

해양 분야의 과학 데이터 수집 사례 분석

이상호* · 이상환* · 김성대**

I. 서론

최근 과학기술 분야에서는 측정기구나 관측기기 등이 대부분 디지털 기기로 개발, 판매되고 있어 이러한 관측 기기에서 생성되는 데이터들도 모두 디지털 매체에 저장할 수 있는 디지털 데이터로 취급할 수 있게 되어 과거에 비해 데이터의 관리가 훨씬 쉬워지게 되었다. 또한 특정 장소에서 실시간으로 관측 대상을 연속적으로 관측함에 따라 생성되는 데이터의 양이 지속적으로 증가하고 있으며 영상 데이터, 시뮬레이션 데이터 등은 데이터 자체의 크기가 매우 클 뿐만 아니라 이러한 원시 데이터를 기반으로 하여 다시 가공, 처리, 재생산된 2차 데이터, 3차 데이터들이 계속 생성되고 있어 과학기술 분야에서 생산되고 있는 데이터는 과거에는 상상할 수 없을 정도의 큰 규모를 가지게 되었다 [1].

한편 이렇게 생성된 데이터 가운데에서 주로 거대한 측정 또는 관측 장비로부터 생성된 데이터들은 해당 장비를 보유하고 있는 특정 분야의 데이터센터들을 중심으로 데이터들이 수집, 관리되어 이용자들에게 활용되고 있으나 개인 연구자나 연구그룹 또는 대학에 소속되어 있는 교수들의 연구 활동에서 얻어진 과학 데이터들은 개인 또는 연구실 차원에서 관리되고 있어서 데이터의 공유가 이루어지지 않고 시간이 경과함에 따라 데이터들은 점차로 유실되어 국가 예산과 연구자의 시간과 노력이 투입되어 얻어진 귀중한 과학 데이터들이 제대로 관리, 활용되지 못하고 연구 현장에서 점차로 사장되고 있는 실정이다.

본 연구의 목적은 한국해양연구원과 부속 극지연구소의 데이터 관리 및 규정에 관한 법제도와 해양 자료의 유통 특성을 과학기술기본법 및 국가연구개발사업 관리 규정과 비교하여 과학 데이터의 관리 측면에서 참고할 수 있는 시사점을 찾아내어 향후에 국가의 과학 데이터 관리체제를 구축하는데 있어 도움이 되도록 하기 위함이다.

II. 본론

1. 해양자료 관리

1) 해양과학조사법

우리나라 해역에서 외국인 또는 국제조직에 의한 해양과학조사의 실시에 필요한 절차를 정하고, 대한민국 국민에 의하여 실시된 해양과학조사의 결과물인 조사자료의 효율적 관리 및 공개를 통하여 해양과학기술의 진흥을 도모함을 목적으로 해양과학조사법이 1995년에 제정되었으며, 1996년과 1999년에 개정되었다. 해양과학조사법과 동법 시행령에는 국내 해양관련 기관의 해양과학 조사자료의 관리에 대하여 규정하고 있다.

해양과학조사법 제21조에서는 국가예산으로 수행된 조사에서 생산된 자료의 원칙적인 공개를 명시

* 이상호, 한국과학기술정보연구원 과학데이터연구팀 책임연구원, 전화 042-869-1765, shlee@kisti.re.kr

* 이상환, 한국과학기술정보연구원 과학데이터연구팀 팀장, 전화 042-869-1715, sanglee@kisti.re.kr

** 김성대, 한국해양연구원 자료정보실 실장, 전화 031-500-4510, sdkim@kordi.re.kr

하고 있으며, 필요에 따라 공개 범위를 조정할 수 있도록 되어 있다. 동법 제22조에서는 관계부처에서 해양과학조사자료의 국가 관리기관을 지정할 수 있도록 되어 있으며, 관리기관의 의무사항 내용을 명문화하였다. 이 규정에 의해 국가해양과학조사자료 관리기관에서는 각 해양관련 기관에 자료목록의 제출을 요청할 수 있는데, 현재, 국립수산물과학원이 관리기관의 업무를 수행하고 있으며, 해양수산부에서는 3년마다 관리기관 평가 및 재지정을 검토하고 있다. 또한 제23조에서는 해양관련기관의 자료제출을 독려할 수 있는 법적 근거를 마련하고 있다.

우리나라에서 수집되는 해양과학조사자료의 관리에 대한 구체적인 사항은 해양과학조사법 시행령 제3장 ‘대한민국 국민의 해양과학조사’에 규정되어 있다. 시행령 제12조에서는 해양과학조사법의 적용을 받는 기관을 명시하고 있는데, 이 법에 의하면 대학들도 포함되어 있으나, 현재 각 대학의 자료는 국가적으로 관리되지 않고 있다. 또한 제13조에서는 관리대상 조사자료의 항목과 함께, 기초자료와 시료자료의 관리방법을 지정하고 있으며, 제14조에서는 자료의 위탁에 대해 규정하고 있으나, 아직 자료를 위탁하는 해양기관은 없다. 제15조에서는 기초자료의 공개 및 공개제한에 대해 규정하고 있는데, 각 기관의 기관장은 자료의 공개를 3년간 제한할 수 있도록 되어 있다. 이 규정은 많은 노력을 투입하여 생산한 해양자료에 대하여 연구자의 자료독점권을 3년간 보장해 줄 수 있는 법적 근거를 제공하고 있으며, 또한, 공개로 인해 공공이익이 훼손되는 경우에도 공개를 제한할 수 있다. 제16조에서는 해양과학조사자료 관리기관 업무를 수행할 수 있는 기관의 범위를 명시하고 있고, 제17조에서는 자료목록 제출에 관한 사항을 규정하고 있는데, 자료관리기관에서는 수집한 자료목록을 즉시 배포하도록 되어 있다. 제18조에서는 법 제23조의 권고이행 이전의 사전조치사항을 명시하고 있다.

