

디지털 도서관 구축을 위한 표준에 관한 고찰

A Study on Standards for Digital Library Construction

안 현 수*, 이 혜 옥**
(Hyun Soo Ahn, Hye Ock Lee)

초 록

최근 정보자원의 효율적인 관리 및 활용을 위해 국내외적으로 디지털 도서관 시스템 구축작업이 활발히 진행중이다. 이러한 시스템 개발시 관련 표준을 적용함으로써 상호운용성을 확보하여 시간과 노력, 비용을 획기적으로 줄일 수 있다. 따라서 본 고에서는 표준의 의의 및 개발과정과 표준의 종류, 주요 표준화기구에 대해 언급을 한 후, 디지털 도서관 시스템 구축을 위한 주요 표준에 대해 설명하고자 한다.

키 워 드

디지털 도서관, 표준, 정보자원 식별, 정보자원 기술, 컴퓨터 응용

ABSTRACTS

Many digital libraries are constructed for effective management and use of information resources around the world now a days. While constructing these systems, it is possible to reduce the time, efforts and expenses through getting interoperability by applying the standards. The authors will describe the benefits of standards, development process, types of standards, standards organization, key standards for digital libraries.

* 한국통신 연구개발본부 정보관리부 전임연구원
(Researcher, Tech. Info. Div., KT R&D Group)

** 한국통신 연구개발본부 정보관리부 선임연구원
(Senior Researcher, Tech. Info. Div., KT R&D Group)

KEYWORDS

digital library, standards, resources identification, resources description, computer application

I. 서 론

국제표준화기구(ISO)의 정의에 의하면 “표준(standard)”이란 “일반인들이 입수할 수 있는 문헌으로서 이 표준에 의해 영향을 받는 모든 이해관계자의 협력과 합의 또는 전체적 승인아래 작성이 되고 과학, 기술 및 경험을 결집시킨 결과의 기반 위에서 최선의 사회이익 촉진을 목적으로 하는 것이며, 표준화 단체에 의해 승인받은 것”을 말한다(한국공업표준협회, 1988).

예를 들어 현재 널리 사용되고 있는 전화 카드나 현금 카드가 국제 표준으로 지정되어 있기 때문에 전세계 어디를 가나 이들이 통용이 되고 있으며, 이에 따라 생활이 편리해지고 제품이나 서비스의 신뢰성이 높아지게 된다. 마찬가지로 도서관/정보센터의 목록 데이터베이스 구축의 핵심 도구인 기계가독형 목록(MARC)의 국제적인 표준 포맷이 개발됨으로써 전 세계적으로 목록 데이터를 공유할 수 있게 되었으며, 이에 따라 많은 사서들의 노력을 줄이고 중복투자를 방지하는 효과를 얻게되었다.

본 고에서는 표준의 의의 및 개발과정에 대해 언급을 한 후, 일반적인 표준의 종류와 이들 표준을 개발하고 관리하는 각종 기구에 대해 설명한다. 그리고 디지털 도서관 구축시에 고려해야 할 표준으로 크게 정보자원의 식별 및 기술과 관련된 표준과 컴퓨터 응용과 관련된 표준으로 구분하여 설명한다.

II. 표준의 의의 및 개발 과정

디지털 도서관 구축에 있어서 표준을 사용하는 의의는 다음과 같다. 첫째,

표준을 사용함으로써 호환성(compatibility)을 갖게 되며, 이에 따라 장비, 데이터, 작업과정 등에서 상호운용성(interoperability)을 확보하여 정보자원을 쉽게 이용할 수 있게 된다. 둘째, 표준을 이용할 경우 시스템을 이용하기가 쉬울 뿐만 아니라 운영하는데 비용이 적게 든다. 즉, 표준을 이용함으로써 기계가독형 레코드를 공유할 수 있으며, 특정 조직내에서 정보기술들을 인터페이스시킬 수 있고 다른 조직의 시스템들과의 인터페이스도 가능하다. 셋째, 다양한 업체에서 제공되는 제품과 서비스가 일정 수준의 품질을 만족시킨다는 보장을 받을 수 있기 때문에 이용자 입장에서 표준을 이용할 경우 잇점이 있다. 표준화에 의해 제품과 서비스들이 상호교환이 가능할 정도로 호환성을 갖게 될 경우, 업체들끼리 경쟁이 발생하여 보다 저렴한 비용으로 원하는 제품이나 서비스를 이용할 수 있다. 넷째, 제품과 서비스가 표준을 따를 경우 국내뿐만 아니라 국제 시장에서의 판매를 증대시킬 수 있다. 즉, 시장이 확대될 경우 생산자 입장에서는 규모의 경제와 효율성을 이룰 수 있기 때문이다. 따라서 다국적 기업의 경우 표준화는 사업 성공의 핵심요소가 된다.

