 촉각적인 형태. 그것은 원시 데이터이거나 클리닝된 데이터이거나, 처리된 데이터이며, 어떠한 포맷 또는 매체에 보관됨
Melbourne Univ.	연구데이터는 주장, 이론 또는 테스트의 기초가 되는 사실, 관찰 또는 경험을 의미. 데이터는 숫자로 되어 있거나, 서술적이거나 시청각적. 데이터는 원시 데이터이거나, 분석, 실험, 또는 관측데이터임. 데이터는 연구노트, Field 노트, 1차 연구데이터(하드카피 또는 컴퓨터 가독 형 형태의 연구데이터 포함), 질문지, 오디오테이프, 비디오테이프, 모델, 사진, 필름, 테스트 검사 등을 포함. 연구 컬렉션(Research Collections)은 슬라이드, 박물관류(artifacts), 표본, 샘플을 포함. 데이터에 대한 출처 정보는 어떻게, 언제, 어디서 그것이 수집되었고, 무엇과 수집되었는지에 대한 것을 포함. 데이터를 생성하고, 주석을 달거나 분석하기 위해 사용된 소프트웨어 코드 또한 포함될 것임
Monash Univ.	연구데이터는 그것의 내용이나 형태와는 상관없이 데이터이자 기록이자, 파일이자, 증거로 1차 자료와 분석 데이터를 포함하여 연구관찰, 발견 또는 결과로 구성됨
Griffith Univ.	연구데이터는 숫자, 상징, 텍스트, 이미지 또는 사운드 형태의 사실정보로 연구의 1차 자료로 사용되며, 연구 커뮤니티에서 공통적으로 연구 결과를 입증하는 데 필요한 것으로 받아들여 짐

* 출처 : ANDS Guide.

NIH의 경우 ‘최종연구데이터(final research data)’라는 명칭으로 “공유 대상이 되는 연구 커뮤니티에서 공통적으로 받아들여지는 사실적인 자료를 기록한 것이며 이는 연구 결과를 기록, 지원, 확인하는 데 필요한 내용을 의미한다”고 규정하고 있다. 영국 STFC(Science and Technology Facilities Council, 과학기술시설위원회)의 정책에서 데이터는 원시 데이터(raw data), 파생 데이터(derived data), 출판된 데이터(published data)의 세 가지 유형으로 구분된다. 원시 데이터는 실험이나 측정, 관측의 결과로서 직접적으로 생성된 데이터를 의미한다. 파생 데이터는 데이터의 양을 줄이거나 물리적으로 의미 있는 좌표계(coordinate system)로

변환하는 과정에서 특정 표준이나 절차의 조건 하에서 생성된 데이터를 말한다. 출판된 데이터는 출판물에 게시되거나 인용된 데이터로서 연구 결과를 도출하는 데 근거로 사용된 데이터를 말한다.¹⁸⁾

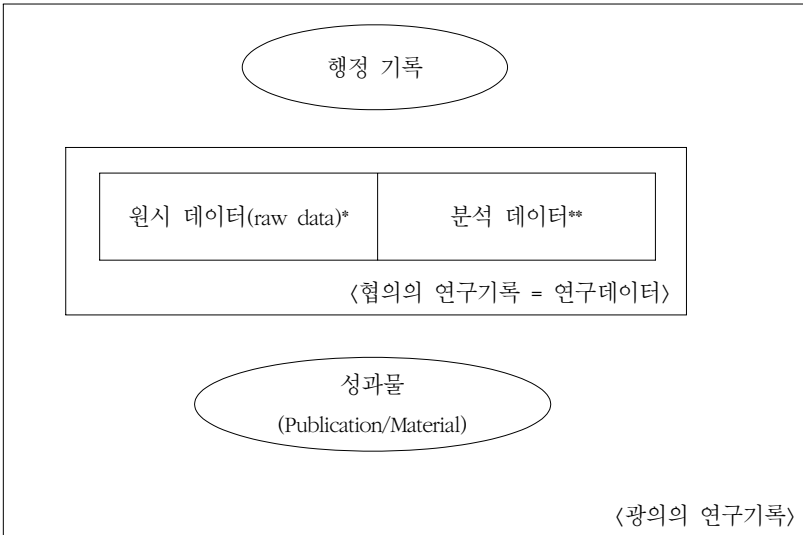
OECD의 『원칙과 가이드라인』¹⁹⁾에서는 연구데이터를 과학연구의 1차 자료로, 수치 스코어, 문자기록, 이미지, 사운드 등의 형태로 된 사실에 기반을 둔 기록(factual records)이자, 일반적으로 과학연구집단에서 연구 발견을 입증하기 위해 필요하다고 인정된 것으로 보고 있다. 이와 함께 연구데이터에는 실험실 노트, 1차 분석, 연구 논문의 초안, 향후 연구 계획, 전문가 검토(peer review) 또는 동료와의 개인적 연락, 물리적 객체(실험 샘플, 박테리아 종, 쥐와 같은 실험용 동물) 등은 포함하지 않는다고 구분하였다. 즉 OECD는 연구발견을 입증하는 기록에만 초점을 맞추고 그 외의 성과물이나 결과물에 대한 접근은 또 다른 고려사항을 다뤄야 하는 문제로 본 것이며, 이런 구분은 일부 국가에서도 적용되고 있다.

다양한 정의를 바탕으로 연구기록의 범위를 광의와 협의로 구분하면 아래 <그림 2>와 같이 정리할 수 있다. 우리가 소위 연구기록이라고 하는 범주에는 연구 관리와 지원을 위한 행정기록과 연구단계별 각종 성과물, 그리고 연구수행과정에서 나온 원시 데이터와 분석 데이터를 모두 포함하나, 협의의 연구기록인 연구데이터는 분석, 실험, 관측 등의 연구과정을 그대로 지시하고 기록한 것으로만 한정하여 1차 자료인 원시 데이터와 이로부터 파생된 분석 데이터를 포함한다.

18) 김지현, 「국의 정부연구비지원기관의 연구데이터 관리정책 분석: 미국, 영국, 캐나다, 호주를 중심으로」, 『한국문헌정보학회지』 47(3), 한국문헌정보학회, 2013, 257-258쪽.

19) OECD, OECD principles and guidelines for access to research data from public funding, 2007, pp.9-10. [cited 2014.6.24.] <<http://www.oecd.org/science/sci-tech/38500813.pdf>> ※ OECD 원칙과 가이드라인은 공개적으로 접근가능한 지식을 생산하기 위한 목적으로 공공의 기금을 사용하여 수집된 연구데이터에 적용하기 위해 만들어 졌다.

〈그림 2〉 연구기록의 범위



* 원시 데이터

- 실험, 조사, 분석, 관측, 탐사, 측정 등 연구가 진행되는 동안 획득한 원래의 자료
- 새로운 형태로의 전환이나 가공이 되기 전 최초의 형태를 지닌 자료(원천자료)

** 분석 데이터: 원시 데이터를 처리(작업)한 데이터

그러나 이와 같은 정리에도 불구하고, 연구분야의 특수성, 연구기록의 생산 프로세스와 기록유형 등에 따라 기관마다 분야마다 연구기록의 정의와 범위가 다를 수 있음은 인정되어야 한다.

