

통합 선거데이터베이스 구축을 위한 기초연구

- 서울대학교 도서관을 중심으로 -

박 선 희

(서울대학교 중앙도서관 사서)

< 목 차 >

1. 들어가는 말
2. 전국 대학도서관 선거 구축 현황
3. 서울대학교 선거화일
 - 3.1 선거 데이터베이스의 Format
 - 3.2 선거레코드의 유형
 - 3.3 선거표목 표기의 원칙
 - 3.4 선거데이터 작성과정
4. 도서관의 선거시스템 구축 현황
5. 맺는 말

1. 들어가는 말

도서관의 시스템은 목록, 수서, 대출 및 검색, 연속간행물, 자료교환 및 기증 등 여러 가지의 서브시스템들로 구성이 되어있으며 이러한 서브시스템들이 하나로 합쳐진 통합시스템으로 운영되고 있다. 모든 업무가 유기적으로 연결이 되어 있어 어느 하나라도 중요하지 않은 업무가 없지만 결국 이 모든 시스템들은 이용자들의 검색의 편의를 도모하고자 운영되는 것이다. 이러한 검색의 효율을 위해 전거DB 구축의 필요성이 제기되는 것이다.

서지데이터에 있는 용어를 그대로 색인화해서 사용할 경우 누락되는 자료를 검색할 수 있게 하고, 이용자는 관련된 용어 중 어느 하나만 알더라도 검색할 수 있도록 해주는 것이 전거DB이다. 즉 관련된 용어들을 하나의 그룹으로 묶어주고 그 용어들 중에서 대표적인 용어(표목)를 선정하여 다른 용어와의 관계를 규정하고 서지 DB와 연결 하므로써 이용자에게 최대의 검색의 편의와 검색 결과를 제공해 주는 것이다. 전거 DB는 한 도서관의 DB 검색에서뿐만 아니라 여러 개의 DB를 공유하는 분담 목록 시스템에서 더욱더 필요한 것으로 DB검색의 필수 요소가 되는 것이다. 즉 서울대 도서관 시스템 뿐만 아니라 대학 도서관 학술 정보전산망을 위한 DB 구축 사업의 기초로서 표준이 될 수 있는 전거 DB의 구축은 매우 중요한 것이다.

도서관에서 모두 목록을 작성하지만 전거 레코드를 작성하는 기관은 그리 많지 않다. 따라서 국내 도서관의 전거화일 구축현황을 알아보는데 그 초점을, 일부 몇몇 대학도서관들에 맞춰 조사해 보고자 한다.

2. 전국 대학도서관 전거 구축현황

대학명	전거항목	개인명 단체명	회의명	통일서명 (총서명)	주제명	계	전거DB 구축 적용형식
가톨릭대학교 중앙도서관		12,116		52		12,168	기타
강원대학교 삼척캠퍼스 중앙도서관		36				36	KORMARC

경성대학교 중앙도서관	15,615				15,615	기타
경희대학교 서울캠퍼스 중앙도서관	143				143	KORMARC
경희대학교 수원캠퍼스 중앙도서관	243,208	2,881	13,190	16	259,295	USMARC
관동대학교중앙도서관	307,331				307,331	USMARC
나사렛대학교 도서관	49			65	114	기타
덕성여자대학교 도서관	43,688	45			43,733	KORMARC
상주대학교 중앙도서관	1,729				1,729	KORMARC
서강대학교 로올라도서관	177,761		20,163	1,271	199,195	KORMARC
서울대학교 중앙도서관	297,626	3,144	19,136	63	319,969	MARC21
신라대학교 합정보센터	12,835		176		13,011	기타
안동대학교 도서관	11,685				11,685	USMARC
연세대학교 중앙도서관	496,870			26,890	523,760	KORMARC
이화여자대학교	139,576	401	12,423		152,400	USMARC
전남대학교 도서관 여수분관			46	46	92	KORMARC
한성대학교 학술정보관	532				532	기타
호서대학교 중앙도서관	12,903				12,903	USMARC

[표 1] 2006년 기준 전국 대학도서관 전거구축 현황

3. 서울대학교 전거화일

이러한 전거화일은 대부분 대학도서관에서 도서관자동화시스템을 구축하면서 운영을 시작하였다. 서울대학교는 1992년부터 전 국립대학 도서관에서 공동으로 사용할 수 있는 도서관 전산화소프트웨어(SOLARS : SeOul Library Automation and Research System)를 한국컴퓨터 주식회사와 공동으로 개발하였다. 2006년 두차례의 확장 개발을 통하여 SOLARS III로 업그레이드 되었으며 전거DB의 구축도 10년을 넘어서게 되었다.

3-1. 전거DB의 Format

전거데이터는 기본적으로 서지데이터와 동일한 레코드 구조를 가지고 있다. 데이터 필드에서는 동일한 인명이나 총서명, 주제명의 다양한 표기를 하나의 그룹으로 형성한 후 목록 규칙에 의거하여 그 중에 대표되는 표목을 1XX 필드로 채택한다. 그 외 표목은 4XX로 기술하여 “~보라” 참조를 해 주고, 관련되는 표목은 5XX에 기술하여 “~도 보라” 참조를 한다.

주요필드는 다음과 같다.

0XX 제어정보, 식별정보 및 분류기호

1XX 표목

4XX 부출 지시 “보라” 참조

5XX 부출 지시 “도 보라” 참조

6XX 데이터 처리 정보, 주기 등

9XX 로컬 필드

- 전거 Format은 언어와 무관하게 단일화하며 USMARC Format for Authority Data를 기본으로 한다.

- 전거레코드 표목의 유형은 Name, Series, Subject 로 대분한다.

- 검색 key가 되는 index는 1XX, 4XX, 5XX 만을 대상으로 한다. 따라서 검색에서 전거레코드의 ‘~보라’ 부출필드 (4XX)와 ‘~도 보라’ 부출필드 (5XX)에 작성된 표기로 검색을 하여도 해당되는 관련 서지레코드를 검색할 수 있다.