대부분의 표준화 기관들은 표준을 개발하는데 있어서 일반적으로 다음과 같은 여섯 단계를 거치게 된다(Boss, 1996). 첫째, 제안단계. 새로운 표준을 개발하거나 기존의 표준을 갱신하고자 할 경우에 제안이 이루어진다. 제안 내용을 검토하는 위원회의 투표가 필요하며 이 작업에 참여하고자 하는 자원자들이 결정된다. 그리고 의장이나 프로젝트 책임자가 지명된다. 둘째, 준비 단계. 의장은 전문가로 구성된 작업반(working group)을 소집한다. 최초의 회의 이후부터는 대부분의 작업이 e-mail, fax, 전화 등을 통해 이루어진다. 초안들이 오고가며 이 과정에서 의견들이 수집된다. 셋째, 위원회 검토 단계. 작업반에서 작성한 최초의 초안이 상위의 위원회에 제출되어 검토 및 의견수렴 작업이 이루어진다. 넷째, 검토 단계. 표준화기구 내부뿐만 아니라 외부에도 표준의 초안을 광범위하게 회람시켜 상충되는 견해나 기타 의견 등을 수렴한 후, 통일된 최종안을 작성한다. 다섯째, 승인 단계. 의견수렴 과정을 거친 초안은 표준화기구의 모든 구성원들을 대상으로 공식 투표에 회부된다. 국제 표준의 경우 일반적으로 2/3 이상의 대다수 찬성이 요구된다. 여섯째, 출판 단계. 일

단 표준이 승인되면 출판이 된다.

Ⅲ. 표준의 종류 및 주요 표준화 기구

표준에는 크게 단체 표준, 국가 표준 그리고 국제 표준이 있다. 단체 표준이란 한 국가안의 사업단체, 학협회 등에 의해 제정되어 원칙적으로 그 단체 구성원 내지 사업분야, 학계에 적용되는 표준을 말한다. 국가 표준은 국가 또는 국가적으로 인정된 표준화 기관이 제정하여 전국적으로 적용하는 표준이다. 우리나라의 국가 표준으로는 KS(Korean Industrial Standard)가 있다. 국제 표준이란 전 세계 각국에 개방되어 있는 국제적 조직에서 제정되고 국제적으로 적용되는 표준이다. 대표적인 예로는 국제전기표준회의(IEC)와 ISO 등이 있다.

디지털 도서관 시스템 구축과 관련된 대표적인 표준화 기구로는 ISO, NISO, ANSI, IETF 등이 있다. 1947년에 정식 발족된 국제표준화기구(ISO)는 지적, 과학적, 기술적 및 경제적 활동분야에 있어서의 국제간 협력을 도모하고 전기분야 이외의 모든 분야에 관한 국제 표준 제정, 심의 및 발행을 촉진하는 비정부간 협의기구이다. ISO는 총회(General Assembly), 이사회(Council), 중앙사무국(Central Secretariat), 정책개발위원회(Policy Development Committees), 특별 자문그룹(Ad Hoc Advisory Group), 기술관리부(Technical Management Board), 기술위원회(Technical Committees), 소위원회(Subcommittees), 작업반(Working Groups) 등으로 구성되어 있다(<http://www.iso.ch/infoe/intro.html>). 정보처리 분야와 관련된 ISO 조직으로 기술위원회 46(TC 46)이 있는데 이것은 미국내 정보처리 분야의 표준화 기관인 국가정보표준화기구(NISO)와 동일하다. ISO/TC 46 /SC 4는 도서관, 정보센터, 색인 및 초록 서비스 센터, 출판사 등과 같은 정보서비스기관에서 사용하는 기술표준을 담당하는 ISO의 소위원회이다. 여기에는 정보검색 및 도서관 상호대차를 위한 프로토콜, SGML 응용, 데이터 요소 사전, 데이터 포맷, 문

자 세트, 코드, 이용자 명령어 등이 포함된다(<http://lcweb.loc.gov/loc/standards/isotc46/>). ISO/TC 46/SC 9는 문헌의 표현, 식별 및 기술에 관한 표준을 담당한다. 여기에서 작성된 표준의 예로 국제표준도서기호(ISBN), 국제표준연속간행물기호(ISSN), 다국어 시소러스 개발지침 등이 있다(<http://www.nlc-bnc.ca/iso/tc46sc9/index.htm>).