지금까지의 내용을 토대로 생산자 측면, 연구 행위의 측면, 기록의 측면에서 연구기록의 특징과 시사점을 정리하면 다음과 같다. 먼저, 생산자의 측면이다. 앞서 정의에서 살펴보았듯이, 연구기록의 생산자는 연구자이다. 2006년 연방 연구개발기관을 대상으로 한 NARA의 연구²⁰⁾는 생산자의 특징을 확인할 수 있는 예이다. 이 연구에서는 실험

실에 있는 연구자와 엔지니어들 스스로가 적절하다고 생각하는 그들 자신만의 기록관리 역사를 가지고 있음을 확인했다. 한 예로 은퇴하거나 다른 직장으로 옮길 때 연구자들은 기록도 함께 가져갔다. 그 이유 중 하나는 연구자와 엔지니어들은 그들의 업무를 단지 사무실에서만 하는 것으로 보지 않았기 때문이다. 많은 이들은 자신의 업무를 자신의 '삶'으로 생각하고 스스로가 생산한 기록을 삶의 일부라고 생각했다. 때때로 연구자들은 특정 프로젝트를 수행하는데 여러 해를 보내는데 자신이 생산한 기록을 은퇴할 때 연방정부에게로 넘겨주고 싶어 하지 않았다. 그 결과, 기록은 사무실에서 사라지고, 때로는 연구자들의 다락, 차고, 집에서 운명이 다하거나 연구자들이 민간 연구기관, 교육기관, 또는 다른 정부기관으로 옮겨 다음 일을 할 때 활용된다. 실험실과 사무실 밖으로 나온 연구기록은 손상, 손실, 오용 등을 경험하게 된다.

이처럼 연구기록은 행정기록과는 달리 업무공간을 떠나서도 생산 관리되고 연구자들의 일상에 자연스럽게 함께 존재하곤 한다. 기록의 생산자가 기록에 대해 지나친 애정을 가진 연구자들이라는 점은 연구 기록관리와 관련한 일련의 프로세스에 연구자가 반드시 관여해야 한다는 명분을 갖게 하지만, 기록에 대한 소유권의 문제와 함께 기록관리자와의 역할구분에 있어 갈등의 소지가 발생할 수 있다는 우려 역시 갖게 한다. 사실 연구자는 과거에도 기록의 생산자이자 관리이자 이용자였고 지금도 마찬가지이다. 연구기록이 기록관리자가 얘기하는 관리의 영역으로 제대로 넘어오지 않은 이유는 계속해서 다음 연구에 활용해야 한다는 연구자들의 주장과 연구자들의 이런 특성에 기인한다고 볼 수 있다. 따라서 연구자들을 연구기록에 대한 전문가

20) National Archives and Records Administration, Scientific records: Preserving the past to protect the future, 2011, p.14. [cited 2014.6.22.] (http://library.bldrdoc.gov/NARA_Pres.ppt)

로 인정하고 기록의 관리활동을 직접 수행하게 하거나 적극적으로 개입시키되 기록관리자와 시너지를 낼 수 있는 협력방법을 모색해야 한다.

두 번째, 행위의 측면에서 연구기록은 행정업무의 과정이 아니라, 연구활동 과정에서 생산된다는 점이다. 연구란 ‘과학적인 방법을 비롯하여 구체적이고 논리적인 체계와 과정을 통해 증거를 얻고자 하는 탐구 과정’이다. 따라서 연구과정과 결과를 입증하기 위해 연구자체를 기록(document)한 것과 연구과정을 지원한 행정기록, 그리고 연구의 결과물은 관리측면에서 구분될 필요가 있다.

세 번째, 기록의 측면이다. 연구기록보다 연구데이터란 용어가 범람하고, 데이터 중심의 과학이라 일컫는 지금, 우리는 이미 기록의 속성 변화에 어떻게 적응할 것인지에 대한 질문을 받았다. 현재 데이터는 연구기록으로써 인식되고 관리되기 보다는 데이터 그 자체로써 생산시스템에 그대로 저장되거나 특정 저장소(repository)에 보존되고 있다. 데이터가 연구의 과정과 결과를 입증하는 열쇠라는 사실이 불변하는 한 우리는 데이터의 라이프 사이클 관리를 위한 여러 장치들의 변화에 대응해야 한다. 그래야만 정확성, 신뢰성, 진본성 등 기록의 속성을 보장해 줄 수 있다.

오늘날 많은 원시 데이터는 컴퓨터에 입력되자마자 전자적으로 획득되고 처리된다. 그리고 처리된 데이터는 연구자가 검토하기 위해 필요한 만큼만 존재하기도 하고, 연구자에 의해 실시간으로 폐기되기도 한다. 지금까지는 이렇듯 새롭고 다양하게 등장하는 연구기록의 특징에 대해 기록관리자로서 적절하게 대응하지 못했음을 인정해야만 한다. 따라서 기술적, 환경적으로 빠른 변화상황에서도 변하지 않는 기록의 가치를 확인하는 ‘평가’는 더욱 중요할 수밖에 없다.

3. 연구기록의 평가

어떤 종류의 기록을 이용자가 필요로 하고 따라서 관리 보존되어야 하는가의 문제는 데이터의 폭발적 증가와 함께 더욱 더 질적 관리의 문제를 야기했다. 기록의 양이 증가한다하더라도 품질이 확보되지 않는다면 더 이상 발전할 수 없기 때문이다.

미국 국립과학원은 디지털시대에 있어 연구데이터의 중요성을 강조하고 데이터의 관리·활용을 위한 3대 원칙과 11개 권고안을 제시하였다. 여기에서 연구데이터는 연구 수행 중 생성되거나 작성된 모든 정보로서 최종결과물뿐만 아니라 중간생성물도 포함한다. 연구데이터의 3대 관리 원칙 중 하나인 관리성(stewardship)은 연구데이터의 관리(management)와 보존(preservation)의 의미를 내포하는데, 연구데이터의 가치판단과 장기적 보존을 위해서는 데이터 보존의 필요성 확립과 데이터 보존계획 수립이 필수라는 의미이다. 또한 가치있는 데이터를 관리하기 위해서는 보존해야 할 데이터와 버려야 할 데이터에 대한 판단이 필요하나 이에 대한 합의가 부족하다고 지적하였다. 권고안 10에서는 연구데이터 관리를 위한 기준 개발의 일환으로 각 연구분야에서 생산되는 데이터를 평가하고 어떠한 데이터를 보존할 것인지에 대한 기준을 수립하는 지침서를 개발해야 한다고 권고하였다.²¹⁾

이처럼 데이터 가치판단에 대한 합의와 평가기준 수립은 연구데이터 관리에 있어 중요한 기본원칙으로 인정받고 있다. 평가는 기록이 지닌 의미와 가치를 판단하는 작업으로 계속적 가치를 지니는 기록을 선별하기 위해서는 합리적인 평가기준과 평가방식 설계, 평가주체들 간의

21) 미국 국립과학원에서 발간한 『디지털 시대 연구데이터의 완전성, 접근성, 관리성 확보 방안』의 내용을 KISTEP이 정리, 분석한 내용을 재인용함. (원서명: Ensuring the integrity, accessibility, and stewardship of research data in the digital age, 2009.10) [cited 2014.6.22.] <http://www.cre.or.kr/board/?board=datamanagement_articles&no=1382849>

협력적 역할수행이 중요하다. 본 장에서는 연구기록의 다양한 평가기준과 평가방식을 살펴보고, 연구기록의 생산자이자 관리자, 이용자인 연구자들의 평가에 대한 시각을 확인하고자 한다.

