

XML/KORMARC 시스템 구현에 관한 연구

조윤희 강사(중앙대학교)

이두영 교수(중앙대학교)

1. 서론

1.1 연구의 필요성 및 목적

도서관 정보시스템에서 서지정보 작성에 사용되는 대표적인 레코드 포맷은 MARC(Machine Readable Cataloguing)로서 현재 전세계적으로 66개국에서 사용되고 있다. 이러한 MARC는 1960년대 중반 미국 의회도서관(Library of Congress)에 의해 개발되어, 1981년 제정된 서지정보 교환용 포맷 ISO 2709의 기반이 되었다.

우리나라는 1983년 KORMARC운영협의회가 구성되어 LC MARCⅡ포맷과 UKMARC 및 ISO 2709를 기반으로 단행본용(1993), 연속간행물용(1994), 비도서자료용(1996), 선거통계용(1999), 소장정보용(1999), 고서용(2000) KORMARC 포맷이 KS로 제정된 바 있다.

도서관에서 MARC 포맷의 도입은 인쇄 카드목록의 배포에서 온라인 목록 서비스에 이르기까지 편목업무의 효율성을 증대시켰고, 검색도구로서 중요한 역할을 수행하면서 도서관 정보서비스 영역에 새로운 전환점을 가져다주었다. 그러나 MARC 포맷은 도서관 위주의 목록에 근거를 둔 레코드 포맷으로 관련업계와의 상호교류나 이 기종

간 호환성이 보장되지 않고, 포맷의 경직성으로 인해 새로운 자원이나 필드정보를 확장하는데 어려움이 있으며, 레코드간이나 레코드내에서의 연결기능이 부족한 것이 문제점으로 지적되어 왔다.

최근에는 인터넷의 급속한 성장에 따른 디지털 정보자원의 급증으로 웹 자원을 도서관장서로 개발, 보존, 이용시켜야 할 필요성이 대두되고 있다. 이는 도서관이 지금까지 축적해 온 MARC 포맷의 서지 정보와 함께 새롭게 도서관 자원으로 웹 자원 메타데이터를 조직하거나 별도로 시스템을 구축하여 통합하는 문제로 발전되었다. 이에 도서관은 기존의 MARC기반 서지데이터베이스와 웹 자원 메타데이터를 통합하는 새로운 정보조직의 메커니즘을 개발할 필요성에 직면하게 되었다.

이러한 필요성에 따라 기존의 도서관 MARC 데이터베이스와 완벽하게 호환성을 유지하면서 웹 자원 메타데이터 조직에도 적합한 표준 기술로 W3C(World Wide Web Consortium)에서 제안한 XML(eXtensible Markup Language)을 채택하였다. 이에 본 연구는 KORMARC 포맷을 XML/KORMARC 포맷으로 재설계하는 시스템을 실험적으로 구현하였다. 이러한 XML/KORMARC 시스템의 구현을 통하여 웹에서 상호운용성을 확보하고 정보서비스 영역을 확장하여 도서관 정보시스템의 성능을 향상시키는 데 그 연구목적을 두었다.

1.2 연구 내용 및 방법

본 연구에서 구현하고자 하는 XML/KORMARC 시스템은 별도의 새로운 시스템 구축이 아니라 인터넷 표준기술로 대두되고 있는 XML기술을 도입하여 MARC기반 서지데이터베이스를 XML/KORMARC 데이터베이스로 재설계하는 시스템이다. 이에 본 논문의 연구내용과 방법은 다음과 같다.

첫째, 목록 기술 및 웹 자원 기술의 표준과 MARC와 XML 통합의 이론적 배경을 문헌적 고찰을 통하여 수행하였다. 둘째, 변환프로그램 개발을 위해 현행 한국문헌자동화목록형식의 구조 분석을 통하여 KORMARC의 DTD(Document Type Definition)를 설계하고, 이를 토대로 KORMARC와 XML/KORMARC 데이터의 양방향 변환 프

로그를 웹에서 변환 가능하도록 개발하였다. 셋째, 변환프로그램을 통해 생성된 데이터를 입력물로 XML/KORMARC 시스템을 구현하였다. 이 시스템은 XML 1.0과 JavaScript, MySQL, PHP 및 phpDOM을 이용하여 구현하였다. 넷째, 시스템의 평가는 도서관의 편목 전문사서들을 대상으로 SERVQUAL도구를 이용하여 측정하였다. 조사는 웹 설문 방법으로 대학도서관 및 전문도서관 편목 전문사서를 대상으로 하였으며, 평가 결과의 분석은 사분면 분석법으로 사서들의 특정서비스 기능에 대한 중요도 인식을 파악하였다.

2. 이론적 배경

2.1 목록 기술의 표준

최초의 기계가독형 목록은 1966년 10월 미국 의회도서관에서 개발한 MARC Pilot Project에서 비롯되었다. 이 프로젝트는 기계가독레코드의 생산과 배포의 표준화된 포맷의 타당성 검토를 위해 수행되었다. 이러한 검토를 통해 포맷을 갱신한 MARC II는 지난 30여년 동안 세계 각국에서 자체 개발한 MARC 포맷의 기본 모델이 되었다.