국가정보표준화기구(NISO)는 미국국가표준연구원(ANSI)에 의해 표준 개발 기관으로 지정된 비영리기관이다. NISO의 투표권을 가진 회원기관에는 도서관, 출판사, 정부기관, 정보서비스제공업체 등이 있다. NISO에서 개발한 표준으로는 정보검색 프로토콜인 Z39.50, SGML 도구인 NISO/ANSI/ISO 12083, 정보상호교환포맷인 Z39.2, 언어표현용 코드인 Z39.53, 과학기술보고서 포맷인 Z39.18 등이 있다(<http://www.niso.org/>).

ANSI는 미국내 국가 표준의 제정 및 승인을 총괄하기 위해 1918년에 설립된 비영리 민간법인이다. ANSI는 정책을 총괄하는 이사회(Board of Directors), 이사회를 보좌하는 부위원회(Board Committees), 회원의 의견을 수렴하여 정책에 반영하는 평의회(Council), 표준 제정을 위한 정책과 절차를 규정하는 표준집행위원회(Executive Standards Council), 기술활동에 대한 조정업무를 담당하는 표준국(Standard Board) 등으로 구성되어 있다(<http://www.ansi.org/whatansi.html>).

IETF(Internet Engineering Task Force)는 인터넷 구조의 발전과 인터넷의 원활한 운영에 관심이 있는 네트워크 설계자, 운영자, 업체, 연구자 등으로 이루어진 국제 협의체이다. IETF의 실질적인 기술 작업은 작업반에서 이루어지는데 각 작업반은 응용, 전송, 보안 등과 같은 주제들에 따라 여러분야로 조직되어 있다. IETF는 1년에 3차례의 회의를 갖는다(<http://www.ietf.org/>).

IV. 디지털 도서관 구축을 위한 표준

디지털 도서관 구축을 위한 표준은 크게 두 영역으로 구분할 수 있는데 첫

째는 정보자원의 식별 및 기술과 관련된 표준이고 두 번째는 컴퓨터 응용과 관련된 표준이다. 전자와 관련된 주요 표준으로는 메타데이터, 디지털 객체 식별자(DOI), 통합자원설명자(URN) 등이 있다. 후자에 속하는 주요 표준으로는 데이터 통신, 데이터 교환용 포맷, 정보검색 및 도서관 상호대차용 프로토콜 등이 있다.

1. 정보자원의 식별 및 기술과 관련된 표준

(1) 정보자원의 식별 관련 표준

정보자원의 식별과 관련된 표준은 네트워크 자원 식별자와 서지 식별자로 구분할 수 있다. 전자의 예로는 통합자원설명자(URN : Uniform Resource Name), 디지털 객체 식별자(DOI : Digital Object Identifier) 등이 있으며 후자의 예로는 연속간행물 상세식별자(SICI : Serial Item and Contribution Identifier), ISBN, ISSN 등이 있다.

① 통합자원설명자(URN)

통합자원설명자(URN : Uniform Resource Name)는 “정보 작성자(publisher)” 혹은 공인 기관에 의해 부여되는 불변의 객체식별자로서 기존 인쇄 자료의 ISBN과 동일하다. 이에 따라 정보자원의 위치와 접근방법과는 무관하게 해당 정보자원에 쉽게 접근할 수 있다. URN의 일반적인 포맷은 “urn:hdl:cnri.dlib/december97”로 나타낸다. 정보 작성자(publisher)가 자신이 작성한 정보에 대한 URN을 직접 만들거나 혹은 다른 기관에서 대신 하도록 권한을 부여한다. 정보 작성자의 이름과 URN에 대한 등록과정이 필요하며 URN을 이용할 경우 URN을 URL로 변환시키는 과정이 수반된다. 현재 URN과 변환 과정(resolution)에 대한 연구가 계속되고 있으며 초기의 URN에 관한 설명은 “RFC 1737(Functional Requirements for Uniform Resource Names)”에 언급되어 있다.

