온라인공동편목을 목표로 한 종합목록데이타베이스 구축과 발전방향

A Construction of Union Bibliographic Database for Online Shared Cataloging and its Perspectives

김창근*

□목 차□

Ⅰ. 서론
Ⅱ. 과학단지장서종합목록데이타베이스
Ⅲ. 과학기술도서종합목록데이타베이스
Ⅳ. 데이타베이스 구축 사업결과 분석
Ⅴ. 발전방향
Ⅵ. 결론 및 제언

초 록

정보화사회의 도래, 기관 고유기능의 전문화, 다양화와 함께 과학기술분야 관련기관들은 독립적으로 전문도서를 비롯한 각종 정보자료를 수집하며 관리·운영하고 있다. 그러나 수집되는 정보량이 연구개발사업 지원에 크게 부족한 것은 차치하더라도 그나마 수집된 정보의 공동활용상 난제는 상존하고 있다. 이에 과학기술정보유통을 위한 첫단계 사업으로 제시된 과학기술종합도서관 사업을 시발로 연구단지내 각 기관에 분산·소장된 도서정보자료의 공동활용을 위한 표준화·집중화 사업이 시행되었으며, 본 사업의 성공적 사례를 바탕으로 수행된 연구단지정보관리협의회 소속기관에서 소장하고 있는 서양서 단행본 데이타베이스 구축 1단계 사업도 완료된 바 지금까지 축적된 종합목록데이타베이스 구축기술을 소개하고 온라인공동편목을 포함한 공동활용체제를 위한 발전방향을 언급함으로써 향후 우리나라에서 구축될 각종 종합목록데이타베이스를 위한 한가지 지침으로 제시하였다.

ABSTRACT

It is described about the construction of Union Bibliographic Database for online shared cataloging and its perspectives in the Dae-duck Science Park technical information centers and libraries.

* 한국과학기술원 과학도서관 / 과학기술도서종합목록데이타베이스구축 연구책임자

Ⅰ. 서 론

1986년 과학기술처는 순수과학기술정보 제공기관의 필요성에 대한 타당성 조사 방안의 하나로 한국과학기술원을 연구수행기관으로 하여 "과학기술종합도서관 설립 타당성 및 추진에 관한 연구[1]"에 착수한 바 있었다. 그 연구결과에 의하면 국가 경제적 차원에서 보았을 때 독립된 기구의 성격을 갖춘 정보센터를 설립하고 그 기관에 종합적인 정보제공 기능을 부여하는 것보다는 기존 도서관 중에서 상당한 수준에 도달한 도서관을 지정하여 이를 바탕으로 육성 발전시키는 것이 경제원칙에 입각해서 즉, 최소한의 투자로 최대의 효과를 거두는 방안을 채택하는 것이 가장 타당성이 있다는 결론에 이르게 되었다.

이와같은 타당성 연구가 수행되게 된 것은 1960년대 말부터 꾸준히 성장해 온 한국과학기술정보센터(KORSTIC)가 정부의 무원칙적 이합집산 정책에 의하여 1982년에 강행된 연구소 통폐합 과정에서 전혀 성격이 다른 국제경제연구원(KIEI)과 통합되어 산업연구원(KIET)으로 재출발하게 됨으로써 당시까지의 순수과학기술정보 공급기관에서 산업기술정보 제공기관으로 탈바꿈하게 되었기 때문에 순수과학기술정보를 심층적이고 광범위하게 제공할 수 있는 기관의 필요성이 다시 대두되었기 때문이다.

1983년 12월 과학기술처는 과학기술처 소관 국가정보자료관리세부시행규칙(과학기술처 훈령 제205호 : 1983. 12. 29)[2]을 공표하였으며, 이어 1986년에 수행된 연구결과에 따라 국가과학기술종합도서관의 설립과 그 추진을 강력히 지원하기 시작했고 동 사업의 질적수준 향상을 위하여 별도 채널을 통하여 긴밀하게 추진되어 왔던 UNDP(유엔개발계획기구)와의 협력을 본 사업과 연계하여 과학기술정보서비스 진흥방안에 대한 국가간 협정[3]을 체결함으로써 정보서비스 기능 발전을 위한 자금 확보에도 성공을 거두었다.

본 사업은 UNDP, UNESCO, 과학기술처 3자간의 상호협력체제하에 수행하되 자금지원은 UNDP가, 국제집행은 UNESCO가, 사업수행은 과학기술처가 하도록 되어있기 때문에 이에따라 과학기술처는 한국과학기술원으로 하여금 세부사업을 수행토록 지정하였다. 이에 한국과학기술원 과학도서관은 UNDP 자금을 활용하여 UNESCO에서 추천받은 정보관

---

1) 박계홍 외, 과학기술종합도서관 설립 타당성 및 추진에 관한 연구, 과학기술처, 1986

2) 과학기술처 훈령 제205호 (제정 1983.12.29) 제3조(각기관별임무) 3항, 훈령 제300호 (개정 90.5.21), 과학기술처소관 국가정보자료관리 세부시행 규칙.