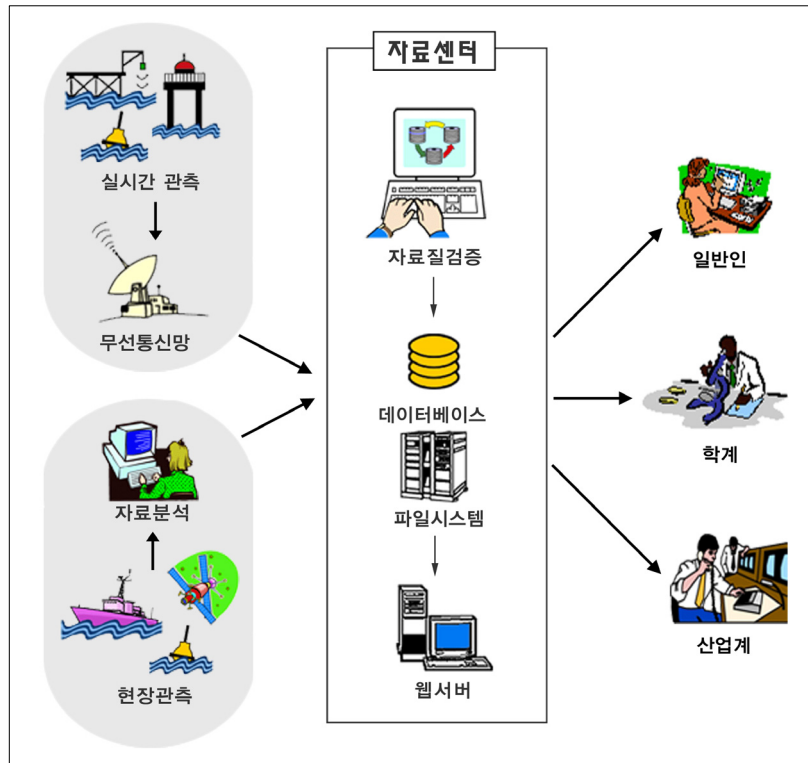
2) 해양자료의 유통

해양에서의 산업, 레저, 군사 활동에서 수온, 파랑, 적조, 해일 등의 해양자료는 꼭 필요한 정보이지만, 접근의 어려움, 관측기술의 부족, 높은 자료수집 비용으로 인해 충분한 해양자료를 확보하는 것은 쉽지 않으므로 한 번 수집된 자료를 유실 없이 보관하고 최대한 활용할 수 있도록 유통시키기 위한 자료관리, 유통체계들이 구축, 운영되고 있다.

선박 등을 이용한 해양조사에서 관측 장비를 이용하여 얻었거나 채취한 시료를 실험실에서 분석하여 얻은 해양자료는 조사가 이루어진 후 일정시일이 지나야 사용 가능한 형태로 정리되는데, 이런 현장 관측 자료의 관리에서는 관측일시, 위치, 관측책임자, 분석책임자, 연구책임자, 관측기기 등의 자료생산과 관련된 정보인 메타데이터(Metadata)를 같이 관리하여야 하는데, 이는 질 좋은 자료의 생산을 유도하고 사용자에게 좋은 자료를 제공하기 위함이다. 해양자료관리 분야에서는 일반적으로 메타데이터 데이터베이스를 일차적으로 구축하고 이를 바탕으로 실제 자료의 수집하고 데이터베이스 구축을 추진한다. 또한 최근에는 해양에서의 관측기술과 자료전송기술의 발달로 인해 관측탑, 해상부이, 연안고정점, 인공위성 등에서 실시간 관측 자료들이 많이 생산되고 있으며, 실시간자료의 종류는 다양하지 않지만 자료의 양이 많은 특징이 있으며, 자료의 생산과 동시에 자료제공을 실시하는 경우가 많다.

또한, 해양현상을 컴퓨터로 재현, 예측하기 위한 수치모델링에서 산출되는 자료는 해양예보정보의 신속한 제공을 위해 관리, 유통되며, 수치모델 산출자료는 2차원, 3차원 격자망에서 컴퓨터로 계산되는 값이므로 자료의 구성이 매우 규칙적인 반면 자료의 양이 방대한 특징이 있다. 이 자료들은 격자자료 저장을 위해 개발된 바이너리 파일포맷인 netCDF, HDF 등의 형태로 저장하여 유통하는 것이 일반적이다.

해양자료 및 정보의 유통 개념도를 <그림 1>에 나타 내었다.