3-2. 전거 레코드의 유형

- 이름 전거 : 개인명, 단체명, 회의명, 총서명, 통일서명이 포함된다.

- 주제명 전거 : 일반 주제명, 지명, 개인명, 단체명, 회의명, 통일서명이 포함된다.

3-3. 전거표목 표기의 원칙

3-3-1. 채택표목(Established Heading)의 표기

- 한국인명과 한글 통일서명 및 총서명 : 한글

- 외국서 (중국서, 일서 포함)는 서지 레코드의 표목선정 원칙과 동일

- 일어 : “수정 Hepburn”식 로마자표기

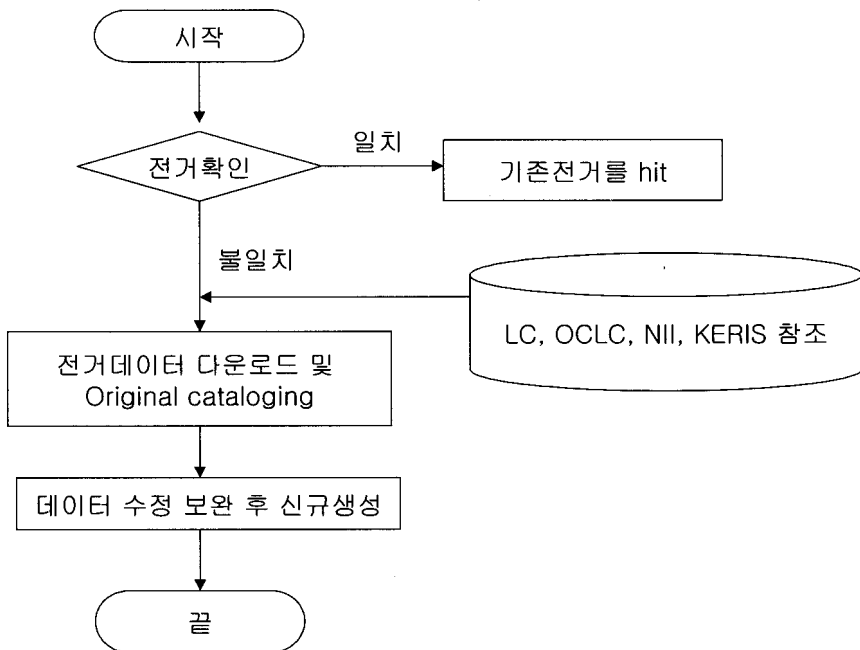
- 중국어 : “Pinyin”식 로마자표기
- 기타 비로마자 : “ALA Cataloging rules for author and title entries (Chicago : ALA, 1949)” p. 243 (IV. Transliteration)

3-3-2. 변형표기

- 문교부 편수자료 II-2 (외래어표기용례)에 따라 표기하되, 변형표기가 있을 때마다 작성한다.
- 로마자의 타 언어 읽기 : 변형이 발생할 때마다 작성
- 원어표기 : 표기가 나올때 마다 추가
- 타 언어표기 : 표기가 나올때 마다 추가

3-4. 전거데이터 작성과정

전거데이터의 작성과정은 전거레코드를 검색하여 기존에 이미 전거레코드가 생성되어 있는 것과 동일한 것이라면 신규로 작성하지 않고 기존 레코드를 따르고, 생성되어 있지 않거나 기존 데이터와 일치하지 않다면 신규로 작성하는 과정을 따른다.

