

# 분담목록시스템을 위한 데이터 표준화(Ⅲ)

서울대학교 중앙도서관  
조수련 · 조순영

## 목 차

- |          |             |
|----------|-------------|
| I. 서 언   | 2. 검색시스템    |
| II. 본 론  | 3. 상호대차 시스템 |
| 1. 목록시스템 | III. 결 언    |

## I. 서 언

1991년 국공립 대학 도서관 전산화 사업 추진의 일환으로 구성된 목록분과위원회는 사실상 제주대학에서의 제17차 회의를 끝으로 일단계 표준화 작업을 마무리하게 되었다. 그간 5년여에 걸쳐 지속적인 표준화 작업을 해오면서 실제 많은 부분에서 표준안을 제시하고 분담목록의 기반을 다진것은 사실이나 도서관 목록 데이터의 성격상 아직도 미결된 부분이 적지 않고 문자 코드 문제나 목록규칙과 같이 실제 일개 분과 위원회에서는 결정할 수 없는 국가적인 차원이나 도서관계 전체에서의 결정이 선행되어야 하는 문제들을 아직도 안고 있는 것이 사실이다. 따라서 이들 부분은 앞으로도 시스템이 가동되는 시기까지 지속적으로 연구 검토하고 예측이 어려운 부분들에 대해서는 실제로 시스템을 운영하면서 부분적인 조정을 해야 될 것으로 생각한다. 두차례에 걸친 시험적인 데이터 업로드 과정에서 최소한의 목록 규칙을 벗어난 데이터의 양과 사례가 당초의 예상을 넘어 상당량 발견되었다. 그래서 표준화 작업이 필요한 것이지만 현실은 그 보다 훨씬 심각했고 급기야는 오류 데이터 및 중복데이터의 기준을 대폭 완화할 것인지 데이터의 품질관리와 운용상의 편의를 위해 DB 공유형의 원칙을 고수하고 이용 기준을 엄격히 제한할 것인지

지 결정을 해야 하는 상황이었다. 거듭된 회의 결과 결론적으로 센터 DB의 양이 어느 정도 많아지고 시스템의 이용이 일정수준으로 오를때까지는 일단 초기에 일괄 업로드되는 데이터의 기준을 대폭 완화함으로써 많은 도서관에서의 시스템 이용을 지원하고 그 이후 온라인으로 업로드하는 데이터 부터는 본래의 오류 및 중복 데이터 기준을 적용하기로 하였다. 이에 세번째 연재되는 본고에서는 분담목록의 변경된 표준안과 기능에 대해 설명하고 이후 시스템의 운용 방향을 제시하고자 한다.

## II. 본 론

분담목록은 크게 목록, 검색, 상호대차의 세가지 기능으로 구분되며 이들 기능간은 물론 모든 분담 목록 가입 기관의 시스템과도 상호유기적으로 운영된다.

### 1. 목록시스템

#### 1) 서지

##### 가) 일반

• 센터에서 확장하여 사용하는 tag

- 016 : 각 가입 기관에서 센터데이터의 업/다운로드 시 참조할 수 있도록 서지, 전자, 소장 모두에 tag 016을 두어 각각의 센터통제번호를 부여한다.
- 019 : 2개의 서지데이터가 통합되는 경우 통합한 서지의 tag019에 통합된 서지의 센터제어번호를 기입한다.
- 940 : 동양서의 읽기 및 정확한 키워드 생성을 위해 띄워 쓰기한 서명을 재 입력한다. 이때 제 1지시자를 사용하여 동양서의 서명 이외에 해당하는 tag은 삭제한다.

제1지시자 ∅ : 동양서의 서명읽기

| : 각 도서관별로 사용하는 local tag

• tag 050, 060, 070, 082을 사용하여 각 기관의 청구기호를 타기관에서도 참

조할 수 있도록 한다.