<그림 1> 해양자료 및 정보의 유통 개념도

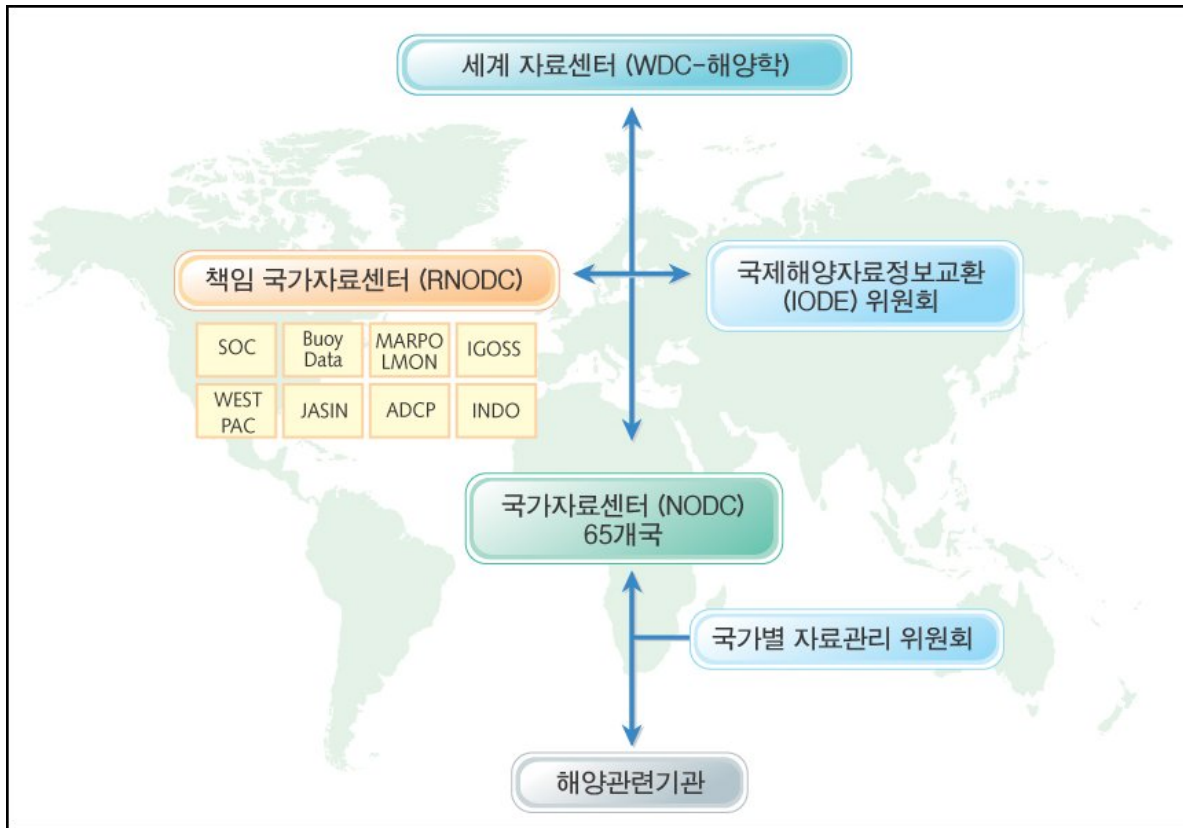
3) 국제 해양자료 및 유통체계

해양과학 분야에서는 연구자간의 자료 교환, 더 나아가 국가 간의 자료 교환의 중요성을 인식하고 오래 전부터 이를 해결하고자 노력하였다. 정부간 해양과학위원회 (IOC, Intergovernmental Oceanographic Commission)에서는 회원국 간의 자료 및 정보 유통망 구축을 위해 1961년에 IODE(International Oceanographic Data and Information Exchange)를 설립하였는데 IODE는 해양자료와 정보 교환 활성화, 국제적 자료와 정보 교환 포맷 및 방법의 표준화, 회원국 자료와 정보 관리기술 향상을 목적으로 하고 있다.

IODE는 자료 및 정보 교환 활성화를 위해 각 회원국의 국가자료센터(NODC, National Oceanographic Data Center), 지역별 자료중계 또는 자료별 관리의 별도 임무를 수행하는 책임국가자료센터(RNODC, Responsible National Oceanographic Data Center), 최종적으로 자료를 수집, 관리하는 세계자료센터(WDC, World Data Center - Oceanography)로 구성되는 국제 네트워크를 구축하였다.

우리나라에서는 국립수산과학원이 KODC(Korea Oceanographic Data Center)를 운영하고 있으며, 일본의 JODC(Japan Oceanographic Data Center)가 동북아 지역의 책임국가자료센터 역할을, 중국의 NMDIS(National Marine Data and Information Service)가 세계자료센터 역할을 수행하고 있고, 전 세계적으로는 2004년 기준으로 3개의 세계자료센터, 65개의 국가자료센터, 8개의 책임국가자료센터가 운영되고 있다.

IODE의 해양자료 유통체계를 <그림 2>에 나타내었다.



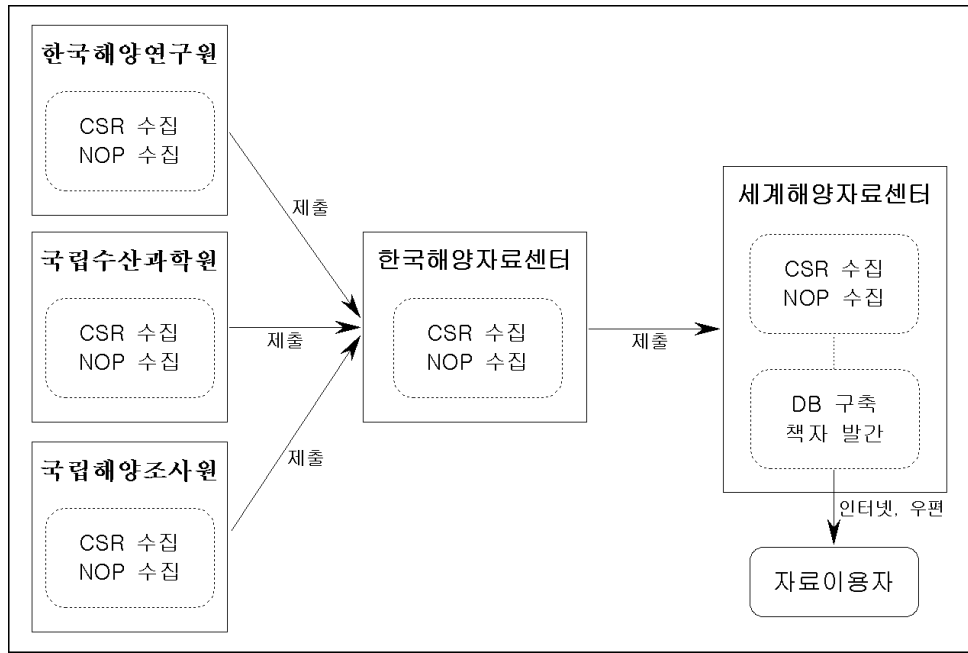
<그림 2> IODE의 해양자료 유통체계

4) 국내 해양자료 관리절차

국내에서 해양자료를 생산, 제공하는 기관에서는 국제 IODE 시스템, 국내 해양과학조사법, 각 기관별 자료관리규정에 따라 해양자료의 관리, 공개 업무를 수행하고 있다. 국제 IODE 시스템에 의하면, 각 국가에서는 매년 초에 전년도 항해보고서(CSR, Cruise Summary Report)와 국가해양프로그램(NOP, National Oceanographic Program)을 세계해양자료센터에 제출하도록 되어 있다. 항해보고서에는 전년도 수행한 해양조사의 기간, 해역, 조사항목, 조사책임자 등에 대한 정보를 수록하도록 되어 있으며, 국가해양프로그램에는 당해 연도에 계획하고 있는 조사의 예상 해역, 기간, 항목들을 수록하도록 되어 있다.

국내에서는 국립수산과학원이 KODC 역할을 수행하므로, 해양관련 각 기관에서는 매년 초에 CSR과 NOP를 수집하여 KODC에 제출하고 있으며, KODC에서는 각 기관으로부터 수집한 자료를 재정리하여 세계해양자료센터에 제출하며, 세계자료센터에서는 전 세계의 CSR, NOP를 수집하여, 책자형태로 만들거나 인터넷에 수록하여 전 세계에 제공하고 있다.

