

## 한국과학기술정보연구원의 정보유통부문 표준화 현황분석 A Survey on the Standardization of Information Service Sector in KISTI

이 윤 석\* · 서 태 설\*\*  
Yun-Seok Lee · Tae-Sul Seo

### 차 례

- |           |                    |
|-----------|--------------------|
| 1. 서 론    | • 참고문헌             |
| 2. 조사결과분석 | • 부록 : 표준화 관련 용어해설 |
| 3. 결 론    |                    |

### 초 록

이 연구의 목적은 KISTI의 정보유통 관련 부서들이 채택하고 있는 지식정보의 표준 및 기술 활용실태와 문제점을 파악하여, 향후의 국가 지식정보 표준화 기반구축의 기반자료를 수집하는 것이다. 데이터의 수집은 설문조사와 인터뷰를 통하여 수행되었고, 주요한 분석의 내용은 데이터 관리, 교환 및 통신 네트워크 운영의 표준화의 측면에서 검토되었다. 총 50종의 원내 DB 운영 체제 및 소프트웨어 활용과 표준의 활용 현황 분석에서 시급히 추진할 표준화의 과제는 과학기술분류, 메타데이터, 데이터 모델링, 데이터 포맷 및 DBMS, 과학기술용어, 검색프로토콜 등으로 나타났다.

### 키 워 드

데이터베이스, 정보 유통, 과학기술, 지식정보, 표준화, 소프트웨어

### ABSTRACT

The purpose of this study is to explore the current situation and problems in standardizing knowledge information and information technology in the KISTI, and

\* 지식정보표준화연구실 연구원(Researcher, Knowledge Information Standardization Dept., KISTI)

\*\* 지식정보표준화연구실장(Senior Researcher, Knowledge Information Standardization Dept., KISTI)

collect the basic data necessary for establishing the national information infrastructure. Through a survey and interviews with nine departments of the Information Services Sector of the KISTI, data were collected regarding the management and exchange of information and the operation of communication networks. According to the analysis of fifty KISTI databases and their operation, tasks to be first standardized include Classification of ST resources, Metadata, Date modeling, Data format, DBMS, ST terminology, and Retrieval protocol.

## KEYWORDS

Database, Information Service, Science & Technology, Knowledge Information, Standardization, Software

# 1. 서 론

## 1.1 조사의 배경과 목적

디지털화, 네트워크화가 급속히 전전됨으로써 서로 다른 지식정보 시스템간의 표준화가 선결과제로 대두되고 있다. 정보의 교환이나 공유를 고려하지 않고 정보 시스템을 구축하면 상이한 시스템 간의 호환성 부족으로 정보자원의 낭비가 초래되므로 정보의 접근성 및 공유성을 향상시키기 위한 지식정보 표준화는 필수적이다. 지식정보 표준화는 정보간의 상호교환을 가능하게 하며, 정보이용의 편의성을 증진시키며, 원활한 정보유통의 구현으로 산업전반에 생산성을 향상시켜 고부가가치를 창출할 수 있다. 또한 국제적인 개방형 표준을 수용하여 정보화 사업의 통일성을 확보하며 운영시스템의 글로벌화를 위한 기반을 조성할 수 있다. 뿐만 아니라 정보화사업간의 연계·통합을 통한 정보의 공동활용으로 유사 또는 중복되는 정보자원의 구축방지

등 효율적으로 정보자원을 관리할 수 있다.

한편, 인터넷의 급속한 확산은 정보 컨텐츠의 구축 및 이용을 기하급수적으로 증가시키고 있으며, 최근에는 생산자 중심의 분산형 DB 구축이 폭발적으로 증가하고 있다. 그런데 각각의 기관이 자신이 필요로 하는 정보 컨텐츠를 임의의 시스템으로 구축하고 있어 DB를 범국가적 차원에서 공동활용 할 수 있는 시스템으로 구축하는 것이 필요하다.

2001년 출범한 한국과학기술정보연구원(KISTI)은 산업기술정보원(KINITI)과 연구개발정보센터(KORDIC)의 통합으로 설립된 기관이다. 통합 이전의 두 기관은 국내외 과학기술정보의 수집 및 유통기관으로서 국가 산업 및 과학기술 발전을 지원하는 역할을 하였으나, 새롭게 탄생한 KISTI는 과학·기술·산업에 관한 정보를 종합적으로 수집·분석·관리하고, 이에 관한 정보유통기술·정책·표준화 등을 전문적으로 조사·연구할 뿐만 아니라, 연구

개발 인프라를 체계적으로 구축·운영하는 확대된 임무를 부여받고 있다. 따라서 본 현황조사의 1차적 목적은 기관통합의 시너지를 최대한 달성하기 위해 통합이전 양 기관에서 각기 구축·운영했던 산업 및 과학기술정보 표준화의 현황을 파악하는 것이나, 향후 한국과학기술정보연구원(KISTI)의 표준화 추진전략 도출, 나아가 국가 과학기술 지식정보 표준화 중심기관으로써의 역할과 계획을 수립하는 데 기초 자료로 사용할 예정이다.