② 디지털 객체 식별자(DOI)

디지털 객체 식별자(DOI : Digital Object Identifier)는 하나의 식별자 이면서 동시에 이 식별자를 구현하는 것이되기도 하는데 여기에서는 출판사와

저작권자에게 URN의 라우팅 서비스를 제공하기 위해 CNRI(Corporation for National Research Initiatives)의 핸들 시스템(Handle system)을 이용한다. DOI는 “/”로 구분되는 접두사와 접미사로 구성되며 최대 128자까지 가능하다. 접두사는 각 출판사에 지정하게 되는데 이것은 디렉토리 관리자가 부여한다. 두 번째 부분인 접미사는 식별하고자 하는 특정 콘텐츠에 대해 출판사가 부여하는 것이다. 많은 출판사들은 공인된 국제 표준을 접미사로 사용하고 있다. 미국 출판협회(AAP)에 의해 처음 시작된 DOI는 1997년 10월에 열린 프랑크프루트 도서전시회에서 만들어진 DOI 재단(DOI Foundation)에 의해 계속 운영될 예정이다.

③ 연속간행물 상세식별자(SICI)

연속간행물 상세식별자(SICI : Serial Item and Contribution Identifier)는 연속간행물의 호수(items)와 기사(contribution)의 식별을 위해 개발되었으며 ANSI/NISO Z39.56-1991로 정의되어 있다. SICI는 ISSN에 호와 기사에서 추출된 상세 내용을 추가하여 만들어진다. SICI는 출판사에서 직접 만들거나 혹은 대행사나 원문제공기관 등과 같은 제3의 기관에서 작성하기도 한다.

(2) 정보자원의 기술 관련 표준

정보자원의 기술과 관련된 표준에는 네트워크 자원의 기술과 서지자원의 기술 두 가지가 있다. 전자의 예로는 TEI, MARC 등이 있고 후자의 예로는 소장정보 기술(Holdings statement) 등이 있다.

① TEI

전산인문과학협회(ACH), 전산언어학협회(ACL), 전산어문학협회(ALLC) 등에 의해 개발된 TEI(Text Encoding Initiative)는 모든 종류의 전자 정보 특히 언어학, 문학, 인문과학 분야 연구 프로젝트에서 생성되고 이용되는 데이터 집합을 표현하는데 이용될 수 있는 SGML 태그 정의의 표준 집합을 제공하고자하는 주요 국제 프로젝트이다. 모든 TEI 문헌에는 해당 문헌을 식별하고 소스의 상세사항을 제공하는 헤더가 포함된다. 시, 연극, 사전 등을 위한 기본 태그 집합들이 제공되며 이들 기본 집합으로부터 필요한 조합을 만들 수

있는 방법도 정의되어 있다. 연결, 분석, 도표 등과 관련된 정보를 수집하기 위해 추가 태그 집합들이 제공된다.

② MARC

도서관/정보센터에서의 서지정보 교환을 위해 미국국회도서관(LC)에서 1960년대 중반부터 개발하기 시작하여 전세계적으로 널리 활용되고 있는 MARC(Machine Readable Cataloging)는 현재 각 국가별로 자국 출판물의 특성에 맞는 포맷을 개발하여 이용하고 있다. 이에 따라 전세계적으로 50개 이상의 MARC 포맷들이 개발되어 비호환성 문제가 발생함으로써 이를 해결하기 위해 UNIMARC 포맷이 개발되기도 하였다. MARC의 국제 표준인 ISO 2709에서는 MARC 레코드에서 적용되는 자기테이프 교환포맷을 정의한다. ISO 2709 파일은 고정장 레코드 레이블이 먼저 나오고 그 뒤에 디렉토리 맵, 레코드 식별자, MARC 유보 필드, MARC 서지 필드가 포함된 가변장 레코드가 나온다. ISO 4873 IS2와 IS3 코드가 각각 필드 및 레코드 구분자로 이용된다. 미국국회도서관은 서지용, 전거용, 소장정보용, 분류용 USMARC 포맷을 관리하고 있으며 MARC 자문위원회에는 ALA 기계가독형 서지정보위원회(MARBI), 미국 국립도서관, 캐나다국립도서관, OCLC, RLIN 등과 같은 대규모 서지 유틸리티, 전문도서관협회(SLA) 등과 같은 도서관협회, 도서관 시스템업체 등이 포함되어 있다.