나) 일괄 업로드

- 일괄서지업로드는 서지신규, 소장추기및 소장수정만 가능하다.  
(일괄서지수정은 time-stamp비교 등의 복잡한 기능이 추가될 뿐 아니라 서지의 질도 일정하게 유지할 수 없으므로 일괄작업에서는 제외한다.)
- 업로드된 DATA는 PC의 OK FILE에 WRITE하여 LOCAL DB의 운영이 쉽도록 한다.
- 중복 1,2차 키 동일서지는 동일서지로 처리한다.
- 기관의 제어번호가 없는 데이터는 받아들이지 않는다.
- 각 서지마다 관련 전거데이터를 자동 생성하고 표목 점검도 동시에 수행한다.

다) 일괄 다운로드

- 센터에 의뢰시 일괄로 다운로드한다.
- 다운로드 형태
  - 서지만 다운로드
  - 의뢰 기관의 소장사항이 포함된 형태의 서지 다운로드
- 검색항목
  - 기본항목: 소장기관, 생성기관, 수정기관, 출판사
  - 제한항목: 생성일자, 최종수정일자, MARC구분, 언어, 자료유형, 출판년
- 기본처리  
센터 제어번호를 tag016에 기입하고 소장사항에 기관의 제어번호가 있으면 tag001에 기입한 후 다운로드 한다.

라) 온라인 작업

- 서지에 소장사항을 포함할 수 있다.
- 센터서지 및 참고DB 검색 후의 작업
  - 신규 : 검색 결과가 없거나 유사한 서지 검색 후 신규로 작성 후 업로드
  - 수정 : 검색된 서지를 수정
  - 추기 : 검색된 서지에 추기

- 다운로드 : 검색한 결과를 PC file에 저장

## 2) 전거

### 가) 일반

- 인명, 주제명 전거를 함께 관리한다.
- 기존 전거에 대한 수정은 센터에서만 가능하다.

### 나) 표목의 검증

- 동일한 주제명표목표를 사용하는 레코드(008/11)에 대해 다음의 경우는 오류로 처리한다.

센터 전거 1XX(5XX) 표목 == 생성된 전거 1XX(5XX) 표목

센터 전거 1XX(5XX) 표목 == 생성된 전거 4XX 표목

센터 전거 4XX 표목 == 생성된 전거 1XX(5XX) 표목

### 다) 일괄 다운로드

- 센터에 의뢰시 일괄로 다운로드한다.
- 서지 검색 후 관련 전거레코드를 다운로드 한다.
- 검색항목 (서지 검색)
  - 기본항목 : 소장기관, 생성기관, 수정기관, 발행기관
  - 제한항목 : 생성일자, 최종수정일자, MARC구분, 언어, 자료유형, 발행년
- 기본처리  
센터 제어번호를 tag016에 기입하여 다운로드 한다.

### 라) 온라인 작업

- 전거레코드 검색 후 다운로드 한다.
- 수정은 할 수 없고 서지 생성시 관련 전거의 신규생성은 허용한다.
- 서지작업 중 전거관련표목에 대해 통제 한다.

## 3) 소장

### 가) 일반

- 센터에서 지원하는 데이터 형태.
  - USMARC Format for Holdings data ( 분리형태 )
  - 센터에서 수정한 삽입형태
- 삽입형태의 소장사항 처리는 서지와 함께 처리하며 분리형태인 경우만 별도로 처리한다.
- 작업은 서지 검색 후 해당 서지에 대해서만 가능하다.
  - 신규
  - 수정( 추기와 동일 )
  - 삭제
- 삽입형태의 소장사항 기입 규칙

자료유형	Tag	Subfield	내 용
모든 유형	841*	a 레코드의 유형	소장레코드의 LDR/06
		b 고정장 데이터	소장레코드의 008
		c 입력수준	소장레코드의 LDR/17
연 속 간행물	866/867/ 868** (반복가능)	a 기관명	센터에 등록된 기관명
		b 분관명	센터에 등록된 분관명
		h 매체	소장레코드의 007***
		t 서술형 소장데이터	실제 소장하고 있는 사항****
단행본	599	a 기관명	센터에 등록된 기관명
		b 분관명	센터에 등록된 분관명
		h 매체	소장레코드의 007***
		t 서술형 소장데이터	실제 소장하고 있는 사항****

