

## 해양관측데이터 관리를 위한 결정요소 도출\*

### Deriving the Determining Factor for the Management of Oceanographic Data

김 선 태\*\* · 이 태 영\*\*\* · 김 용\*\*\*\*  
Sun-Tae Kim · Tae-Young Lee · Yong Kim

#### 차 례

1. 서 론	5. 결정요소 도출
2. 해양관측 분석	6. 결 론
3. 해양관측데이터 유형별 특징	· 참고문헌
4. 해양관측데이터 검색요소 분석	

#### 초 록

해양관측데이터의 관리와 서비스를 위한 결정요소를 두 가지 측면에서 도출하였다. 1) 해양관측데이터가 생성되는 관측의 유형들과 해양과학의 기본적 구분인 해양물리, 해양화학, 해양생물, 해양지질 분야에서 생산된 대표적인 유형의 샘플 데이터를 분석하였다. 2) 국내에서 서비스되고 있는 해양관측데이터 서비스(KODC, NFRDI, KHOA)를 조사하여 검색에 사용되는 검색요소를 분석하였다. 분석결과 9개 영역(일반, 관측선, 위성, 관측기기, 관측소, 공간정보, 프로젝트, 관측데이터, 자료처리)에서 42개 결정요소를 도출하였다.

#### 키 워 드

연구데이터, 과학데이터, 메타데이터, 해양관측데이터

\* 본 논문은 전북대학교 대학원 박사학위논문(2012.2)의 일부를 수정·요약한 것임.

\*\* 한국과학기술정보연구원 과학데이터연구팀 선임연구원 (제1저자)  
(Senior Researcher, Scientific Data Research Team, Korea Institute of Science and Technology Information, stkim@kisti.re.kr)

\*\*\* 전북대학교 문헌정보학과 교수 (공동저자)  
(Professor, Library and Information Science, Chonbuk National University, taehyun@jbnu.ac.kr)

\*\*\*\* 전북대학교 문헌정보학과 조교수, 인문영상연구소 연구원 (교신저자)  
(Professor, Library and Information Science, Chonbuk National University, yk9118@jbnu.ac.kr)

• 논문접수일자: 2012년 6월 8일  
• 최종심사(수정)일자: 2012년 7월 14일  
• 게재확정일자: 2012년 7월 16일

## ABSTRACT

This paper derives determining factor for the management of oceanographic data in two ways, 1) The type of oceanographic observation and the raw data which were collected from marine physics, marine chemistry, marine biology, marine geology area were analyzed, 2) The services of the KODC(Korea Oceanographic Data Center), NFRDI(National Fisheries Research & Development Institute), KHOA(Korea Hydrographic and Oceanographic Administration) were analyzed to derive metadata elements for retrieval. After analyze, the 42 deciding factor were derived in the 9 areas (general, Observer, satellites, observation instruments, observatories, space, information, projects, and observational data, data processing).

## KEYWORDS

Research Data, Scientific Data, Metadata, Oceanographic Data

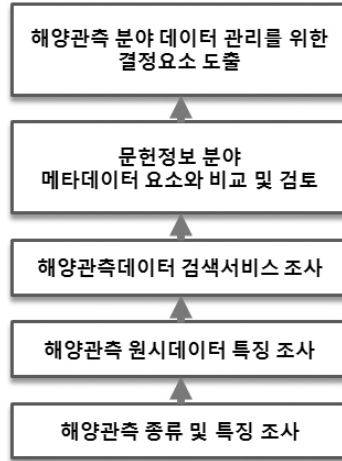
## 1. 서론

최근 들어 하드웨어의 눈부신 발전과 고성능 네트워크, 센서장비의 확산이 여러 연구 분야에서 다양한 포맷의 데이터를 발생시키고 있다. 이는 마크 와이저(1988)의 주장대로 '보이지 않는 컴퓨팅'이 실현되고 있음을 반증한다. 이러한 현상은 연구패러다임의 변화를 가져왔다. 가설을 수립하고 수립된 가설을 증명하기 위해 실험 방법론을 설정하고 그 조건에 맞는 데이터를 수집하여 분석 처리한 후 결과와 결론을 도출하는 일반적인 방법론과 별도로 데이터가 연구의 중심도구로 활용되는 제 4세대 연구패러다임이 도래한 것이다. 데이터를 수집하는 시간이 연구의 대부분을 차지(Green 2009)하기 때문에 수많은 관측, 실험 장치 등에서 생산되는 데이터를 사장시키지 않고 효과적으로

이용하는 방법이 모색되어야 한다.

특히 해양관측 분야에서 생산되는 데이터는 전 지구적 문제를 해결하기 위한 주요한 정보 원으로써 보존 및 재활용될 수 있는 인프라가 필요하다. 이를 위해서는 해양관측 분야에서 생산되는 데이터의 특성과 기존 서비스 시스템들의 분석을 통해 해양관측데이터의 관리와 서비스를 위한 결정요소를 도출하는 연구가 우선적으로 수행되어야 한다.

이에 본 연구에서는 해양관측의 종류와 특징을 살펴보고, 해양관측 분야에서 생산되는 원시데이터의 특징 및 데이터를 관리하기 위해서 필요한 기본적인 메타데이터 요소를 분석하였다. 원시데이터 분석의 경우 해양물리, 해양생물, 해양지질, 해양화학 분야의 샘플 원시데이터를 한국해양연구원에서 제공받았으며, 한국해양자료센터, 국립수산과학원, 국립해양조사



〈그림 1〉 연구절차

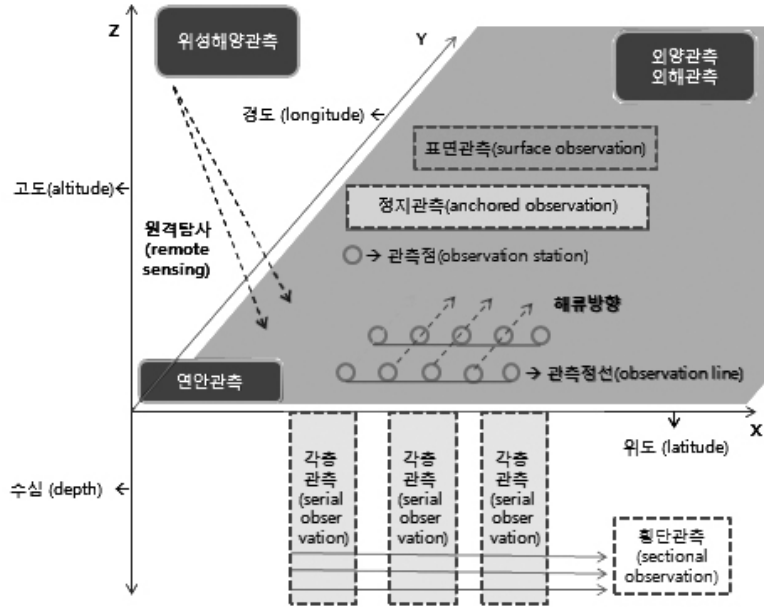
원의 웹서비스를 분석하여 해양관측데이터의 검색을 위해 활용되고 있는 메타데이터 요소를 분석하였다. 이상의 과정을 거쳐 결정요소를 도출하였으며, 마지막으로 문헌정보를 기술하는데 사용되는 메타데이터 표준들이 분석된 메타데이터 요소들에 적용될 수 있는지 검토하였다.