③ 소장정보 기술표준

ISO 10324(Information and documentation—Holdings statements—Summary level)로 정의된 소장정보 기술표준은 ISO TC 46 / SC 9에 의해 개발되었다. 이 표준은 소장정보의 상호교환에 있어서의 일관성을 높이기 위해 연속간행물과 비연속간행물의 요약 수준에서의 소장정보기술을 위한 디스플레이 요구사항을 언급하고 있다. 즉, 한 개 이상의 도서관/정보센터에서 소장 정보의 목록작성을 위해 필요한 연속간행물과 비연속간행물의 요약수준의 소장정보기술을 위한 데이터 요소들을 정의하고 있다.

2. 컴퓨터 응용과 관련된 표준

디지털 도서관 구축에 있어 컴퓨터 응용과 관련된 표준은 크게 데이터 교환,

데이터 통신, 정보검색 및 도서관상호대차 분야 등으로 구분할 수 있다.

(1) 데이터 교환 관련 표준

1) 아스키 포맷

ASCII는 정보교환용 7비트 문자 집합으로된 표준 코드로서 ANSI에 의해 제정되었으며 ISO/IEC 646으로도 정의되어 있다. 영어 문서 교환에 적합하도록 모두 128개의 문자로 구성되어 있으며 대부분의 PC 및 워크스테이션 컴퓨터 코드 집합의 기본을 이루고 있다. ASCII 코드 리스트는 “ftp://dkuug.dk/il8n/charmmaps.all/ ANSI-X3.4-1968”에서 구할 수 있다.

2) 이미지 포맷

① GIF

GIF(Graphic Interchange Format)는 컴퓨서브사(CompuServe Inc)에서 개발한 레스터(raster) 이미지 파일 포맷으로서 시스템에 무관하게 그래픽 데이터를 교환하기 위한 것이다. GIF 그래픽은 RGB 컬러 값을 갖는 픽셀들의 열로 저장된다. 각 픽셀의 위치를 별도로 나타낼 필요가 없는데 그 이유는 픽셀들이 좌측 상단에서부터 우측 하단으로 연이어서 저장이 되기 때문이다. GIF는 하나의 파일내에 한 개 이상의 이미지를 갖을 수가 있기 때문에 이들 이미지를 연속해서 빨리 보여줄 경우 동작 효과를 보여줄 수도 있다. 이것이 바로 최근들어 Web 상에서 널리 활용되는 “animated GIF”다.

② JPEG

JPEG(Joint Photographic Experts Group)은 흑백이나 완전 컬러, 그레이스케일(greyscale), 디지털 정지 이미지의 압축용 포맷으로서 주로 사진 이미지 저장용으로 사용되고 있다. JPEG 포맷은 ISO/IEC 10918 Information Technology-Digital compression and coding of continuous-tone still images으로 정의되어 있다.

③ TIFF

Aldus사와 마이크로소프트사에 의해 개발된 TIFF(Tagged Image File Format) 포맷의 특징은 아스키나 이진 코딩을 이용하여 거의 모든 2차원의

래스터 데이터 형태를 갖는다. 이 포맷은 탁상용 전자출판(DTP)과 탁상용 어플리케이션에서 널리 사용되고 있다. TIFF는 스캐너에서 디지털 이미지 교환용 포맷으로 사용될 뿐만 아니라 일부 응용 소프트웨어 개발자들에 의해 메모리 포맷으로 이용된다. 또한 TIFF는 DTP와 고급 응용소프트웨어 사이의 이미지 파일 인터페이스로 사용되고 있다. 다양한 TIFF 포맷들을 표준화하기 위한 노력으로 ANSI IT8.8/TC 130이 있다.

3) 페이지 기술 포맷

① 포스트스크립트

포스트스크립트(PostScript)는 벡터(vector)와 래스터(raster) 그래픽을 포함한 형식화된 텍스트 파일로서 아도브 시스템사(Adobe Systems Inc.)에서 개발한 표준이다. 포스트스크립트의 특징은 다음과 같다. 첫째, 강력한 그래픽 기능을 내장하고있는 하나의 범용 프로그래밍언어이다. 둘째, 프로그램 기능을 갖는 하나의 페이지 기술 언어이다. 셋째, 모니터와 프린터 등과 같은 래스터 출력 장치를 조작할 수 있는 대화형 시스템이다. 1985년에 Level 1이 발표된 이후 1997년 현재 Level 3이 나와 있으며 이 표준에 관한 완전한 사양은 “PostScript Language Reference Manual 2nd edition, Addison Wesley Longman, December 1990”에서 구할 수 있다.