3) UNDP ROK/87/010 사업은 당초 "Promotion of Scientific and Technological Information Services and Establishment of Popular Science Resource Center"라는 표제로 1988년 7월을 기점으로 하고 4년간에 걸쳐 정부지원 47억원, UNDP 지원 40만불의 계획예산으로 본 사업이 시작되었으나 정부지원금 부족과 사업수행기관의 내부 기구 분할 및 통합 등 여건 변화에 따라 1990년 9월 사업관련기관 3자간 회의 (Tripartite Review Meeting:MOST, UNDP, UNESCO) 결과에 따라서 표제를 "Promotion of Scientific and Technological Information Services"로 변경하여 종합도서관사업에 치중하도록 조치되었음.

리 전문가를 초빙하고 정보관리 각 분야의 선진기술에 대한 그들의 지식을 전수 받았으며 국제적으로 공인된 정보서비스 제공기관에 대한 구체적이고도 심도있는 조사분석을 통하여 2개 기관[4]을 지정하고 Bibliographic Services, Union Catalog, Library Planning, Library Automation, Library Networking 등의 내용으로 연구단지정보관리협의회[5] 회원기관 소속 직원 총15명에 대해서 실무자 연수교육을 실시함으로써 정보서비스 요원의 능력계발을 도모한 바 있다.

한편 정부는 과학기술종합도서관업무 수행을 위한 핵심 자원이라고 볼 수 있는 장서확충에 소요되는 막대한 예산 투입에 대한 부담때문에, 또한 최고 책임자의 잦은 교체에 따른 정책전환과 일관성 없는 시각에 따라 원문정보(Resources) 보다는 정보유통망에 대한 투자비중과 효용성을 더 높게 평가했던 것이다.

막대한 자료 구입비를 투입할 수 없게 된 이와같은 상황에서도 한국과학기술원 과학도서관은 UNDP와의 협정 도서관분야 6개항[6] 이행의 일환으로 대덕연구단지내 모든 연구기관 소속의 기술정보실 또는 자료실 (이하 기술정보실[7]에서 소장하고 있는 정보자료를 하나로 묶음으로써 분산되어 있는 정보자료의 소재정보를 파악할 수 있도록 하여 한정된 지역내에서나마 정보자료를 공동으로 활용할 수 있는 방안을 모색하지 않을 수 없게 되었던 바 우선적으로 과학기술도서중 서양서 단행본을 대상으로 종합목록데이타베이스 구축사업[8]을 수행하게 되었으며 본 사업을 수행하는 과정에서 축적된 기술을 바탕으로 과학기술정보유통망 사업[9]을 주도하게 된 시스템공학연구소(SERI)로 부터 자금지원을 받아서 대덕연구단

4)-OCLC (Online Computer Library Center):미국 오하이오주 더블린 소재의 세계 최대의 도서관 협력망 운영기관.
-UTLAS (University of Toronto Library Automation System):카나다 토론토대학 도서관자동화시스템을 바탕으로 발전된 세계 4대 도서관 협력망 운영기관 중 하나.

5)1962년 창설된 서울연구개발단지 도서관실무자 협의회가 대다수 연구기관의 대덕연구단지 이전을 계기로 1986년 그 명칭을 개정한 것으로서 과학기술분야 정보관리부서간의 상호협력과 정보교환을 주 목적으로 운영되는 협의체로서 1992년 4월 현재 32개 기관이 가입되어 있음.

6)(1)단행본 300,000권과 정기간행물 5,000종을 확보하고 온라인으로 정보서비스를 제공할 수 있는 과학도서관의 설립.
(2)연구단지내에서 가용한 모든 정보가 수록된 종합목록 제작.
(3)선정된 분야에 대한 서지데이타베이스 구축.
(4)선진형 정보서비스 제공 체제 구축.
(5)관련된 국내 데이타베이스를 탐색할 수 있는 통신망의 구축.
(6)온라인 도서관 자동화시스템의 구축과 그에 의한 정보서비스와 대덕연구단지내의 상호대차체제 구축, 그리고 1991년에는 전국 규모로 확대

7)연구단지정보관리협의회 소속 32개 회원기관 중 15개의 기관에서 정보관리부서명칭을 기술정보실 또는 기술정보부, 기술정보센터로 사용하고 있음.

8)과학단지장서종합목록 데이타베이스 구축사업 :1. 사업기간—1989년 11월~1991년 6월(1년 8개월). 2. 소요예산—1억원. 3. 예산지원—경제기획원.

9)정부관련부처 최고책임자의 지시로 1990년 초에 긴급 구상된 후 소관부처인 과학기술처내에 구성된 과학기술정보유통추진위원회의 의결 등 계획부터 확정에 이르기까지 1년 미만의 기간에 만들어진 국가과학기술정보정책 사업으로서 본

지 뿐만 아니라 연구단지정보관리협의회에 소속된 모든 기관으로 그 대상을 확대함으로써 향후 전국에 산재되어 있는 과학기술관련 도서의 총체적 종합목록데이타베이스를 구축할 수 있는 기초를 확립하게 된 것이다.

Ⅱ. 과학단지장서종합목록 데이타베이스

1. 사업의 개요

한국과학기술원에서는 과학기술발전을 지원할 과학기술정보서비스의 충실화, 선진첨단과학기술의 토착화를 위하여 과학기술정보유통망을 형성한 후 장차 구축될 교육연구전산망을 활용하여 전국 네트워크에 편승함으로써 국내외 과학기술정보자료를 경제적이고도 효과적으로 활용하고 향후 국가과학기술정보자료에 대한 총체적 데이타베이스의 구축과 구축된 데이타베이스는 상호 활용하도록 한다는 것을 대전제로 한 과학기술종합도서관사업의 일환으로 과학단지내 각 기관에서 수집·관리되고 있는 과학기술도서 중 서양서 만을 우선적으로 취합하여 하나의 표준화된 데이타베이스로 구축하기위한 사업을 추진한 것이다.