### 1.2 조사의 범위, 방법 및 내용

KISTI는 국내외 과학기술 및 산업과 관련된 종합적인 정보수집·관리·보급, 슈퍼컴퓨터 운영·서비스, 초고속망 운영·서비스, 정보분석 업무 등을 수행하고 있으나, 본 조사에서는 KISTI 내의 정보유통관련 9개실(지식정보관리실, 컨텐츠개발실, 정보기술지원실, 정보인프라지원실, 전자정보개발실, 정보시스템연구실, 정보유통사업실, 사실정보사업실, 해외정보사업실)만을

대상으로 실시하였다.

이 연구에서는 설문조사, 전화 및 방문 인터뷰 등을 통해 DB부문의 표준화 현황을 조사하였으며, 포함된 내용은 1) 실별 데이터베이스 운영, 2) 실별 데이터베이스 관련 소프트웨어 활용, 3) 기반표준 활용, 4) KISTI의 가장 시급한 표준화 대상, 5) 부서별 중점 표준화 추진 대상, 6) 표준화 관련 애로사항, 7) 표준화 관련 교육 수요, 8) KISTI의 공유가능한 표준 및 자원, 9) 표준화 관련 대외활동의 현황 등이었다.

### 1.3 표준화 대상 분야

본 연구에서는 데이터베이스 표준화의 범위를, 데이터베이스를 구축하기 위해 필요한 데이터(정보자원)를 관리·가공하고 정보 제공을 위해 데이터를 유일하게 식별하는 기술·데이터의 저장위치·저장형태 등의 정보를 관리하는 데이터관리분야와 데이터와 데이터·데이터베이스와 이용자, 데이터베이스와 데이터베이스 간의 정보교환을 위한 데이터교환분야와, 최종이용자

<표 1> 데이터베이스 표준화 구분

| 표준화 항목 |         |                   | 관련조직 및 분야     |
|--------|---------|-------------------|---------------|
| 분류     | 세부분류    | 기술/표준 예시          |               |
| 데이터 관리 | DBMS    | RDBMS, CORBA      |               |
|        | 메타데이터   | Dublin Core, MARC | JTC1/SC32/WG2 |
|        | 식별기호    | URN, URL, DOI     | ISO/TC46/SC9  |
|        | 데이터 모델링 | UML, RDB          |               |
| 데이터 교환 | 데이터 포맷  | ASCII, GIF        | JTC1/SC32/WG1 |
|        | 표현언어    | HTML, XML         | JTC1/SC32/WG3 |
|        | 용어      | JST시소러스           | ISO/TC46/SC3  |
|        | 분류표     | JST분류표            | ISO/TC46/SC9  |
| 통신네트워크 | 통신 프로토콜 | TCP/IP            | JTC1/SC6      |
|        | 검색 프로토콜 | Z39.50            | ISO/TC46/SC4  |

가 데이터베이스를 이용하기 위한 검색 및 통신과 관련된 통신네트워크분야 등 3분야로 나누어서 조사했다(〈표 3〉 참조).

## 2. 조사 결과 분석

### 2.1 실별 데이터베이스 운영 현황

운영하는 데이터베이스가 KISTI 원내 서버와 도메인을 통해 서비스되는지 여부에 따라 KISTI 내부 데이터베이스와 KISTI 외부 데이터베이스로 크게 나누어진다. 먼저, 컨텐츠개발실은 과학기술문헌정보(BIST) 등 5종의 문헌정보 DB와 한국공개특허(KUPA) 등 7종의 특허정보 DB를 제작하고 있는데, 이것은 과학기술 및 산업기술 개발에 필수적인 핵심정보들이다.

다음으로 지식정보관리실은 정보학문현

정보(INFO) 등 2종의 문헌정보 DB와 외국학술잡지종합목록(UCAT) 등 4종의 목록 DB를 제작하고 있고, 전자정보개발실은 연구보고서 등 2종의 문헌정보 DB와 과학기자재정보 DB 1종, 과학기술인력정보 DB 1종 등을 제작·운영하고 있다. 특히 전자정보개발실에서 제작하는 연구보고서정보 DB는 국가의 연구개발사업을 효과적으로 관리·운영하는 시스템의 근간을 이루는 정보이다. 사실정보사업실은 한국산 고등균류정보 등 29종의 사실정보 DB를 제작·운영하고 있는데, 이 중 정보 특성상 전문성이 요구되는 9종의 DB는 외부 기관에서 직접 운영하도록 하고 있으며, 나머지 20종의 DB는 자체적으로 운영하고 있다.