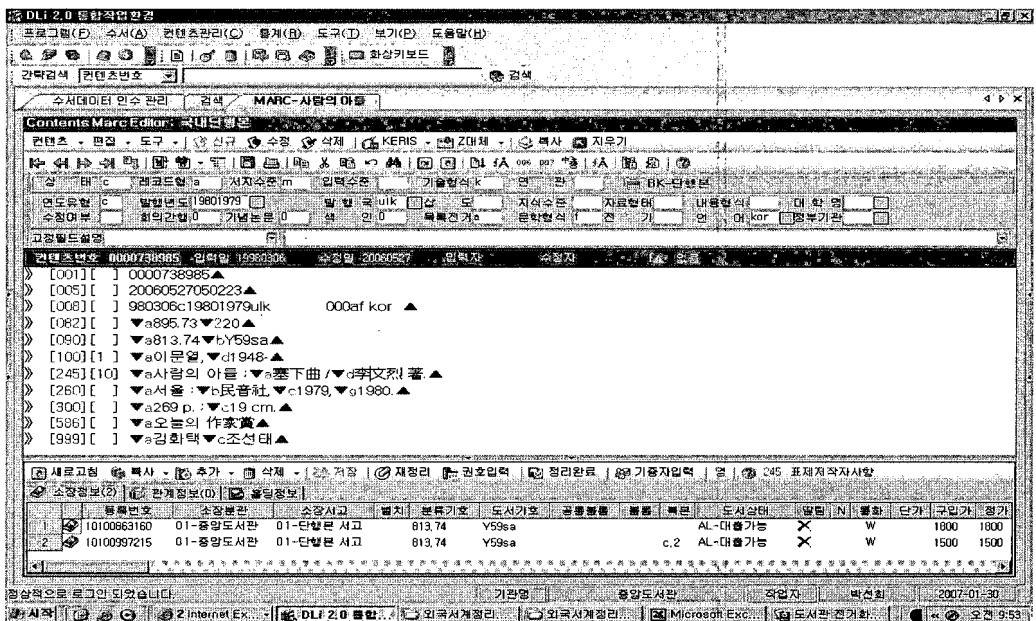


[그림 1] 전거레코드 작성 과정

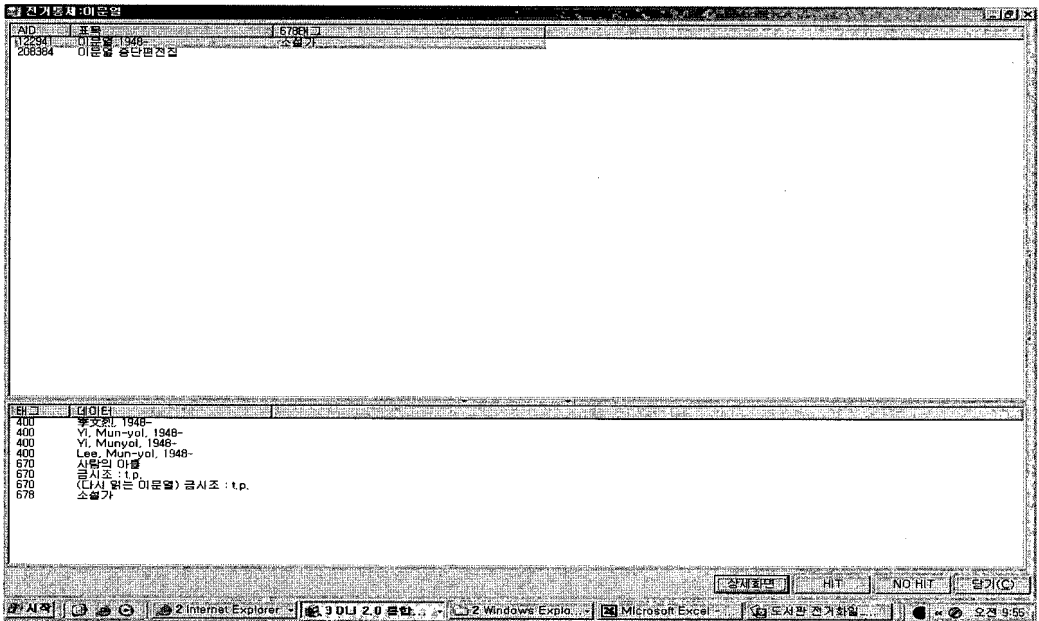
3-4-1. 한글인명

표목(100)은 한글을 채택하였으며 동명이인과의 차별화를 위해서 서지를 작성하는 대상에서 찾을 수 있다면 생년월일을 \$d를 앞세워 기술한다.

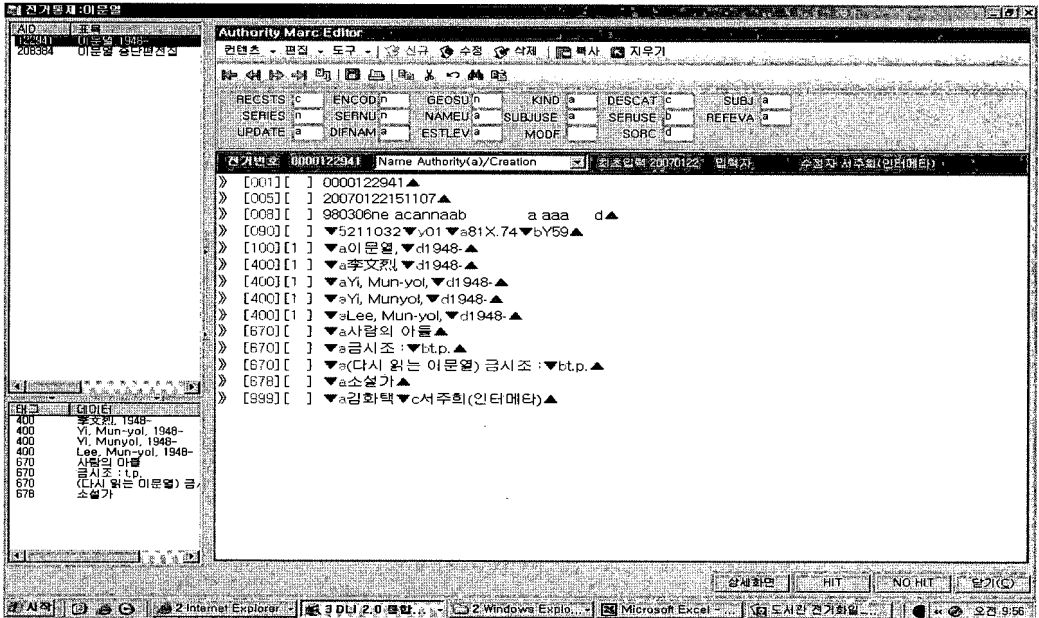
또한 보라참조(400)을 통하여 실물에서 찾을 수 있는 다른 저자명(李文烈, Yi Mun-yol, Lee Mun-yol)을 기술한다. 670 tag에는 1XX 표목에 대한 인용정보, 정보 위치를 기술(사람의 아들, 금시조)하고, 678 Tag에는 1XX 표목에 대한 전기적, 역사적 정보, 이력서에서 발췌한 정보 등을 기술(소설가)한다.



[화면 1] 전거레코드를 작성하려고 하는 100 tag에 F2 key로 조회하여 “이문열”이라는 표목으로 기구축된 전거레코드를 조회한다.