\* 추후 좀더 상세한 기관별 소장정보를 제공하기 위해 사용

\*\* 서지의 007이 복수개면 해당 매체마다 소장을 분리, 소장 분관이 다른 경우 분리

\*\*\* 소장포맷에서 사용하는 코드를 쉽게 알아 볼 수 있도록 아래 표와 같이 기입

\*\*\*\* 서술시 구두점 사용은 ANSI Z39.44에 따른다.

007/00-01	Display text	007/00-01	Display text
ha	Aperture Card	tb	Parge Print
hb	Microfilm Cartridge	tc	Braille
hc	Microfilm Cassette	tz	Other Text
he	Microfilm Reel	zu	Unspecified
hf	Microfiche	zz	{None}
hg	Microopaque	dc	Audio Cassette-Tape
hz	Microform	ad	Audio Disc
mm	Multiple Media	vf	Video Cassette
ta	Text	co	CD-ROM

기호	이름	목적	예
-	Hyphen	연속되는 소장사항 표기	1953-1957
,	Comma	소장사항 전체에서 결호 표기	1942-1945, 1953-1955
/	Diagonal	두 연도를 하나의 기간으로 하던가 두 권은 한 권으로 표시하기 위해 사용	1957/1958
?	Question mark	날자에서 모르는 부분을 나타내기 위해 사용	197?
:	Colon	여러 단계별로 표시되는 소장의 단계 구분	v.1:no.1:pt.1
;	Semicolon	출판되지 않은 결호 표기	v.1-4:v.6
	Space	소장데이터 요소를 구분	CaOOn Aero
()	Parentheses	계수적, 연대기적 표기가 함께 사용될 때 구분을 위해 사용	v.1 (1950)
=	Equal signs	계수적표기시 대체번호를 구분	v.2:no.5=no.11
{} []	Square brackets	연대기적 표기에서 현행 태양력으로 변경된 날자를 추가하는데 사용	Showa 56 nendo [1981 1982]

#### 나) 일괄 업로드

- 분리형태의 소장 추가 및 수정만 가능
- 센터에 있는 서지와 연결될 수 있는 항목으로 센터제어번호, 기관별제어번호가 있으며 이 가운데 기관별제어번호는 반드시 존재하여야 한다.

다) 일괄 다운로드

삽입형태의 경우는 서지 다운로드시 함께 처리되며 분리형태만의 다운로드  
는 불가능하다.

라) 온라인 작업

- 센터서지DB 검색 후 연결된 서지에 대해 다음과 같은 작업  
선택 가능
  - 신규 : 해당 서지에 소장 기관 및 사항을 추가
  - 수정 : 기존의 소장사항을 수정
  - 삭제 : 기존의 소장사항 삭제

## 2. 검색시스템

검색은 일반이용자를 위한 기능으로 상호대차시스템으로 연결된다.

1) 검색대상 Database

가) 센터 기본 서지DB

나) 참조서지DB : 최신데이터로 유지

- LC MARC 데이터
  - 수록기간 : 1968-1994
  - 종 수 : 5,676,000
- DDB(독일)
  - 수록기간 : 1976-1994
  - 종 수 : 1,972,000
- BNB(영국)
  - 수록기간 : 1984-1993
  - 종 수 : 523,000
- National Diet Lib. (일본)
  - 수록기간 : 1993-1994
  - 종 수 : 141,000

2) 검색시스템 고려사항

가) 각 주제명표목표의 분리 검색

LC, NLMC, NALC등의 주제명표목을 별도의 검색항목으로 구분하여 검색하게 하되 주제명 키워드 검색시는 통합검색이 되도록하여 주제명표목표에 따라 달리 사용되는 용어로 인한 검색 실패를 방지하고 각 표목표에 익숙하지 않은 이용자를 위해 통합된 키워드 환경을 동시에 제공한다.