1) 평가의 주체

제2차 세계대전이 끝난 후 ‘거대과학(Big Science)’²²⁾시대를 거치면서 연구의 규모와 예산, 진행방식이 드라마틱하게 변화했음에도 불구하고 연구기록에 관한 국가적 규범은 2000년대 이전에는 발견하기 어렵다. 오스트리아의 경우, 1990년대 후반 연구데이터의 부적절한 이용 때문에 바람직한 연구 행위를 보장하기 위한 『행위 규칙(Rules of Conduct)』을 채택하기 시작했는데 이는 독일의 ‘과학전문자율규제위원회(Commission on Professional Self Regulation in Science)’에서 권고한 『바람직한 연구수행을 수호하기 위한 제안(Proposals for Safeguarding Good Scientific Practice)』에 기초한 것이다. 이 규칙에서는 연구데이터가 생산된 실험실 또는 기관에서 연구의 원시 데이터를 다른 관련 기록 및 문서와 함께 8년에서 10년간 관리할 것을 권고하고 있다. 오스트리아의 가장 중

22) 우주개발·원자력·해양개발·MHD 발전(電磁流體發電)·대형 컴퓨터·중유탈황(重油脫黃)·공해문제 해결 등의 연구가 그 예이며, 연간 수십억에서 수백억 원, 때로는 수천억 원의 예산을 쓴다. 거대과학이라는 말이 처음으로 사용된 것은 1957년경 미국의 오클리지 국립원자력연구소를 중심으로 해서 행해졌던 원자폭탄 제조계획을 추진하던 때이다. 18~20세기에는 개인의 재능에 의한 발명이나 발견에 따른 사업이 많았지만, 최근에는 국가가 지출하는 거대한 자금과 많은 과학자들의 대형의 계획적 연구가 많아지는 경향이다. 여기에는 이론(異論)도 있으며, 개인의 과학적 연구나 기초과학이 경시된다는 우려를 표명하는 견해도 있다. 그러나 사회 또는 인류의 생존을 위해, 많은 사람이 협력하고 방대한 예산을 들이지 않으면 해결할 수 없는 문제가 있다는 것은 사실이다. 그러한 문제가 무엇이며 그것을 해결하기 위해서는 어떤 조직으로 추진해야 하는가 등이 논의의 대상이 되고 있다. • 출처: [네이버 지식백과] 거대과학 [big science, 巨大科學] (두산백과).

요한 공공기금지원기관인 연구기금(Fund for Scientific Research)은 2005년에 이 권고안을 채택했을 뿐만 아니라, 각 대학이 유사한 규칙을 이행하도록 요구했다. 오스트리아 대학에서 이행되는 『행동 강령(Codes of Conduct)』은 특정 기간, 보통 10년 동안 원시 데이터를 유지하도록 하는 조항을 포함하고 있으나 이것을 어떻게 실현해야 하는지에 대한 상세한 규칙 또는 가이드라인은 없으며, 연구자들이 준수해야 한다는 공식적인 통제도 없다. 즉, 연구기록관리와 관련한 가이드라인이나 지휘 체계가 부재하며, 바람직한 연구 행위를 보장하기 위해 규제보다는 자유의사에 따라 관리가 이루어지고 있는 것이다.²³⁾

이런 상황에서 연구데이터를 행정기록의 시각으로 본 정부기관과 연구기록 자체의 특수성에 이해의 기반을 둔 연구자들의 충돌은 매우 흥미롭다.

먼저 “연구기록에 관한 규칙을 넘어선 충돌(A clash over standards for scientific records)”이라는 제목으로 『Science』에 실린 Eliot Marshall(1990)의 보고에 따르면²⁴⁾, 1990년대만 하더라도 정부당국이 과학자들에게 보고서(published report)에 대한 백업자료(backup material)를 생산할 것을 요구했을 때 연구자들의 심한 거부와 반대에 부딪혔다.

1990년, 미 공중위생국(Public Health Service, 이하 PHS)이 ‘연구 진실성 검토(Scientific Integrity Review)’라는 주제로 연구데이터 저장에 관한 규칙을 논의하기 위해 개최한 연구자들과의 회의에서 연구자들은 그들 스스로가 더 잘 해결할 수 있는 문제에 정부가 개입하는 것에 비판의 목소리를 냈다. David Korn(스탠포드 대학)은 모임의 목적 자체에 거부

23) R. Arovelius et al., Management and preservation of scientific records and data, International Council on Archives: Paris, 2010, p.9. [cited 2014.6.24.] (<http://paarl.wikispaces.com/file/view/Management+and+Preservation+of+Scientific+Records+and+Data.pdf>)

24) Eliot Marshall, “A clash over standards for scientific records” Science Vol.248, 1990, pp.544-545.

감을 나타내며, “외부의 지시없이 대학과 전문가협회가 이 이슈에 더 잘 대처할 수 있으며 국가정책은 필요없다”고 주장했다. Babara Hansen (메릴랜드 대학) 역시 문제가 있다면 PHS 스스로가 만든 조치에 기인한 것이라고 지적하며, “미국 법은 이미 연구자금을 지원받을 경우 최종보고서 제출 이후, 3년 동안 연구데이터를 유지할 것을 요구(PHS는 파일이 존재한다면 연방자금을 지원받은 데이터를 어느 때나 조사할 권한을 가지고 있다고 말한다)할 수 있다”는 PHS의 시각 자체에 우려를 나타냈다. 그리고 지적 진실성(intellectual integrity) 영역에 PHS가 회계규칙을 투영하는 것은 전반적으로 새로운 양상이며 잠재적으로 중요한 문제라고 지적하였다.²⁵⁾ Drumond Rennie(캘리포니아 대학)는 “조치를 취할 정도로 심각한 (기록관리)문제는 결코 없다”며 연구사기(fraud)를 방지하기 위함이라는 PHS의 목적은 데이터 보존 관행을 변경하려는 충분한 근거가 되지 않는다고 하였다. 그는 또한 PHS가 3년 보존 규칙을 회계기록뿐만 아니라 연구기록에도 적용하는 것은 부적절하고 부당한 것으로 수정되어야만 한다고 주장했다. 연구자들과 반대 입장에 선 PHS의 정부 측 인사 가운데 Robert Charrow는 “법률이 연방연구비 수행자에게 그들이 생산한 데이터의 독점권을 주진 않았다”며 그 누구도 절대적 소유권을 가지지 못한다고 지적하였다.

이 연구에서 나온 공통된 의견은 첫째, 연구자들이 생산한 원 자료(raw materials)는 분야마다 매우 다양하여 광범위하게 적용될 수 있는 규칙을 만들려고 시도하는 것은 소용없는 일이라는 점, 둘째, 공유와 아카이빙 간에는 차이점이 존재한다는 것으로 PHS는 실제로 법규적 아카이빙에 관심을 가지고 있어 법령이 요구하는 기록은 최소 3년간 유지되면 되지만 더 필요한 경우도 있다는 점이였다. 마지막으로 대부분의 발표자들은 연구자 개개인이 데이터를 통제할 수 있어야 하는데 대학

25) 연구의 본질적 가치에 대한 평가가 아니라, 회계규칙 적용을 통해 연구데이터를 3년간 유지해야 한다는 논리에 반대하는 것으로 해석된다. (저자 주)

과 관리 의무를 함께 해야 한다면 공개부분에 있어 비밀을 지켜야하는 상황들이 침해받을 수 있다는 우려를 나타냈다. 이 회의의 핵심 이슈는 분야마다 다른 연구특수성에 대한 재천명과 연구자들이 스스로 제기한 행정기록과 연구기록에 대한 보존기준의 차별적 적용, 연구기록의 관리 권한은 연구자의 것이라는 주장이다.