본 사업의 수행을 위하여 경제기획원으로부터 1억원의 특별예산 지원을 받았으며 1989년 11월부터 1991년 6월까지 1년 8개월간을 사업기간으로 설정하였으며 1989년 11월 "도서관자동화시스템에 관한 국제세미나/워크샵"을 시발점으로 하여 한국과학기술원이 주관기관으로, 연구단지정보관리협의회 회원기관인 국방과학연구소(ADD), 산업기술정보원(KINITI: 구 KIET), 통신개발연구원(KISDI), 한국과학기술연구원(KIST), 한국국방연구원(KIDA), 한국동력자원연구소(KIER), 한국원자력연구소(KAERI), 한국인삼연초연구소(KGTRI), 한국전기연구소(KERI), 한국전력공사기술연구원(KEPCO), 한국전자통신연구소(ETRI), 한국통신연구개발단(KTRC), 한국표준연구소(KSRI), 한국해사기술연구소(KRISO), 한국해양연구소(KORDI), 한국화학연구소(KRICT) 등 16개 기관을 데이타 협력기관으로 결정하면서 우리나라 최초의 Union Database 사업은 시작되었다.

사업주관기관인 KAIST는 각 데이타 협력기관 중에서 본 사업을 위한 목록데이타의 표준화 도구로서 지정된 CD-ROM 데이타베이스인 CAT CD450 또는 BiblioFile을 보유한 7개 기관(KAIST, KAERI, KIER, KIST, KRISO, KSRI, KTRC)을 과제수행기관으로 지정하였으며 7개 기관 과제공동분담처리 방법을 통해 다 함께 참여하는 연구 즉, 과학기술정보자료 공동활용 기반조성을 위한 기초를

사업의 주요 골자는 첫째, 전문분야별 정보 수집처리와 DB의 구축은 해당 출연연구기관이 수행하되 사업의 총괄과 관련 기술의 표준화를 위한 중앙조정기관으로 연구전산망사업추진기관인 시스템공학연구소를 지정하여 중앙정보센터 기능을 부여하고 둘째, 구축된 DB는 각종 통신망과 연결하여 대외서비스를 실시한다는 것으로서 이는 1989년까지 KAIST 도서관을 중심으로한 과학기술종합도서관 설립 사업과는 정면으로 배치되는 것이며 이와같은 정책전환이 계기가 되어 종합도서관 사업지원은 완전히 단절 되었음.

확립하였다.

본 사업을 효과적으로 수행하기 위하여 주관기관인 KAIST는 정보관리협의회 임원을 중심으로 한 사업추진위원회를 구성하고 실무를 담당하는 사서 및 프로그래머로 실무반을 구성하여 표준 데이타 포맷을 위한 표준화 작업을 하였고, 소급변환을 위하여 각 기관별로 소장된 목록데이타를 수집하였으며, 목록데이타의 검증(데이타의 누락, 불일치, 표기 오류 등)과 수정작업, 소급변환이 완료된 로칼데이타 화일간의 중복 점검과 통합작업, 참여 17개 기관 전체 데이타의 상호 중복 점검과 통합작업, Union Database 생성, 책자형 목록발간 등의 순서로 진행되었다.

2. 사업추진 내역

2.1 사업추진위원회와 실무반 구성

종합목록데이타베이스 구축사업을 효율적으로 추진하기 위하여 1)사업추진계획의 입안과 수행 2)사업비 결산과 회계 3)기타 사업수행에 필요한 주요사항 협의 등의 업무를 관장할 수행기구로써 사업추진위원회가 1990년 1월에 구성되었으며 사업추진위원회의 하부조직으로써 데이타베이스 구축 실무를 담당할 실무반은 사업수행기관인 7개 기관을 중심으로 정보관리부서 실무사서와 전산 분야에 경험이 풍부하고 본 사업에 적극적으로 참여할 사람을 추천받아 1990년 5월에 구성되었는데 본 실무반의 주요 협의내용은 1)데이타베이스 구축에 필요한 기술적 검토요망 사항 즉, 데이타 포맷 선정, 로칼데이타 범위, 데이타 수집방법 및 처리방안과 2)데이타베이스 구축을 위한 구체적인 방안 협의 등 사업추진을 위한 제반 안건의 처리이다.

2.2 데이타 수집

각 기관에서 소장하고 있는 정보자료중 본 사업을 위한 수집 대상으로는 단행본 서양서로 국한하였으며 목록데이타 송부시 반드시 수록해야 할 서지사항은 ISBN 또는 LCCN을 포함하는 가능한 가능한 모든 서지사항이 요구되었으며 데이타 송부형식은 각 기관별 환경차를 고려하여 무리한 요구는 피하여 가능한 형식으로 하되 다만 3가지 형식 즉, 1)이미 구축된 목록데이타가 있는 경우 그로부터 출력한 형태 (Printed Form), 2)카드목록 복사형태 (Shelf Card), 3)표준 MARC 형식을 갖추지 않았더라도 이미 구축된 데이타베이스로부터 다운로드 하였거나 PC용으로 입력된 형태 (5¼ inch Floppy Diskette)중 한가지를 선택하여 송부하도록 하였다. 송부요구 기한내 수집된 데이타량은 〈표1〉과 같다.