나) Hyphen( - )의 처리

서명, 저자명에 따라 달리 적용하던 hyphen의 처리를 스페이스 처리로 통일하여 서명에 포함된 저자명과 일반 저자명의 색인이 달리 생성되어 야기되는 혼란을 피한다. 단, 동양서 인명의 로마나이즈로 인해 음절 단위로 분리된 색인이 생성되지 않도록 tag 9XX를 사용하여 변형한다.

다) 색인 생성시 2 Byte 특수문자의 삭제

기존 서지에 사용되던 2byte 특수문자를 색인생성시 삭제하여 소급작업을 줄인다.

A1.A4, A1.AD, A2.B0, A3.DB, A3.DD, A1.B4, A1.B5,

A1.B6, A1.B7, A1.B8, A1.B9, A1.BA, A1.BB, A1BC, A1BD, A1.D7

라) 동자이음 한자처리

2가지 이상으로 읽히는 한자의 경우 한자생성을 위해 선택된 한글이 다르면 생성된 한자는 동일하나 한글읽기 색인 생성시 다른 색인이 생성되므로 동자이음 한자에 대한 데이터를 구축하여 다양하게 색인을 생성하는데 참조한다.

마) 형태소분석

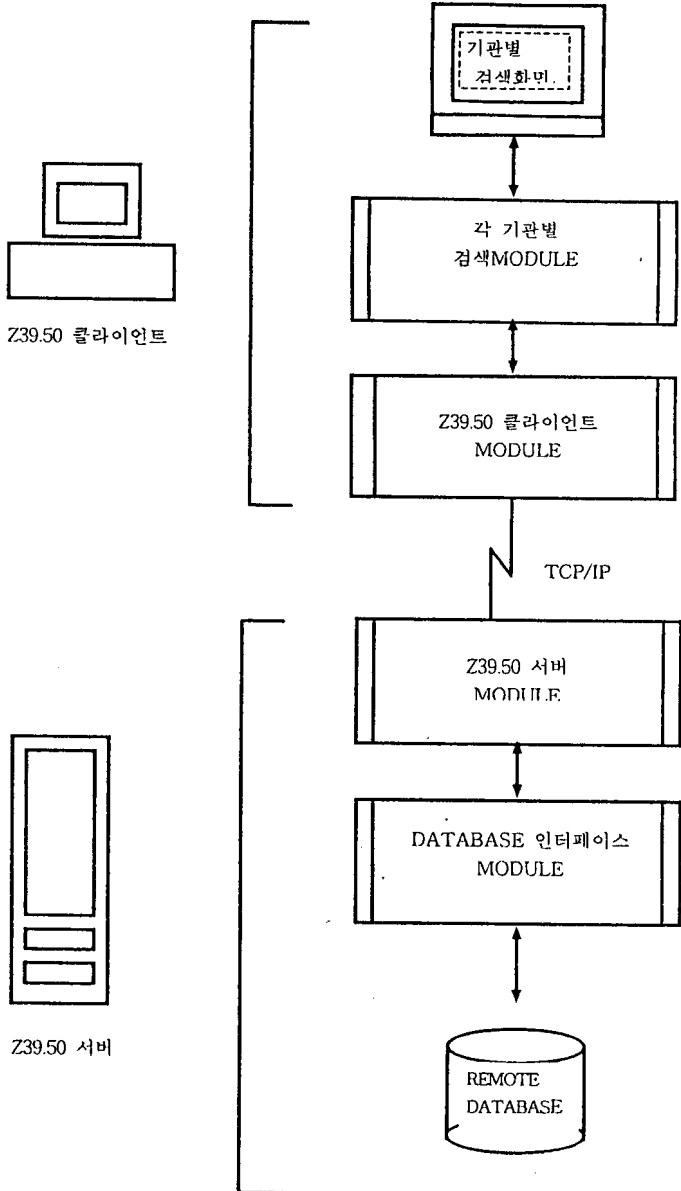
숙어사전에 등록되지 않은 단어의 경우 끝부분의 음절들이 조사사전에 등록되어 있으면 조사에 해당하는 부분이 자동 삭제되고 색인이 생성되는데 이 경우 숙어사전이 완벽하지 않아 잘못 생성되는 색인을 방지하기 위해 형태소분석기를 도입한다.