## 2. 해양관측 분석

해양관측을 구분하는 방법에는 여러 가지가 있을 수 있다. 외양관측은 외해관측(offing ocean observation)이라고도 한다. 외해관측은 항해 중에 표면수 성질의 관측, 즉 표면수의 수온·수색·해조류·파랑·해빙·생물 등의 관측을 한다. 또 채취한 해수의 성질을 측정하는 것을 표면관측(surface observation)이라 한다.

표면관측과는 달리 관측선을 멈추고 때로는 위치를 정하여 각 깊이에 대하여 수온·수질·해조류·수색·투명도·생물·수심·저질 등을 측정하는 것을 각층관측(serial observation)이라고 한다. 특정 관측지점에서 관측하는 것을 정점관측이라 하며, 관측점을 이은 줄을 따라하는 관측을 정선관측이라 한다(박용안 2005; 조규대 2009). 〈그림 2〉는 이상의 관측구분을 그림으로 표현한 것이다. X선은 위도, Y선은 경도, Z선은 고도와 수심을 의미한다. 연안관측이 의미하는 것은 내륙과 가까운 연안지역에서 이루어지는 관측을 의미한다. X, Y의 좌표값이 구성하는 평면은 수면을 의미한다. 각층관측은 Z좌표의 수직적인 관측을 의미한다. 횡단관측은 정선을 따라서 연속적으로 수행하는 각층관측을 의미한다.

이상의 내용을 생성되는 관측데이터와 관측선의 측면에서 살펴보면 네 가지 내용이 도출



〈그림 2〉 해양관측의 종류

된다. 첫째, 해양관측데이터의 생성에 있어 관측선, 관측기기, 관측위치, 관측시간, 관측방법은 맥락정보로 매우 중요한 의미가 있으며, 데이터와 함께 관리되어야 한다. 둘째, 하나의 관측데이터에 여러 개의 관측기기가 사용될 수 있으며, 각각의 관측요소에 대한 관측기기에 대한 설명이 필요하다. 셋째, 하나의 관측데이터에는 하나 이상의 다양한 관측요소의 값이 기술되어 있다. 넷째, 하나의 관측데이터는 하나 이상의 관측선에서 생성된다. 이상에서 도출된 사항은 해양관측데이터를 기술하기 위한 메타데이터 요소를 선정할 때 반드시 고려되어야 할 사항이다.

### 3. 해양관측데이터 유형별 특징

#### 3.1 해양물리 분야

해양물리 분야의 관측요소로는 수온, 염분, 해류, 조류, 조석, 파랑, 해면변화, 해수광학특성 및 수중음향 등이 있다. 여기서는 수온과 염분데이터, 해류데이터, 위성데이터의 특징을 분석하였다.

##### 3.1.1 수온·염분 관측데이터 특징

〈표 1〉은 수온과 염분의 측정을 위한 관측기구에서 생산된 원시데이터의 일부를 보여준다. 첫 번째 '10, 20, ...32'가 의미하는 것은 수심을 의미하며, 두 번째와 세 번째 항목은 각각

수온과 염분의 값을 의미한다. 관측기기의 특성상 원시데이터에 관측행위와 관련된 정보가 없이 <표 1>에서 보여주는 것과 같이 관측된 값만으로 구성된 데이터도 존재한다. 따라서 이러한 데이터의 검색을 위해서는 메타데이터가 반드시 필요하다. <표 2>는 <표 1>의 원시데이터를 관리하기 위해서 한국해양연구원에서 관리하고 있는 메타데이터 형식이다. 원시데이터를 기술하기 위해 사용되고 있는 메타데이터 요소는 <표 3>과 같다. 식별자(obs\_id)는 관측 정점의 식별자로서 관측데이터에 부여되는 식별자가 아니다. 프로젝트ID는 관측데이터가 생성되는데 지원을 하고 있는 프로젝트 식별자이다. 수심(bot\_depth)은 관측정점에서 해저로의 수심을 의미한다. 관측선(vessel)은 관

측시 사용한 조사선의 식별자를 의미한다. 관측장비(method\_instrument)는 수온과 염분을 측정할 장비를 기술하기 위한 요소로서, <표 2>에서는 'Smart.EG&C' 회사의 CTD 장비를 사용했음이 기술되어 있다. 마지막 항목인 정밀도(data\_precision)은 소수점 이하 유효숫자를 의미한다.

### 3.1.2 해류관측데이터 특징

<그림 3>은 해류관측 원시데이터의 모습을 보여준다. ①번은 관측을 실시한 일시로서 1989년 5월 4일 14시 40분에 관측이 실시되었음을 의미하며, 10분 간격으로 관측이 이루어졌음을 알 수 있다. ②번은 수온을 의미하며, ③번은 염분을 의미한다. ④번은 해류의 속도(cm/sec)

<표 1> 해양물리 분야 수온·염분 원시데이터

10 25.802 32.334	20 22.163 32.485	30 14.926 32.563
12 25.282 32.340	22 21.721 32.464	32 11.286 32.761

<표 2> 해양물리 분야 수온·염분데이터의 메타데이터 작성 예

obs_id	project_id	start_date	start_latitude	start_longitude	bot_depth	vessel	data_group	method_instrument	data_precision
P00000 6362	G0300014	19861119 1415	34.5008333	128.5022222	66	8	1001	CTD (Smart, EG&G)	3