LC가 MARC Pilot Project를 수행하던 시기에 영국의 BNB(British National Bibliography)는 국가 서지생산을 위한 기계가독레코드 기술의 가능성에 관한 조사를 시작했다. BNB의 MARC Office와 LC의 Information Systems Office는 두 조직간의 레코드 교환에 대한 미래 가능성을 인지하여 포맷간의 호환성 보장을 협의하였다. 그러나 상당한 노력에도 불구하고 두 개의 포맷으로 분리되었다. BNB, 프랑스, 독일, 이탈리아, 캐나다, 오스트리아를 포함한 국가 서지기관에서 수행된 Pilot프로젝트는 LC MARC II를 모델로 하여 국가 포맷의 개발을 추진하게 되었다. 각국의 자체 포맷 개발은 다른 국가 포맷과의 호환성 유지에 중점을 두었으나 서지 약정과 국가적 특성의 차이로 상세 사항에 있어 차이가 있는 포맷을 개발하게 되었다.

1980년대 중반 국제도서관연맹(IFLA : International Federation of Library Associations and Institutions)의 International MARC Programme에서 조사한 결과에

의하면 국가 서지 레코드를 MARC 포맷으로 마그네틱 테이프 서비스를 제공하는 국가가 21개국에 달하였다. 그러나 10년 후인 1990년대 중반 IFLA의 세계서지통정(UBC : Universal Bibliographic Control)프로그램에 의해 수행된 MARC 포맷 사용에 대한 조사 결과는 전세계 66개 국가 도서관이 MARC기반 포맷을 채택하고 있는 것으로 나타났다. 이들 기관의 약 50%는 USMARC를 기반으로 하고 있으며, 25%는 UNIMARC, 나머지 25%는 각 국가 포맷이나 기타 포맷으로 개발되었다. 이와 같이 30여년에 걸친 MARC 포맷의 확산은 국가 서지 기관간에 레코드 교환이나 상호운용 데이터베이스 개발에 대한 가능성을 증가시켰다.

IFLA의 UBC 프로그램은 국가 서지기관간 데이터의 국제교환에 대한 필요성을 인식하여, 1972년에 MARC데이터 국제교환을 위한 내용지정명세와 관련한 워킹그룹을 결성하였다. 이 그룹이 개발한 포맷은 강제성을 띄지 않는 표준인 'SuperMARC'으로 국가 커뮤니케이션 포맷으로 권장되는 UNIMARC이다. UNIMARC를 바탕으로 국가 MARC포맷을 개발한 대표적인 국가로는 일본과 대만이 있다. 두 나라가 이를 채택한 주요한 이유는 LC MARCⅡ보다 일반화된 포맷으로 설계되어 있고, 비 로마자 및 문자처리를 위한 메커니즘을 제공하고 있었기 때문이다.

최근 미국 의회도서관과 캐나다 국립도서관은 공동으로 USMARC과 CAN/MARC 형식을 통합한 MARC 21을 제안하였다. 이 포맷은 새로운 MARC 포맷이 아니라 21세기로의 이동과 세계적인 확장을 제시하는 국제적 특성을 지닌 MARC포맷의 연장이다. 1994년부터 1997년까지 USMARC과 CAN/MARC 사용자협의회는 유사한 형식을 제거하고 호환성을 증점적으로 서지, 전거, 소장, 분류, 커뮤니티 정보의 5개 영역으로 개발하여 1999년부터 MARC 21 서비스를 제공하고 있다. MARC 21은 프랑스어를 비롯한 9개국어로 번역되어 제공되고 있으며, 현재 40여 국가에서 사용하고 있다.

우리나라는 1976년 국립중앙도서관 업무전산화 타당성 검토에서 최우선적으로 편목업무의 전산화가 선정되었으며, 이것이 KORMARC개발의 태동이 되었다. 1980년 편목업무 전산화의 기본이 되는 서지데이터 포맷으로 LC의 MARCⅡ 포맷과 UKMARC 포맷 그리고 ISO 2709로 제안된 서지정보 교환용 포맷을 토대로 우리의 문헌적 특성을 수용하여 KORMARC 실험용 포맷(단행본)이 개발되었다. 이 포맷은

1981년 '표준용 포맷' 으로 수정, 보완되어 컴퓨터가 도입된 1982년부터 국립중앙도서관에 납본된 단행본을 중심으로 목록데이터 입력에 사용하기 시작하였다. 1983년 KORMARC운영협의회를 구성하여 표준용 포맷에 대한 전반적인 개정작업을 착수하였으며, 이미 정해진 표준용 포맷에 의거하여 작성된 인쇄카드를 전국의 도서관에 보급하였다. 1984년에는 한국문헌자동화목록법 : 표준포맷(단행본용)을 확정하였다. 1991년 단행본용과 연속간행물용 포맷의 개정과 비도서자료용이 개발되었으며, 1993년에 단행본용 포맷이 KSC 5867로 지정되었고, 1994년에는 연속간행물용 포맷이 KSC 5795로 지정되었다. 이어서 비도서자료용(1996), 전거통제용(1999), 소장정보용(1999), 고서용(2000)이 KS로 지정되었다. 현재 모든 자료의 MARC 포맷을 통합하는 KORMARC 통합서지용 포맷을 개발하는 중이다.