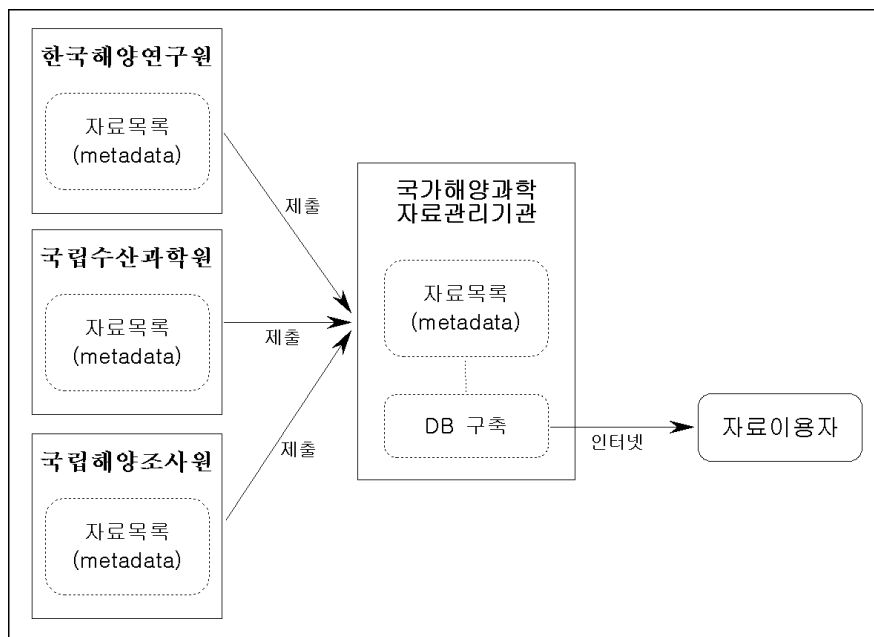
<그림 3>에는 IODE의 해양자료 관리절차를 나타내었다.



<그림 3> IODE의 해양자료 관리절차

국내 해양과학조사법에 의하면, 각 해양기관에서는 전년도에 수행한 해양조사결과의 목록(metadata)을 국가 해양과학조사자료 관리기관에 제출하도록 되어 있으며, 국가 해양과학조사자료 관리기관은 해양수산부에서 3년 단위로 신청기관에 대한 평가를 통해 지정하고 있는데, 현재 국립수산과학원이 자료센터 역할을 담당하고 있다. 한국해양연구원을 비롯한 해양기관에서는 매년 자료목록을 정리하여 자료 관리기관에 송부하고 있으며, 국립수산과학원에서는 해양과학정보시스템의 목록(metadata) 데이터베이스에 입력하여 관리하고 있다.

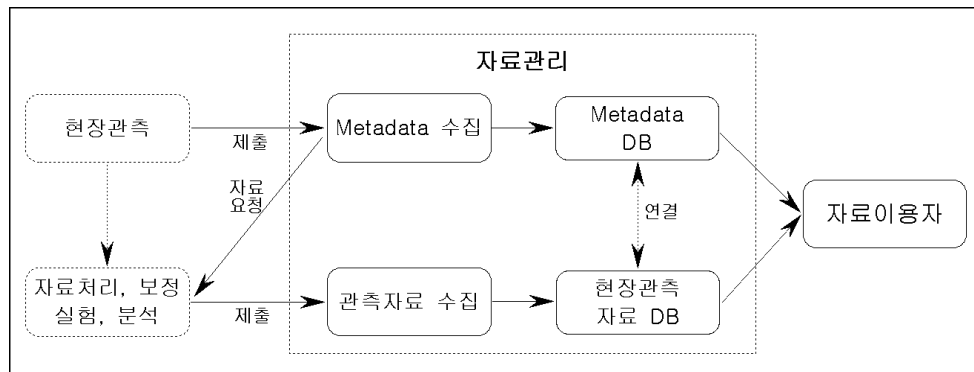
해양과학조사법에 의한 국가 해양자료 관리를 <그림 4>에 나타내었다.



<그림 4> 해양과학조사법에 의한 국가 해양자료 관리

또한, 각 해양관련 기관에서는 자체 기관에서 생산되는 자료의 관리를 위해 자체 자료관리규정을 운영하고 있으며, 필요에 따라 DB 시스템을 구축, 운영하고 있는데 한국해양연구원에서는 기관차원의 자료관리 업무 이 외에도 각 사업별로 자료관리 업무를 수행하고 있다. 특히, 최근 들어 다학제 간 연구수행으로 인해 연구개발 사업들의 규모가 커지면서, 연구참여자 간의 자료교환도 중요하게 인식되고 있으며, 연구결과를 집대성하기 위한 자료 DB 또는 GIS 구축들이 추진되고 있다. 이 경우에도 기존 자료관리 시스템과 마찬가지로 메타데이터를 일차적으로 수집하여 DB를 구축하고, 사업종료 후에 실제 자료를 수집하는 방식을 따르고 있다.

<그림 5>에는 대형 연구사업의 자료관리 절차에 대해 나타내었다.



<Figure 그림 5> 대형 연구사업의 자료관리 절차

2. 해양자료관리를 통해 본 시사점

1) 해양과학조사법과 과학기술기본법의 목적

“과학기술기본법”과 “국가연구개발사업 등의 성과평가 및 성과관리에 관한 법률 및 규정”은 범부처 차원에서 논문, 특허, 보고서 원문, 연구 장비 등의 수집과 관리를 통해 국가의 연구개발과제의 관리, 연구장비의 관리, 중복 과제의 탐색, 연구 현황의 분석, 연구자의 관리와 연구개발 통계 분석 등 국가의 연구개발사업을 전체적으로 관리하기 위한 목적이라고 할 수 있다. 이는 범부처적으로 추진되는 것으로서 국가 전체의 연구개발사업을 총괄 관리한다는 측면에서 의미가 있으며 현재 대부분의 국가 연구개발사업은 이 체제하에서 관리되고 있다고 할 수 있다.