<표 3> 수온·염분데이터를 관리하기 위한 메타데이터 요소

메타데이터 요소	표시상수	메타데이터 요소	표시상수
obs_id	식별자	BOT_DEPTH	관측지역 수심
PROJECT_ID	프로젝트ID	VESSEL	관측선ID
START_DATE	관측시작일시	DATA_GROUP	데이터 그룹
START_LATITUDE	위도	METHOD_INSTRUMENT	관측장비
START_LONGITUDE	경도	DATA_PRECISION	정밀도

```

89 5 41440① 214 41.17 40.02② 13.31 33.81③ 25.43④ 38.18 105.8⑤
89 5 41450 214 42.30 40.02 13.44 33.70 25.32 38.46 105.1
89 5 415 0 214 42.47 39.79 13.13 33.76 25.43 77.94 103.3
89 5 41510 214 43.35 39.87 13.22 33.75 25.40 80.18 105.8
89 5 41520 214 43.52 40.02 13.35 33.78 25.40 84.66 100.2
89 5 41530 214 43.52 40.33 13.69 33.77 25.32 81.02 97.8
89 5 41540 214 43.52 40.02 13.47 33.67 25.29 83.26 97.8
89 5 41550 214 43.87 40.10 13.44 33.77 25.37 85.78 91.1
89 5 416 0 214 44.13 40.18 13.50 33.79 25.37 86.06 89.7
89 5 41610 214 44.48 39.87 13.35 33.64 25.29 83.54 97.1
    
```

〈그림 3〉 해류관측데이터

〈표 4〉 해류데이터를 관리하기 위한 메타데이터 작성 예

OBS_ID	Project_id	start_date	end_date	s_latitude	s_longitude	bot_depth	vessel	data_group	Method_instrument	data_precision
P10000001	G0300013	198506171150	19850621500	34.184722	125.595833	67	13	1006	Aanderaa RCM-4, cm/sec	2

를 의미하며, ⑤번은 북쪽을 0도로 한 각도 값을 의미한다. 〈표 4〉는 〈그림 3〉의 원시데이터를 기술하기 위한 메타데이터 요소로서 한국해양연구원에서 관리하고 있는 메타데이터 형식이다. 해류데이터의 특성상 수온·염분데이터의 기술항목과의 차이점은 메타데이터 요소 중에 관측시작일시와 관측종료일시가 함께 기술되어 있는 점이다.

### 3.1.3 위성해양관측데이터 특징

위성해양관측데이터는 제공되는 내용에 따라 원시자료와 분석 자료로 구분된다. 위성해양관측데이터는 데이터의 형식에 따라 이미지 영상데이터와 ASCII로 구성된 텍스트 데이터로 구분된다. ASCII파일 형식은 영상데이터보다 용량이 큰 특징을 가지고 있다.

〈표 5〉는 국립수산과학원에서 제공하는 프로그램을 활용하여 동북아해양위성정보시스템에서 수집한 샘플파일을 대상으로 ASCII파일로 변환한 결과를 보여준다. 해양관측영상데이터가 ASCII 파일변환 후 데이터의 크기가 22배 커진 것을 확인할 수 있다. ASCII로 구성된 파일은 〈표 6〉과 같이 헤더정보와 데이터 항목 정보로 구성되어 있다. ASCII로 변환 테스트를 통해 생성된 데이터 파일의 실제 내용은 〈그림 4〉와 같다. 〈그림 5〉는 테스트를 통해 영상데이터를 ASCII로 변환한 파일을 텍스트 편집기에서 불러온 화면이다. 마지막 레코드의 행 번호가 '400004'임을 확인할 수 있다. 하루에 한 장의 해양위성사진을 촬영하고 1년 동안 촬영된 데이터를 모두 ASCII로 변환한다고 가정했을 때 14억개 이상의 레코

〈표 5〉 위성영상데이터와 ASCII변환 파일용량 비교

구분	용량
해양관측 영상데이터 샘플데이터 <sup>1)</sup>	3,909KB
ASCII 변환 데이터 샘플데이터 <sup>2)</sup>	88,679KB

〈표 6〉 위성관측 영상데이터의 ASCII 파일의 메타데이터 요소

구분	메타데이터 요소
Header	일시(date), 위성(satellite), 생성자(maker)
데이터 항목	경도(longitude), 위도(latitude), 수온(temp)

Header 정보		데이터항목정보			
Date : sample_2		0	10	20	
SATLLITE : NOAA AVHRR (Advanced Very High Resolution Radiometer)		3999996	142.098	25.010	27.44
MAKER : National Fisheries Research and Development Institute		3999997	142.110	25.010	27.44
longitude latitude temp		3999998	142.122	25.010	27.44
122.114 30.92017.08		3999999	142.135	25.010	27.30
122.126 30.92013.86		4000000	142.147	25.010	27.16
122.138 30.92013.86		4000001	142.159	25.010	27.30
122.150 30.92099.99		4000002	142.171	25.010	27.30
.....		4000003	142.183	25.010	27.44
122.211 30.92016.10		4000004	142.196	25.010	27.44
122.224 30.92016.52					
122.236 30.92016.10					
122.248 30.92016.80					

〈그림 4〉 위성영상데이터

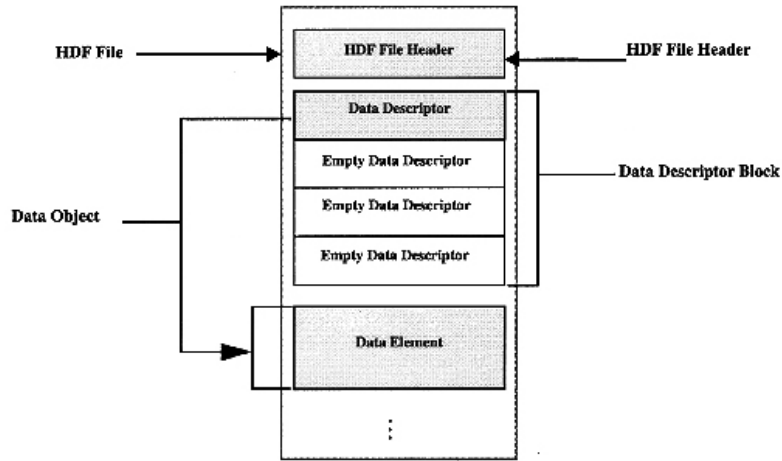
〈그림 5〉 ASCII 변환 데이터

드가 생성된다. 이는 대용량 데이터를 대상으로 특정 지역의 수온이 일정 온도 이상인 위성 사진을 검색하고자 할 때는 많은 기술적인 문제들이 고려되어야 함을 의미한다.