3) Z39.50 서버/클라이언트 개발

분담목록시스템은 각 기관에서 사용하는 일반 검색과 달라 표준화 방식을 제공하여 이미 다른 시스템을 사용하고 있는 이용자들이 친숙하게 사용할 수 있는 환경을 제공해야 한다. 따라서 센터에서는 Z39.50 서버를 개발하여 각 기관이 사용하고 있는 여러 종류의 검색시스템을 그대로 수용할 수 있도록 환경을 제공할 뿐아니라 기관 고유의 검색시스템을 위한 Z39.50 클라이언트 개발도 지원한다.

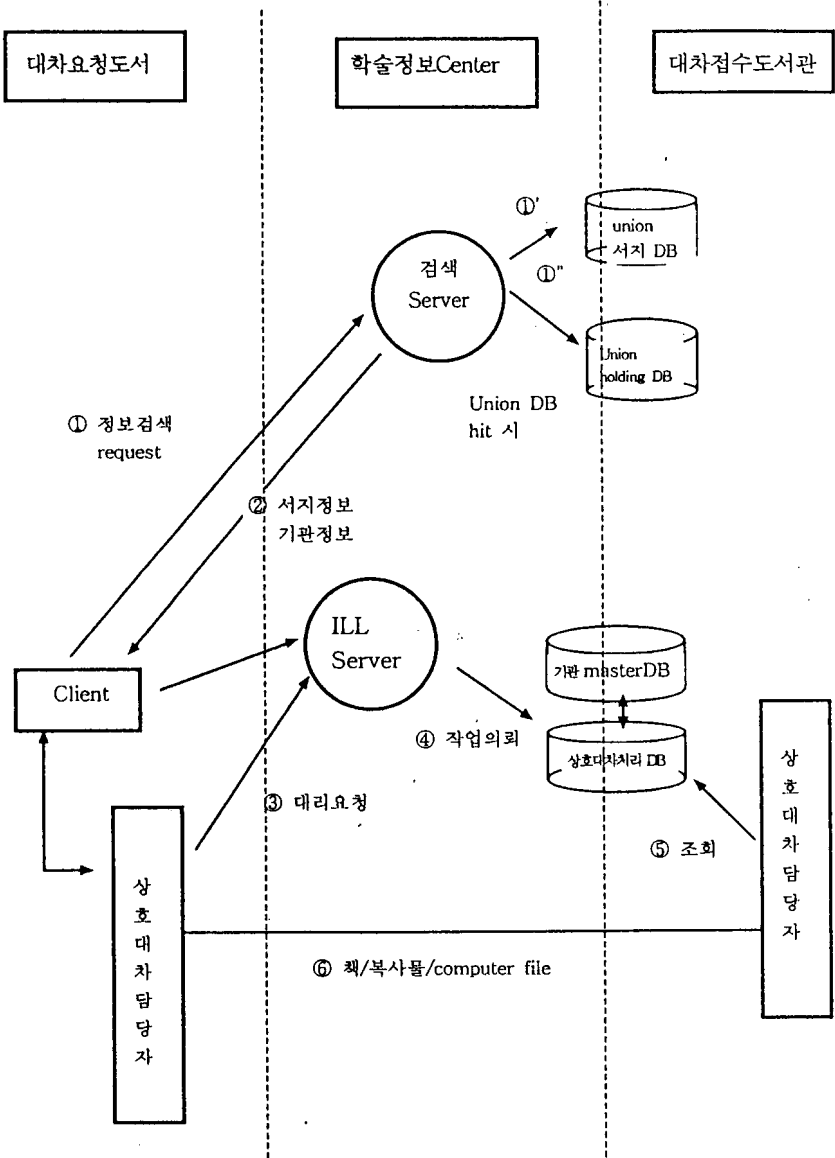


(Z39.50 클라이언트 서버 개념도)