정보유통사업실은 KISTI 홈페이지 운영을 담당하는 부서로 자체적으로 DB를 제

〈표 2〉 KISTI가 직접 운영하는 데이터베이스 및 관련 현황

| 담당 부서   | 데이터베이스   | 서버   | 도메인         |
|---------|--|--|-------------|
| 컨텐츠개발실  | 문헌 : 5종 · 특허 : 7종  | eagle(HP D390)   |             |
| 지식정보관리실 | 문헌 : 2종 · 목록 : 4종  | tiger(HP V2500)  | kisti.re.kr |
| 전자정보개발실 | 문헌 : 2종 · 장비 : 1종 · 인력 : 1종  | i2s(SUN E-5000)<br>Linux Server                            |             |
| 사실정보사업실 | 사실 : 20종   | ruby(Silicon Graphics)                                     | kodic.re.kr |
| 정보유통사업실 | 학회논문정보   | society(SUN C-2000)  |             |
| 정보기술지원실 | 기업지원 : 6종 · 장비 : 1종<br>인력 : 1종   | spider(SUN E-3500)<br>dbe(SUN E-5500)<br>penta(SUN E-5000) | innonet.net |
| 해외정보사업실 | 해외과학기술 뉴스 · 정보*  | kosen(SUN E-4000)  | oasis.or.kr |
| 7실      | 50종(학회, 해외 제외)<br>문헌 : 9종, 특허 : 7종, 목록 : 4종<br>장비 : 2종, 인력 : 2종, 사실 : 20종<br>기업지원 : 6종 | 10대  | 4개          |

\* 학회논문정보와 해외과학기술 뉴스 · 정보는 데이터베이스 특성상 종수 구분하기가 어려워 제외되었음.

작하지 않는 대신 160여개 학회의 학회정보 DB 제작 및 운영지원을 하고 있다.

정보기술지원실은 종합기업 서비스 정보망(INNO-NET) 홈페이지를 운영하면서 제품별 시장동향 및 수요예측정보 등 6종의 기업지원정보DB와 시험연구장비정보 DB 1종, 전문인력정보DB 1종 등을 운영하고 있다.

해외정보사업실은 한민족 과학기술자 네트워크(KOSEN)와 해외 과학기술정보 네트워크(OSTIN) 홈페이지를 각각 운영하면서 해외과학기술 뉴스·정보를 제공하고 있다. 이 정보들은 수시 갱신되는 게시판 형태로서 대부분의 정보를 기사 작성자가 자율적으로 게시할 수 있는 체제로 운영되고 있다(<표 2> 참조).

KISTI 외부 도메인을 사용해 서비스되고 있는 정보로는 정보기술지원실에서 운영하는 종합기업서비스정보망(INNO-NET)과 문헌 정보를 제외한 다양한 정보를 다루는 사실정보사업실의 사실정보 데

이터베이스가 있다. 정보기술지원실은 유휴설비/해외자본재 등 11종의 사실정보를 이노넷을 통해 서비스하고 있으며, 사실정보사업실은 KS재료열물성 등 9종의 정보를 KISTI 홈페이지를 통해서 서비스하고 있다(<표 3> 참조).

KISTI에서 제작·운영하고 있는 실별 데이터베이스 현황을 요약하면 지식정보관리실, 컨텐츠개발실은 데이터베이스를 제작하는 업무는 수행하고 있지만 운영은 정보인프라지원실에서 수행하고 있으며, 정보기술지원실, 전자정보개발실, 정보유통사업실, 해외정보사업실 등 4개 실은 데이터베이스의 제작 및 운영업무를 직접 관리하고 있다.

데이터베이스 수는 내부 50종과 외부 20종 총 70종의 데이터베이스를 제작·운영하고 있으며, 이를 운영하기 위한 도메인은 내부 4개와 외부 17개 총 21개의 도메인이 관련되어 있는 것으로 나타났다. 데이터베이스를 운영하기 위한 서버에는 EAGLE

&lt;표 3&gt; KISTI에 링크되어 있는 외부 데이터베이스 현황

| 정보기술지원실    |                  | 사실정보사업실   |                     |
|------------|------------------|-----------|---------------------|
| 데이터베이스     | 도메인              | 데이터베이스    | 도메인                 |
| 유휴설비/해외자본재 | koami.or.kr:6001 | KS재료열물성   | jungwonsystem.co.kr |
| 투자유망기업     | kisc.org         | 어류/해양오염물질 | chungnam.ac.kr      |
| 가스 에너지     | kgs.or.kr        | 충청GIS/광산  | kigam.re.kr         |
| 유통물류       | kcci.or.kr       | 곰팡이       | kribb.re.kr         |
| 외국인공장설립    | kicox.or.kr:7999 | 담수어류      | chonbuk.ac.kr       |
| 에너지 절약     | kemco.or.kr:8001 | 식용어류      | yosu.ac.kr          |
| 수출전환기업     | 203.233.236.66   | 유전질환      | ajou.ac.kr          |
| 벤처카탈로그     | 147.46.76.68     |           |                     |
| 신기술사업화     | kitech.re.kr     | 외부 DB 합계  | 외부 도메인 합계           |
| 사이버 전시회    | eiak.org         | 20종       | 17개                 |

〈표 4〉 DBMS/IRS 활용현황

| 구 분      | 지식정보<br>관리실 | 컨텐츠<br>개발실 | 정보기술<br>지원실 | 정보인프<br>라지원실 | 전자정보<br>개발실 | 정보시스<br>템연구실 | 정보유통<br>사업실 | 사실정보<br>사업실 | 해외정보<br>사업실 |
|----------|-------------|------------|-------------|--------------|-------------|--------------|-------------|-------------|-------------|
| ORACLE   |             |            |             |              |             |              |             |             |             |
| SQL      |             |            |             |              |             |              |             |             |             |
| JDBC     |             |            |             |              |             |              |             |             |             |
| KRISTAL  |             |            |             |              |             |              |             |             |             |
| SEARCH97 |             |            |             |              |             |              |             |             |             |

(HP D390) 등 총 10대의 서버를 운영하고 있는 것으로 나타났다.