② PDF

PDF(Portable Document Format)는 포스트스크립트 포맷의 발전된 형태로서 페이지 기술언어의 일종이다. PDF의 주요 특징으로는 직접 연결(hot links), 각 페이지에 대한 요약 아이콘(thumbnail icons of pages), 각 장에 대한 요약정보(chapter outlines), 각 페이지에 대한 주석(page annotations) 등이 있다. 아도브사의 Acrobat 패키지 소프트웨어의 일부인 변환(Distiller) 프로그램을 이용할 경우 포스트스크립트 파일을 PDF 파일로 쉽게 바꿀 수 있다. 이러한 특징을 갖는 PDF 파일을 볼 수 있는 강력한 뷰어 소프트웨어(Acrobat Reader)가 무상으로 제공되기 때문에 거의 모든 플랫폼에서 PDF 파일을 손쉽게 활용할 수 있다. PDF는 포스트스크립트와 마찬가지로 아도브 시스템사(Adobe Systems Inc.)에서 개발한 표준이며 상세한 설명은 “Porta-

ble Document Format Reference Manual, Addison Wesley Longman, November 1996”에서 구할 수 있다.

4) 레이아웃 정보를 포함하는 파일 포맷

① ODA

ODA(Open Document Architecture)는 파일의 내용과 두 개의 계층적 구조(논리적 구조 및 레이아웃 구조)에 의해 사무용 문헌들을 기술하는 포맷이다. 문헌들은 포맷된 형식(여기서는 레이아웃 구조만 이용함)이나 처리가능한 형식(여기서는 논리적 구조만 이용함) 혹은 포맷되어 처리가능한 형식(두 구조를 상호 교환함) 등으로 상호교환될 수 있다. 특정 범주에 속하는 문헌의 주요 특징은 문헌 응용 프로파일(DAP)에 의해 정의되는데 이들은 국제표준 프로파일(ISP)의 일부로 정의된다. 현재 ISP의 3 계층은 단순문헌구조와 확장판 문헌 구조에 맞도록 정의되어 있다. ODA에서는 문헌 교환용의 두 개의 부호화(encoding) 방법을 정의하고 있는데 가장 일반적으로 이용되는 것으로 ISO의 ASN.1(Abstract Syntax Notation One)과 SGML이 있다. ODA의 초기 표준은 “ISO/IEC 8613 Information Technology—Open Document Architecture(ODA) and interchange format”에 정의가 되어 있다.

② LaTeX

LaTeX는 수학과 전산학 등과 같은 과학기술분야에서 널리 이용되는 문서 작성용 시스템으로서 수학 공식을 작성하기 위한 훌륭한 도구로 이용되고 있다. 대부분의 워드프로세서와 달리 LaTeX에서는 이용자가 LaTeX 코드를 이용하여 직접 아스키 파일에 마크업을 한 후, 디스플레이나 인쇄를 위해 LaTeX 파일을 포스트스크립트 등과 같은 적절한 포맷으로 변환을 시킨다. 이러한 변환은 두 단계를 거치게 되는데 먼저 파일을 LaTeX 포맷터를 통과시킴으로써 DVI 형태로 변환을 시킨 다음 이것을 포스트스크립트 파일로 변환 시킨다. LaTeX 마크업도 SGML과 마찬가지로 논리적인데 그 이유는 문헌의 디스플레이 양식보다는 논리적 구조를 중점적으로 기술하기 때문이다. LaTeX는 미국수학학회의 지원을 받아 스탠포드대학교의 Donald Knuth에 의해 개발되어 배포되고 있다.