위의 표에서 볼 수 있는 바와 같이 인쇄물로 데이타를 제출한 기관이 8개 기관, 카드목록으로 제출한 기관이 5개 기관, 인쇄물·카드 혼합기관이 1개 기관, 미제출 1개 기관이며, 플로피디스켓으로 제출한 기관은 한국해양연구소 1개 기관뿐인 점을 보더라도 각 기관의 정보관리부서마다 뚜렷하게 전산화하지 못했다는 것을 반증해 주고 있는 것을 볼 수 있다. 따라서 본 사업이 참여기관 모두의 로칼 데이타베

〈표1〉 소장 목록데이타 수집 현황

기 관 명	데이타 수집 형태	데이타 건수
국방과학연구소	인쇄물(가제본 35권)	25,000건
산업기술정보원	인쇄물	17,168건
통신개발연구원	인쇄물	2,938건
한국과학기술연구원	카드목록	25,000건
한국과학기술원	카드목록, 인쇄물	60,000건
한국국방연구원	인쇄물	8,425건
한국전력공사기술연구원	인쇄물(등록원부복사)	1,777건
한국전자통신연구소	인쇄물(가제본 3권)	8,500건
한국동력자원연구소	카드목록	11,500건
한국원자력연구소	카드목록	20,000건
한국통신연구개발단	인쇄물	5,000건
한국표준연구소	카드목록	8,000건
한국해사기술연구소	카드목록	8,000건
한국해양연구소	플로피 디스켓	10,866건
한국화학연구소	인쇄물	1,023건
계		213,197건

이스 구축에도 기여하는 바 클 것으로 예상되었다.

2.3 데이타 소급변환 분담처리

사업추진실무반은 수집된 데이타의 소급변환을 위한 데이타 포맷의 표준화용 목록데이타 변환 통일안 작성을 협의하면서 선정된 표준안에 따른 소급변환의 분담처리를 논의한 바 KAIST가 60,000건, KTRC가 KIDA, KINITI, KISDI의 데이타를 포함한 30,000건을 처리하고 기타 5개 과제수행기관은 나머지 총 66,000건을 과제수행기관별로 공히 분배하여 처리하기로 하며 타기관과의 수탁계약이 가능한 5개 기관(KAERI, KIER, KIST, KRISO, KSRI)에게는 데이타 변환작업 관련 비용조로 일정액을 지원하도록 함이 좋다는 제안을 하였으며 이에따라 과제수행기관선정, 과제분담 및 예산지원 방안 등이 사업추진위원회에서 확정되었으며 과제공동분담처리를 위한 용역계약을 체결함으로써 종합목록대이타베이스 구축을 위한 데이타 소급변환작업이 구체적으로 시작되었다.

목록데이타 변환작업 기준에 따라 각 기관에서 처리하는 도중 자관의 소장목록데이타 중 할당량 변환작업을 완료한 기관은 분담처

〈표2〉 데이타 분담처리 기관별 데이타 분배내역

데이터 분담 기관명	처리 대상 데이타	데이타 송부 형태	데이타 건
한국표준연구소	국방과학연구소분	인쇄물 (가제본 35권)	25,000건
한국동력자원연구소	한국전자통신연구소분	인쇄물 (가제본 3권)	7,934건
한국해사기술연구소	한국해양연구소분	플로피 디스켓 (2HD 2장)	10,866건
한국통신연구개발단	산업기술정보원분	인쇄물	17,168건
	통신개발연구원분	인쇄물	2,938건
	한국국방연구원분	인쇄물	8,420건
한국과학기술원	한국화학연구소분	인쇄물	1,023건
	한국전력공사기술연구원분	인쇄물(도서원부복사)	1,177건
계			74,531건

리기관 간의 미처리 데이타 분담처리보다는 자관의 데이타를 자관에서 처리하는 것이 효율적이라고 판단되었기 때문에 계속 진행하기로하고 협력기관에서 입수된 소장목록데이타는 할당량에 미달되는 각 분담처리기관에 분배 〈표2〉하여 처리하였다.

2.4 소급변환 완료 데이타

과제분담기관은 1990년 12월과 1991년 2월에 각각 42,844건, 40,448건, 도합 83,292건의 변환된 데이타를 MARC 통신형태로 송부하였으며 주관기관인 KAIST에서 변환한 54,116건을 합하여 총 137,408건의 데이타가 소급완료됨으로써 로칼데이타의 통합을 위한 기초자료 준비가 끝나게 되었다.