또 다른 예로는 1992년 1월 미국 국립기록청(National Archives and Records Administration, 이하 NARA)이 후원한 연방정부의 자연과학분야 연구·기술데이터의 장기보존과 관련한 계획 수립 회의이다. 미국 국립과학원(National Academy of Science)에 의해 설립된 국가연구위원회(National Research Council)는 국립기록청(NARA), 해양대기관리청(NOAA), 항공우주국(NASA) 등의 지원을 받아 자연과학(Physical Universe) 분야의 연구데이터 보존에 관한 전략을 구상하였다.²⁶⁾

이 연구에서 확인된 사항으로는 우선 모든 연구분야에서 데이터 관리와 보존에 부여된 업무의 우선순위가 낮다는 점이였다. 또한 전자 및 다른 형태로 NARA에 소장된 연구·기술데이터는 연방정부기관 및 그들로 부터 지원을 받는 기관이 소장한 데이터와 비교해 매우 작은 규모였다. 더군다나 NARA의 전자기록센터(Center for Electronic Records)는 모든 형태의 연방전자기록을 아카이빙할 책임을 가지고 있음에도 이에 대한 예산은 많은 개별기관들의 예산보다 적었다. 따라서 NARA가 연구데이터세트(scientific data set)의 대부분을 관리하기에는 불가능하다는 것이 이 회의에서 나온 지적이였다. 연구데이터 관리에 있어 NARA의 역할에 대한 문제제기와 함께 또 다른 중요한 지적은 바로 평가와 관련한 것이였다. NARA는 기관의 활동에 대한 기록의 정보적, 증거적 가치에 기초해서 기록을 평가하는 반면, 연구·기술데이터는 기본적으로 정

26) 이 회의와 관련한 내용은 다음 자료를 정리한 것이다. National Research Council, Preserving scientific data on our physical universe: A new strategy for archiving the nation's scientific information resources, National Academy Press: Washington D.C., 1995. [cited 2014.6.2.] (http://www.nap.edu/download.php?record_id=4871#)

보적 가치를 지니며 그 가치는 기록을 수집하고 생산한 기관의 활동과 관련한 증거보다는 기록에 담긴 연구내용에 기초한다는 것이다. 평가자와 관련해서는 모든 이해당사자(연구자, 연구관리자, 정보관리전문가, 아키비스트, 주요 이용자 그룹)가 해당되어야 하나 개별 데이터세트의 평가는 특정 데이터에 대해 가장 많은 지식을 가진 사람, 기본적으로 연구책임자(Principal Investigator)와 프로젝트 매니저에 의해 평가되어야 한다고 하였다.

이 회의의 중요 결론 중 하나는 비록 NARA가 연방기록 보존에 대한 법률적인 권한을 가지고 있지만, NARA가 대다수 자연분야 연구데이터(physical science data)에 대한 보관자로서의 역할을 하려고 하지도 할 수도 없다는 것이었다. 그 근거로 NARA에게 배정된 매우 낮은 예산에 비해 연구데이터의 볼륨은 너무나 크고, NARA의 직원들은 특별한 연구 지식을 가지고 있지도 않으며, 연구기관들 간의 연계관계도 부재하다는 점을 들었다. 그러나 NARA에게 기존의 국립 아카이브로서의 전통적인 역할과는 다른 연구·기술데이터세트를 위한 아카이빙 프로세스의 파트너로서의 역할을 당부하였다.

이 두 사례는 모두 1990년대에 발생한 것이지만 20년이 지난 지금도 연구기록관리 현실은 큰 변화가 없는 듯하다. 연구기록의 특성상 생산자와의 협력관계가 중요하다는 측면에서 기록의 생산자인 연구자들의 의견이 전면에서 드러난 사례를 살펴보았다. 이 두 사례 모두 연구기록 평가에 있어 연구의 특수성과 전문성에 바탕을 둔 연구자의 역할을 강조하였다. 기록관리자의 전문영역이라 여겼던 평가분야에서 연구자들이 주체여야 한다는 이 주장들에 기록관리자들은 이견을 제기할 수 있다. 그러나 연구기록의 특성상 기능이 아닌 내용적 가치를 평가해야 한다면 지금의 기록관리자가 어떻게 전문성을 주장할 수 있을지는 큰 고민이 아닐 수 없다. 물론 기관전체의 연구기록 평가를 총괄해야 하는 입장에서 기록관리자는 전체 평가프로세스의 통제자이자 조정자로서

제 몫을 할 수 있으며 분야별 전문가로 구성된 평가위원회와의 협력은 효율적이고 내실있는 평가진행의 해결책이 될 수도 있다.

2) 평가기준

기록은 연구 개발을 수행하는데 중요한 정보를 포함하고 있으므로 적절한 평가는 행위의 증거로써 기록을 보호하고 보존하기 위해 극히 중요하다. 연구기록이 생성되는 시점에서부터 기록관리에 적절한 관심을 두지 않으면 연구팀이 연구를 완료한 뒤에 해당 과제의 전체 데이터 컬렉션은 고아가 될 수밖에 없고, 이것은 다음 세대에는 쉽게 생성하지 못할 수도 있는 연구데이터를 재사용하는 일이 불가능할 수도 있음을 의미한다. 가령 수많은 과학 관측은 결코 정확하게 반복될 수 없는 사건(event)이고, 한번 없어지면 그러한 데이터는 결코 대체될 수 없다. 관측데이터의 경우 특별한 사건의 발생 빈도와 변화율을 결정하는 기준을 제공한다. 데이터는 생성된 것과 다른 목적으로 장래에 이용될 수도 있고, 데이터 재분석은 심지어 먼 미래에 대한 새로운 이해를 가져올 수 있다. 관측데이터는 자연발생의 역사이자 인간 성취의 기록이다. 다시 말해 일시적인 자연 현상(예를 들어, 홍수, 지진 등)은 재현할 수 없다는 점에서도 연구기록관리는 매우 중요하다.

혹자는 연구기록의 가치는 현재의 유용성이 큰 비중을 차지하고, 정보기술의 발달로 대용량 저장시스템에 용량과 관계없이 기록을 저장 관리할 수 있기 때문에 평가의 중요성을 간과하려 하기도 하나 그에 따른 관리비용부담은 물론 데이터의 품질과 이용효율을 고려했을 때 평가는 역시 연구기록관리의 핵심 영역이다.

연구기록 평가기준은 연구분야의 연구특성, 연구집단의 다양한 요구와 규칙은 물론 법률과도 밀접한 관련이 있다. Arovelius는 “비록 다양한 요건에 의해 차이점이 존재하지만, 보존을 위한 많은 기본적 기준이 있

으며, 이 기준은 대학과 연구기관의 가장 오랜 전통인 개방성(openness)에 있고 이 개방성은 연구결과의 광범위한 전파와 평가를 겨냥한다”고 하였다. 그리고 그는 ‘연구결과 증명의 필요성, 접근 가능성, 데이터의 재활용 가능성 및 새로운 연구 기획의 가능성’을 평가기준으로 제시하였다.²⁷⁾

미국 국가연구위원회는 관측데이터와 실험데이터가 각기 다른 특성을 가지고 있음을 설명하면서 데이터세트의 저장 여부를 결정하기 위해 양쪽 모두에게 적용될 수 있는 다음의 선별기준을 제시하였다. 유일성 측면에서 진본카피가 존재하고 적절한 백업도 되고 있다면 보존대상이 아니며, 대체비용의 경우 관측데이터는 재획득될 수 없으나 실험데이터는 원칙적으로 재획득 가능하다는 특징이 있기 때문에 관측인지 실험인지 연구유형도 고려되어야 한다고 하였다.