5) 구조화된 정보를 갖는 파일 포맷

① HTML

HTML(hyper Text Markup Language)은 Web상에서의 데이터 교환을 위한 마크업언어로서 특정 플랫폼에 종속되지않고 하이퍼텍스트 문헌을 작성하여 이용할 수 있다. HTML은 1990년 이후로 Web에서 사용되기 시작했으며 현재는 RFC 1866으로 정의된 버전 2가 주로 이용되고 있다. 다중언어 데이터 표현기능 등과 같은 새로운 기능이 포함된 HTML 4.0 규격의 초안이 1997년 7월에 발표되었다. HTML의 마크업을 이용하여 하이퍼텍스트 뉴스, 메일, 문헌, 하이퍼미디어, 선택 메뉴, 데이터베이스 질의 결과, 그래픽이 포함된 간단한 구조화된 문헌 등을 표현할 수 있다.

② SGML

SGML은 1969년에 미국 IBM에서 사내의 주요 문서들을 쉽게 교환하고 관리할 수 있도록 전자화하기 위해 만든 GML(General Markup Language)로부터 시작되었다. SGML은 1986년에 국제표준화기구(ISO)의 표준(ISO 8879)으로 제정되었고 1993년에는 한국산업규격(KSC 5913)으로 제정되었다. SGML이 본격적으로 이용되기 시작한 것은 1980년대 중반 이후부터 미국 국방성(DoD)의 광속거래(CALS) 프로젝트와 미국 출판협회(AAP)의 전자원고작성 및 마크업 프로젝트(Electronic Manuscript Preparation and Markup) 등에서 대규모로 적용되기 시작하면서 이다. 정보를 주고 받는 시점에서 정보의 내용과 구조 등을 상세하게 기술함으로써 정보를 효율적으로 교환하기 위한 일종의 프로토콜인 SGML은 기술적 마크업 언어(descriptive markup language), 모듈화에 의한 융통성과 확장성 등과 같은 특징들을 가지고 있다. SGML은 “ISO 8879:1986 Information processing—Text and office systems—Standard Generalized Markup Language(SGML)”에 정의되어 있다.

(2) 데이터 통신 관련 표준

1) HTTP

HTTP(HyperText Transfer Protocol)는 TCP 위에서 작동하는 데이터 의

세스용 프로토콜로서 현재 Web의 기반이 되고 있다. 시스템들이 HTTP 프로토콜을 이용하여 서로 교신을 하고자 할 경우에, 가장 효율적으로 통신할 수 있는 방법을 설정하기 위한 정보가 포함된 일련의 메시지들을 서로 교환한다. 이 표준의 버전 1은 1996년 5월에 RFC 1945로 발표되었으며 이 프로토콜에 대한 상세한 자료는 “ftp://ftp.ietf.org/internet-drafts/draft-ietf-http-v11-spec-07.txt”의 IETF 초안 정보저장소에서 구할 수 있다.

2) SMTP

IETF RFC 821로 정의되어 있는 SMTP(Simple Mail Transfer Protocol)는 인터넷상에서 서버들 사이에 전자우편 메시지를 상호 교환하기 위한 프로토콜이다. 각각의 메시지는 표준화된 헤더를 가지고 있는데 이 헤더는 메시지를 받는 사람의 전자우편 주소, 메시지를 보내는 사람의 전자우편 주소 및 이름, 메시지가 통과한 네트워크상의 노드들에 관한 정보 등을 식별하는데 이용된다. 이 표준에 관한 상세한 정보는 “http://www.ietf.cni.reston.va.us/”에서 구할 수 있다.

(3) 정보검색 및 도서관상호대차 관련 표준

1) Z39.50

Z39.50은 두 대의 컴퓨터가 정보검색을 위해 서로 통신을 하는 표준화된 방법을 정의하고 있는데 Z39.50을 이용할 경우 정보 검색 절차와 방법이 표준화가 되기 때문에 대규모 정보 데이터베이스를 이용하는데 훨씬 쉬워진다. 1988년에 Z39.50-1988이 표준으로 승인된 이후에 곧바로 Z39.50 개발자들의 모임에서 Z39.50 표준프로토콜의 이용을 촉진시키는 작업이 시작되었으며 NISO는 1992년에 Z39.50의 개정판에 대한 투표를 실시하였다. 시스템 개발자들이 Z39.50 표준을 계속 발전시킴에 따라 버전 3 개정판이 만들어졌다. 버전 3인 Z39.50-1995는 1995년 7월에 완성이 되었으며 버전 4는 1995년 가을부터 개발이 시작되어 현재 개발이 진행중이다.