또한 주관연구기관의 장 또는 전문기관의 장은 3개월 이내에 연구개발결과물을 해당되는 전담기관에 등록 또는 기탁해야 할 의무가 있으며 적용대상으로서는 국가연구개발사업을 통하여 창출된 연구성과 중에서 논문, 특허, 보고서원문(전자원문 포함), 연구시설·장비, 기술요약정보, 생명자원 중 생명정보, 소프트웨어, 생명자원 중 생물자원 및 화합물이며, 등록대상의 경우는 원문 및 전자문서 등 그 연구결과에 대한 정보를 등록하며, 기탁대상의 경우에는 연구결과에 대한 정보와 실물을 성과물 전담기관에 기탁한다고 되어 있다.

이와 같은 사실로 미루어 볼 때 과학기술기본법에서는 연구개발결과물로서 논문, 특허, 연구보고서, 연구시설, 기술요약정보 등을 관리하고 있으며 과학데이터의 성격을 띠는 것으로서 생명자원 중 유전자, 단백질 등의 생명정보와 화합물, 그리고 소프트웨어가 있는데 이는 과학기술 분야의 모든 데이터를 포함하지 않을 뿐만 아니라 현실적으로 데이터의 수집도 활발히 이루어지지 않고 있다 [2].

또한 연구자의 데이터에 대한 권리를 인정하기 위해 일반적으로 데이터의 공개를 제한하고 있는데 반해 과학기술기본법에서는 3개월 이내에 연구개발 결과물을 등록해야 한다는 사실은 연구과정에서 생

성된 과학 데이터에 초점을 맞추기보다는 연구결과물로서 산출된 논문, 특허, 보고서, 샘플 등을 수집, 관리하기 위한 목적이라 할 수 있다.

그러나 해양과학조사법에서는 국가예산으로 수행된 조사에서 생산된 자료의 원칙적인 공개와 필요에 따라 공개 범위를 조정할 수 있도록 되어 있다는 점과 데이터 관리기관에 자료의 목록을 제출할 의무가 있다는 점, 3년간 데이터 보유자에게 자료 독점권을 부여한다는 점, 관리기관은 제출받은 목록을 지체 없이 공개해야 한다는 점 등의 내용을 미루어 볼 때 해양 분야의 연구 조사 활동에서 얻어진 과학 데이터의 빠른 공개와 활용에 목적이 있다고 생각된다.

2) 과학 데이터의 범위 정의

과학기술기본법에서 취급하는 대상 자료는 논문, 특허, 보고서 원문, 연구장비, 기술요약정보, 생명자원 중 생명정보, 소프트웨어, 생명자원 중 생물자원 및 화합물을 해당 전담기관에 등록 또는 기탁하도록 되어 있다. 해외에서 과학 데이터를 정의할 경우에는 기관, 조직에 따라서 다르나 일반적으로는 학술논문과 특허는 제외하지만 연구 활동에서 생산된 데이터뿐만 아니라 그 밖의 회색문헌, 연구보고서, 연구과제 목록 등 관리되지 않는 학술정보는 모두 연구 데이터의 영역에 넣어 취급한다 [3][4][5]. 따라서 과학기술기본법에서 과학데이터를 정의한다고 할 경우에는 논문, 특허, 보고서 등의 학술정보 영역과 이미 과학기술기본법에서 기술하고 있는 유전자, 단백질 등의 생명정보, 화합물, 소프트웨어 등을 포함한 모든 과학기술 분야의 연구 활동 중에 생산된 모든 유형의 과학 데이터 영역으로 나누어 정의하는 것이 좋다. 그 이유는 현재의 국가과학기술종합정보시스템(NTIS) 체제하에서 이미 논문, 특허, 보고서, 연구장비 등은 범부처적으로 잘 관리되고 있는데 비해 생명정보, 소프트웨어 등 과학기술기본법에서 기술한 과학 데이터 영역의 해당 데이터의 수집과 관리 활동은 아주 미흡하기 때문이다 [2]. 따라서 국가과학기술종합정보시스템에서 취급하는 학술정보 영역과는 별도로 과학기술 전 분야의 과학 데이터 영역에 대해 다시 정의하고 목록 작성, 관리기관, 데이터 저작권, 공개 범위 등 과학 데이터 공유에 특정한 내용이 반영된 수정된 과학기술기본법이 다시 제정되어야 할 것으로 생각된다. 이를 위해서는 아래에 기술하는 해양과학조사법의 내용을 참조하는 것이 좋을 것으로 생각된다.

한편, 한국해양연구원의 경우, 연구보고서, 논문 등의 학술정보는 연구원 내의 해양도서관에서 수집과 관리 및 배포를 담당하고 있으며 해양 데이터는 전담부서인 해양자료센터에서 연구원 내에서 연구 활동 중에 생성된 모든 데이터를 수집, 관리, 배포한다. 또한 해양과학조사법에서는 해양 자료의 범위를 따로 정의하고 있지는 않으나 해양과학조사법에서 표현한 “국가예산으로 수행된 조사에서 생산된 자료”는 해양 데이터를 의미한다고 볼 수 있는데 그 이유는 해당 기관이 목록을 작성하고 관리기관에 제출해야 할 의무를 가지고 있다는 조항이 있는데 이는 논문이나 연구보고서를 대상으로 할 수 있는 작업이 아니기 때문이다. 따라서 해양과학조사법에서의 목록 작성은 현장관측 자료를 관리하기 위해 관측 일시, 위치, 관측책임자, 분석책임자, 연구책임자, 관측기기 등의 자료생산과 관련된 정보인 목록(메타데이터)를 작성하는 것이므로 해양과학조사법에서 취급하는 대상 자료는 현장 관측에서 생산된 과학 데이터를 의미한다고 할 수 있다.

3) 데이터 관리기관과 데이터 공유

국내의 해양 데이터의 경우를 살펴보면 국립수산과학원이 국내 해양 데이터의 관리기관으로 지정되어 다른 해양기관인 한국해양연구원, 국립해양조사원으로부터 매년 정리된 해양 자료의 목록을 입수하여 데이터베이스로 관리하고 있다. 이러한 사실은 동일 분야를 연구하는 기관들은 서로의 데이터를 공유할 필요가 있으며 따라서 과학 데이터의 경우에도 분야별로 데이터의 관리기관이 지정 또는 육성되

고 이를 중심으로 하여 동일 분야의 데이터들이 수집, 관리, 공유되도록 과학 데이터의 구축체제를 유도하는 것이 좋은 방향이 될 것으로 생각된다.