〈표 4〉 미국 국가연구위원회의 보존대상 선별 기준²⁸⁾

항목	의미
유일성 (Uniqueness of data)	NARA의 기준을 충족시키는 진본카피의 존재 여부
접근성-도큐멘테이션의 적절성 (Accessibility)	해당 분야 연구자들이 데이터세트를 이용하기에 충분한 메타데이터가 있는지 여부
접근성-하드웨어의 이용가능성	데이터 이용에 필요한 장치 여부
대체비용 (Cost of replacement)	향후 데이터에 대한 국가적 요구가 발생할 때, 재획득 가능 여부와 비용. 보존비용과의 비교
해당 분야 전문가심사 (Peer review)	무결성, 완전성, 데이터 활용 등에 대한 평가 정도

27) R. Arovelius et al., Management and preservation of scientific records and data, International Council on Archives: Paris, 2010, pp.28-29.

28) National Research Council, Preserving scientific data on our physical universe: A new strategy for archiving the nation's scientific information resources, National Academy Press: Washington D.C., 1995, p.34. [cited 2014.6.2.] <http://www.nap.edu/download.php?record_id=4871#>

NOAA의 기록 처분 핸드북(Records Disposition handbook) 『Chapter 1200. 연구 및 기술 기록(Scientific and Technical Records)』²⁹⁾에서는 NOAA의 연구프로젝트를 기록화(documenting)하기 위해 특히 중요한 영구기록 선별 기준(permanent records selection criteria)을 다음과 같이 제시하고 있다. <표 5>의 기준을 하나 또는 그 이상 충족시키는 프로젝트 케이스 파일(과 관련 연구노트들)은 영구기록으로 선별되고 관리될 것이다. <표 5>를 살펴보면 영구보존 대상으로 선별하는 기준에는 연구의 권위와 공익성, 사회 기여 정도, 역사성과 증빙성이 포함된다.

〈표 5〉 NOAA의 연구프로젝트 기록 중 영구기록 선별 기준

선별 기준
1. 노벨상과 같이 국가 또는 국제적으로 탁월한 상을 수상한 경우
2. 전문가로서의 위상을 널리 인정받은 NOAA의 중요 연구자의 저작 또는 자신의 전문 분야 이외의 국가 또는 국제적인 인정을 받은 경우
3. 공중 보건, 안전 또는 다른 중요한 공공의 이익에 상당한 개선 결과를 가져온 경우
4. 새로운 국가 또는 국제적 환경 정책에 눈에 띄는 공헌을 하거나 국가/국제적으로 우선순위에 있는 과학, 정책, 경제, 또는 사회적 이슈의 발전에 중요한 영향을 미친 경우
5. 국가 또는 국제적으로 광범위하게 언론의 주목을 받은 경우
6. 중요한 사회적, 정치적, 또는 과학적 논쟁의 결과인 경우
7. 의회 또는 정부기관의 면밀한 검토 및 조사의 대상이었던 경우
8. NOAA의 연구 또는 운영 정책을 변화시키는 중요한 선례를 확립한 경우

영국 DCC(Digital Curation Center)의 Whyte와 호주 ANDS의 Wilson(2010)은 공동으로 작성한 보고서에서 평가 및 선별 정책에서 데이터세트나 정보자원의 가치를 평가할 수 있는 기준을 세워야 하며, 이에 따라 얼마나 오랫동안 보존(keep)해야 하는지, 또는 언제 그것을 파기할 것인지가 정해진다고 하였다. 기준은 분야별 특수한 요인에 따라 달라질 수

29) NOAA, Records Disposition Handbook, Chapter 1200: Scientific and Technical Records, [cited 2014.6.20.] <http://www.corporateservices.noaa.gov/audit/records_management/schedules/chapter-1200-science-research.pdf>

있다고 전제한 뒤 다양한 정보원으로부터 뽑은 7가지의 일반기준을 제시하였다.³⁰⁾

〈표 6〉 데이터세트 및 정보자원 평가 선별 기준

항목	의미
1. 미션과의 관련성 (Relevance to Mission)	즉각적인 이용을 넘어 데이터를 보존(retain)하는데 대한 법적 요건을 포함해서 조직의 역할 및 연구기관 또는 자금지원기관의 현재 전략에 명시된 우선순위에 부합해야 함
2. 연구 또는 역사적 가치 (Scientific or Historical Value)	데이터가 연구적으로, 사회적으로 또는 문화적으로 중요한가의 여부. 이 평가는 현재 연구의 증거적 교육적 가치로부터 예상되는 미래의 이용을 추정하는 것을 포함
3. 유일성 (Uniqueness)	자료의 내용이 유일하거나 그것으로부터 파생될 수 있는 정보의 가장 완벽한 출처임을 의미. 수용되지 않을 경우 손실의 위험이 있거나 다른 곳에서 보존될 수 있음
4. 재배포 가능성 (Potential for Redistribution)	데이터 파일의 신뢰성, 무결성, 이용가능성 결정. 부여된 기술 기준을 충족시키는 포맷으로 받은 데이터 파일이며, 지적재산권 또는 사람을 대상으로 한 이슈가 다루어져야 함
5. 복제불가능 (Non-Replicability)	데이터/정보자원을 복제하거나 또는 그렇게 하는 것이 경제적으로 실행 불가능해야 함
6. 경제적 상황 (Economic Case)	정보자원의 관리와 보존을 위한 비용이 추산되고, 미래의 잠재적인 이익을 평가했을 때 정당하면 적절한 자금 확보 가능
7. 충분한 기록화 (Full Documentation)	정보자원의 출처와 그것이 생산, 활용된 맥락에 관한 메타데이터를 포함하여 미래에 탐색, 접근, 재활용을 용이하게 하기 위해 필요한 정보는 포괄적이고 정확해야 함

USGS는 연구기록 컬렉션 평가(appraisal)를 위한 일련의 질문 세트를 활용하고 있는데, 미션과의 부합성, 액세스 및 배포 특성, 부가적인 특성, 물리적 특성, 메타데이터 특성, 경제적 특성 등 총 7개의 섹션으로

30) Angus Whyte & Andrew Wilson, How to Appraise and Select Research Data for Curation, 2010. [cited 2014.6.23.] <<http://www.dcc.ac.uk/resources/how-guides/appraise-select-data#sthash,aHMqZeYV.dpuf>>

구분되어 있다. 질문들을 살펴보면 ‘유일성, 연구에 대한 기여, 기록의 속성(신뢰성, 무결성, 이용가능성)과 가치(본질적, 역사적 가치), 크기 (Volume), 충분한 메타데이터의 보유 여부, 경제적 비용’ 등이 주요 항목임을 확인할 수 있다.