2.2 웹 자원 기술의 표준

웹에서 문서나 데이터를 교환하고 배포하는 표준이 될 수 있는 중요한 마크업 언어로 XML(Extensible Markup Language)이 대두되고 있다. 마이크로소프트사는 XML을 웹에서 데이터를 교환하고 배포하기 위한 표준포맷으로 정의하고 있다. 그러나 이러한 XML의 개발은 국제표준기구인 ISO가 1986년에 표준으로 제정한 문서기술언어 SGML을 웹에서 사용할 수 있도록 하는 것을 목적으로 검토가 시작되었다.

SGML은 메타언어로서 80년대에 대두된 디지털로 만들어진 문서간 호환성 결여의 문제를 해결한 국제표준이다. 1990년대에 들어서면서 SGML은 국방, 제조, 출판, 금융, 교육 등 거의 모든 분야에서 전자문서의 표준으로 선택되는 등 폭발적인 성장세를 보였다. 미국의 경우 SGML에 관한 연구가 활발하여, 주요 대학의 디지털도서관 사업, 주요 학술 잡지의 SGML 디지털화, 미 국회도서관의 전자도서관 사업 등이 진행되었다. 일본의 경우 1992년 SGML의 일본 표준(JIS)이 제정되었고, SGML에 의한 특허공문 전자화, 학술 논문공유를 위한 데이터베이스화, 전자도서관화가 본격적으로 추진되었다. 국내에서도 SGML의 중요성이 인식되어 「KS X 6011 : 문서기술언어 SGML」 과 「KS X 6010 : SGML 문서교환형식」 을 한국산업표준(KS) 문서양식으로 제

정한 바 있다.

SGML은 데이터의 공유와 재사용이 가능하고 정보의 정확한 검색과 구조적 검색을 지원하는 장점을 지니고 있어 웹에서 전자문서를 배포하거나 교환하는 기술로 적합하지만, DTD와의 조합이나 생략태그의 복원을 동반해야 하므로 처리결과가 나쁘고, 규격서의 문법정의가 복잡하고 어려워 웹에서 실제 구현하는데 어려움을 가지고 있다. 반면에 웹 콘텐츠 기술언어로서 1992년 W3C에서 제정된 HTML은 표현지향의 텍스트 외형을 규정하는 데 초점을 두고, 스펙이 단순하면서 웹 브라우저를 통한 단방향 데이터 공유가 가능하도록 설계된 SGML의 응용규격이다. 그러나 HTML은 미리 정해놓은 제한된 태그만을 사용해야 하는 기능적 제한과 작성 문서가 단순하고 구조적 기술이나 문서의 재사용이 어렵다는 단점을 지니고 있다. 따라서 HTML은 전문(full-text) 데이터의 구조적 마크업을 할 수 없고, 브라우저에 따라 서로 다른 결과를 생성하기도 한다.

이상에서와 같이 SGML의 복잡함과 HTML의 제한을 해결하는 해법으로서, SGML 규격에서 불필요한 기능을 제거하고 웹에 필요한 기능을 추가하여 XML이 개발되었다. XML은 SGML과 다르게 ISO에 의해 개발되지는 않았으나 W3C의 지도아래 약 300여개 회사들의 컨소시엄으로 1996년 11월에 표준 초안이 제안되고, 1998년 2월에 XML 1.0 표준 권고안과 제2판이 2000년 10월에 권고되었다. XML은 SGML을 인터넷에서 사용할 수 있도록 개발한 것으로서, 다양한 목적에 적합하도록 확장될 수 있는 확장성과 유연성을 갖고 있다. XML 명세는 SGML의 10분의 1정도로서, 그 구현도 더욱 간단한데 LC에서 개발한 MARC DTD의 경우 SGML은 459페이지에 달하지만, XML의 서지와 전거 DTD는 모두 11페이지에 불과하다. 이러한 이유로 XML을 "SGML-Lite"로 명명하는 학자도 있다.

2.3 MARC와 XML의 통합

도서관은 다양한 정보자원의 기술과 네트워크 환경에서 여러 응용분야로의 확장에 대한 방법론으로 SGML을 도입하였다. 도서관에서 Markup 언어의 도입은 전문데이

터와 서지데이터 영역으로 대별할 수 있다. 전자는 SGML이 갖는 문서 구조화의 장점을 이용하여 보다 향상된 색인 및 검색을 수행할 수 있고, 후자는 DTD를 통해 다양하고 융통성 있는 표현과 데이터 포맷간의 상호교환을 용이하게 할 수 있다는 장점을 지니고 있다. 그러나 웹의 발전과 함께 최근 개발된 XML 표준기술은 도서관에서 서지정보와 웹 정보자원을 통합하는 시스템으로의 유용성을 제시하는데 중요한 계기가 되고 있다. 도서관에 있어 XML은 웹을 통해 이용 가능한 다양한 전자정보원의 조직 방법론으로서 정확한 검색과 구조정보를 제공하는 새로운 분야를 창출하고 있다.