이와는 별도로 정보검색 표준에 관한 국제표준화기구(ISO)의 노력 결과 1991년에 Z39.50과 호환이 되는 두 가지의 국제표준이 완성되었는데 이들은 ISO 10162(Search and Retrieve Application Service Definition)과 ISO

10163(Search and Retrieve Protocol Specification)이다. ANSI/NISO Z39.50-1992가 미국 표준과 국제 표준을 모두 수용함으로써 미국 표준이 국제 표준과 호환성을 가지면서 동시에 국제 표준을 포함하는 관계가 되었다. 결국 1994년 초에 국제 표준 프로토콜 개발 담당자들이 Z39.50을 국제 표준으로 받아들임으로써 1996년의 ISO 10162/10163 SR 프로토콜의 버전 2는 Z39.50-1995와 동일하게 되었다. 이에 따라 Z39.50을 기반으로 만들어진 시스템은 SR을 기반으로하는 시스템과 상호 운용성(interoperability)을 가지고 있다.

Z39.50 관련 단체 및 기관으로는 Z39.50 개발자 모임(ZIG : Z39.50 Implementor's Group)과 Z39.50 관리국(Z39.50 Maintenance Agency)이 있다. 관리국은 인터넷상의 Web 서버를 운영하고 있는데 이 서버를 통해 Z39.50 표준관련 다양한 정보를 제공하고 있으며 주소(URL)는 "http://lcweb.loc.gov/z3950/agency/"이다.

2) ILL

ILL(Interlibrary Loan) 표준은 도서관 상호대차와 관련된 전자 메시지로서 여기에는 ISO 10160 : 1996과 ISO 10161-1, ISO 10161-2 등이 있다. ISO 10160 : 1996에서는 도서관 상호대차 요구, 요구 발송, 통지 발송, ILL 응신, 조건부 응답, 취소, 취소 응답, 접수, 체크인, 연체, 갱신, 분실 통보, 상태 조회 등과 관련된 메시지의 규격을 정의한다. ISO 10161은 ISO 10160과 호환되는 서비스 제공을 위한 ASN.1 규격을 정의한다. ISO 10161-1은 프로토콜의 규격을 정의하고 ISO 10161-2는 프로토콜 구현의 호환성에 관한 내용을 정의하고 있다. ISO 10161 표준의 등록업무는 ISO로부터 업무를 이관받아 캐나다국립도서관에서 수행하고 있으며 1995년에 이 표준의 구현 그룹(IPIG)이 결성되었다.

V. 결 론

도서관/정보서비스센터, 정보서비스, 출판 분야 등에서 표준을 적용하는 주된

이유는 호환성을 확보함으로써 장비, 데이터, 작업 절차 등에서의 상호운용성을 갖도록 하는 것이다. 이에 따라 시간과 노력을 획기적으로 줄일 수 있을 뿐만 아니라 비용 측면에서도 많은 절감을 가져올 수 있다.

특히, 초고속정보통신망을 기반으로 하여 전세계를 대상으로 하는 디지털 도서관 환경하에서는 이러한 표준의 필요성은 더욱 크다 하겠다. 즉, 국가와 지역을 초월한 효율적인 정보 활용을 위해서는 범세계적인 표준의 개발과 적용이 필수적이라 하겠다. 따라서 최근 국내에서 활발하게 추진되고 있는 디지털 도서관 시스템 개발시 무엇보다도 먼저 관련된 여러 표준에 관한 연구와 관심을 기울일 필요가 있다.

〈참 고 문 헌〉

- 한국공업표준협회. 서울 : 사내표준화편람. 1988. pp. 65~66.
- "An introduction to ANSI", <http://www.ansi.org/whatansi.html>
- Berners-Lee, T. et al. Uniform Resource Locators(URL). RFC 1738, 1994.
- Boss, Richard W. "Standards for automated library systems and other information technologies." *Library Technology Reports*, July-August 1996. pp. 459~567.
- "Introduction to ISO", <http://www.iso.ch/infoe/intro.html>
- "ISO TC 46/SC 4 Home Page", <http://lcweb.loc.gov/loc/standards/isotc46/>
- "ISO TC 46/SC 9". <http://www.nlc-bnc.ca/iso/tc46sc9/index.htm>
- "The Internet Engineering Task Force", <http://www.ietf.org/>
- The URN Implementors, "Uniform Resource Names," *D-Lib Magazine*. Feb 1996.
- "What is the National Information Standards Organization?", <http://www.niso.org/>