그러나 해양 분야인 경우에는 국내에 유사한 분야의 기관들이 있어서 정부가 하나의 관리기관(국립수산과학원)을 정하고 이를 중심으로 해양 데이터에 대한 목록을 수집, 공개함과 동시에 IODE의 한국의 국가자료센터(NODC, National Oceanographic Data Center) 역할을 수행하며 세계해양자료센터(WDC, World Data Center - Oceanography)에게도 목록을 제공할 의무가 있기 때문에 하나의 지정된 관리기관이 한국을 대표하는 창구 역할을 수행해야 할 필요가 있지만 다른 분야인 경우에는 그렇지 않을 수도 있다.

따라서 해양 분야와 같이 이미 국제적인 데이터의 교환체제를 갖추고 있는 분야는 그 체제하에서 지정된 관리기관을 중심으로 현재와 같이 데이터의 수집과 공유 활동을 지속해 나가면 좋을 것으로 생각된다. 다만 그렇지 않은 분야인 경우에는 국가에서 하나의 관리기관을 지정하고 이를 중심으로 데이터를 수집하기 보다는 각 기관이 스스로 자신들의 분야에 대한 과학 데이터를 자신들의 리퍼지토리에 목록과 함께 저장하고 다른 기관들이 시스템적으로 서로 연동하여 서로의 목록을 공유하는 체제를 취함으로써 각각의 데이터 리퍼지토리는 스스로 자신들의 데이터 목록을 작성하여 공개하며 같은 분야의 다른 리퍼지토리로부터 시스템적으로 서로 목록을 교환하는 형태가 바람직할 것으로 생각된다. 그러나 특정 분야의 국내 연구소가 유일하게 하나만 존재하며 그 분야의 대표적 성격을 띠는 경우에는 국가가 그 분야의 데이터 관리기관으로 지정하고 이를 중심으로 해당 분야의 데이터를 수집하여 데이터 센터로서 육성하는 것도 고려할 수 있을 것으로 생각된다.

한편 과학기술기본법에서는 연구개발 성과물을 관리하는 전담기관을 지정하고 있으며 논문, 특허, 연구보고서 등의 학술정보 영역은 전담기관에의 제출의무가 법제도적으로 규정되어 있기 때문에 전담기관 지정이 효과가 있다고 생각된다. 그러나 생명정보, 소프트웨어 등 과학 데이터 영역은 데이터의 다양성, 신뢰성, 저작권, 가이드라인 미비, 데이터 공유 문화 등 아직까지 국내에서는 연구자들이 쉽게 데이터를 제출할 수 있는 여건이 마련되어 있지 않기 때문에 생명정보 등에 대한 전담기관의 데이터 수집은 그리 활성화되어 있지 않다. 따라서 관리기관 지정에 앞서 과학 데이터의 제출과 공유를 방해하는 여러 가지 원인을 분석하여 다양한 각도에서 데이터의 공유 및 활용의 활성화를 촉진할 수 있는 방안을 검토해야 한다.

4) 자료 목록의 작성

해양과학조사법에서 가장 특징적인 요소는 국가 관리기관이 해양 관련 기관들에게 연구조사 활동을 통해서 수집된 해양 자료에 대해 목록 제출을 요구할 수 있는 법적 근거를 마련하고 있다는 점이라 할 수 있다. 데이터에 대한 목록 작성은 과학 데이터를 수집하기 위한 절차에 있어 첫 번째 단계라 할 수 있으며 데이터를 외부에 공개하고 데이터의 내용을 이해할 수 있는 가장 기본적인 작업이다. 특히 탐사나 관측에 있어서는 탐사 대상과 일정이 정해진 계획대로 이루어지는 경우가 많기 때문에 얻어질 데이터를 미리 예상할 수가 있어서 목록 작성과 제출이 비교적 용이하다고 할 수 있으며 따라서 국내와 해외의 관련기관들이 우선 데이터 수집 계획에 따른 목록을 서로 공개하는 경우가 많다.

그러나 과학기술기본법에서는 자료의 목록 작성이나 제출에 대한 내용이 없어서 과학 데이터의 수집을 위한 목적은 없다고 보아야 한다. 따라서 위에서 기술한 것처럼 학술정보 영역과는 별도로 과학 데이터의 영역을 정의하고 이러한 데이터를 수집, 공유, 활용하기 위한 여러 가지 시책들을 검토해야 한다.

또한 현재 해양 관련 기관들이 작성하여 관리기관에 제출하고 있는 목록도 조사 대상이나 구역, 실험장비 등을 파악하기 위한 목록이며 데이터 자체를 서로 공유하기 위한 표준화된 메타데이터 항목과

카탈로깅 시스템을 가지고 있지 않다. 따라서 해양 분야도 관련 기관들이 모여서 데이터 제출과 공유를 위한 표준화된 데이터 리포지토리용의 카탈로깅 시스템을 개발하고 함께 사용할 필요가 있다.

5) 자료 공개 범위 제한

해양과학조사법에서는 각 기관의 기관장이 자료의 공개를 3년간 제한할 수 있도록 되어 있다. 이는 연구를 수행한 연구자의 데이터에 대한 자료 독점권을 3년간 보장해 줄 수 있는 법적 근거를 제공하고 있으며, 또한 공개로 인해 공공이익이 훼손되는 경우에도 공개를 제한할 수 있다. 연구를 수행한 연구자에게 일정한 기간 동안 자신의 데이터에 대한 독점권을 부여한다는 사실은 데이터의 획득을 위한 연구자의 노력을 인정한다는 의미이며 이러한 노력이 인정되지 않을 경우 대부분의 연구자들은 자신들의 데이터를 전혀 공개하지 않을 것이다.

한편 국가연구개발사업의 관리 등에 관한 규정 제18조(연구개발결과의 공개) 규정에서는 보안과제로 분류된 경우에는 최대 3년, 지식재산권의 취득을 위하여 공개 유보를 요청한 경우 1년 6개월 이내, 참여기업이 영업비밀 보호 등의 사유로 비공개를 요청한 경우 1년 6개월 이내로 연구개발 결과의 공개를 제한할 수 있으며 중앙행정기관장의 승인이 있으면 최대 3년 범위에서 비공개 기간을 연장할 수 있다고 하고 있다.