데이터베이스 현황 조사를 통해서 나타난 문제점으로는 일부 문헌정보와 장비 및 인력 정보의 데이터 항목이 달라서 통합 검색에 한계가 있는 것으로 나타났다.

## 2.2 데이터베이스 관련 소프트웨어 활용현황

### 2.2.1 DBMS/IRS

DBMS(DataBase Management System)는 ORACLE을 정보기술지원실 등 6개 실에서 활용하고 있는 것으로 나타났고, 그 다음으로 SQL, JDBC 순으로 활용하고 있으며, IRS(Information Retrieval System)는

한국과학기술정보연구원에서 자체개발한 KRISTAL을 전자정보개발실 등 4개 실에서 활용하고 있고, SEARCH97은 정보기술 지원실과 정보인프라지원실에서 활용하고 있다. 데이터베이스를 운영하고 있는 대부분의 부서에서 DBMS로 ORACLE을 암도 적으로 선호하고 있는 것으로 나타났다(〈표 4〉 참조).

### 2.2.2 프로그램 언어

프로그램 언어는 JAVA를 정보기술지원실 등 6개실에서 활용해 가장 많이 활용되고, 그 다음으로 PHP, C언어, ASP, HTML 순으로 활용하는 것으로 나타났다(〈표 5〉 참조).

〈표 5〉 프로그램 언어 활용 현황

| 구 분  | 지식정보<br>관리실 | 컨텐츠<br>개발실 | 정보기술<br>지원실 | 정보인프<br>라지원실 | 전자정보<br>개발실 | 정보시스<br>템연구실 | 정보유통<br>사업실 | 사실정보<br>사업실 | 해외정보<br>사업실 |
|------|-------------|------------|-------------|--------------|-------------|--------------|-------------|-------------|-------------|
| JAVA |             |            |             |              |             |              |             |             |             |
| ASP  |             |            |             |              |             |              |             |             |             |
| PHP  |             |            |             |              |             |              |             |             |             |
| HTML |             |            |             |              |             |              |             |             |             |
| C언어  |             |            |             |              |             |              |             |             |             |

〈표 6〉 데이터 가공·입력용 소프트웨어 활용 현황

| 구 분            | 지식정보 관리실 | 컨텐츠 개발실 | 정보기술 지원실 | 정보인프라지원실 | 전자정보 개발실 | 정보시스템연구실 | 정보유통 사업실 | 사실정보 사업실 | 해외정보 사업실 |
|----------------|----------|---------|----------|----------|----------|----------|----------|----------|----------|
| TOAD           |          |         |          |          |          |          |          |          |          |
| LIVE PICTURE   |          |         |          |          |          |          |          |          |          |
| COSMO PLAYER   |          |         |          |          |          |          |          |          |          |
| 포토샵            |          |         |          |          |          |          |          |          |          |
| 일러스트           |          |         |          |          |          |          |          |          |          |
| 엑셀             |          |         |          |          |          |          |          |          |          |
| HWP            |          |         |          |          |          |          |          |          |          |
| KISTI 자료관리 시스템 |          |         |          |          |          |          |          |          |          |

### 2.2.3 데이터 가공·입력용 소프트웨어

데이터 가공·입력과 관련된 소프트웨어는 각 실에서 제작하는 데이터베이스의 특성에 따라 여러 소프트웨어들을 다양하게 사용하고 있는 것으로 나타났다. 비교적 여러 종의 DB를 제작하고 있는 정보기술 지원실과 사실정보사업실, 그리고 학회지 DB를 제작하고 있는 정보유통사업실에서 주로 사용하는 것으로 나타났으며, 데이터 베이스 가공·입력 관련 소프트웨어 활용 현황을 요약하면 〈표 6〉과 같다.

### 2.2.4 요약

부서별로 다양하게 사용되고 있는 DBMS와 IRS 분야는 데이터베이스의 특성에 맞게 최적화할 필요성이 있으며, 각각의 데이터베이스 특성별로 제작 공정 및 도구를 표준화하여 효율적으로 데이터베이스를 구축하고 운영하는 것이 필요한 것으로 사료된다.

### 2.3. 지식정보표준 활용현황

데이터 관리/데이터 교환/네트워크 등 표준화 대상이 되는 거의 전분야를 모두 활용하고 있으며, DBMS/IRS, 메타데이터, 용어관련, 분류표, 검색 프로토콜 등은 일부 자체 개발하여 활용하고 있는 것으로 나타났다. 지속적으로 지식정보 표준화의 국내외 추세를 모니터링 하여 최신의 기반 표준 입수·활용 및 보급이 필요할 것으로 보인다(〈표 7〉 참조).