MARC와 XML의 통합 필요성은 첫째, Markup 언어인 XML은 기존 MARC의 서지정보를 인터넷 환경으로 이식하는데 용이한 표준기술이다. 이는 초기 MARC 포맷 개발시에 인지하지 못했던 변화하는 정보환경에 적응하는 확장성을 의미한다. 둘째, 기존 MARC 시스템의 활용가능성과 호환성의 문제, 목록을 비롯한 다양한 메타데이터에 대한 새로운 접근방법을 제공하는 융통성 있는 플랫폼 환경을 제공한다. 셋째, 계속적으로 새롭게 개발되는 검색도구의 활용, 메타데이터의 통합에 용이하여 보다 포괄적인 시스템의 구축을 촉진할 수 있는 기술이 될 수 있다.

한편 MARC 포맷은 기술적 Markup 언어로서 XML로의 변환이 용이하다. 도서관 서지정보의 기술은 텍스트 형태로서 목록규칙에 의한 엄밀한 논리적 구조를 가지고 있기 때문에 이를 XML 구조정보로 표현하는 DTD 설계 또한 용이하다. 나아가 MARC는 기술 내용이 상세한 메타데이터로서 비교적 정교하고 복잡한 구조를 가지므로 XML 구조정보로 표현하는 것이 더욱 효과적이다. 지금까지 MARC 포맷이 지니고 있는 대부분의 문제점들은 XML을 기반으로 한 정보기술로 개선될 수 있다. 특히 계층적 서지 구조나 복잡한 특성의 정보자원을 구조화하고 다양한 접근점을 제공하는 측면에서 MARC의 문제점이 상당부분 개선될 수 있다.

현재 많은 도서관들은 XML을 기반으로 시스템을 계획하고 있다. 그 예로 영국의 NLM은 1999년 MEDLINE 데이터베이스를 XML로 변환시키려는 계획을 완료하였다. 프랑스 정부기관은 XML로 UNIMARC를 변환하는 BiblioML을 개발했다. 미국의 LC는 1995년~1998년 사이에 수집한 문헌을 MARC에서 SGML로 전환하는 작업을 수행하였다. 이 작업에서 웹 자원과 서지 데이터 통합 방법론으로 MARC 포맷의 제한점

을 인식하여 1999년부터 XML에 대한 연구를 착수했다. 현재 MARC에서 XML로 변환하는 DTD와 MARC데이터를 XMLMARC 데이터로 변환하는 응용 프로그램을 제공하고 있다.

미국의 Lane Medical Library는 도서관 정보시스템에 있어 XML이 강력한 장점을 갖는 것으로 MARC의 변화를 주장하고 있다. 또한 MARC 형식에서 XML로의 변환과 도서관 정보시스템 벤더가 사용하는 독점 'Library Information'의 XML 대체 및 가까운 미래 XML/MARC 통합도서관시스템(ILS : Integrated Library System) 출현을 예측하고 있다.

2.4 선행연구

전세계 66개국에서 채용하여 국가 도서관 데이터베이스 시스템의 레코드 포맷으로서 사용하고 있는 MARC와 관련한 연구는 1968년 이래로 지속적으로 끊임없이 연구되고 있는 분야이다. 또한 도서관의 환경변화에 따른 미래의 목록과 관련된 연구도 다수 수행되었다. 그러나 인터넷의 표준언어로 새롭게 제기되고 있는 XML을 기반으로 한 MARC 관련 연구는 매우 미흡한 실정이다. 특히 XML을 기반으로 한 KORMARC시스템 구축과 관련된 연구는 전무한 실정이다.

MARC와 관련하여 미래 목록을 언급하는 도서관 목록의 패러다임 전환에서 XML 도입에 이르는 선행연구를 MARC 옹호론, MARC 발전론과 그리고 MARC와 XML의 통합론적 관점으로 대별하여 살펴보고자 한다.

첫째, MARC 옹호론적 관점은 MARC 포맷에서 제기되고 있는 몇 가지 문제점만 보완하면 도서관의 미래 목록으로 MARC 포맷은 변함없이 계속적으로 사용될 것이라는 견해를 가진 주장이다. MARC 옹호론적 입장의 대표적인 학자로는 Tom Delsy와 Alan Hopkinson으로서, 도서관에서 MARC 포맷은 지난 30여년 동안 변함없이 사용되었고, 앞으로도 번성할 유일한 포맷이다. MARC 포맷은 미래에도 서지 데이터의 커뮤니케이션 도구로서 계속적으로 사용될 것이지만, 고정길이 필드의 신축성 부족과 관련한 구조적 문제와 다양한 수준별 레코드의 포맷처리 문제, 그리고 레코드간의 연

결에 관한 문제는 여전히 해결해야 할 과제로 지적되고 있다. MARC 옹호론적 관점은 도서관에서 MARC는 지난 30여년동안 기술과 환경변화에 잘 적응해왔으며 몇 가지 단점들을 보완하면 안정적인 환경을 제공하는 21세기 데이터교환 포맷으로 계속적으로 사용될 것이라는 주장이다.