〈표 7〉 USGS 연구기록 컬렉션 평가 질문 세트

구분	질문
Sec.1	미션에 부합하는 특성(Mission Alignment Characteristics)
	0.기록이 수집 정책의 범위에 적합한가?
	1.데이터의 예상되는 현재/미래의 유용성은 EROS의 미션에 적합한가?
	2.본 기록은 원격탐사, 지도제작, 지구과학 데이터 이용자 집단에게 얼마나 중요하고, 차별성 있으며, 유일무이한가? 즉, 현재 아카이브 소장 자료를 업그레이드 시키는데 어떤 중요하고 고유한 기여를 하는가?
	3.현재 아카이브 소장 자료를 보완하거나 차이를 매우는데 컬렉션이 어떻게 기여하는가?
	4.데이터는 시간이 지남에 따른 지구물리학적 변화에 관한 연구를 지원하는가?
Sec.2	액세스 및 배포 특성(Access & Distribution Characteristics)
	0.기록의 진본성을 어떻게 판단할 수 있는가? 즉, 어떻게 기록이 진본인 것으로 간주될 수 있는가?
	1.기록의 신뢰성은 어떠한가?
	2.기록의 무결성은 어떠한가?
	3.기록의 이용가능성이 기록의 정보적 가치를 예상하는데 도움이 되는가?
	4.데이터가 정부 또는 개인의 법적 권리와 관련되는가? 또는 데이터사기나 허위 진술혐의에 대해 기관이나 정부를 방어하는데 필요한가?
	5.컬렉션에 대한 액세스가 제공되는 경우, 일부 이용자들은 본래의 원시 데이터(raw data)를 요구할 것인가?
	6.컬렉션이 NARA를 비롯한 다른 기관의 이용자들도 이용가능하게 만들어 졌는가?
	7.컬렉션은 어떻게 배포 또는 액세스되는가?
	8.기록에 액세스하는데 있어 물리적, 지적, 법적 장벽은 무엇인가?
9.예상되는 이용자그룹은 누구이며, 예상되는 요구는 무엇인가?	
Sec.3	부가적인 특성(Additional Characteristics)
	0.이 컬렉션에서 커버하는 공간영역(spatial area)은 무엇인가?

	1.컬렉션의 시간 범위는?
	2.컬렉션은 전체 인구 또는 통계적으로 유효한 샘플을 대표하는가?
	3.누가, 어떤 목적으로 기록을 생산했으며, 과거에 이 컬렉션을 소유했던 사람은 누구이며, 현재 소유자로 간주되는가? 저작권 표시 등의 법적 태그를 전송 (transfer)전에 제거할 수 있는가?
	4.고유하거나, 차별화되거나, 어려운 아카이빙, 배포 또는 고객서비스요건을 어떻게 수용할 것인가?
	5.데이터 컬렉션이 지속적으로 증가한다면, 매년 증가가 예상되는 추가적인 기록의 양은 얼마인가?
	6.기록의 본질적 또는 역사적 가치는?
	7.컬렉션의 현재 소유자 또는 생산자로부터 받을 수 있는 교육 기술은?
Sec.4	물리적 특성(Physical Characteristics)
	0.기록이 저장된 매체는 무엇인가?
	1.양(volume)적 측면에서 컬렉션의 전체 크기는?
	2.기록은 현재 어떤 순서로, 어떤 수준의 프로세싱/포맷으로 되어 있는가?(최선의 보존 수준과 연구자에게 가장 유용한 것이 무엇인지에 주목)
	3.기록의 현재 상태 및 전반적인 품질
	4.기록에 사용된 압축 기술
	5.파일 또는 이미지의 네이밍 형식은?
	6.컬렉션이 검색이미지를 포함하는 경우, 포맷과 가용성 기술
Sec.5	메타데이터 특성(Metadata Characteristics)
	0.컬렉션을 기술하는 메타데이터의 양, 품질, 정확성, 수준(level), 가용성
	1.유용한 추가정보의 정도
Sec.6	경제적 특성(Economic Characteristics)
	0.기록을 획득, 보존, 접근가능하게 하는 것과 관련된 비용을 후원하기 위한 프로그램이나 자금지원처를 확인
	1.자본투자 및 반복되는 비용에 대한 비용 분담 기회를 확인
	2.다른 누군가가 컬렉션을 재생산하기 위한 비용을 평가하거나 컬렉션의 연구적 또는 2차가치가 보존 및 접근제공을 위해 사용된 비용을 얼마나 초과하는가를 평가
	3.컬렉션을 확인하고 평가하고 입수하고 처리하고 접근가능하게 하기 위한 대략적인 비용은?
	4.요구되는 보존기능에 필요한 자원을 확인
	5.원본 기록을 소장하는데 소요되는 대략적인 연간 비용은?
	6.기록을 읽거나 처리하는데 필요한 고유한 장비를 확인
	7.컬렉션을 처분, 제거하는 비용을 추정
Sec.7	추가적인 필름 특성(Additional Film Characteristics)
	전체적인 평가 의견(Overall appraisal comments)

이외에 사우샘프턴 대학에서는 연구데이터 보존에 영향을 미치는 요인을 다음과 같이 제시하고 있다.³¹⁾ 이 대학에서는 중요한 연구데이터가 최소 10년 이상 유지되어야 할 필요가 있으며 데이터를 적극적으로 사용하는 곳에서는 더욱 길어질 수 있다고 보고 있다.

- 연구의 영향력(Research impact)
- 학문적 명성(Academic reputation)
- 파생/연계된 출판물(Derived and linked publications)
- 법적 의무(Statutory/legal obligations)
- 대학 또는 자금지원기관의 정책 요구사항(University and/or Funder policy requirement)
- 지속적인 또는 추가 연구(On-going or further research)
- 다른 사람에 의한 검증 및 테스트(Validation and testing by other)

이와 같은 다양한 사례를 통해 우리는 연구기록의 평가기준이 행정 기록의 그것과는 다르게 구성될 수 있음을 확인했다. 평가기준마다 포함되어 있던 ‘연구의 영향력’을 비롯한 ‘학문적 가치와 연구 기여도’는 연구기록 평가시 반드시 고려되어야 할 중요한 항목이다. 참고로 이 이외에도 또 다른 측면, 즉 ‘연구유형’에 따른 평가기준의 차이를 보여주는 사례가 있다.

미국 국가연구위원회는 연구자입장에서 관측데이터와 실험데이터 간 특성 차이가 존재하고 이에 따라 기록의 가치가 구별될 수 있다는 중요한 지적을 했다.

31) Southampton Univ., Research Data Management: Guidance on Retention Periods, [cited 2014,6,24.] (<http://library.soton.ac.uk/researchdata/retention>)

일시적인 자연 현상은 반복될 수 없는 현상이고, 따라서 짧은 순간의 스냅샷과 같은 관측데이터는 때론 정보적 가치 이외에도 역사적, 증거적 가치를 가진다. 관측데이터는 자연현상에 대한 계속되어 온 일련의 시간의 기록이고 따라서 변화의 흐름을 확인하고 비교하기 위해 다음 세대에 매우 중요하다. 게다가 많은 관측데이터세트는 중요한 엔지니어링 또는 노동집약적인 수집활동이고 재수행될 수 있다고 보장할 수 없다. 한편, 실험데이터는 기기가 더 나아진다면 더욱 잘 재생산될 수 있다. 많은 실험에서 원시 데이터는 폐기되거나 다음 처리를 하기 전에 아주 잠시 동안만 존재한다. 심지어 원시 데이터가 획득되거나 저장될 때 연구책임자(PI)는 적절하게 기록화하지도 않는다. 왜냐하면 그들은 또 다른 누군가가 이 데이터를 사용할 것이라고 기대하지 않기 때문이다.³²⁾

중복되지 않고 신뢰할만하며 이용가능한 모든 관측데이터는 영구적으로 보존되어야 하는 반면, 쉽게 재생산 가능한 실험데이터라면 원시 데이터라 할지라도 즉각적으로 폐기될 수 있다는 연구자들의 이와 같은 생각은 앞서 살펴 본 다양한 평가기준을 연구유형에 따라 다르게 적용하는데 있어 세부 주의사항으로 고려될 수 있다.