한편, 위성데이터의 표준으로 인정받고 있는 HDF 형식의 데이터의 특징을 살펴보면 다음과 같다. NASA는 계층적 데이터포맷(Hierarchical Data Format; HDF)을 기본으로써 선택하였다. 따라서 위성관측 분야에서 HDF파일형식의

원시데이터가 생산되고 유통되고 있다. HDF 파일포맷은 기존의 JPEG, TIFF, BMP 등과 같이 하나의 이미지만을 저장하는 것이 아니라 적절한 HDF 자료구조를 사용하여 하나의 HDF 파일에 문자데이터, 수치데이터, 이미지 데이터 등의 여러 종류의 데이터를 동시에 저장할 수 있다. HDF 파일은 한 개의 파일 헤더(file header)와 적어도 한 개의 데이터 설명자 블록(data descriptor block), 및 0개 이상의 테

1) 다음 위치에서 다운로드 받아 확인할 수 있다. <http://bit.ly/v5eaSi>  
 2) 다음 위치에서 다운로드 받아 확인할 수 있다. <http://bit.ly/rCJOiY>



〈그림 6〉 HDF 파일의 물리적 구조

이터 구성요소(data elements)로 이루어진다. HDF 파일의 첫 번째 구성요소는 파일 내에서 첫 4바이트를 차지하는 파일 헤더로서, 차례로 Control-N, Control-C, Control-S, Control-A (^N^C^S^A)인 ASCII 문자로 구성되어 있다. 이 파일 헤더를 통하여 현재의 파일이 HDF 파일인지를 확인할 수 있게 된다. 데이터 설명자 블록은 여러 개의 데이터 설명자를 가지고 있는데, 이들 데이터 설명자와 데이터 구성요소가 합쳐져서 HDF 파일에서 데이터를 묶는 가장 기본적인 집합체인 데이터 객체(data object)를 구성한다. 데이터 설명자는 데이터 구성요소의 형식, 위치, 그리고 크기에 관한 정보를 가지고 있다. 데이터 구성요소는 실제 데이터를 가지고 있다. 이렇게 HDF 데이터 조직이 구성되어 있으므로, HDF파일은 파일에 저장된 데이터 구성을 스스로 설명해 줄 수 있다. 이는 문헌정보학 분야에서 전통적으로 사용했던 기

계가독형목록(MARC)의 구성과 유사한 특징이다. 〈그림 6〉은 HDF파일의 물리적 구조와 데이터 설명자를 도식화하여 보여준다.

### 3.2 해양생물 분야

〈그림 7〉은 일정량의 물을 떠서 플랑크톤의 수를 조사한 해양생물관측데이터 샘플을 보여준다. ①번 'obs\_id'는 관측행위가 발생한 정점의 식별자를 나타내는 ID를 의미한다. ②번 's\_n'은 하나의 관측정점ID내에서 일련번호를 의미한다. 데이터 샘플의 6번 행부터 8번 행까지의 's\_n' 데이터는 관측ID 'B000001024'의 일련번호 1, 2, 3을 의미한다. ③번 'taxo'는 'taxonomy'를 의미하는 것으로서 생물종을 의미한다. ④번 'abun'은 'abundance'를 의미하는 것으로 플랑크톤의 양을 의미한다. ⑤번 'dep'는 'depth'를 의미하는 것으로서 수심을 의미한다



① obs_id	② s_n	③ taxo	④ abun	⑤ dep	⑥ ab_unit	⑦ or_g
B000001023	1	70001	36	0	cells/ml	1
B000001023	2	70001	55	10	cells/ml	1
B000001023	3	70001	20	30	cells/ml	1
B000001023	4	70001	5	50	cells/ml	1
B000001024	1	70001	36	0	cells/ml	1
B000001024	2	70001	44	10	cells/ml	1
B000001024	3	70001	8	30	cells/ml	1
B000001025	1	70001	22	0	cells/ml	1
B000001025	2	70001	29	10	cells/ml	1

〈그림 7〉 해양생물 플랑크톤 관측데이터

〈표 7〉 해양생물 관측데이터를 관리하기 위한 메타데이터 작성 예

obs_id	project_id	obs_date	latitude	longitude	bot_depth	vessel	method_instru
B000001023	G0300001	198611191415	34.500833	128.502222	66	8	Sampling: Rosette sampler, Fixation: Acid Lugol's solution

〈표 8〉 해양생물 플랑크톤 관측데이터를 관리하기 위한 메타데이터 요소

메타데이터 요소	표시상수	메타데이터 요소	표시상수
obs_id	식별자	longitude	경도
project_id	프로젝트ID	bot_depth	관측지역 수심
obs_date	관측시작일시	vessel	관측선ID
latitude	위도	method_instru	관측방법

다. ⑥번 'ab\_unit'은 플랑크톤 수의 단위를 의미한다. 본 샘플데이터에서는 1ml당 발견된 플랑크톤의 수를 의미한다. ⑦번 'or\_g'는 생물종의 상세구분을 의미한다. 이상의 내용으로 샘플데이터의 내용을 살펴보면 다음과 같다. 관측정점ID가 'B000001024'인 데이터들은 모두 3개이며, 수심 0m, 10m, 30m에서 플랑크톤이 1ml당 각각 36, 44, 8마리가 출현했음을 의미한다. 〈표 7〉은 〈그림 7〉의 플랑크톤 관측데이터를 관리하기 위한 실제 메타데이터 작성 예시로서, 원시데이터를 기술하기 위해서 한국해양연구원에서 사용하는 메타데이터 형식이다.

〈표 8〉은 메타데이터를 기술하기 위해 도출된 메타데이터 요소들을 보여준다. 특히 사항은 관측방법에 대한 상세한 기술로서, 샘플데이터의 값이 의미하는 것은 'Rosette Sampler'로 물을 채집하였음을 의미한다. 따라서 채집 방법 또한 데이터의 관리항목으로 중요하다고 판단된다.

### 3.3 해양지질 분야

〈표 9〉는 해양지질 분야에서 생성된 원시데이터의 모습을 보여준다. 'obs\_id'는 관측행위가 발생한 정점의 식별자를 나타내는 ID를 의미

〈표 9〉 해양지질 관측데이터의 원시데이터

obs_id	gravel	sand	silt	clay	sed_type	mean_size	std_dev
G000000043	5.74	83.41	10.36	0.49	gmS	1.85	2.04
G000000044	0.87	92.2	6.24	0.69	(g)S	2.76	1.42
G000000045	0.05	91.5	5.81	2.64	(g)S	2.79	1.68

〈표 10〉 해양지질 관측데이터를 관리하기 위한 메타데이터 작성 예

obs_id	project_id	obs_date	latitude	longitude	bot_depth	vessel	data_group	method_instru
G000000039	G0300007	199305280545	33.250000	128.500000	140	1	4001	Shipeck or McIntyre Grab

한다. “gravel, sand, silt, clay”는 각각 자갈, 모래, 침적토, 점토의 비율을 의미한다. ‘sed\_type’은 ‘sediment\_type’을 의미하는 것으로서 침전물을 구분하는 코드를 값으로 갖는다. ‘mean\_size’는 입자들의 평균사이즈를 말하며, ‘std\_dev’는 이의 표준편차를 의미한다. 〈표 10〉은 해양지질관측데이터를 관리하기 위해서 한국해양연구원에서 사용하는 메타데이터 형식이다. ‘method\_instru’ 요소로 기술된 내용은 ‘Grab’이라는 장비로 샘플데이터가 채집되었음을 의미한다.