둘째, MARC의 발전론적 관점은 현재의 MARC 포맷에 보다 실질적이고 긍정적인 변화를 수용하는 발전적 포맷으로 HYPERMARC이나 MARC III 등을 제안하는 견해이다. MARC 발전론적 관점의 대표적인 학자로는 Barbara B. Tillet과 Michael Gorman을 들 수 있다. 미래 목록으로 MARC 시스템은 진보된 서지통정시스템의 요구에 맞게 개정되어야 하며, MARC 시스템의 대체물로 HYPERMARC 시스템이 필요하다. 미래 HYPERMARC 시스템은 표준화를 통해 형식화된 서지레코드들이 예견 가능한 서지통정시스템의 중요한 부분이 되어야 한다. 이에 도서관은 MARC 포맷을 더욱 정교한 시스템으로 수정, 발전시켜 보다 안정적인 시스템으로 사용해야 한다. 이는 현재보다 더욱 정확하게 데이터를 기술할 수 있는 특수한 코딩수단의 개발이 필요하며, MARC데이터의 검색과 표현 측면에서 더욱 용이하고 효율적인 방법을 강구할 필요성이 있음을 의미한다. 이상과 같이 MARC의 발전론적 관점은 MARC 포맷이 환경의 변화에 적합하게 수정되어 확장된 새로운 포맷으로 발전되어야 할 뿐 아니라 미래 목록으로 MARCIII와 HYPERMARC을 제안하고 있다.

셋째, MARC와 XML 통합론적 관점은 인터넷 환경에서 제공되는 OPAC 서비스는 MARC 포맷을 더욱 개방되고 호환성이 우수한 포맷으로 확장, 개선해야 한다. 이에 따른 그 개선의 새로운 방법론으로 XML 표준기술의 도입을 주장하는 이론이다. 대표적인 학자로는 K. T. Lam과 David Dorman 그리고 Dick R. Miller 등이 있다. 인터넷 환경에서 도서관의 기계가독목록 형식은 더욱 개방되고 호환성이 우수한 포맷으로 변환되어야 한다. XML은 서지레코드를 생성하고, 다양한 포맷으로 출판할 수 있다. 또한 XML은 서지레코드를 변환 없이 도서관 시스템으로 즉시 반입할 수 있는 인터넷 표준기술이다. 이는 서지레코드가 데이터 손실없이 XML과 MARC 상호교환이 가능함을 의미하며, 도서관은 XML 표준기술을 통해 언어 문제, 로마자, 전거통제와 관련한 많은 문제점들을 해결할 수 있게 될 것이다. 따라서 도서관은 지금까지의 경험

을 향상시키는 풍부한 포맷으로 가능한 한 조속한 시일내에 MARC 포맷은 교체할 필요성이 있는데, 그 교체 형식은 인터넷 표준기술 XML이 될 것이다. 이에 따라 도서관에서 개방된 인터넷 표준언어 XML기술을 이용하여 완전 통합도서관시스템을 3~4년 이내에 구현 가능할 것으로 예견하고 있다.

3. XML/KORMARC 시스템 구축을 위한 변환프로그램 개발

3.1 변환프로그램 설계 원칙

변환프로그램 개발을 위한 설계 원칙은 프로그램 설계의 전반적인 사항과 관련한 일반 설계원칙과 프로그램 세부 설계와 직접적으로 관련된 세부 설계원칙으로 나누어 설정하였다.

XML/KORMARC 변환프로그램의 설계원칙에 있어 가장 중요한 점은 한국문헌자동화목록형식의 구조를 분석하여 작성된 DTD를 이용하여 KORMARC 형식과 XML/KORMARC 형식간의 양방향 변환과정에서 정보의 손실이나 오류가 없이 데이터 변환이 이루어지도록 하는 것이다.

설계의 기본원칙으로는 첫째, 현재 KS 표준으로 제정된 한국문헌자동화목록형식 : 단행본용(1993)과 전자통제용(1999), 소장정보용(1999)을 토대로 DTD를 개발한다. 이는 개발된 DTD가 KORMARC 형식에서 XML/KORMARC 형식으로의 이행과정에서 두 형식간의 데이터 변환에 대한 무결성과 인터넷을 통한 One-Stop 변환 서비스가 가능하도록 이용자 인터페이스 환경을 중점적으로 고려하기 위한 것이다. 둘째, 변환프로그램은 KORMARC와 XML/KORMARC 양 시스템의 데이터가 완전하게 호환성을 유지할 수 있도록 개발한다. 이는 KORMARC에서 XML/KORMARC로 시스템이 변환하는 데는 일정기간이 필요하거나 일정기간 동안 두 시스템은 공존이 필수적이기 때문이다. 이에 따라 본 연구에서 개발된 DTD는 KORMARC 표준에 따라 지속적인 수정 보완을 전제로 하였다.

본 연구에서 사용한 DTD 편집기 TagFree 2000은 전체 구조도의 제공과 함께

DTD 내용을 작성하고 수정할 수 있도록 편집기능을 제공하고 있다. 한국문헌자동화 목록형식의 DTD 구조도는 KORMARC의 전체적인 내용을 대상으로 구조 정보를 분석, 설계하고, 이를 도식화하여 표현하였다. 이 구조도를 바탕으로 전체 구조정보를 파악할 수 있으며, 편집기의 DTD 자동생성기능을 이용하여 DTD 모듈의 내용을 수정하거나 추가할 수 있다.