3. 데이타베이스 구축 기술

3.1 시스템 환경

UNDP 자금[10]은 1)전문가 초청 2)관련자 해외훈련 3)도서관자동화 소요장비 구입 등으로 일정액을 할당하여 년차별로 제공되기 때문에 세부사업수행기관인 KAIST는 우선 CD-ROM 서지데이타베이스를 운용할 수 있는 워크스테이션을 비롯하여 1989년 11월에 개최된 바 있는 도서관자동화 국제세미나/워크샵에서 제시되고 사업추진실무반에서 사용결정한 바 있는 표준 MARC 데이타베이스인 CAT CD450를 우선 구매하였으며 경제기획원으로부터 할당받은 사업비로는 여타 소요장비

10)UNDP 자금은 UNDP에서 매년 수정집행이 가능하며 장비구입의 경우 세부사업수행기관에서 정부측 수행기관인 과학기술처로 소요장비내역을 보내면 과학기술처는 본 건을 심사한 후 UNDP로 발송하게 되며 UNDP는 국제집행기구인 UNESCO에 구매의뢰하게 된다. 구매요구를 접수한 UNESCO는 이를 국제입찰에 부쳐 구입한 후 UNDP 해당 국가 사무소로 이첩하며 도착 통보를 받은 세부사업수행기관에서 인수 및 검수를 하게 된다.

〈표3〉 하드웨어 일람표

기 종	수 량	도입일자	용 도 및 기 능	비 고
◎ M310 Workstation -CPU i80386/SX (20MHz) -HDD (42MB FORMATTED) -FDD (5.25″ 1.2MB) -HITACHI CDR2500 1pack 4 drive -14″ EGA COLOR MONITOR	1set	'89. 11.	UNION DB 구축용 (CD-ROM 서지데이타베이스 CAT CD450 운영)	구 입
◎ 386 ULTRA CD SERVER -CPU i80386/33MHz -HDD(100MB FORMATTED) -FDD'S (5.25″:1.2MB or 3.5″:1.44MB) -HITACHI CDR3600 CD DRIVE 4대 내장 -14″ PAPER-WHITE MONITOR	1set	'90. 8.	UNION DB 구축용 (CD-ROM 서지데이타베이스인 CAT CD450 운영)	구 입
◎ 386 OPTIMA CD PLATFORM -CPU i80386/SX (20MHz) -HDD (40MB FORMATTED) -FDD (5.25″ 1.2MB) -HITACHI CDR3600 CD DRIVE 1대 내장 -14″ VGACOLOR MONITOR	2set	'90. 8.	UNION DB 구축용 (CD-ROM 서지데이타베이스인 CAT CD450 운영)	구 입
◎ 286 OPTIMA CD TERMINAL -CPU i80286/SX(20MHz) -HDD (30MB FORMATTED) -FDD (5.25″ 1.2MB) -14″ PAPER-WHITE MONITOR	2set	'90. 8.	UNION DB 구축용 (CD-ROM 서지데이타베이스인 CAT CD450 운영)	구 입
◎ 486 PC -CPU i80486/SX (33MHz) -HDD(340MB FORMATTED) -FDD(5.25″ 1.2MB) -14″ PAPER-WHITE MONITOR	1set	'91.5.~6	UNION DB 구축용 (목록데이타 통합 작업 및 책자형 종합목록 출력)	임 대
◎ TGLBP 600A	1set	'91.5.~6.	UNION DB 구축용 (책자형 종합목록 출력)	임 대
◎ 시스마시스템사의 내장형 TAPE BACKUP STREAMER	1set	'91. 6.	UNION DB 구축용 (M/T 수록용)	임 대
◎ DPT사의 SMARTCACHE PLUS SCSI CONTROLLER	1set	'91. 6.	UNION DB 구축용 (광디스크 수록용)	임 대

일체를 구입하였으며 로칼 목록데이타 중복 체크, 통합작업, 책자형 종합목록 출력용 소프트웨어는 오롬컴퓨터(주)로 부터 구입했다. 본 사업추진에 소요되는 장비 즉, 로칼데이타베이스 통합과 통합된 데이타의 백업 및 광디스크 수록용 장비는 일회성이기 때문에 구입대신 임대하는 방식을 채택했다. 관련 장비와 소프트웨어 내역은 〈표3〉과 〈표4〉와 같다.

3.2 데이타 처리 표준

3.2.1 표준 MARC 포맷 선정

사업추진위원회와 실무반은 목록데이타 변환작업을 효율적으로 추진하기 위하여 로칼데이타 처리를 위한 데이타의 수준향상과 공동편목의 효과를 고려해서 표준 MARC을 사용하는 것이 소급변환의 효율성을 가장 높일 수 있다고 판단, 결정하고 다양한 MARC 포맷 중에서도 세계적으로 표준 포맷이라 할수 있는 US MARC과 그 내용면에서 99%이상 호환성이 있는 LC MARC 또는 OCLC MARC을 본 사업을 위한 목록데이타 소급변환작업의 표준 데이타 포맷으로 선정하였다.

그리고 각 도서관의 장서구성 성격, 관종별 특성, 현재 운용 및 계획중인 시스템에서 사용하는 MARC과의 호환성뿐만 아니라 소속 도서관간의 협력 효용성을 고려해서 LC MARC 또는 OCLC MARC 레코드 전체가 대용량의 데이타베이스인 CD-ROM에 수록되어 있는 BiblioFile 또는 CAT CD450를 공통으로 선정하여 목록데이타 변환작업에 활용했다.