### 2.4. 시급한 표준화 대상

KISTI에서 가장 시급히 추진되어야 할 표준화 과제로는 과학기술분류, 메타데이터, 데이터 모델링, 데이터 포맷 및 DBMS, 과학기술용어, 검색 프로토콜 순으로 나타났다. 이중에서 지식정보 표준화의 핵심이 되는 과학기술분류와 메타데이터에 대해서는 KISTI 내에 연구 그룹(study group)을 운영함으로써 표준기반연구를 수행하여 빠른 시일 내에 분류체계 및 메타데이터 표

〈표 7〉 분야별 지식정보표준 활용현황

| 구 분    |          | 활용표준 내역  |
|--------|----------|--|
| 데이터관리  | DBMS/IRS | ORACLE, SQL, MY-SQL, KRISTAL                             |
|        | 메타데이터    | MARC, Dublin Core, XML RDF                               |
|        | 식별기호     |  |
|        | 데이터모델링   | RDB, KRISTAL   |
| 데이터교환  | 데이터포맷    | ASCII, UNICODE, GIF, JPEG, TIF, MPEG, AVI, PDF, ISO 2709 |
|        | 표현언어     | HTML, C&C++, JAVA, PL/SQL, CGI, PHP, ASP                 |
|        | 용어       | KISTI 영한대역집, 개발계획  |
|        | 분류       | LC 주제분류, KISTI 과기문현DB분류, 과학재단분류, 부서자체분류                  |
| 통신네트워크 | 통신 프로토콜  | TCP/IP, X.25, HTTP                                       |
|        | 검색 프로토콜  | K-PROTOCOL, 메타검색   |

준화를 수행할 예정이다.

#### 2.5. 부서별 중점 표준화 추진 대상

정보유통 관련 각 실별 표준화 추진계획에 따르면, 지식정보관리실을 비롯한 데이터베이스 제작과 관련된 부서에서는 메타데이터 표준화 추진계획을 가지고 있는 것으로 나타났다. 데이터포맷의 경우는 TIMS(Total Information Management System)을 통해 중소기업과 데이터 교환이 빈번한 정보기술지원실에서 주로 관심을 갖고 있었다. 용어표준화와 과학기술분류표는 전문 분야 컨텐츠를 다루어야 하는 컨텐츠개발실과 사실정보사업실에서 추진 의향을 갖고 있거나 추진중에 있었다. 검색

프로토콜은 정보시스템연구실에서 향후 중점 표준화 추진 대상으로 계획하고 있는 것으로 나타났다(〈표 8〉 참조).

#### 2.6. 표준화 관련 애로사항

컨텐츠개발실, 정보기술지원실, 사실정보사업실, 해외정보사업실 등 4개실에서 응답했으며, 과학기술분류 표준화에 대한 애로가 가장 많은 것으로 나타났고 제도화된 사내 표준의 필요성도 제기되었다(〈표 9〉 참조).

#### 2.7. 표준화 관련 교육 수요

정보기술지원실, 전자정보개발실, 사실정보사업실, 해외정보사업실 등 4개실에서

〈표 8〉 부서별 중점 표준화 추진대상 현황

| 표준화 추진 대상 | 부 서                               |
|-----------|-----------------------------------|
| 메타데이터     | 지식정보관리실, 컨텐츠개발실, 전자정보개발실, 정보유통사업실 |
| 데이터포맷     | 정보기술지원실                           |
| 용어표준화     | 컨텐츠개발실                            |
| 분류표준화     | 지식정보관리실, 컨텐츠개발실, 사실정보사업실, 해외정보사업실 |
| 검색프로토콜    | 정보시스템연구실                          |

〈표 9〉 부서별 애로사항

| 부 서     | 애로 사항  |
|---------|--|
| 컨텐츠개발실  | 사내 표준이 제도화가 안되어 있어서 표준의 신뢰성이 낮음              |
| 정보기술지원실 | 검색 프로토콜, 어플리케이션 간 데이터 통합 및 통합 인터페이스 개발이 어려움. |
| 사실정보사업실 | 사실 정보를 분류할 수 있는 통일된 분류표가 없음.                 |
| 해외정보사업실 | 과학기술분야 분류 표준화 및 검색엔진                         |

XML(4실), Dublin Core(2실), DBMS, 데이터 구조 및 웹 연동기술(1실)의 교육을 희망했으며, 원내 세미나 및 교육을 통해서 수요에 대한 지속적인 교육이 필요한 것으로 보인다.

#### 2.8. 공유 가능한 표준 및 자원

KISTI가 자체 개발한 표준 및 자원 중에는 KISTI 외부의 기관들과 공유함으로써 활용도를 높일수 있는 가능성이 있는 것도 조사되었다. 그중에는 컨텐츠개발실의 과학기술문헌분류표와 과학기술용어대역집, 정보기술지원실의 종합정보관리시스템(TIMS)의 웹버전, 정보시스템연구실의 KRISTAL, K-PROTOCOL, 그리고 일부대학에 보급하여 사용중인 전자정보개발실의 연구보고서/학위논문수집시스템이 있는 것으로 나타났다. 그리고, 지식정보관리실의 자료관리 시스템도 여러 기관에 보급한 실

적을 가지고 있다.