### 3.4 해양화학 분야

〈표 11〉은 바닷물에 어떤 화학물질이 용존해 있는지를 실험실에서 분석한 원시데이터이다. 원시데이터 각 항목의 의미는 다음과 같다. ‘obs\_id’는 관측행위가 발생한 정점의 식별자를 나타내는 ‘ID’를 의미한다. ‘DEPTH’는 수심을 의미한다. 따라서 샘플의 내용은 같은 관

측정점의 4단계 수심의 데이터를 보여준다. 온도 이후의 항목들은 각각 염분(SALINITY), 수소이온(PH), 용존산소(DO), 이산화질소(NO2), 질산이온(NO3), 인산염(PO4), 이산화규소(SiO2)를 의미한다. 〈표 12〉는 해양화학관측데이터를 관리하기 위해서 한국해양연구원에서 사용하는 메타데이터 형식이다. 메타데이터 요소 중에 〈method\_instrument〉 요소의 값은 관측에 사용된 장비들의 종류를 의미하며, 장비명과 장비의 버전도 함께 기술되어 있다. 이는 관측데이터를 생성한 장비의 종류와 버전이 데이터의 신뢰도를 판단하는 기준이 되기 때문이다. 해양화학관측데이터의 메타데이터 요소 내용은 해양물리, 해양생물, 해양지질 분야의 요소와 거의 동일하다. 이는 관측활동이 이루어지는 해양의 특성을 반영하는 것으로 판단된다. 또한 원시데이터의 기술내용도 관측정점, 관측위치, 관측시간, 수심 등의 요소로 기본요소로 가지고 있다. 차이점은 관측의 목적에 따라 관측요소가 다르다는 점으로 판단된다.

〈표 11〉 해양화학 관측데이터의 원시데이터

obs_id	DEPTH	TEMPERATURE	SALINITY	PH	DO	NO2	NO3	PO4	SIO2
C000001527	0		31.91	7.89		0.4	0.8	0.48	4.2
C000001527	5		31.86	7.88		0.6	1.2	0.62	5.2

〈표 12〉 해양화학 관측데이터를 관리하기 위한 메타데이터 작성 예

obs_id	project_id	start_date	start_latitude	start_longitude	vessel	data_group	method_instrument
C000001526	G0300009	198208	37.000000	126.166667	19	2001	Reversing Thermometer (Yoshino Keiki Co., Y 101), Salinometer (Yeo-Kal Environmental Electronics,

## 4. 해양관측데이터 검색요소 분석

### 4.1 한국해양자료센터 서비스

한국해양자료센터(KODC)에서 메타데이터 검색 항목으로 제공하는 요소는 〈표 13〉과 같다. 1단계 제한으로 “공간적 범위, 시간적 범위, 해역명, 파라미터, 색인어, 프로젝트, 자료센터”를 선택할 수 있으며, 2단계 검색제한을 제공하고 있다. 이는 ‘해양과학메타정보’ 메뉴에서 제공하는 검색 항목들로서 국립수산과학원, 국립해양조사원, 한국해양연구원 등 우리나라의 주요 해양관련 기관에서 1961년부터 수행한 해양관측 및 조사/연구자료를 대상으로 메타데이터를 검색할 수 있다. 한편, KODC에서는 메타데이터 검색뿐 아니라 해양관측자료로서 연안정지관측자료, 정선해양관측자료, 한반도 주변해역의 CTD 자료를 제공하고 있다.

〈표 14〉는 정선관측데이터의 검색을 위한

메타데이터 요소를 검색제한 단계별로 보여준다. 연안정지관측데이터와 공통된 메타데이터 요소는 “해역, 관측일시, 수심, 수온”이며, 정선관측에 특화된 검색요소는 “정선-정점, 년도, 염분, 용존산소”이다. 연안정지관측에 특화된 검색요소는 “연안정지명, 연안정점, 기온”이다. 〈표 14〉, 〈표 15〉, 〈표 16〉, 〈표 17〉은 정점관측과 정선관측데이터의 검색을 통해 수집되는 샘플데이터를 각각 보여준다. 〈표 15〉의 해양조사자료 통계 데이터에는 검색에 사용된 메타데이터 요소이외에 “인산염인, 아질산질소, 일산질소, 규산규소, 수소이온농도” 등이 있다. 이러한 요소들은 관측요소로서 새로운 관측요소를 관측할 수 있는 장비가 나올 때 마다 관측데이터의 기술항목으로 새롭게 추가될 수 있는 특징이 있다. 한편 KODC에서 제공하는 메타데이터 검색의 상세 결과는 〈표 18〉에서 제시하는 메타데이터 요소로 서비스 되고 있다.

〈표 13〉 KODC 메타데이터 검색을 위한 단계별 검색 제한자

관측구분	1단계 제한	2단계 제한
해양자료 검색	공간적 범위	위도, 경도
	시간적 범위	From, To
	해역명	동해, 서해, 남해, 동중국해
	파라미터 (관측요소)	수온, 염분, 해류, 조류, 조석, 파랑, 해면변화, 해수광학특성, 너울, 비중, 수색, 투명도, 수소이온농도 등
	색인어	-
	프로젝트 자료센터	'황해의 물질균형 및 생지화학적 순환 연구'를 포함한 약 5,400여개의 프로젝트 '국립수산과학원'을 포함한 49개 기관

〈표 14〉 정선관측 데이터

해역	정선-정점	위도	경도	관측일시	관측수심(m)	수온(°C)	염분	용존산소(mg/L)
동해	102-00	36.077	129.5883	2011-01-08 15:45	0	13.7492	34.3067	5.8593

〈표 15〉 해양조사자료 통계 데이터

순번	검색 년도	정선 정점	수심	수온	염분	용존 산소	인산 염인	아질산 질소	질산 질소	규산 규소	수소이온 농도
16	2008	102-05	0	16.4	34.0	5.18	0.17	0.19	1.76	3.71	-

〈표 16〉 연안정지관측 데이터\_1

해역	연안정점	영문명	위도	경도	관측 시작일	관측 종료일
동해	거진 (10801)	Geojin	38.475	128.4333	1967-04-01	1996-04-30

〈표 17〉 연안정지관측 데이터\_2

해역	연안정점	관측일자	수온	기온	풍향	풍력등급	운량	염분	비중수온	비중	천기
동해	감포 (12001)	2011-11-29	16	16	N	2	71%- 80%			1.026	흐림 (운량:8-10)

## 4.2 위성해양정보서비스

위성해양관측 분야는 해양환경 부분과 관련이 있다. 국립수산과학원에서 운영하는 '위성

해양정보서비스<sup>3)</sup>'의 검색요소를 분석하였다. 국립수산과학원에서는 해양위성자료에 대하여 "메타데이터정보, 배포정보, 식별정보, 위성정보, 자료처리정보"의 5가지 카테고리로 분류하

3) 국립수산과학원에서 제공하는 위성해양정보서비스이다.