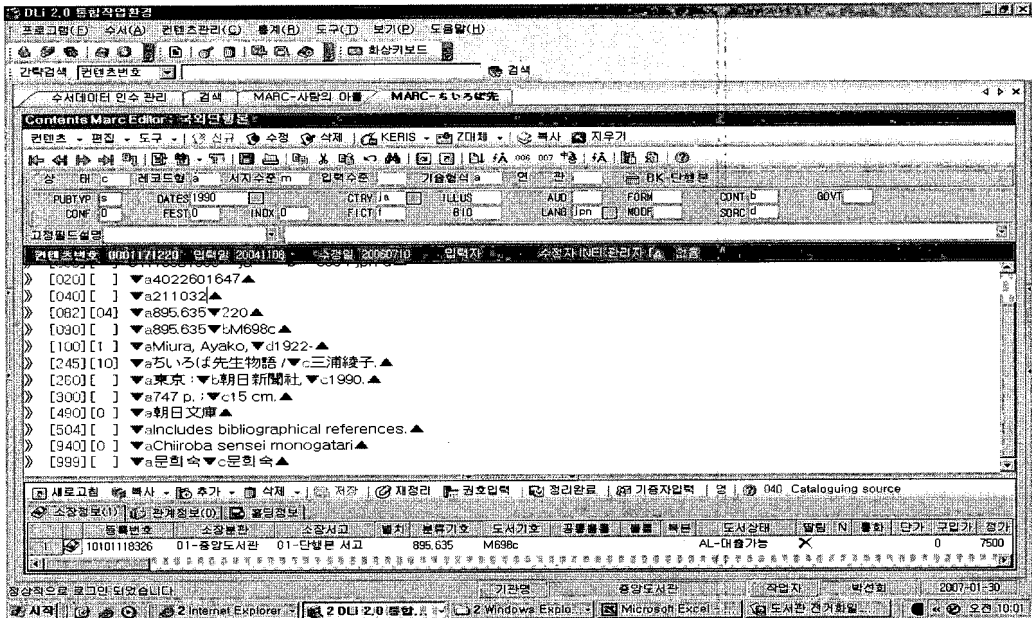
[화면 2] 선거레코드 조회결과 표목부분에 “이문열”이 포함된 선거레코드 간략결과



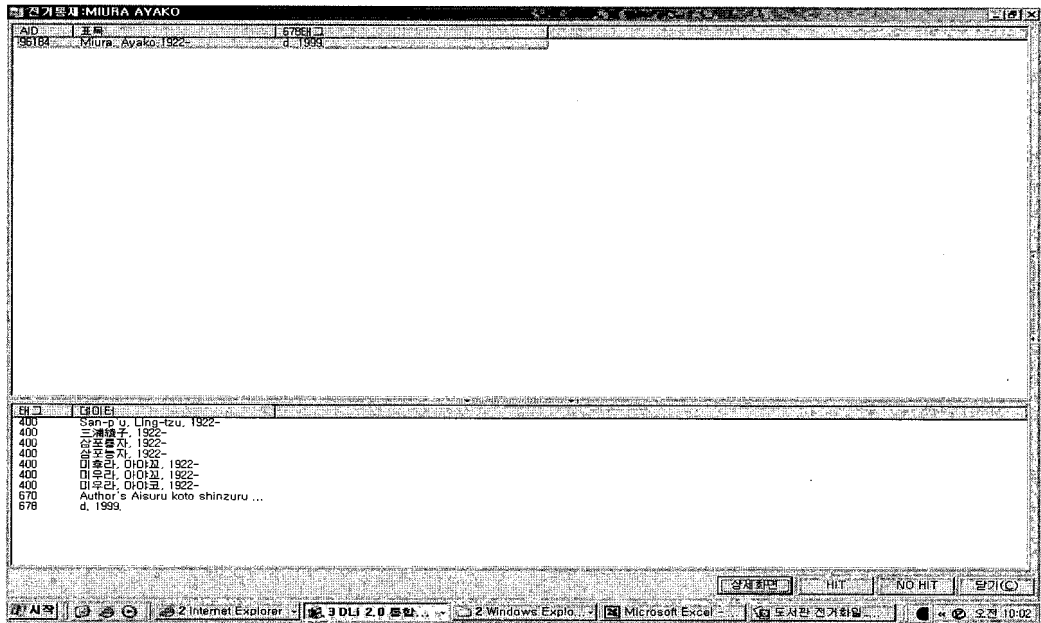
[화면 3] 선거레코드 표목 “이문열, 1948-”의 상세화면의 내용을 살펴본 후 목록자가 기구축된 선거레코드를 따를지 아니면 다른 선거레코드를 구축할지의 여부를 결정한다.

3-4-2. 일본인명

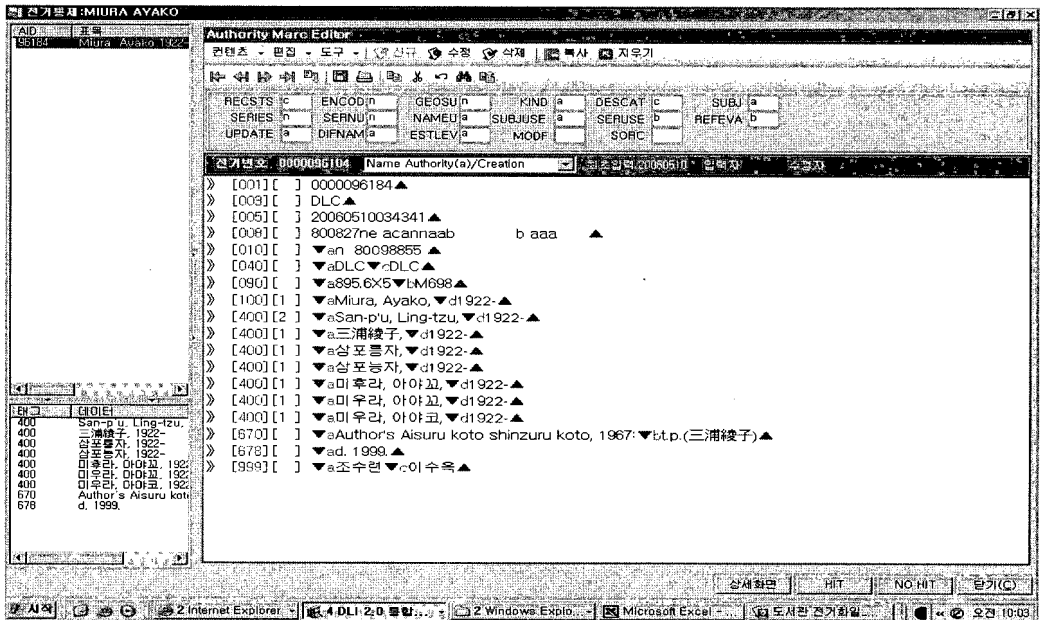
표목(100)은 “수정 Hepburn식 로마자 표기(Miura, Ayako)를 채택하였으며 동명이인과의 차별화를 위해서, 서지를 작성하는 대상에서 찾을 수 있다면 생년월일을 \$d를 앞세워 기술한다. 또한 보라참조(400)을 통하여 실물에서 찾을 수 있는 다른 저자명(San-p'u Ling-tzu, 三浦綾子, 삼포능자, 미우라 아야꼬, 미우라 아야코) 등을 기술한다. 신규로 일본인명 전거를 생성할 때 서지레코드 245 tag의 \$c의 저자의 한자명 정보를 전거레코드 400 tag로 끌고 온다. 670 tag에는 1XX 표목에 대한 인용 정보, 정보위치를 기술(Author's Aisuru koto shinzuru koto, 1967)하고, 678 Tag에는 1XX 표목에 대한 전기적, 역사적 정보, 이력서에서 발췌한 정보 등을 기술(d. 1999)한다.