요소별 DTD는 모듈화된 설계방식을 채택하여 15개 그룹으로 설계하였다. 총 요소는 447개로서 KORMARC에서 제공하는 그룹화 기준에 따른 최상위 요소 15개, KORMARC 표시기호로 구분되는 기본 필드에 대한 요소 93개, KORMARC 식별기호에 의해 구분되는 하위필드에 대한 요소 316개, 그리고 리더를 비롯한 고정길이 '필드'의 각 자수위치에 대한 요소 23개로 구성된다.

엔티티 선언부에서는 KORMARC 표시기호 뒷자리 두자리 숫자로 식별되는 개인명, 단체명, 회의명, 통일서명과 총서명, 주제명, 연관저록과 소장데이터 등 반복적으로 출현하는 하위필드를 그룹화하여 엔티티로 선언하였다. 엔티티로 선언된 속성(*.att) 요소는 해당 요소에 참조문(%*.att)으로 기술함으로써 DTD의 단축화를 꾀하였다.

3.2 변환프로그램 개발

KORMARC의 요소별 DTD를 기반으로 하여 KORMARC 레코드가 XML/KORMARC 데이터로의 완전한 변환과 함께 역변환이 가능하도록 양방향 변환 프로그램 개발을 목적으로 한다. 입력데이터로 KORMARC 원시데이터를 사용하여 변환을 원하는 요소를 선택할 수 있도록 설계하고, 데이터의 변환은 텍스트형식의 입력을 통한 변환과 파일형식의 변환을 모두 지원할 수 있도록 개발하였다. 본 연구에서 개발하고자 하는 변환프로그램은 두 가지 기본 목적을 전제로 한다.

첫째, 설계된 KORMARC DTD를 기반으로 KORMARC 데이터가 정확하게 XML/KORMARC 형식으로 변환되고, 변환과정에서 정보의 손실이나 오류가 없어야 한다. 둘째, 데이터 변환이 용이하도록 인터페이스를 설계하여 레코드 개별 변환과 일괄적 파일 변환을 모두 지원하도록 개발하며, 변환작업은 실시간으로 신속하고 효율

적으로 수행될 수 있도록 한다.

두 가지 기본목적을 달성할 수 있도록 변환 프로그램 개발은 KORMARC 형식에서 XML/KORMARC 형식으로 변환하는 'MARCTOXML' 프로그램과 XML/KORMARC 형식에서 KORMARC 형식으로 변환하는 'XMLTOMARC' 프로그램으로 구성하였다. 또한 데이터 변환시 사용 도서관에 적합하도록 레코드내의 변환 요소들을 선택할 수 있도록 설계하였다.

변환 프로그램은 두 형식간 변환기능과 함께 간단한 편집기능을 추가하여 필요한 경우 변환 결과를 수정하여 저장할 수 있도록 설계하였다. 또한 변환프로그램을 통해 생성된 XML/KORMARC 데이터는 시스템 구현의 입력물로 사용하여 변환데이터의 정확성과 유용성을 확인하도록 하였다.

본 변환 프로그램은 웹 환경 Internet Explorer 5.0과 Netscape Navigator 5.0 이상에 적합하도록 개발하였으며, 프로그래밍 언어로는 JavaScript와 PHP(Professional HTML Preprocessor)를 이용하여 개발하였다.

4. XML/KORMARC 시스템 구현 및 평가

4.1 XML/KORMARC 변환프로그램 평가

변환프로그램의 개발은 도서관의 현존 KORMARC 시스템과 새롭게 구축되는 XML/KORMARC 시스템과의 호환성을 보장하기 위한 선행작업의 일환이다. 따라서 변환프로그램의 평가 목적은 KORMARC 레코드가 완전하게 XML/KORMARC 데이터로 변환되고 역으로도 완전하게 변환되어 두 시스템간 호환성이 보장되도록 하는데 있다. 이러한 변환프로그램은 개발 목적과 설계원칙에 입각하여 다음의 4가지 평가요소를 추출하여 그 성능을 평가하였다. 변환프로그램의 평가요소로는 데이터 변환의 완전성, 변환 수행의 신속성, 변환 데이터의 신뢰성, 변환 방법의 용이성을 고려하였다. 평가는 사전에 준비한 KORMARC 원시데이터 100건의 데이터변환을 통한 연구자의 직접적 평가와 대학 및 전문도서관 사서들에게 웹 환경에서 직접 변환을 수행한

후 평가하도록 한 이용자 설문평가의 두 가지 방법으로 진행하였다.

변환프로그램의 평가결과 분석결과를 요약하면 편목 전문사서들의 변환 프로그램 평가 결과는 데이터변환의 완전성, 신뢰성과 신속성 측면에서 모두 6점을 선회하는 점수로 우수한 것으로 나타났으며, 변환방법의 용이성 측면에서 다소 낮은 5.40을 나타냈다. 이는 현 시스템에서 MARC 원시데이터를 반출하고, 이를 다시 복사하거나 파일을 불러 변환하는 방법이 다소 불편했던 것으로 추정된다.