였으며, 카테고리별 세부항목을 정의하여 메타 데이터를 작성하고 있다.

해양위성자료는 처리수준에 따라 원시자료(Level 0, Level 1), 영상변환자료(Level 1), 수온분석영상(Level 2), 수온분석영상정보(Level 3)로 구분하고 있다. <표 19>는 국립수산과학원에서 자료 관리를 위해 설계한 카테고리별 메타 데이터 요소를 보여주며, <표 20>은 원시영상과 분석영상의 관리항목을 보여준다. 식별정보의 “센서명, 채널수, 화소수(행), 투영법” 등과 위성정보의 “위성명, 궤도방식”, 자료처리정보의

“처리수준, 분석적용이론” 등의 메타데이터 요소들은 문헌정보를 관리 및 활용하기 위한 메타 데이터 요소로 기술하는데 한계가 있다고 판단된다.

이상의 위성해양관측데이터를 분석한 결과 다음과 같은 특성을 도출하였다. 첫째, 관측위성에 따라 제공하는 해상도가 다르기 때문에 각 해상도별 자료제공이 가능하도록 해야 한다. 따라서 메타데이터를 관리할 때 관측된 위성사진 데이터의 해상도를 관리하여야 한다. 둘째, 관측위성에 따라 제공하는 분석영상의

<표 18> KODC의 카테고리별 메타데이터 요소

카테고리	항목
요약사항	Metadata ID, Dataset 제목, Dataset 요약, 일련번호, Dataset 언어
인용 및 참고사항	제목, 최초생산자, 출판물명, 출판 ID, 출판일, 출판자, 데이터 표현형태, URL 주소, 참고사항
Dataset의 범위	시간적범위(개시일, 종료일), 고(古)시간적범위(paleo개시일, Paleo종료일, 단위), 공간적범위(위도, 경도), 해역명, Dataset처리상태, Dataset 유지보완
Dataset의 속성	파라미터, 자료소스명, 센서명, 프로젝트, 일반 색인어
자료센터/조사기관	자료센터/조사기관, DatasetID, 담당자이름, 전화, 팩스, E-Mail, 주소, 접근제한, 사용제한, 품질
Data의 배포	
Data URL 주소	
메타데이터 정보	자료입력일자, 최종개정일자, 개정경과
상위연계레코드	

<표 19> 국립수산과학원의 자료관리를 위한 카테고리별 메타데이터 요소

카테고리	메타데이터 요소
메타데이터정보	자료구분, 파일식별자, 언어, 적용대상명, 책임 기관명, 책임부서명, 전화, 주소, 작성일, 버전
배포정보	파일포맷, 파일용량(MB), 배포자정보, 배포기관명, 배포부서명, 연락처
식별정보	위성명, 센서명, 채널수, 해상도(km), 촬영일시, 좌표범위(좌상), 좌표범위(우하), 화소수(행), 화소수(열), 분석항목, 투영법
위성정보	위성명, 발사연도, 고도(km), 궤도방식, 관측주기, 센서명, 채널수, 해상도(km), 관측폭(km), 밴드정보, 응용분야
자료처리정보	처리수준, 포맷변환, 기하보정 프로그램, 분석 적용이론

\* 출처: <http://portal.nfrdi.re.kr/sois/intro/intro05.jsp>

〈표 20〉 위성영상데이터에 관한 메타데이터 사례

카테고리	메타데이터 요소	원시영상	분석영상
메타데이터 정보	자료구분	NO-L00(원시자료/Raw Data)	NO-L21(분석자료/수온)
	파일식별자	20050101.1800.n18.pass	20050101.1800.n18.sst
	언어	한국어	한국어
	적용대상명	위성영상	위성영상
	책임기관명	국립수산과학원	국립수산과학원
	책임부서명	수산해양종합정보과	수산해양종합정보과
	전화	051-720-2220	051-720-2220
	주소	부산시 기장군 기장읍 시랑리 408-1	부산시 기장군 기장읍 시랑리 408-1
	작성일	2006-06-16	2006-06-16
	버전	제1판	제1판
배포정보	파일포맷	TAF(Terascan Data Format)	TDF(Terascan Data Format)
	파일용량(MB)		
	배포자정보		
	배포기관명	국립수산과학원	국립수산과학원
	배포부서명	수산해양종합정보과	수산해양종합정보과
	연락처	051-720-2220	051-720-2220
식별정보	위성명	NOAA-18	NOAA-18
	센서명	AVHRR	AVHRR
	채널수	5	5
	해상도(km)	1.1	1.1
	촬영일시	20050101 18:00 GMT	20050101 18:00 GMT
	좌표범위(좌상)		1174754E 445942N
	좌표범위(우하)		1421206E 250018N
	화소수(행)		2000
	화소수(열)		2000
	분석항목		해수면온도(해수면온도)
	투영법	경위도	경위도
자료처리 정보	처리수준	Level0	Level2
	포맷변환	Terascan ARCHIVE	Terascan HRPTIN
	기하보정 프로그램		Terascan TeraNav v1.26
	분석 적용이론		MCSST(수온분석)

종류가 다양하다. 따라서 관측된 위성영상 데이터의 분석영상정보를 관리해야 한다. 셋째, 원시자료와 분석자료를 기술하기 위한 메타데이터 요소는 차이가 없으며, 제공되는 데이터의 형태도 동일하다. 또한 ASCII형식의 데이

터도 파일의 형식만 차이가 있을 뿐 동일한 메타데이터 요소로 기술이 가능하다. 넷째, 위성해양관측 데이터의 검색방법으로 총 3단계의 검색제한 인터페이스가 필요하다. 1단계 검색 제한자로는 “자료종류, 관측기간, 위성종류”가

〈표 21〉 위성해양정보 검색을 위한 단계별 검색 제한자

관측구분	1단계 제한	2단계 제한	3단계 제한
위성해양 정보검색	자료종류 (원시자료 or 분석자료)	-	-
	관측기간	Pass(format) 분석영상	센서, 해상도, 고도 Terra/Aqua 분석영상 (ANGSTROM_531 외 22종) NOAA 분석영상(1회 수신 외 4종)
	위성종류 (GMS, GOES, Terra/Aqua, MTSAT, NOAA, OrbView)		

사용되며, 2단계 검색 제한자로는 “형식(format)과 분석영상”, 3단계 검색 제한자로는 “센서, 해상도, 고도, 세부 분석영상”이 사용될 수 있다. 〈표 21〉은 위성해양정보 검색을 위한 단계별 검색 제한자의 구성을 보여준다.