컨텐츠 자원으로는 해외정보사업실의 해외과학기술동향정보가 전자부품연구원, 국세청, 포항공대 등 여러 곳에서 정보를 받아 활용하고 있고, 하이텔을 통해서도 서비스되고 있으며, 사실정보사업실의 사실 정보는 CD로 제작하여 교육기관에 배포한 사례가 있었다.

#### 2.9. 표준화 관련 대외활동

데이터베이스 관련 국내표준화기구는 산업표준을 제정하는 한국산업표준원(KISI: Korean Industrial Standards Institute)과 정보통신표준을 제정하는 한국정보통신기술협회(TTA: Telecommunications Technology Association), 데이터베이스 단체표준을 제정하는 한국데이터베이스 진흥센터(KDPC: Korea Database Promotion Center)와 그 밖에 한국전산원

〈표 10〉 공유 가능한 표준 및 자원 현황

| 자 원      | 부 서  |
|----------|--|
| 분류 및 용어집 | 컨텐츠개발실(과학기술문헌분류표, 과학기술용어대역집)                 |
| 소프트웨어    | 정보기술지원실(TIMS WEB), 정보시스템연구실(KRISTAL)         |
| 시스템      | 전자정보개발실(연구보고서/학위논문수집시스템), 지식정보관리실(자료 관리 시스템) |
| 검색 프로토콜  | 정보시스템연구실(K-PROTOCOL)                         |
| 컨텐츠      | 해외정보사업실(해외과학기술동향정보), 사실정보사업실(사실정보 CD)        |

(NCA), 한국전자통신연구원(ETRI), 기술 표준원(ATS), 한국표준협회(KSA) 등이 있으며, 국제표준화기구로는 ISO(International Organization for Standardization), IEC(International Electrotechnical Commission), ITU(International Telecommunication Union) 등이 있다.

특히 데이터 관리 서비스 전문위원회(ISO/IEC JTC1/SC32)와 정보 및 도큐멘테이션 전문위원회(ISO/TC46) 등의 국제 표준화 활동에 적극 참여하여 국내의견을 적극적으로 반영하고, 국제표준을 신속하게 국내에 보급해야 할 것이다.

KISTI의 표준관련 대외활동은 <표 11>과 같으며, 각 부서에서 산발적으로 수행되고 있는 표준화활동의 조정이 필요하며 표준화협의회를 통한 정보공유와 체계적으로 국내외 표준화 활동을 수행하기 위한 전략을 수립해야 할 것이다.

### 3. 결 론

이번 KISTI의 정보유통부문 표준화 현황조사 결과에 따르면, 부서별로 약간의 차

이는 있지만 전부서가 자료수집부터 데이터베이스 제작·운영 등 거의 전분야에 걸쳐서 업무를 수행하고 기반기술을 활용하고 있는 것으로 나타났다. 각 부서에서는 표준화에 대한 인식은 많이 하고 있으나, 부서마다 개별적으로 기술 및 표준을 적용하고 있어서 기관 차원의 표준화 추진이 절실히 확인할 수 있었다.

연구원 내에서조차 데이터베이스별 인터페이스와 검색 시스템이 상이한 것이 많은 것으로 나타났다. 데이터베이스별 특성을 유지하면서 통일된 표준 인터페이스로 이용자들의 검색에 대한 편의성을 증진시켜야 하며, 데이터베이스의 특성별 공정 및 도구 표준화로 데이터베이스 제작의 효율성을 제고가 필요하다. 또한 인력, 장비 등 중복 정보의 조정도 대두되었다.

우선, KISTI의 가장 시급한 표준화 대상으로서는 과학기술분류와 메타데이터 표준화로 조사되었다. 대책으로서 과학기술분류와 메타데이터의 원내 연구회 운영을 통하여 표준기반 구축을 위한 기초연구 및 개발을 수행하고 지속적인 지식정보 표준의 국내외 추적조사, 표준화기구와의 협력

<표 11> KISTI의 표준 관련 대외활동

| 표준 활동           | 수행 부서          |
|-----------------|----------------|
| TTA08 워킹그룹      | 정보사업부, 정보기술지원실 |
| TTA URI 포럼      | 지식정보표준화연구실     |
| TTA 전자상거래 통합포럼  | 지식정보표준화연구실     |
| 한국전산원 지식관리표준화회의 | 정보유통사업실        |
| 산자부 참조표준정보센터    | 사실정보사업실        |
| CODATA, GBIF    | 사실정보사업실        |
| 국가전자도서관사업       | 정보인프라지원실       |

체계 구축을 통해 최신 표준정보의 해당부서 보급과 KISTI의 공유 가능한 소프트웨어(TIMS, KRISTAL)와 시스템(연구보고서/학위논문수집 시스템)의 원내외 공유체계가 필요한 것으로 나타났다.