상기의 내용으로 유추해 볼 때 해양과학조사법은 과학 데이터에, 국가연구개발사업의 관리 규정은 논문, 특허, 보고서 등의 학술정보에 초점을 두고 있다고 생각되며 공개 제한의 대상이 되는 데이터는 메타데이터가 아니라 원시 데이터라 생각된다. 또한, 과학 데이터의 경우에는 국가가 데이터의 공개를 강제하거나 공개 시점을 정하게 되면 연구자들이 데이터를 제출하지 않거나 신뢰성이 떨어지는 데이터를 제출할 가능성이 높으며 최근에는 데이터의 제출의 활성화를 유도하기 위해 연구자 자신이 자신의 데이터에 대한 공개 권한을 가지도록 데이터 카탈로깅 시스템을 설계하는 경우도 있다. 그러나 이 경우에도 데이터에 대한 메타데이터는 제출 즉시 공개되며 원시 데이터에 대한 공개 권한을 연구자가 가진다고 할 수 있다.

6) 자료의 관리기관에의 위탁

현재 해양과학조사법 시행령 제12조에서는 해양과학조사법의 적용을 받는 기관으로서 대학을 명시하고 있으나 현재 대학의 자료는 관리되지 않고 있으며 제14조에서는 자료의 위탁을 규정하고 있으나 관리기관에 자료를 위탁하는 해양기관은 없다고 한다. 이는 주로 국내 3개 해양기관(한국해양연구원, 국립수산과학원, 국립해양조사원)을 중심으로 수집된 자료의 목록이 공개되고 있음을 뜻하며 대학에서 수집된 해양 자료는 공개 또는 관리가 되지 않고 있다고 할 수 있다. 따라서 해양 관련 연구 활동이 대학에서도 매우 활발하게 이루어지고 있음을 볼 때 각 대학의 해양 연구자들이 자신들이 수집한 데이터를 등록 내지 제출할 수 있는 국가적인 데이터 리포지토리가 필요할 것으로 생각되며 해양과학조사법에서 지정한 관리기관이 해양 분야의 데이터 리포지토리를 개발, 관리하면 효율적으로 데이터의 목록뿐만 아니라 각 대학의 원시 데이터 자체도 제출받아 관리할 수 있을 것으로 생각된다.

과학기술기본법에서는 논문, 연구보고서 등의 학술정보 원본과 화합물 등의 샘플 기탁은 전담기관(화합물의 경우 한국화학연구소 화합물은행)이 담당하고 있으며 농업, 의학, 생물자원 등도 부처별로 샘플을 기탁 받는 전담기관이 지정되어 있으나 샘플 또는 학술정보에 포함되지 않는 원시 데이터는 등록, 제출하거나 국가적으로 관리할 수 있는 세부 사항을 규정하고 있지 않다.

7) 자료의 저작권

국가연구개발사업 관리 규정 제20조에서는 연구개발사업 수행에서 얻어지는 연구기자재, 연구시설, 시제품, 연구노트 등의 유형적 결과물은 주관연구기관의 소유로, 지식재산권, 판권 등 무형적 결과물은 주관연구기관의 소유 또는 협약에 따라 참여기관 단독이나 주관, 참여기관 공동 소유로 한다고 명시하고 있다. 따라서 국가연구개발사업 관리규정에서는 연구자가 연구과정 중에서 취득한 데이터는 일반적으로 소속된 연구기관의 소유로 한다고 볼 수 있다.

한편 해양과학조사법에서는 자료의 소유권에 대해서는 따로 규정하고 있지 않으나 국가기관 또는 지방자치단체의 장 및 대통령령이 정하는 법인의 대표자가 국가 예산으로 실시한 해양 조사 자료를 성실히 관리할 의무가 있고 공개 및 공개 제한의 권리를 가진 것으로 명시되어 있다.

연구자들이 얻은 데이터에는 연구과정에서 얻은 것들과 그렇지 않은 것들도 있고 연구자가 저술한 학술논문은 연구자의 창작성이 들어가 있으므로 연구자에게 저작권을 인정하는 것처럼 데이터의 특성으로 볼 때 연구자의 창작성이 들어간 데이터도 있을 수 있으므로 데이터의 다양한 생성 형태, 생성 동기, 생성 주체 등을 고려해 볼 때 데이터의 저작권을 일률적으로 정하기는 어렵다고 볼 수 있다. 그러나 과학기술기본법에서 학술정보 영역과는 별도로 과학 데이터의 영역에서 정의할 경우에는 해양과학조사법에서 자료의 소유권을 언급하지 않았다고 하더라도 과학 데이터의 소유권에 대해 명시할 필요가 있다고 생각되며 특히 데이터 리파지토리에 과학 데이터를 제출할 경우에는 제출자에게 데이터의 저작권 보유 여부를 묻고 있기 때문에 데이터에 대한 저작권이 없는 경우에는 데이터를 제출할 수 없도록 되어 있다. 또한 자신이 소유권을 가진 데이터를 공개할 경우에도 타인들이 어떤 범위로 자신의 데이터를 이용할 수 있다는 저작권의 이용 범위를 자신이 결정하기 때문에 기본적으로는 데이터의 소유권을 가진 사람이 데이터 리파지토리에 자신의 데이터를 올릴 권한이 있다고 할 수 있다. 따라서 국가 또는 소속기관이 데이터에 대한 소유권을 가질 경우 데이터를 취득한 연구자가 어떤 권리로 데이터 리파지토리에 올릴 수 있는지를 정의할 필요가 있다고 생각된다.

데이터 소유권이 어떻게 정의되든지 가급적이면 연구자가 스스로 자신의 데이터를 쉽게 올릴 수 있는 환경을 만들어서 많은 예산과 연구자의 노력으로 얻어진 귀중한 과학 데이터가 보다 많이 공개되고 다시 활용되어 국가의 과학기술 발전의 데이터 인프라로서 기능하는 체제를 만들 필요가 있다고 생각된다.