〈표4〉 소프트웨어 일람표

기 종	수 량	도입일자	용 도 및 기 능	비 고
◎ CAT CD450 (CD-ROM 서지데이타베이스)	1 set	'89. 11.	UNION DB 구축용 (데이타 소급변환용 표준 포맷 수록 CD-ROM DB)	구 입
◎ 오롬 2000 목록 모듈	1 set	'91. 5.	UNION DB 구축용 (로칼목록데이타 중복체크 및 통합작업-로칼 DB생성)	구 입
◎ Union Database Merge 모듈	1 set	'91. 5.	UNION DB 구축용 (로칼DB 통합작업-UNION DB 생성)	구 입
◎ Union Database 종합목록 출력 모듈	1 set	'91. 5.	UNION DB 구축용 (책자형 종합목록 출력)	구 입

3.2.2 소급변환용 매체 CD-ROM 데이타베이스

LC MARC에 기초한 CD-ROM 서지데이타베이스인 BiblioFile은 1900년 부터 1964년까지는 많이 쓰이는 레코드만을 수록하였고 1965년 이후는 전체 레코드를 수록하고 있으며 4장의 디스크로 된 LC MARC:English language (약 200만건 이상)와 2장의 디스크로 된 LC MARC:Foreign language(약 100만건 이상:index 포함)으로 구성되어 있다.

CAT CD450는 Recent Books (128만건/2장)와 Older Books/Nonbooks(128만건/2장)로 구성되어 있으며 Recent Books는 OCLC의 Online Union Catalog (OLUC)에서 최근 6년 이내에 출판된 도서중 이용율이 높은 레코드를 수록하고 있으며 Older Books/Nonbooks는 OLUC에서 출판년도가 6년이 넘는 도서중 가장 빈번하게 이용되는 레코드와 출판년도에 관계없이 이용율이 높은 비도서자료 레코드가 수록되어 있다.

이러한 레코드를 수록하고 있는 두 종류의 데이타베이스는 제공하는 로칼데이타 처리방법이 각 시스템의 성격에 따라 첫째, 화면에 디스플레이 되는 모양이나 사용가능한 키 스트로크, 둘째, 소급변환방법, 셋째, MARC 통신화일로 변환시키는 방법 등 기술적인 측면에서 3가지가 다르지만 종합목록데이타베이스에 활용할 화일인 MARC 통신형태는 동일하기 때문에 각각에서 다운로드된 MARC 레코드의 혼합사용이 가능하다.

3.2.3 협력기관간 변환작업 통일기준안 채택

CD-ROM 데이타베이스를 활용한 목록변환작업을 위해 선택된 2종의 데이타베이스를 모두 소유하고 있는 기관에서는 처리 능률성을 고려하여 혼용해서 사용할 수 있다. 즉, 목록데이타의 ISBN 또는 LCCN이 확인되는 경우에는 CAT CD450가 효율적이지만 그렇지 않은 경우 다시 말하자면 서명이나 저자명 정도의 데이타만을 알 수 밖에 없는 경우에는 BiblioFile이 효과적인 것으로 증명됨으로써 후자의 경우는 일단 BiblioFile로 1차 변환작업을 한 후 hit 되지 않는 데이타만 CAT CD450로 다시 검색하여 변환작업을 수행했다.

소급변환되는 목록데이타는 선정된 2종의 데이타베이스에서 공통적으로 사용할 수 있는 태그(Tag)를 정하고 로칼데이타의 처리범위는 각 기관의 청구기호, 소장기관표시 코드, 자료의 소장사항, 등록번호로 정하였으며 추후 계속 추가되는 목록데이타도 동일한 범위내에서 처리하도록 하였으며, 본 사업이 종합목록데이타베이스 사업인 관계상 각각의 기관에서 기존에 사용하던 시스템간의 차이때문에 발생할 수 있는 여러가지 변수 즉, 로칼데이타의 범위, 입력필드, 입력방법, 목록레코드 선정 및 편집, 로칼세이브화일 생성, MARC 통신형태 화일 생성과정인 Validate와 Export 작업방법 등의 기본사안이 통일되어야 하기때문에 여러 차례에 걸친 실무반의 회의과정을 거

치면서 "Union Database 구축을 위한 Manual"이 마련됨으로써 협력기관 간에 일관성 있는 목록데이타 변환작업 통일기준(안)을 채택하게 되었다.

3.2.4 로칼데이타 처리

3.3.4.1 청구기호처리

데이타 협력기관으로 부터 수집된 목록데이타가 7개 과제수행기관에 분배되어 소급변환되는 과정에서 가장 중요한 것은 서로 다른 7개 기관이 공통된 기준하에 변환되어야 하는 것이다. 전술한 바와 같은 목록데이타 변환작업 통일기준안에 의하여 필드가 지정되고 입력방법이 통일된 내역은 다음과 같다. 우선 로칼데이타 입력필드는 태그 049로 지정하고 소장기관(서브필드 a), 자료소장사항(서브필드 v,y,c), 자관의 등록번호(서브필드 l)를 입력하고 태그 09X에는 청구기호를 입력하는데 청구기호의 입력필드와 입력방법은 분류기호의 유형에 따라 즉, 090은 LC, 092는 DDC, 098은 UDC, 099는 Local Free Text Number를 입력하도록 하였다. 이때 각각의 경우를 도표로 나타내면 아래와 같다.