[화면 4] 전거레코드를 작성하려고 하는 100 tag에 저자 “三浦綾子”를 LC, OCLC, NII 등을 검색하여 일본 가나문자를 수정Hepburn 식으로 로마나이즈 한 후 F2로 조회하여 “Miura, Ayako”라는 표목으로 기구축된 전거레코드를 조회한다.



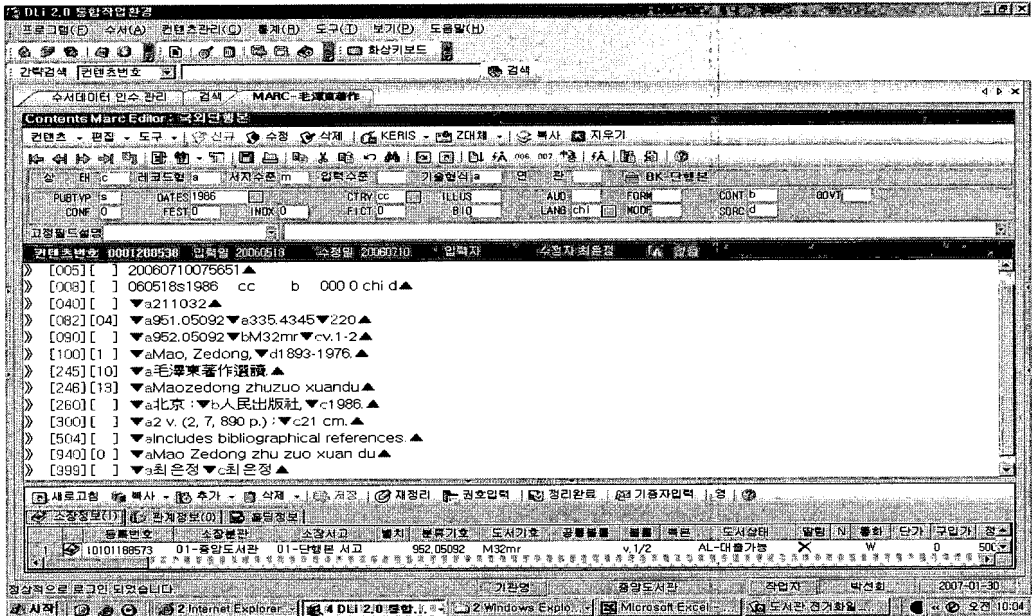
[화면 5] 전거레코드 조회결과 표목부분에 “Miura, Ayako”가 포함된 전거레코드 간략결과



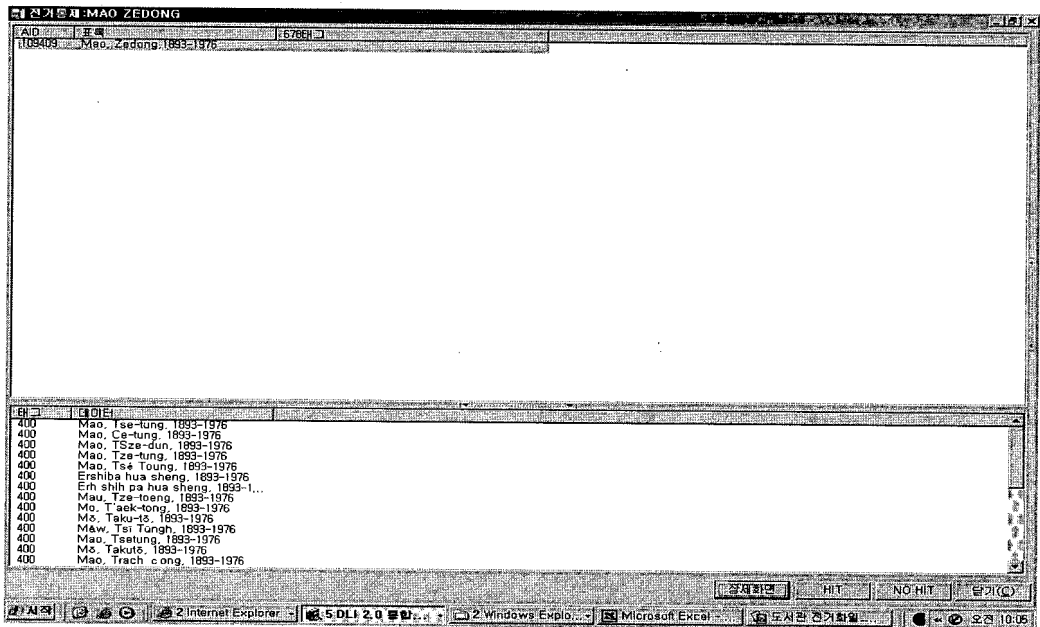
[화면 6] 전거레코드 표목 “Miura, Ayako, 1922-”의 상세화면의 내용을 살펴본 후 목록자가 기구축된 전거레코드를 따를지 아니면 다른 전거레코드를 구축할지의 여부를 결정한다.