8) 데이터관리규정

현재는 과학기술기본법이나 국가연구개발사업 관리 규정 또는 해양과학조사법에서는 데이터관리규정에 대해 언급하고 있지는 않다. 그러나 최근 한국해양연구원 산하 극지연구소에서 극지데이터 관리규정과 요령을 만들었으며 관리규정의 중요한 내용을 살펴보면, 제2조(용어의 정의)에서 데이터를 원시 데이터, 처리 데이터, 결과 데이터로 나누고 있으며, 제3조(소유권)에서 데이터 소유권은 극지연구소에 있으며, 제6조(데이터관리계획의 수립 및 이행)에서는 연구책임자가 데이터관리계획을 수립하고 데이터 관리 의무를 다하기 위해 메타데이터, 원시, 처리, 및 결과 데이터를 등록해야 한다고 하고 있다. 또한 제11조에서 16조에서는 데이터의 원칙적 공개와 공개 제한, 데이터 활용자의 활용 신청서 제출 및 결과 보고 의무를 명시하고 있다.

극지데이터 관리 요령 제8조에서는 관리 대상 데이터로서 디지털, 아날로그 데이터와 원본 및 복제 데이터, 미생물, 동식물, 시료, 그리고 연구과제 필요에 의해 생성된 연구 데이터와 연구 데이터의 획득을 위한 활동과 인프라 운영과정에서 부가적으로 획득한 인프라 및 연구환경 데이터도 관리 대상으로 하고 있다.

그러나 제9조(관리 예외 데이터)에서는 아래의 어느 하나에 해당하는 데이터는 해당 법령이 정하는 바에 따른다고 하여 관리 대상에서 제외하고 있다.

1. 극지연구 중 작성되는 기록 및 연구노트
2. 특허·실용신안·기술·노하우 등의 지식재산
3. 논문·보고서·초록·발표데이터 등 최종 연구성과
4. 그 밖에 별도의 법령 및 규정에 의해 관리되는 데이터

그리고 극지데이터의 관리를 위해 제10조에서 제12조에 걸쳐 한국극지데이터센터와 극지데이터관리위원회의 운영을 명시하고 있으며 제14조에서 제17조에서 연구책임자는 데이터관리계획을 수립하고 데이터센터장은 이를 심의, 승인하고 연구관리부서에 통보하며 데이터관리계획이 이행되는지 점검하여야 한다고 명시하고 있다. 또한 제19조에서는 데이터센터장과 연구책임자가 데이터의 보존 기간 지정과 사용자 및 데이터별 접근 권한을 관리한다고 되어 있으며, 제21조에서 제26조에서는 연구책임자의 관리 아래 데이터의 획득과 메타데이터 작성, 데이터 보호, 데이터의 등록과 품질관리를 행하도록 되어 있다. 제29조에서 제31조에서는 데이터의 공개 원칙과 공개 제한범위, 공개 방법이 기술되어 있고 제32조에서 제34조까지는 데이터의 사용 신청과 데이터 제공 및 데이터 사용에 따른 의무 사항을 명시하고 있으며 제36조에서 38조에는 데이터의 폐기 사항이 명시되어 있다.

이 외에도 제39조에서 제47조까지는 데이터 관리체계의 구축 및 운영, 데이터의 보존 및 재난대책, 외부 데이터의 수집 및 관리, 교육 훈련, 보상 및 벌칙, 그리고 부칙이 기술되어 있으며 별지에는 각종 양식과 가이드라인이 소개되어 있다.

상기의 극지데이터 관리규정과 요령은 국내의 연구기관이 연구과정에서 생성된 자신들의 과학 데이터를 관리하기 위한 훌륭한 지침이라 생각되며 과학기술기본법에서 과학 데이터의 영역을 정의하는데 참고가 될 것으로 생각된다 [6]. 또한 현재 대학에서 생성되는 데이터에 대해서는 전혀 관리가 이루어지지 않고 있어서 국내의 분야별 연구기관이 자신의 분야에서 데이터센터(데이터 리포지토리)로서의 역할을 충실히 수행할 경우 대학에서 생성되는 과학 데이터도 모두 이 분야별 데이터센터를 중심으로 관리될 수 있을 것으로 생각된다.

III. 결론

본 연구에서는 한국해양연구소와 부속 극지연구소의 과학 데이터 수집 사례를 과학기술기본법과 비교하여 분석함으로써 국내의 과학 데이터 관리체제 구축에 참고가 될 수 있는 중요한 시사점을 제시하였다.

아직까지 국내의 막대한 연구개발 예산이 투입되어 수행된 연구개발 과정에서 생성된 귀중한 공공재로서의 과학 데이터는 빅데이터의 시대를 맞아 데이터의 수집과 분석의 중요성이 더욱 부각되었음에도 불구하고 본 연구에서 살펴본 바와 같이 국가적 차원에서 관리되지 못하고 있는 실정이다. 더구나 연구개발 성과물을 관리하는 과학기술기본법 및 국가연구개발사업 관리 규정에서도 과학 데이터를 관리하기 위한 세부 규정은 매우 미흡한 실정이어서 이에 대한 대책이 요구되고 있다.

그러나 한국해양연구원과 부속 극지연구소에서는 오래 전부터 전담부서를 두어 데이터의 수집과 관리 및 데이터 공개 활동을 꾸준히 추진해 오고 있으며 관련 법과 규정이 정비되어 있어서 국가의 과학 데이터 관리체제를 구축하는데 크게 참고가 될 것으로 생각된다.

참고문헌

- [1] Clifford Lynch, “Jim Gray’s Fourth Paradigm and the Construction of the Scientific Record”, in Tony Hey et al. (eds.), *The Fourth Paradigm*, 2009, p. 177.
- [2] “국가과학데이터의 효율적 관리 및 활용을 위한 법제도 기본 연구”, 윤종민 외, KISTI 위탁연구보고서 p. 213, 2012년 9월,
- [3] 한국과학기술정보연구원, “국가과학데이터 공유·융합체제 구축에 관한 연구”, 2011, p. 4
- [4] NSF, *Proposal and Award Policies and Procedures Guide, Part I – Proposal Preparation & Submission Guidelines GPG, Chapter II, B.2.d.(i)*.
- [5] ANDS, “ANDS Guide: What is Reserach Data?”
- [6] NSF, “*Proposal and Award Policies and Procedures Guide, Part I – Proposal Preparation & Submission Guidelines GPG, Chapter II, C.2.j*.”