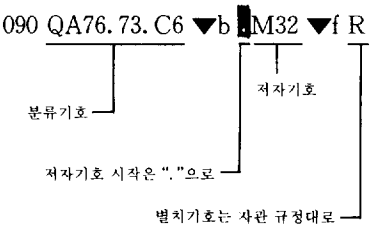
A. 청구기호의 입력 필드 및 입력 방법

1) LC Number의 경우

가) 입력 필드

태그	지시자	서브필드
090	XX	▼a 분류기호 ▼b 저자기호 (도서기호) ▼f 별치기호

나) 입력 방법

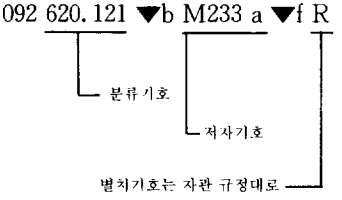


2) DDC Number의 경우

가) 입력 필드

태그	지시자	서브필드
092	XX	▼a 분류기호 ▼b 저자기호 (도서기호) ▼f 별치기호

나) 입력 방법

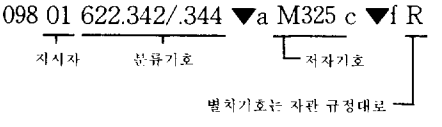


3) UDC Number의 경우

가) 입력 필드

태그	지시자	서브필드
098	01	▼a 분류기호 ▼b 저자기호 (도서기호) ▼f 별치기호

나) 입력방법



4) Local free text call number의 경우

보고서 또는 conference proceedings 등의 성격을 가진 자료에 대해서 Local library에서 고유한 번호를 청구기호의 의미로 부여한 경우에 사용하는 필드이다.

가) 입력 필드

태 그	지시자	서 브 필 드
099	XX	▼a Local free text call number (▼a Local free text call number) ▼f 별치기호

나) 입력 방법

099 IEEE80CH1624-6

Local free text call number

또는

099 UNESCO ▼a SR-15 ▼f R

별치기호는 각관 규정대로

Local free text call number

5) Local Call Number (LC, DDC, UDC, Local free text call number) 입력 필드에 volume, number 등의 기입을 원하는 경우가 있으면 다음과 같이 입력한다.

090 QA73 ▼b .B32 v. 14

3.2.4.2 소장기관코드 처리

소장기관코드의 입력필드와 입력방법은 태그 049 필드에서 첫번째 서브필드 a에 알파벳 대문자 4자리로 입력한다. 이때 알파벳 대문자 4자리를 시스템의 "SET UP"모드에서 지정하면 EDIT 화면에서는 자동으로 디스플레이된다. 동일한 기관코드가 발생하는 것을 방지하기 위해 주관기관에서는 소장기관의 코드 〈표5〉를 등록해야 하며 목록데이타 변환작업시에는 주관기관에 등록된 소장기관코드만을 사용해야 한다.

〈표5〉 소장기관 코드

기 관 명	소장기관 코드
국방과학연구소	ADDD
산업기술정보원	KIET, KINI
통신개발연구원	KISD
한국과학기술원	KAIS
한국과학기술연구원	KIST
한국국방연구원	KIDA
한국동력자원연구소	KIER
한국전기통신공사연구개발단	KTRC
한국전력공사기술연구원	KEPC
한국전자통신연구소	ETRI
한국원자력연구소	KAER
한국표준연구소	KSRI
한국해사기술연구소	KRIS
한국해양연구소	KORD

3.2.4.3 소장사항과 등록번호 처리

volume, part, section, number 등 소장분책, 내용년도, 복본표시 등의 소장사항과 등록번호는 태그 049 필드에서 소장기관표시가

입력되는 첫번째 서브필드 다음에 해당 서브필드를 입력한다. 이때 총서, 복본, 연감인 경우에는 각 item에 대해 발생하는 서브필드를 계속 연결해서 입력한다. 단, 한 item에 대한 서브필드 중 '▼1 (자관의 등록번호)'이 마지막에 입력되어야 하며 그 실례를 들면 다음과 같다.

가) 입력필드

태그	지시자	서브필드
049	XX	▼a 소장기관 코드 ▼v volume, part, section, number 등 소장분책 표시 ▼y 내용년도 표시 ▼c 복본표시 ▼1 자관의 등록번호

나) 입력방법

(1)낱권 도서인 경우

049 KAIS ▼1 1432

(2)총서 및 복본인 경우

049 KIST ▼v v. 3 ▼1 2215 ▼ v v. 3 ▼c c. 2 ▼1 2216 ▼v v. 4 ▼1 2315

(3)연감인 경우

049 KSRI ▼y 1988 ▼1 2214
또는
049 KSRI ▼y 1986-1987 ▼1 2216 ▼y 1988-1989 ▼1 2217

(4)volume과 part (or number, section등)가 함께 있을 경우에는 소장사항 입력 필드에 다음과 같이 입력한다.