### 3. 상호대차시스템

#### 1) 상호대차시스템의 흐름



## 2) 기본기능

- 소장확인

검색을 통해 소장기관과 소장사항을 확인한다.

- 대차요청 및 접수

- 원문복사서비스

FAX를 이용한 원문복사 및 디지털자료의 통신을 통한 전송 등의 기능이 포함된다.

- 상호대차 규약 관리 및 회계처리

- 실물 자료의 상호대차

## 3) 기관별 협조사항

각 기관의 소장사항과 상태를 정확히 파악하는 방법에는 각 기관의 모든 소장정보를 센터에서 유지하거나, 요청이 있을 때마다 기관의 소장 데이터를 조회하는 두 가지가 있다. 그러나 첫번째 방법은 센터에 지나치게 많은 정보가 치중될 뿐 아니라 기관에서 변경이 일어나는 내용을 즉시, 또는 일단위로 갱신해야 하므로 센터나 기관시스템 모두에 영향을 미치게 되어 두 번째 방법을 채택한다. 이를 위해 다음과 같은 정보를 센터에서 유지해야 하며, 각 기관의 대출 상태DB구조를 통합하거나 검색에서의 Z39.50과 같은 인터페이스 프로그램의 개발이 필요하다. 이러한 프로그램의 개발은 각 기관의 협조 없이는 불가능하며 인터페이스 프로그램이 완성되기 전까지는 소장데이터에 포함되어 있는 정보만을 기준으로 운영해 나아가야 할 것이다.

〈센터에서 관리하는 기본 데이터〉

- 기관/분관명 : 센터에 등록된 것으로 기관(분관)에 대한 유일한 키로 사용된다.
- 시스템의 IP 주소 : 각 기관에서 대출상태데이터를 관리하는데 사용되는 시스템의 IP주소로 TCP/IP를 통한 원거리 접속을 위해 사용
- LOGIN/PASSWORD : 사용이 제한된 ID를 부여하여 간단한 조회나 작업이 가능하게 한다.
- 담당자명
- EMAIL 주소

### III. 결 언

분담목록 시스템은 어느 한 도서관의 시스템이 아니고 처음 운용시부터 많은 기관에서 동시에 운용해야 하는 시스템이기 때문에 더더욱 철저한 준비 과정 및 테스트가 요청된다. 각 도서관 전산 시스템의 초기 운용시 경험했던 많은 시행착오나 오류가 분담목록시스템에서 만큼은 최소화 되어야 할 것이다. 따라서 센터에서는 운용시기가 다소 지연될지라도 최대한의 준비 기간과 테스트 기간을 갖고 작업에 임하려고 한다. 또한 시스템의 모든 기능을 동시에 가동하지 않고 단계별로 가동하여 안정적인 출발을 하고자 한다. 1단계로 서지데이터의 목록, 검색기능을 가동하고 각 회원 도서관의 소장정보가 충분히 확보되는 시기에 2단계로 상호대차시스템을 가동하여 대학의 일반 이용자에게 직접적인 혜택이 주어지도록해 나갈 것이다. 물론 그동안 운영해오던 국공립대학 중심의 표준화 작업은 더이상 안할지라도 국·사립 대학의 구별 없이 실제 모든 회원 도서관을 대표할 수 있는 표준화 위원회를 재구성하여 지속적인 표준화 작업을 추진하고 보다 훌륭한 도서관 협력 시스템으로서의 발전을 위해 해외의 우수한 분담목록센터와의 협력 관계를 계속적으로 추진해 나갈 것이다.