지속적 세미나 개최 및 정기적인 연구원 내의 표준화협의회 운영을 통해 KISTI 표준화 사업의 객관성 제고와 표준정보 공유를 도모해야 할 것이다. 이제, 새롭게 출범하게 된 한국과학기술정보연구원이 국가 과학기술 지식정보 인프라 중심기관으로서 국가적인 소명을 수행하기 위한 첫 시도로서 실시한 이번 현황조사결과가 향후 과학 기술 지식정보 표준화 기반구축을 위한 기초자료로서 활용되길 바란다.

### 참고문헌

1. 백두권. 1999.“데이터기술분야의 표준화 동향분석 및 표준화 연구.”1999년도 정보통신표준화 활동지원사업 소과제 연구 보고서 : 한국정보통신기술협회.
2. 안계성. 1998. “데이터베이스 표준화 현황 및 추진방향”. 정보관리연구, vol. 29, no. 4 pp.1~ 22.
3. 한국데이터베이스진흥센터. 2001.“국가 지식정보자원 관리 관련 기술·표준 활용실태 및 수요 조사서.”
4. 한국데이터베이스진흥센터. 2001.“데이터베이스 표준화 연구 최종 연구개발보고서”: 정보통신부.

## 부록

### 규정하기 위한 메타데이터

#### <표준화 관련 용어해설>

ASCII(American Standard Code for Information Interchange):

컴퓨터나 인터넷상에서 텍스트 파일을 위한 가장 일반적인 포맷이다. 아스키 파일에서는 각각의 알파벳이나 숫자, 그리고 특수문자들이 7비트의 2진수(7개의 0 또는 1의 조합으로 이루어진 스트링)로 표현되며, 총 128개의 문자가 정의되어 있다.

아스키는 미국규격협회인 ANSI(American National Standards Institute)에 의해 개발되었다.

CORBA(Common Object Request Broker Architecture):

네트워크에서 분산 프로그램 객체를 생성, 배포, 관리하기 위한 구조와 규격이며, 네트워크 상의 서로 다른 장소에 있고 여러 벤더들에 의해 개발된 프로그램들이 “인터넷페이스 브로커”를 통해 통신하도록 해주는 것으로 가장 쉽게 설명한다면, 서로 다른 운영체제와 서로 다른 프로그래밍 언어에서 상호간에 운용(통신)을 용이하게 만들어주는 기술

Dublin Core:

더블린 메타데이터 코어 요소집합(Dublin Metadata Core Element Set, 이하 DC)을 일컫는 말로서, 1995년 미국 더블린에서 열린 1차 워크숍을 통해 정의되었으며, DC는 네트워크 환경에서의 각종 전자 정보를 기술하는데 핵심적인 데이터요소를

DBMS(DataBase Management System):

데이터베이스 관리 시스템의 약어로 데이터를 효과적으로 이용할 수 있도록 정리, 보관하기 위한 기본 소프트웨어

DOI(digital object identifier : 디지털 객체 식별자):

인터넷 주소가 변경되더라도 사용자가 그 문서의 새로운 주소로 다시 찾아 갈 수 있도록, 웹파일이나 인터넷 문서에 영구적으로 부여된 식별자이다. 중앙에서 관리되는 디렉토리에 DOI를 제출하고 나서, 정식 인터넷 주소 대신 그 디렉토리의 주소에 DOI를 더하여 사용하게 된다. DOI 시스템은 Corporation for National Research Initiatives와 협력관계에 있는 미국 출판업 협동조합에 의해 고안되었으며, 현재는 세계 DOI 재단에 의해 관리되고 있다.

GIF(Graphics Interchange Format):

JPEG과 함께 월드와이드웹에서 지원되는 두 가지 그래픽 파일 형식 중의 하나이다. 웹이나 기타 인터넷의 다른 서비스(예를 들면 게시판 서비스 등)에서 GIF는 이미지 형식에 있어, 사실상의 표준이 되어가고 있다.

HTML(Hypertext Markup Language : 하이퍼텍스트 생성 언어):

웹 브라우저 상에 정보를 표시하기 위한 마크업 심볼 또는 파일 내에 집어 넣어진 코드들의 집합이다. 마크업은 웹 브라우

저에게 웹 페이지 내의 글자나 이미지를 어떻게 나타낼 것인가에 관해 알려주는 역할을 한다. HTML은 SGML에 기반을 둔 공개된 포맷이며, 단순한 일반 텍스트 편집기에서부터 WYSIWYG 작성도구에 이르기까지 다양한 방법으로 작성하거나 편집할 수 있다. 제목, 문단, 리스트, 하이퍼링크 등은 모두 태그를 사용하여 구조화 시킨다.

MARC(Machine Readable Cataloging Record):

도서관 자료의 서지정보를 컴퓨터가 처리할 수 있는 형태로 기록한 기계가독형 목록을 의미하는 것으로 도서관간 레코드의 상호교환을 위하여 1965년에 미국 국회 도서관(LC)의 도서관 자동화계획에 따라 처음으로 개발됨.