4.2 XML/KORMARC 시스템 구현

4.2.1 구축 환경

XML/KORMARC시스템은 PentiumIII 600MHz Dual에서 Linux Redhat 7.0을 OS로 구현하였다. Web Server는 Apache 1.3.19를 사용하고 PHP 스크립트언어를 사용하여 XML데이터를 MySQL 2.2.3 DBServer에 저장하도록 하였다. XML/KORMARC 시스템 개발을 위한 도구로 phpDOM을 사용하였고 MySQL 데이터베이스와 연동에는 PHP 언어를 사용하였다.

4.2.2 시스템 설계

XML/KORMARC 시스템은 입력, 검색, 화면, 출력 설계 등 기능을 중심으로 설계하였다. 단행본 도서의 서지정보는 한국문헌자동화목록형식에 의거하여 작성하였으며, XML/KORMARC 시스템의 입력물은 앞서 개발한 변환프로그램을 통해 생성된 데이터를 입력물로 사용하였다.

4.2.3 시스템 구현

본 연구에서 개발한 XML/KORMARC 시스템은 3-Tier 구조로서 Client Tier,

Middle Tier, DataBase Tier로 전체 3개 층으로 구성하였다. 서지정보를 저장할 DBMS로 구성된 DataBase Tier, XML 문서 저장 및 검색을 수행하고, 이용자 층으로부터 정보를 입력받아 DataBase Tier로 XML데이터를 넘겨주는 Middle Tier, 웹 브라우저를 사용하여 시스템에 접근하는 사용자인 Client Tier로 구성되어 있다.

XML/KORMARC 시스템의 3-Tier 구조에서 이용자는 웹 브라우저를 통해 검색 조건을 입력한다. 이용자의 질의는 XML언어로 Middle Tier에 넘겨져 하부의 XML 검색기에서 처리할 수 있는 질의어를 생성해서 XML DataBase Tier에 넘겨진다. 이에 따라 XML DataBase Tier는 질의어로 검색을 수행한 결과를 Middle Tier로 넘겨주어 이용자가 검색 결과를 XML 데이터로 볼 수 있도록 개발되었다. 본 시스템은 도서관 서지정보와 인터넷 서점에서 제공하는 도서 구입정보를 함께 공유함으로써 좀더 확장된 기능을 이용자에게 제공할 수 있도록 설계하였다.

4.3 XML/KORMARC 시스템 평가

본 연구에서 구현한 시스템의 평가는 시스템 성능측면에서의 평가와 편목 전문사서들의 시스템 평가를 병행하여 수행하였다. 시스템의 성능 평가는 KORMARC 와 XML/KORMARC 시스템 성능을 그 특성별로 비교 평가하였고, 만족도 평가는 대학 및 전문도서관 편목 전문사서들을 대상으로 시스템을 사용 후 SERVQUAL 도구로 작성된 설문지를 통해 평가하였다.

XML/KORMARC 시스템의 평가 목적은 다음과 같다. 첫째, 지금까지 MARC 포맷의 문제점으로 제기되어 왔던 몇 가지 문제점이 XML/KORMARC 시스템에서 개선되었는가의 여부를 평가한다. 둘째, 웹 기반 XML/KORMARC 시스템을 최소 서비스수준, 이상적 서비스수준, 현행 서비스수준을 의뢰하여 정보시스템에 있어 특정서비스 기능에 대한 편목 전문사서들의 중요도 인식을 조사하여 평가하고 시스템의 개선점을 추출하는데 그 목적이 있다.

시스템의 만족도 평가는 각 기능에 대하여 7점 척도로 작성되었다. 이상적 기대치에서 산정된 평균 기대치와 실제 제공된 서비스에서 산정된 제공서비스 평균치는 사

분면의 기준선을 나타낸다. 사서들의 이상적 서비스 기대치는 수직축에 나타나며, 실제 제공된 서비스 인식에 대한 평균은 인식된 기대치에서 제공된 서비스를 뺀 차이를 각 기대치로 계산해서 수평축에 나타낸다.

1사분면에 해당되는 기대들은 사서들에게 매우 중요한 요소이며, 2사분면에 해당되는 기대들은 사서들에게 매우 중요하지만 중요한 서비스로 인식되지 않은 것들이다. 3사분면에 나타나는 기대치들은 사서들에게 그다지 중요하지 않지만, 서비스와 연결하여 생각하고 있는 요소들이다. 4사분면은 실제 제공되는 서비스 질도 낮고 사서들도 중요하게 인식하고 있지 않는 요소들이다. 따라서 4사분면은 새로운 서비스로 개발할 필요성이 있는 분야로서 이용자들을 충족시키기 위한 기대치로 무엇이 있는가를 검토하는 도구로 활용할 수 있다.