3) 평가방식

현재 공공기록의 평가방식은 ‘기록관리기준표 작성 → 단위과제 신설 시 보존기간 책정 → 해당 기록의 보존기간 만료시 재평가 → 처분(보

32) National Research Council, Preserving scientific data on our physical universe: A new strategy for archiving the nation's scientific information resources, National Academy Press: Washington D.C., 1995, p.35. [cited 2014.6.2.] (http://www.nap.edu/download.php?record_id=4871#)

류/폐기/재책정'의 과정으로 진행되고 있다. 기본적으로 기능평가방식이 적용되고 있는데 기능평가는 기록의 계속적 가치를 판단하는데 있어 기록을 생산한 조직의 기능이 적절하게 기록화되었는지의 여부를 중심에 둔 평가방식이고 따라서 그 기능이 가장 완전하고 간결하게 담긴 기록을 선별하는 방식³³⁾이기 때문에 개별 기록의 내용적 가치가 중요한 연구기록 평가에 적용하기에는 한계가 있다.

최현옥, 이해영(2010)은 연구업무와 생산·관리되는 기록의 종류를 분석하여 연구기록의 분류체계를 마련하고 분류된 연구기록의 종류별로 보존기간을 책정할 것을 제안하였다. 그리고 보존기간 책정은 규정, 기록의 활용도와 가치에 따라 영구, 준영구, 30년, 5년, 계속사용으로 나눌 수 있다고 보고 보존기간 책정 기준의 예를 <표 8>과 같이 제시하였다.³⁴⁾ 이는 현 보존기간 책정체계의 한계를 넘어서려는 시도로 해석될 수도 있으나, 아쉽게도 행정기록에 적용되는 방식을 크게 벗어나지 못했고 연구기록의 다양한 평가기준을 고려하지 않았다는 점에서는 한계가 있다.

<표 8> 연구기록 보존기간 책정 기준의 예(최현옥, 이해영)

보존기간	대상기록
영구	• 정부 또는 국고지원금을 받은 연구기록
준영구	• 현재는 활용되지 않지만 언젠가는 활용될 가능성이 있는 연구기록 • 현재 법적 문제 등으로 인해 지속적으로 보존해야 하는 연구기록
30년	• 연구일지(아이디어 일지, 실험방법 일지, 실험데이터 일지, 프로젝트 관리 일지)
5년	• 최소 의무 보존기간을 지켜야 하는 연구기록 • 연구비와 관련한 연구회계 기록
계속 사용	• 연구일지, 기록 목록 및 일정표, 계산 등을 위해 사용되는 자료, 스크랩 등 연구수행을 위해 매일 사용해야 하는 기록

33) 한국기록학회, 『기록학용어 사전』, 역사비평사, 2008, 43쪽.

34) 최현옥, 이해영, 「우리나라 공과대학 연구실 연구기록의 체계적 관리 방안에 관한 연구」, 『한국기록관리학회지』 제10권 제1호, 한국기록관리학회, 2010, 132-133쪽.

우리의 평가방식과는 다소 다른 해외의 사례 가운데 USGS의 EROS 연구기록 평가과정³⁵⁾은 9단계로 되어 있다: ① 평가되어야 할 데이터세트의 확인 → ② 평가팀 구성 → ③ 데이터세트 팩트 수집 → ④ 연구 리뷰 → ⑤ 온라인 Tool을 이용, 기록된 사실과 연구 리뷰 → ⑥ 아키비스트와 프로젝트 관리자에게 브리핑 → ⑦ 작성된 평가의견(recommendation)을 상급관리자(senior management)에게 송부 → ⑧ 상급관리자의 질의응답 → ⑨ 최종 결정

좀 더 자세히 살펴보면, 평가팀은 기록관리 담당자와 연구자(Archive staff, Science Staff, Project manager, Archivist)로 구성되며 기록관리 담당자(Archive staff)가 데이터세트와 관련한 여러 정보를 수집하고 연구자(Science team members)가 데이터의 본래 가치 및 미래가치를 평가한다. 데이터의 보유/보존여부를 결정하기 위해 USGS는 S/W Tool을 개발하여 평가정보도 획득한다. 기록관리 담당자(Archive staff)는 아키비스트와 프로젝트 관리자에게 이 내용을 간략하게 보고하고, 아키비스트는 USGS 프로그램 코디네이터와 EROS 상급관리자에게 검토 의견을 송부한다. 이 코디네이터와 상급관리자의 의견은 물론 다를 수 있다. 그들이 EROS의 장에게 자신들의 의견을 보내면 그가 최종 결정한다. EROS에서 보존되는 모든 컬렉션은 이 EROS 연구기록 평가프로세스를 통해 검토된다.

이 프로세스에서 눈에 띄는 점은 일단 많은 단계를 거칠 뿐만 아니라, 연구자와 기록관리자가 평가과정 전반에 함께 참여한다는 점이다. 현재 우리의 평가방식은 생산자와 기록관리자의 역할이 단계마다 분리되고, 기록관리자에게 평가의 모든 업무를 맡긴 채 생산자는 마치 옵저버와 같은 지위를 차지하려는 경향이 있다. 그러나 기록 생산의 특징과 기록유형을 고려했을 때 연구기록은 행정기록을 평가하는 기능평가방식보다 더욱 생산자의 평가가 중요해진다.

미국 국가연구위원회는 대부분의 연구자들이 기록관리 목적을 위한

35) EROS(Earth Resources Observation and Science) Scientific Records Appraisal Process

평가자로서의 역할은 인식하지 못하나, 이 전문가들의 관점은 획득된 데이터의 장기보존 또는 영구보존 가치와 관련된 결정에 있어 큰 영향을 미친다고 지적하였다. 주요 연구자와 프로젝트 관리자는 분명히 수집, 분석하는 데이터가 그들 자신의 연구 목적을 위해 얼마나 오랫동안 가치가 있을지 잘 알고 있다. 주요 이용자 역시 그들 분야에서의 데이터 이용에 관한 상세한 정보를 제공할 수 있지만, 주요 이용자와 데이터 수집가들은 자신의 필요를 넘어 생각하지 않을 것이다. 따라서 기관은 연구 프로젝트의 시작에서부터 제2, 제3의 이용자에게 유용한 기록 평가를 위해 모든 이해당사자(과학자, 연구관리자, 정보관리 전문가, 아키비스트, 주요 이용자 그룹)가 대표되고 특히 개별 데이터세트의 평가는 데이터에 관한 지식이 가장 많은 이들, 기본적으로 연구책임자(PI)와 프로젝트 관리자에 의해 수행될 것을 권고했다. 그리고 평가는 각 프로젝트를 위해 수립된 데이터관리계획(Data Management Plan, 이하 DMP)에 따라 실행되어야 한다는 제안도 덧붙였다.³⁶⁾

이런 제안과 USGS의 평가 프로세스를 고려했을 때 우리의 평가방식은 전체적인 연구기록관리계획하에서, 다양한 이해당사자와의 협의를 통해 설계되고, 프로세스 전반에 걸쳐 기록관리자와 연구자가 협력하는 형태여야 할 것이다. 특히 전자기록환경에서 개별기록의 내용적 가치보다 기능에 따른 평가가 이루어지는 지금, 연구기록의 평가방향과 프로세스 설계는 매우 중요하다. 또한 연구기록이 가진 내용적 가치의 중요성만큼이나 기록의 양적 증가가 중요한 이슈인 만큼, 현실적인 평가 방안은 무엇일지 많은 고민이 필요하다.