RDB(relational database : 관계형 데이터베이스):

관계형 데이터베이스는 1970년에 IBM의 E. F. Codd에 의해 개발되었다. 관계형 데이터베이스는 일련의 정형화된 테이블로 구성된 데이터 항목들의 집합체로서, 그 데이터들은 데이터베이스 테이블을 재구성하지 않더라도 다양한 방법으로 접근하거나 조합될 수 있다. 사용자와 관계형 데이터베이스를 연결시켜주는 표준 검색언어를 SQL(Structured Query Language)이라고 하는데, SQL 문장은 관계형 데이터베이스에 있는 데이터를 직접 조회하거나 또는 보고서를 추출하는데 사용된다.

RDBMS(relational database management system : 관계형 데이터베이스 관리 시스템):

관계형 데이터베이스를 만들거나, 수정하고 관리할 수 있게 해주는 프로그램으로 RDBMS는 사용자가 입력하거나, 또는 응용프로그램 내에 포함된 SQL 문장을 취하여 데이터베이스의 생성, 수정 및 검색 등의 서비스를 제공한다. 잘 알려진 RDBMS로는 마이크로소프트의 액세스, 오라클의 오라클7, Ardent의 UniData 등이 있다.

THESAURUS(시소러스):

용어의 사용법과 용어들 사이의 관계에 대한 정보를 제공하는 어휘 도구를 말한다. 용어의 관계성은 일반적으로 상위 개념(BT: broader term), 하위 개념(NT: Narrower Term), 용례 혹은 동의어(UF: Use For Or Synonymous), 관계어(RT: Related Term), 대체어(USE) 등으로 분류되는데, 시소러스는 이러한 관계성을 이용, 템색시 질의에 포함된 용어의 의미를 확대하기 위해 주로 사용된다.

TCP/IP(Transmission Control Protocol/Internet Protocol):

인터넷의 기본적인 통신 프로토콜로서, 인트라넷이나 엑스트라넷과 같은 사설망에서도 사용된다. 사용자가 인터넷에 접속하기 위해 자신의 컴퓨터를 설정할 때 TCP/IP 프로그램이 설치되며, 이를 통하여 역시 같은 TCP/IP 프로토콜을 쓰고 있는 다른 컴퓨터 사용자와 메시지를 주고받거나, 또는 정보를 얻을 수 있게된다.

URL(Uniform Resource Locator):

인터넷에서 접근 가능한 자원의 주소를 일관되게 표현할 수 있는 형식을 말한다. 여기서 말하는 자원의 형태는 인터넷 응용 프로토콜에 따라 다를 수 있다. 즉, 월드와이드웹의 프로토콜인 HTTP를 사용하는 경우, 자원이란 HTML 페이지, 이미지 파일, CGI나 자바 애플리케이션 같은 프로그램, 기타 HTTP가 지원하는 다른 파일들이 될 수 있다.

URN(Uniform Resource Name):

정보자원의 소장 위치, 프로토콜, 호스트 등과는 상관없이 고유의 기호로 표현한 것으로 영구적으로 남을 수 있는 불변의 의미를 가지는 이름으로 인터넷 자원을 찾는 방식이다.

UML(Unified Modeling Language):

소프트웨어 개발 과정에서 산출되는 산출물들을 명시, 개발, 문서화하기 위한 모델링 언어이다. UML은 모델링 언어일뿐 메소드(또는 방법론)는 아니다. 메소드는 프로세스에 대한 정의와 각각의 업무들에 대한 지침과, 업무들 간의 순서들을 명시해야 하는 반면, 모델링 언어는 표기법(또는 다이어그램)들만을 제시하는 것이다. 따라서 UML은 소프트웨어 개발에 사용하기 위한 여러 다이어그램들을 정의하고 있으

며, 또 다이어그램들의 의미들에 대해 정의하고 있다.

XML(Extensible Markup Language):

1996년 W3C(World Wide Web Consortium)에서 제안한 것으로서, 웹 상에서 구조화된 문서를 전송 가능하도록 설계된 표준화된 텍스트 형식이다. 이는 인터넷에서 기존에 사용하던 HTML의 한계를 극복하고 SGML의 복잡함을 해결하는 방안으로서 HTML에 사용자가 새로운 태그(tag)를 정의할 수 있는 기능이 추가되었다고 이해하면 쉽다. 또한, XML은 SGML의 실용적인 기능만을 모은 부분집합(subset)이라 할 수 있으며, 인터넷상에서 뿐만 아니라 전자 출판, 의학, 경영, 법률, 판매 자동화, 디지털도서관, 전자상거래 등 매우 광범위하게 이용될 전망이다

Z39.50:

온라인 데이터베이스 내의 도서목록 데이터 검색에 관한 표준 통신 프로토콜이다. Z39.50은 인터넷에서 도서관장서의 OPAC (Online Public Access Catalogues)을 검색하는데 사용되며, 때로는 서로 다른 OPAC들을 하나의 연합된 OPAC으로 결합시키는 데에도 사용된다. Z39.50은 ANSI/NISO 표준이다.