3-4-3. 중국인명

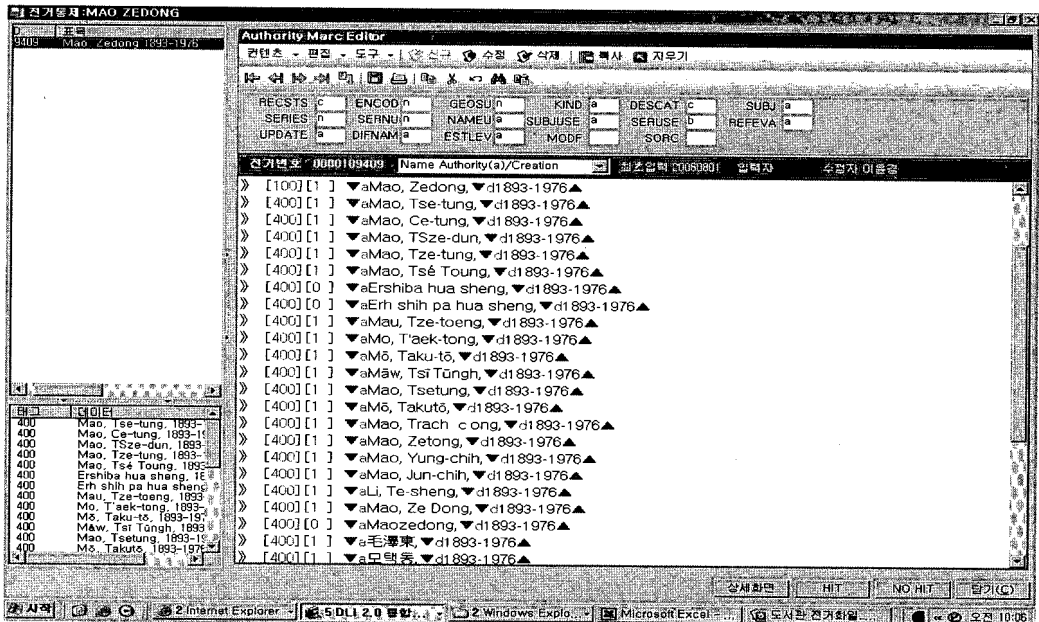
표목(100)은 Pinyin식 로마자 표기(Mao, Zedong)를 채택하였으며 동명이인과의 차별화를 위해서, 서지를 작성하는 대상에서 찾을 수 있다면 생년월일을 \$d를 앞세워 기술한다. 또한 보라참조(400)을 통하여 실물에서 찾을 수 있는 다른 저자명(Mao Tse-tung, 毛澤東, 모택동, 마오쩌둥)을 기술한다. 신규로 중국인명 전거를 생성할 때 서지레코드에 있는 Pinyin식 영문표기를 표목으로 하여 한자명과 한글명, Pinyin 한글음 표기, Wade-Giles 한글음 표기를 400 tag에 자동으로 생성한다. 670 tag에는 1XX 표목에 대한 인용정보, 정보위치를 기술하고(Mao Tse-tung ch'ing ch'un sui yu@eh, 1992), 678 Tag에는 1XX 표목에 대한 전기적, 역사적 정보, 이력서에서 발췌한 정보 등을 기술한다.



[화면 7] 전거레코드를 작성하려고 하는 100 tag에 저자 “毛澤東”를 Pinyin식으로 로마나이즈 한 후 F2 key로 조회하여 “Mao, Zedong”이라는 표목으로 기구축된 전거레코드를 조회한다.



[화면 8] 전거레코드 조회결과 표목부분에 “Mao, Zedong”이 포함된 전거레코드 간략결과



[화면 9] 전거레코드 표목 “Mao, Zedong, 1893-1976”의 상세화면의 내용을 살펴본 후 목록자가 기구축된 전거레코드를 따를지 아니면 다른 전거레코드를 구축할지의 여부를 결정한다.

4. 도서관의 전거시스템 구축 현황

	서울대 도서관	연세대 도서관	이화여대 도서관	국립중앙 도서관
Format	MARC21	KORMARC	MARC21	
전거통제의 대상	개인명, 단체명, 회의명, 인명주제명, 단체주제명 통일서명, 총서명	개인명, 단체명,회의명 인명주제명, 단체주제명, 지명주제명	개인명, 단체명, 회의명, 인명주제명, 통일서명, 총서명,	
전거레코드 구축 방식	LC, OCLC, NII 참조 및 다운로드	LC, OCLC, NII 참조	LC, OCLC, NII 참조	
동양서(일서, 중국서)의 대표표목	*영문표목채택 (*일서: 수정 Hepburn식 *중국서: Pinyin식)	한자의 한글음과 \$h한자표기	한자의 한글음과 \$q한자표기\$d생물년	한자의 로마자 표기 한글음과 \$h한자표기(국립국 어원 외래어 표기법)
동양서 로마자의 변형표기 방식		*일서: 수정 Hepburn식 *중국서: Pinyin식	실물에서 찾을 수 있는 로마자만 기입	
서지레코드와의 연결	서지레코드와 별도로 유지. 표목을 통해서만 연결됨.	서지레코드의 전거대상요소들과 해당 전거레코드를 전거레코드에 대한 그룹기호로 연결됨. (100%)	서지레코드의 전거대상요소가 전거레코드의 고유전거제어번호와 연결됨	
고정장필드 32번째 자리(동명이인)	'a'	*국내서, 중국서, 일서 : 'b' *외국서 : 'a'	'n'	