36) National Research Council, Preserving scientific data on our physical universe: A new strategy for archiving the nation's scientific information resources, National Academy Press: Washington D.C., 1995, p.39. [cited 2014.6.2.] <http://www.nap.edu/download.php?record_id=4871#>

4. 결론 및 제언

2011년 기준, 우리나라의 정부연구개발예산 규모는 OECD 회원국가 중 미국, 일본, 독일, 프랑스에 이어 5위에 해당하며, 2002년 이후 연평균 9.9%의 높은 증가율(미국의 연평균 증가율 3.8%, 일본 3.5%, 영국 0.3%, 독일 5.7%, 프랑스 1.4%)을 기록하고 있다. 또한 우리나라 GDP 대비 정부연구개발예산 비중은 1.05%로, 이 역시 핀란드(1.09%)에 이어 두 번째로 높은 수준이다.³⁷⁾

전 세계적으로 연구기록 공유에 대한 국가차원의 강제조항들이 생겨나고 공공재로 인식하는 정도가 확산됨에도 불구하고 연구기록이 가진 중요성 및 가치에 비해 관리는 미흡하다. 동일한 연구에서도 전혀 다른 데이터가 생산되기도 하고, 데이터에 따라서는 재연이 어렵거나 생산에 천문학적인 비용이 소요되기도 하기 때문에 연구데이터(scientific data)는 존재 그 자체만으로도 의미있음에도 불구하고 데이터 범람의 한편에서는 생산된 데이터가 체계적으로 축적되거나 관리되지 못해 데이터의 유실이 발생하고 있다.³⁸⁾ 앞서 지적한 바와 같이 드라마틱하게 변한 기록의 속성과 정보기술의 발전은 연구기록관리방식 변화에 촉매로 작용할 것이다.

본 연구에서는 연구기록의 특성을 생산자, 연구행위, 기록의 측면에서 살펴보고 그 범위를 정리하였다. 그리고 공공기록물 관리에 관한 법률 및 관련 규정을 연구기록에 적용시키는 것은 한계가 있고 평가 기준 설계시 연구기록의 특성을 고려하여 행정기록과는 다른 기준이

37) 한국과학기술기획평가원, 『OECD 자료로 살펴본 주요국 정부연구개발예산 현황』 한국과학기술기획평가원, 2013, 2쪽 재인용.

38) 김선영, DB구축의 사각지대, 과학데이터, 2010. [cited 2014.6.17.] (<http://blog.naver.com/PostView.nhn?blogId=dbstory2009&logNo=120107578549&redirect=Dlog&widgetTypeCall=true>)

적용되어야 함, 평가방식에 있어 기록관리자는 계획단계부터 연구자와 철저한 협력관계를 가져야 한다는 점을 확인하였다. 좀 더 구체적인 방안으로는 현실적인 평가도구인 처분일정표(Records Schedule)를 지금의 연구기록 특성에 맞춤형된 것으로 전면 재설계해야 하며, 평가주체, 평가기준, 평가프로세스를 포괄하는 연구기록관리계획((Research Records Management Plan, RMP)을 수립해야 한다. 해외 사례처럼 전문가로 구성된 평가위원회를 구성하는 것도 서로의 한계를 극복하는 좋은 방법이 될 수 있다.

이젠 연구자들의 자유의사에 따라 이뤄지던 연구기록관리가 일정한 지휘체계하에서 규제나 가이드라인을 바탕으로 체계적으로 관리되어야 할 시점이다. 극단적으로 연구자가 원시 데이터의 보존가치를 발견하지 않는다면 이후에 데이터는 아무도 사용할 수 없을 것이다. 무엇보다 연구기록의 생산자인 연구자가 가진 연구기록에 대한 시각과 관리행위의 역사, 연구특성 등을 살펴봤을 때 연구기록의 평가주체, 평가기준, 평가방식에 있어 연구자가 기록관리자와 깊은 협력관계를 맺어야 하는 대상이라는 점은 더욱 분명해졌다. 물론 연구기록관리에 있어 기록관리자의 역할에 대한 고민은 더 많이 이루어져야 한다. 연구자와의 협력방식이나 현실적 평가 프로세스를 어떻게 설계할지는 분명 기록관리자의 몫이다. 기록관리자들이 얘기하는 기록관리 전반에 걸친 각 프로세스의 의미가 각별하지만, 연구자들에게 주 업무는 연구이지 데이터 관리가 아니다. 연구과정에서 기록의 생산, 관리, 보존, 활용의 주체인 연구자들에게 합리적이고 용이한 프로세스를 제시하지 않는 한 연구기록의 품질을 담보할 수 없으며 협력은 불가능하다. 따라서 연구 프로세스 분석과 연구기록 특성에 대한 깊은 이해가 선행되어야 하며 이에 따른 기록관리 프로세스와 업무 모형 등은 연구자들의 검증을 받아 완성되어야 한다. 기록관리에 대한 지식을 바탕으로 연구자들과의 합의된 기록관리 모형을 만드는 것, 그것이 '전문가'로서 기록관리자의 역할이 아

닐까 싶다.

연구기록에 대한 연구자들의 시각과 다양한 조직의 평가기준과 방식은 살펴보았으나 실용학문인 기록학에 기여할만한 구체적인 모형이나 프로세스를 제시하지 못했다는 점은 본 연구의 한계이다. 일차적으로 연구기록은 연구자가 배제된 어떤 기록관리 행위도 어렵다는 점에서 그들의 특성과 시각을 확인하고자 한 것이 본 연구의 큰 의미라 할 수 있으며, 다양한 평가영역과 영역별 평가기준 도출, 행정기록과 차별화된 처분일정표와 평가프로세스 설계는 다음 연구로 넘기고자 한다.

ABSTRACT

A Study on the Appraisal of Research Records in Science and Technology : Focusing on Foreign Cases

Lee, Mi-Young

With the quantitative growth of research data, the issue of enormous preservation cost and sharing expansion, the organizations should prioritize the collections then select the data that are worthy of save. Therefore, today, it is important for the organizations to appraise the continuing value of produced records.

Considering the universities and the public institutions such as government-funded research institutes as the heavy producer of the data, it becomes a rising problem for the records management that it does not go beyond the framework of “administrative records” and “public records”.

In this study, I looked into the background of the contention that the research records must be managed in a different perspective and checked the characteristics of research records in terms of the producers, research activities and records. Based on this analysis, I suggested the main issues and considerations about the subjects, criterias and methods in research records appraisal.

Key words : research records, records of science, scientific data, research data, appraisal, archival appraisal, research records appraisal, appraisal criteria, appraisal method, research data management, research records management