### 4.3 국립해양조사원 서비스

국립해양조사원(KHOA)에서 제공하는 파랑, 수온, 염분, 해양기상관측데이터에 대한 검색요소를 분석한 결과 데이터를 검색하기 위한 메타데이터 항목이 〈표 22〉와 같이 분석되었다. 관측항목의 차이가 있을 뿐 대표적인 검색제한은 모두 동일하다. 공통 검색제한으로

“관측유형, 관측소, 관측항목, 통계기간”이 도출되었다. 〈표 22〉, 〈표 23〉은 각각 파랑관측, 해양기상관측데이터의 샘플을 보여준다. 관측항목의 구분에 따라 제공되는 데이터의 필드가 구성이 되는 특징을 가지고 있으며, 특이 사항은 관측요소 자체에 대한 검색이 아니며 제공되는 데이터에 해당 관측요소의 값이 출력되도록 하는 특징을 가지고 있다. 즉 관측요소의 경우 검색제한으로 지정하는 관측항목이 검색의 제한자로 사용되는 것이 아니고, 검색 결과의 출력항목을 지정하기 위해 사용되고 있는 것을 확인할 수 있다. 이것은 실제 생산되는 원시데이터의 검색을 위해서 메타데이터의 중요성이 매우 더욱 높음을 시사한다.

〈표 22〉 파랑관측데이터

통계 일시	최대파주기(sec)	최대파고(m)	유의파주기(sec)	유의파고(m)
2011/01/01 00시	7.3	1.92	5.93	1.18

〈표 23〉 해양기상관측데이터

통계 일시	시정(m)	돌풍(m/s)	풍속(m/s)	풍향(16 point)	기압(hPa)	기온(℃)
2011/01/01 00시		11.57	9.62	NW	1013.08	-2.38
2011/01/01 01시		11.98	10.02	NW	1012.67	-2.35

## 5. 결정요소 도출

### 5.1 결정요소 분석

해양물리, 해양생물, 해양지질, 해양화학분야의 원시데이터 특징과 검색서비스를 분석한 결과 도출된 특이사항은 아래와 같다.

첫째, 각 분야별 원시데이터는 관측과 관련된 문맥정보를 포함하고 있는 경우와 그렇지 않은 경우가 모두 존재한다. 예를 들어 수온·염분데이터의 경우는 단순하게 수심과 수온, 염분에 대한 정보만 존재하며, 해류데이터에도 관측일시와 관측요소에 대한 정보만 존재하며, 텍스트로 변환된 위성데이터의 경우는 관측일시, 위성, 생성자, 위도, 경도 정보 등이 관측요소 정보와 함께 존재한다. 또한 플랑크톤 관측데이터의 경우 문맥정보로서 관측행위가 이루어진 정점의 식별자 정보와 생물종, 플랑크톤의 양, 수심 등의 관측요소 값이 함께 존재한다. 둘째, 관측요소는 관측행위를 통해 실제 관측하고자 하는 요소들을 의미한다. 하나의 관측정점과 관측정선에서 상이한 관측의 목적을 가지고서 상이한 관측방법을 사용하면 관측되는 데이터의 내용도 다를 수밖에 없다. 또한 관측장비에 따라서 생성되는 원시데이터의 형식도 다양하다. 셋째, 해양관측의 특성상 관측정점과 관측정선 식별자, 공간정보(위도, 경도), 수심정보는 필수 요소로 판단되었다. 넷째, 관측기와 관측방법, 관측데이터 생성을 지원한 프로젝트, 관측소 정보 등도 관측데이터를 관

리하고 활용하는데 주요한 요소로 판단된다. 다섯째, 관측요소는 관측목적에 따라 원시데이터에 기술되기 때문에 검색 접점으로 사용되기 보다는 검색결과와 항목 제한으로 사용되고 있다. 여섯째, 관측데이터의 종류에 따라 원시데이터를 기술하기 위한 상이한 메타데이터 요소가 존재한다. 예를 들어 위성해양관측의 경우 '채널수, 해상도, 좌표범위, 화소수, 투영법, 자료처리수준, 포맷변환, 기하정보 프로그램, 분석 적용이론' 등의 메타데이터 요소가 반드시 필요하다. 이는 관측요소의 다양성과는 다르게 위성관측데이터를 관리하고 활용하기 위한 맥락정보로서 해양관측 분야 메타데이터 선정 시 반드시 고려되어야 할 요소들로 판단된다.

### 5.2 결정요소 표준적용 검토분석

해양관측 분야의 원시데이터 자체를 기술하는 요소들을 분석하여 도출된 요소들이 문헌정보학 분야에서 메타데이터 표준으로 사용되고 있는 MODS의 요소들로 기술이 될 수 있는지 검토하였다. 또한 다양한 분야에서 웹 자원의 기술을 위해 사용되고 있는 DC와 데이터를 기술하기 위해 개발된 OECD 메타데이터 표준요소들로 기술이 될 수 있는지 검토하였다. <표 24>는 해양관측데이터의 원시데이터 기술요소들과 DC, MODS, OECD의 요소 매핑을 보여준다. 표의 '구분' 중 '일반'은 해양관측데이터를 기술하는데 사용될 수 있는 문헌정보



〈표 24〉 해양관측데이터 기술을 위한 기 표준요소 적용 검토

구분	기술요소	DC	MODS	OECD
일반	데이터 이름	title	titleInfo(title)	Main Title
	데이터 요약	description	note	X
	데이터 기술언어	language	language(languageTerm)	language
	데이터 최초생산일	date	originInfo(dataCreated)	Period
	데이터 출판일	date	originInfo(dateIssued)	Publication date
	데이터 출판자	publisher	originInfo(publisher)	X
	데이터 포맷	Format	internetMediaType	X
	데이터 용량	extent	extent	size
	데이터접근 URL	identifier	location	DOI
	데이터 색인어	subject	subject	Theme(s)
	데이터 생성기관	Creator	name(namePart)	Author(s)
	담당자 이름	Creator	name(namePart)	Author(s)
	데이터 접근제한	rights	accessCondition	X
	데이터 품질	X	reformattingQuality	X
	입력일자	date	dateCreated	X
	최종개정일자	date	dateModified	X
개정경과	version	edition	X	
관측선	관측선ID	identifier	X	X
	관측선명	X	X	X
위성	위성ID	identifier	X	X
	위성이름	X	X	X
	해상도	X	X	X
	화소수	X	X	X
	투영법	X	X	X
관측기기	장비명	X	X	X
	모델명	X	X	X
	버전	X	X	X
관측소	관측소ID	identifier	X	X
	관측소명	X	X	X
공간정보	위도	coverage	X	X
	경도	coverage	X	X
	수심	X	X	X
	해역정보	X	X	X
프로젝트	프로젝트ID	identifier	X	X
	프로젝트명	X	X	X
관측 데이터	관측데이터ID	identifier	identifier	doi
	관측요소	X	X	X
자료처리	처리수준	X	X	X
	포맷변환	X	X	X
	버전	originInfo	X	X
	기하정보프로그램	X	X	X
	분석적용이론	X	X	X

관련 메타데이터 요소들을 그룹화하여 표현한 것이다. 도출된 특이사항은 다음과 같다.