[표 2] 4개 도서관의 전거레코드

	서울대	연세대	이화여대	국립중앙
한글 인명	이문열,\$d1948-	이문열\$h李文烈	이문열\$q李文烈,\$d1948-	
	李文烈,\$d1948-	Yi, Mun-yol,\$d1948-	Yi, Munyol	
	Yi, Mun-yol,\$d1948-		Lee, Mun yol	
	Lee, Mun-yol,\$d 1948-			
일본 인명	Miura, Ayako,\$d1922-	삼포능자\$h三浦綾子	삼포능자\$q三浦綾子,\$d1922-1999	미우라 아야코,\$d1922-1999
	三浦綾子,\$d1922-	Miura, Ayako	Miura, Ayako	Miura, Ayako, \$d1922-1999
	삼포능자,\$d1922-	삼포릉자		삼포능자=\$h三浦綾子, \$d1922-1999
	삼포릉자,\$d1922-			
	미우라, 아야꼬,\$d1922-		미우라 아야코	삼포릉자=\$h三浦綾子, \$d1922-1999
	미우라, 아야코,\$d1922-	미우라, 아야꼬	미우라 아야꼬	ミウラ アヤコ, \$d1922-1999
	미후라, 아야꼬,\$d1922-	미우라, 아야코		미우라 아야꼬,\$d1922-1999
	San-p'u, Ling-tzu,\$d1922-	ミウラ,アヤコ		
중국 인명	Mao, Zedong,\$d1893-1976	모택동\$h毛澤東	모택동\$q毛澤東,\$d1893-1976	마오쩌둥, \$d1893-1976
	Mao, Tse-tung,\$d1893-1976	Mao, Ze-dong	Mao, Zedong	모택동=\$h毛澤東, \$d1893-1976
	毛澤東,\$d1893-1976	Mao, Tse-tung,\$d1893-1976	Mao, Tse-tung	윤지=\$h潤之, \$d1893-1976
	모택동,\$d1893-1976	마오쩌둥	마오저둥	룬즈, \$d1893-1976
	마오쩌둥,\$d1893-1976	마오쩌둥	마오쩌둥	Mao, Zedong, \$d1893-1976
	마오쩌둥,\$d1893-1976			Mao, Tse-tung, \$d1893-1976

[표 3] 4개 도서관의 개인명 전거레코드

5. 맺는 말

비용효과적인 측면에서나 품질 관리의 측면에서 볼 때 개개의 도서관에서 전거 데이터베이스를 구축한다는 것은 어려움도 따르고 효율적이지도 않다고 본다. 또한 서두에서도 언급한 바 있지만 목록의 궁극적인 목적은 검색의 효율을 높이는 것이다. 따라서 각 기관의 협력을 통하여 앞으로는 통합 전거데이터베이스를 구축하는 방향으로 나아가야 할 것이며, 이러한 통합을 하기 위해 여러 방면의 규칙들을 정비해 나가는 일이 선행되어야 하리라 본다.

5-1. 전거 데이터베이스의 품질관리

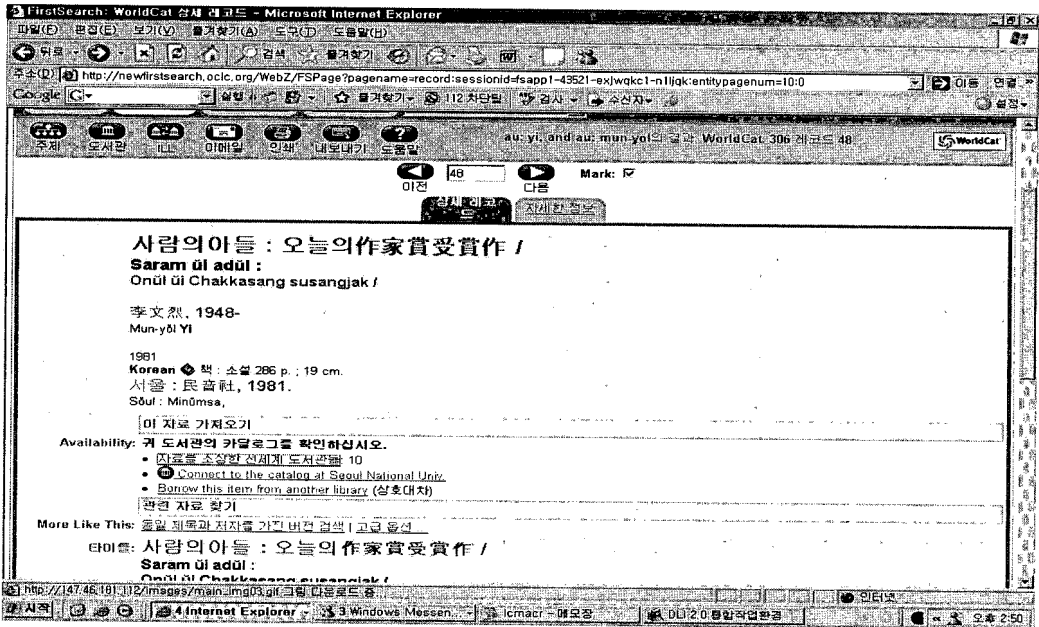
어느 한 기관, 혹은 2~3 기관이 중심이 되어 전거 데이터베이스의 품질을 관리하고 나머지 기관들은 이에 협력하는 방안, 또는 각 기관에 권한 가진 전문가만이 전거 데이터베이스의 품질을 관리하고 이 사람들을 제외한 나머지 사람들이 협력하는 방안 등등이 있을 수 있겠다.