다음은 편목 전문사서들에게 SERVQUAL 설문지를 통해 받은 이용자 만족도에 대한 인식도 차이를 분석한 자료이다. 서비스에 대한 기대치와 제공된 서비스간의 차이를 토대로 기대된 서비스에 대한 10개의 평균치는 수직축을 따라 위치하고, 이러한 기대치와 실제 서비스에 대한 인식도 간의 차이는 수평축을 따라 위치한다. 응답자 20명에 대한 10개의 기대치 점수를 합하고, 그 결과를 10으로 나누어 6.17의 평균치를 얻었다. 실제 서비스에 대한 평균치는 4.77이다. 서비스에 대한 기대치와 실제 서비스 제공과의 차이는 평균치 1.40의 차이를 보이고 있다.

이상의 결과를 요약하면 편목 전문사서들은 XML/KORMARC 시스템에서 정보제공의 풍부성과 데이터 변환의 편의성, 서비스 확장의 용이성에 중요도를 인식하고 있는 것으로 나타났다. 그러나 다양한 레코드의 지원과 검색의 신속성 요소는 다소 중요도가 떨어지는 것으로 나타났다.

한편 편목 전문사서들은 시스템의 품질 평가요소로서 신뢰성, 안전성에 가장 중요도를 두고 있으며, 다음으로 효율성, 접근 편의성 순으로 중요도를 두고 있는 것으로 나타났다. 또한 시스템의 품질 평가요소로서 미적이고 감각적인 시스템의 디자인적 외양성에 가장 낮은 중요도를 두고 있는 것으로 나타났다.

5. 결론

현재 도서관 정보시스템을 구성하고 있는 MARC 포맷은 경직성으로 인해 새로운 정보의 수록이나 통합이 어려워 표현형식의 개선이 필요한 패러다임 전환의 시점에 놓여있다. 따라서 본 연구에서는 KORMARC 포맷 개선을 위한 방법론으로 XML 기술을 도입하여 XML/KORMARC 시스템을 구현하였다.

XML/KORMARC 시스템을 구현한 본 연구의 연구결과를 요약하면 다음과 같다.

첫째, KORMARC 구조를 분석하고 DTD를 설계하여 KORMARC 데이터와 XML/KORMARC 데이터의 양방향 변환 프로그램을 개발하였다. 변환프로그램을 통한 변환 데이터는 완전성, 신속성, 신뢰성 측면에서 우수한 것으로 평가되었다.

둘째, 변환 프로그램을 통한 XML/KORMARC 데이터의 생성은 지금까지 도서관이 구축한 대량의 KORMARC 데이터를 완전하게 활용할 수 있고, 양방향 데이터 변환을 통해 XML/KORMARC와 KORMARC 두 시스템간 호환성을 확보할 수 있다.

셋째, XML/KORMARC 시스템은 도서관간 상호운용성을 기반으로 MARC 포맷의 경직성과 도서관간 수준별 레코드 처리, 시스템간 호환성, 데이터 공유 측면에서 개선된 성능을 확인하였다.

넷째, 도서관 서지정보에 인터넷 서점의 수서정보를 연계함으로써 정보서비스 영역의 확대를 시도하였으며, 향후 출판사, 밴더 등과 연계하는 서비스 확장이 가능함을 확인하였다.

다섯째, 도서관시스템간 연계방안을 제안하기 위해 본 연구에서 설계된 KORMARC DTD를 웹상 공유로 설정하였다. 본 시스템을 통해 제공되는 XML/KORMARC 데이터는 모든 도서관에서 재작성 없이 사용 가능하며, XSL의 설계만으로 자관에 적합한 다양한 인터페이스를 구현할 수 있다. 이는 별도의 변환 프로그램이나 추가적인 노력없이 웹상에서 직접 데이터 상호호환과 자원을 공유할 수 있을 뿐만 아니라 도서관 서비스를 확장할 수 있음을 의미한다.

여섯째, 시스템 품질 평가 요소를 측정된 결과 사서들은 시스템의 신뢰성과 안정성에 가장 중요도를 두며, 정보의 풍부성, 연결의 다양성, 서비스 확장의 편의성 등의 서비스 요소를 중요하게 인식하고 있는 것으로 나타났다.

이상과 같이 본 연구에서는 지금까지 도서관 정보서비스의 주류를 이루고 있는 서지데이터베이스의 KORMARC 포맷을 재설계하여 XML/KORMARC 포맷으로 변환하고, 도서관 자원으로 통합되어야 할 다양한 자원의 정보조직의 문제를 해결하는 기술로 XML 표준기술을 제안하였다. 이는 지금까지 도서관에서 서지데이터베이스로 축적해온 MARC 데이터를 그대로 활용하면서 향후 도서관 정보시스템에 통합되어질 서지정보 이외의 사실정보, 참조정보 및 전문정보의 통합 기술로 XML 표준기술 도입이 가장 효율적인 방법임을 제시하고자 한 것이다.

XML 표준기술을 도입한 XML/KORMARC 시스템은 KORMARC의 장점을 그대로 유지하면서 도서관 자원의 공유와 호환성 및 확장성 측면에서 지금보다 더 큰 이익을 도서관에 가져다 줄 것이며, 도서관은 더욱 확대된 서비스 영역을 개발하여 이용자 만족을 극대화할 수 있게 될 것이다.