첫째, 해양관측데이터를 기술하는데 기본적으로 필요한 데이터 식별정보, 배포정보, 메타데이터정보 등은 문헌정보학 분야에서 사용하는 메타데이터 표준요소의 적용이 가능하다. 예를 들어, <표 24>에서 '일반'으로 구분된 기술요소들은 문헌정보학 분야의 메타데이터 요소들로 기술이 가능하다. 특히 DC의 경우 기술 대상이 되는 자원의 형식에 무관하게 사용할 수 있기 때문에 그 적용범위가 매우 넓다. 예를 들어 DC의 식별자(identifier) 요소는 "관측선 ID, 관측소ID, 관측데이터ID"에 모두 사용될 수 있다. 또한 관측일시 요소와 자료의 크기에 대응하는 요소는 모든 검토대상 표준에서 제공한다. 또한 MODS의 경우도 '일반'으로 구분된 기술요소들을 모두 지원한다. 둘째, 검토대상 표준들은 "관측선명, 관측장비명, 관측장비모델명, 관측기기버전, 관측소명, 수심, 해역정보, 프로젝트명, 관측요소, 해상도, 화소수, 투영법, 처리수준, 포맷변환, 기하정보프로그램, 분석적용이론" 등의 요소와 대응되는 요소를 제공하지 못하고 있다.

이상의 검토결과 해양관측데이터를 기술하는데 있어 문헌정보를 기술하는데 사용되는 메타데이터 표준 요소들의 검토가 반드시 필요하다. 이는 해양관측데이터를 가장 잘 설명하고 있는 것은 문헌이기 때문에 문헌과 데이터를 통합관리하고 서비스하기 위해서는 반드시 문헌을 기술하는 메타데이터 표준

요소들을 검토해야 함을 의미한다. 한편 해양관측데이터의 특징상 기존의 문헌관련 메타데이터 표준요소들만으로는 데이터를 기술하는데 한계가 있기 때문에 해양관측 메타데이터 표준요소를 선정하기 위해서는 데이터 유통과 관련된 메타데이터 표준과 해양관측 분야의 메타데이터 표준들에 대한 종합적인 검토가 수반되어야 한다.

한편, 해양관측데이터의 검색요소를 분석한 결과 세 가지의 특이사항을 도출하였다. 첫째, 수온, 염분, 풍향 등의 원시데이터의 관측요소들과는 달리, 데이터를 검색하기 위한 메타데이터 요소들은 문헌정보를 기술하는 메타데이터 표준요소들의 활용이 가능하다고 판단된다. 따라서 데이터의 유통측면에 있어서 데이터는 생성부터 수집, 보존, 출판까지 문헌정보와 유사한 생명주기를 갖기 때문에 데이터의 관리 및 활용을 위해서도 문헌정보 분야에서 사용하는 메타데이터 표준요소들의 검토가 반드시 필요하다. 예를 들어 데이터의 제목은 DC의 <Title>, OECD의 <MainTitle>, MODS의 <TitleInfo(title)> 요소 사용이 가능하다. 데이터의 출판기관은 DC, OECD, MODS의 <Publisher> 요소를 사용할 수 있다. 또한 데이터의 주제는 DC와 MODS의 <Subject> 요소와 OECD의 <Theme(s)> 요소를 사용할 수 있다. 이외에도 다양한 메타데이터요소를 기존의 문헌정보 분야의 메타데이터 요소로 기술 가능하다. 둘째, 기록물 보존과 서비스 측면에서 메타데이터 기술요소 검토가 필요하다. 즉 데

이더의 생산과 책임소재, 데이터 배포기관 및 연락처 정보 등도 검토되어야 한다. 셋째, 관측 기술과 관측기기, 관측방법이 다양한 해양관측 분야에서는 데이터의 생산주체인 연구자와 유통 주체들의 의견수렴이 반드시 필요하다.

## 6. 결론

본 연구에서는 해양관측데이터가 생성되는 관측의 유형들과 해양과학의 기본적인 구분인 해양물리, 해양화학, 해양생물, 해양지질 분야에서 생산된 대표적인 유형의 샘플 데이터를 분석하였다. 또한 국내에서 서비스되고 있는 해양관측데이터 서비스를 조사하여 검색에 사용되는 검색요소를 분석하였다.

분석결과 9개 영역(일반, 관측선, 위성, 관측기기, 관측소, 공간정보, 프로젝트, 관측데이터, 자료처리)으로 구분되는 42개 결정요소가 도출되었다.

본 연구는 해양관측데이터의 관리 및 서비스를 위해서 필요한 결정요소 도출을 위하여 원시데이터(raw data)에 대한 분석과 현재의 데이터 서비스 사례를 분석한 것으로서 해양관측데이터를 관리하고 서비스하는 시스템을 개발하는데 있어 기초자료로 활용될 수 있다. 향후 연구로는 본 연구 성과를 바탕으로 해양관측 메타데이터 표준요소 선정 연구가 필요

하다. 특히, 결정요소 표준적용 검토에서 도출된 특징을 고려하여 문헌정보 분야의 메타데이터 요소 검토와 관측분야에서 표준으로 인정받고 있는 메타데이터 요소 검토가 필요하다. 아울러 데이터도 문헌정보와 같은 생명주기를 갖기 때문에 현재의 문헌정보 서비스에 대한 검색요소를 조사하고 관련된 메타데이터 요소를 도출하는 과정도 필요하다. 또한 기록물 측면에서의 메타데이터 요소검토와 전문가들의 메타데이터 요소 제안 및 도출된 요소들에 대한 검토가 필요하다.

## 참고문헌

- 김선태. 2012. 『해양관측 분야 메타데이터 표준요소 선정 및 설계에 관한 연구』. 전북대학교.
- 박용안. 2005. 『바다의 과학』. 서울대학교출판부.
- 조대규, 홍철훈 외. 2009. 『물리해양학』. 부산: 부경대학교출판부.
- Green, Toby. 2009. "We Need Publishing Standards for Datasets and Data Tables." *OECD Publishing*.

### [참조사이트]

- 유비쿼터스. [cited 2011. 2. 8].  
<<http://bit.ly/oHlj9w>>.