5-2. 전거 데이터베이스 구축의 범위

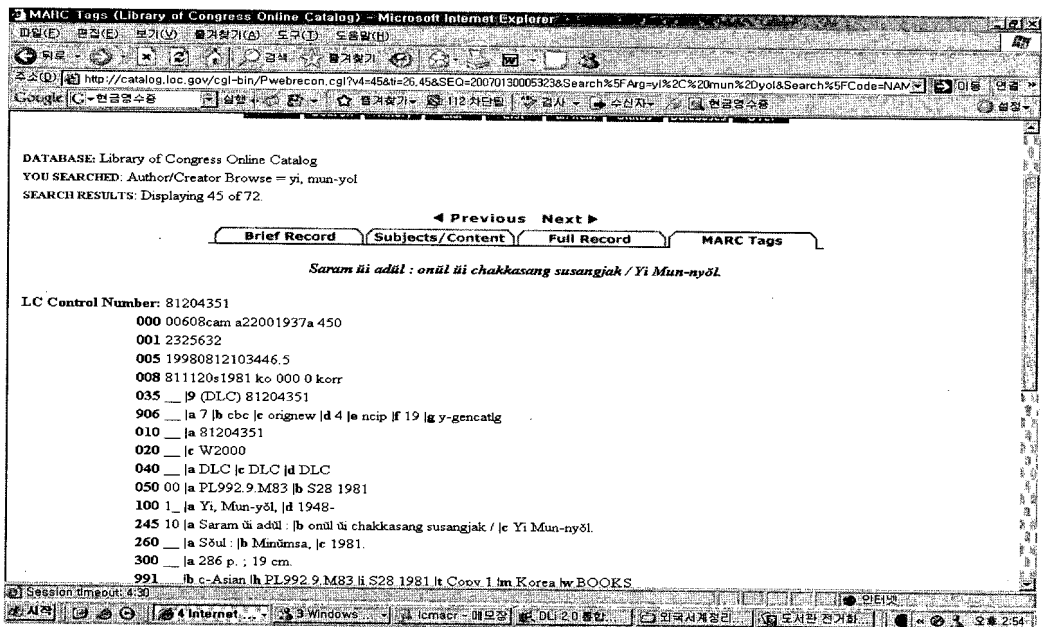
국내 도서관에서는 아직 전거 데이터 베이스를 구축하지 않은 곳도 많다. 이를 전거를 우선적으로 시작하고 차츰 주제명 전거를 하는 방향으로 나아가야 할 것이다.

5-3. 표목의 선정과 형식의 통일

5-3-1. 위의 3개 대학 도서관의 전거레코드의 비교표에서도 알 수 있듯이 3개 대학 도서관의 전거레코드가 각기 다른 것을 볼 수 있다. 따라서 각각의 기관이 서로 전거 데이터베이스를 같이 구축하고 이용하기 위해서는 표목의 규칙을 정비하는 일이 먼저 이루어져서 표목 선정과 형식에 일관성을 유지해야하고, 이형에 어떤 것들을 표기할지는 논의가 되어야 하겠지만, 일본인명과 중국인명에는 수정Hepburn식과 Pinyin 식 로마나이즈가 기입되는데 한국인명에 대해서는 통일된 로마나이즈 규칙의 로마자가 기입되는 것이 아니라 실물 자료에 쓰여진 대로만 로마자를 참조로 기입하는 것은 차후 여타의 서지세계에 융화될 때 장애물이 될 수 있다고 생각한다. 그러므로 한국어 표목에 대한 보라 참조에 McCune-Reischauer (MR) 로마나이즈를 추가하는 것을 고려해 보았으면 한다. (앞서 예에서 본 '이문열'은 보라 참조에 'Yi Mun-yol', 'Lee Mun-yol'을 로마나이즈 하였는데 '이문열'의 McCune-Reischauer (MR) 로마나이즈는 'Yi Mun-yöl'이다).



[화면 10] OCLC FirstSearch WorldCat의 저자‘이문열’의 서지레코드



[화면 11] LC의 저자‘이문열’의 서지레코드

5-3-2. 전거레코드 008 tag 32번째 자리(동명이인) 'a', 'b' 를 채택하는 문제로 a(채택표목레코드나 참조레코드의 필드 100에서 개인명이 한사람으로 사용됨)를 사용하는 경우 동명이인을 식별하기 위해 생몰년을 사용하고 있는데, 생몰년은 식별력이 뛰어나다는 장점을 가지고 있는 반면 정리 업무를 하는 사람의 입장에서 볼 때는 저자의 생몰년의 정확한 파악이 어렵고, 생몰년이 같은 표목의 중복처리 문제가 발생되고 이용자의 측면에서는 본인이 찾고자 하는 자료의 저자의 생몰년을 파악하고 있는 사람이 거의 없을 것이므로 검색의 변별력으로써 저자의 생몰년은 그리 유용하지 않다고 본다.

b(채택표목레코드나 참조레코드의 필드 100에서 개인명이 두사람 이상으로 사용됨)를 사용하는 경우 정리 업무를 하는 사람의 입장에서 볼 때는 저자의 생몰년을 찾는 수고를 할 필요가 없어 전거 데이터베이스를 구축하는 시간을 단축할 수 있는 반면 각각의 표목을 식별하기 위한 Unique한 레코드가 될 수 없다고 본다.

이러한 동명이인의 전거레코드 처리여부도 차후 논의가 되어야 하리라 본다. 한국 인명의 경우 식별요소로 사용되는 생몰년이나 한자를 파악할 수 있는 레코드에 대해서는 'a'를, 생몰년을 파악할 수 없는 레코드에 대해서는 'b'를 사용하는, 두 가지 방법을 모두 채택하는 방식도 제안해 볼 수 있겠다.

또한 이러한 동명이인의 처리방법은 'a',나 'b'를 채택하는 문제 뿐만 아니라 생몰년이나 한자 외에 동명이인을 식별할 수 있는 다른 요소를 찾아보는 것도 함께 논의 되어야 하겠다.

【참고문헌】

- 연세대학교 중앙도서관. 연세대학교 중앙도서관 선거시스템 및 선거DB의 특성. 2002
- 국립중앙도서관. 한국문헌자동화목록형식 : 선거통제용. 국립중앙도서관, 1999
- 윤정옥. 한국과학기술정보연구원의 회의명, 단체명 선거파일 구축 연구, [정보관리연구] 33권 3호(2002). p.69-85
- 기민도와외 4명. 국가선거데이터베이스 구축을 위한 기초연구, [도서관연구회 논문집] (2001). p.45-64
- 미국의회도서관 편 ; 오동근 역. MARC 21 선거 레코드의 이해, 2006
- 이용효 ; 오동근 ... [등]. 연구학술정보 선거DB시스템의 개발과 구축에 관한 연구, 2000
- 선우중호...[등저]. 도서관 전산화 S /W 개발에 따른 선거제어시스템 연구개발보고서 : 선거포맷 및 인명 선거리스트 편 : 제1차년도, 1992