▼v v. 8pt. 1

(5) 1책에 volume이 여럿인 경우 표기방법은

▼v v. 1-3

〈그림1〉 편집완료된 로칼데이타 화면

```
  Search   Edit  Display  Batch  Print  Options Windows
 OCLC - 01461200  (667     )   Rec stat: c     Entrd  750720
 Type: a   Bib lvl: s  Govt pub:    Lang: eng  Source: d  S/L ent: 0
 Repr:     Enc lvl: I  Conf pub: 0  Ctry: nyu  Ser tp:    Alphabt: a
 Indx: u   Mod rec:    Phys med:    Cont:      Frequn:     Pub st: c
 Desc:     Cum ind: u  Titl pag: u  ISDS: 1    Regulr: x    Dates: 1967-1971
-------------------------------------------------------------------------------
>>010      67-22786
>>040      MUL ▼c MUL ▼d FUL ▼d HUL ▼d NSD ▼d OCL ▼d NSD ▼d OCL ▼d m/c
>>022  0   0065-2741
>>050  0   QD515 ▼b .A3
>>082  0   541/.3687/08
└ 090      QD515 ▼b .A3
└ 049      KAIS ▼v v.  1 ▼1 1023757 ▼v v.  2 ▼1 1023758 ▼v v.  3 ▼1 1023759
▼v v.  4 ▼1 1023760
>>220  00  Advances in high temperature chemistry
>>245  00  Advances in high temperature chemistry.
>>260  00  New York, ▼b Academic Press, ▼c 1967-1971.
>>300      v.  ▼b ill. ▼c 24 cm.
>>570      Editor: v.1-   L. Eyring.
>>650   0  High temperatures. ▼w cmc.
>>650   0  Thermochemistry. ▼w cm
>>700  11  Eyring, LeRoy, ▼e ed. ▼w cn
 SEND = Save and Close  F5 = Cancel and Close  Ctrl-F6 = Save and Continue
R:  9   C:1   Size:718
```

3.2.4.4 처리된 목록레코드의 실례

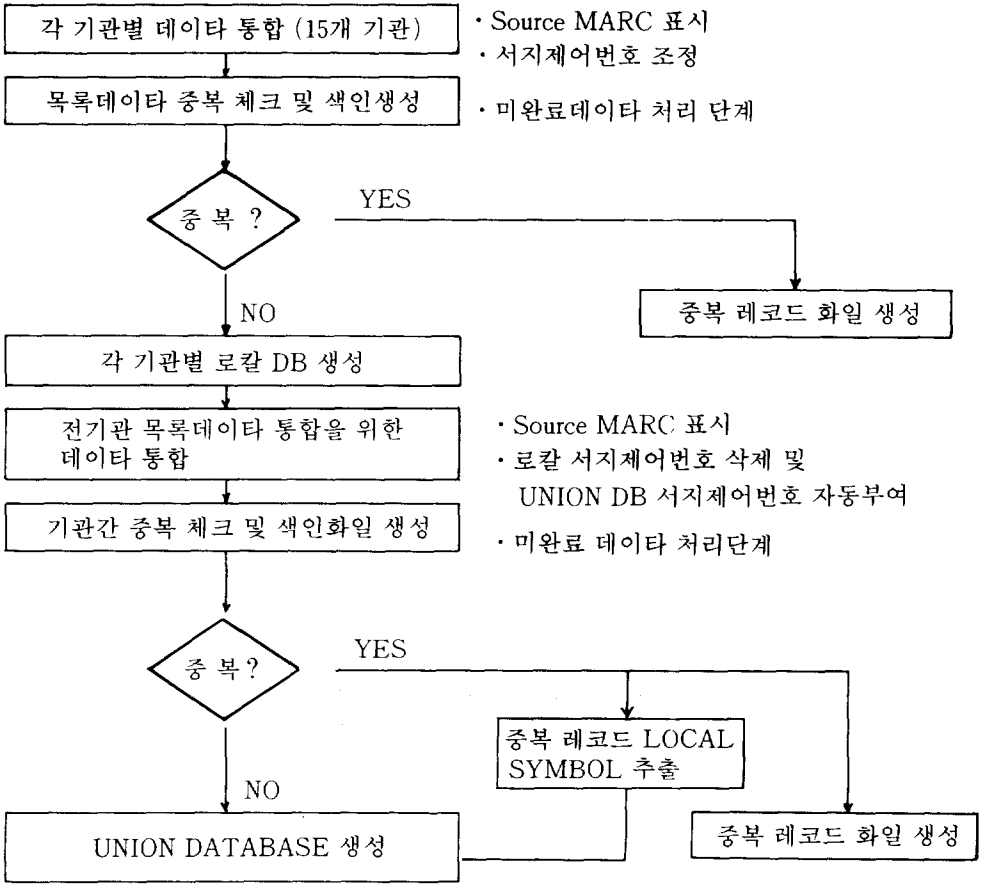
로칼데이타 처리기준에 따라 편집완료된 로칼데이타가 디스플레이 되는 화면은 <그림1>과 같다. 화면내의 태그 082 밑에서 볼 수 있는 "L"표시사항이 로칼편집 데이타이다.

3.4 목록데이타 통합

각 과제수행기관에서 소급변환작업을 완료한 자관소장도서의 목록데이타는 정보교환용 MARC 통신형태로 변환시키고 두단계의 통합작업을 거쳐서 Union Database를 구축해야 한다. 우선 첫단계로 각 기관별 화일간의 목록데이타 중복점검과 점검된 데이타를 통합하면 로칼데이타베이스를 생성할 수 있고 두번째 단계로 생성된 로칼데이타베이스간의 중복점검을 한 뒤 점검된 데이타베이스를 통합하여 완전한 Union Database를 구축해야 한다.

본 사업수행의 최종 단계인 중복점검과 통합작업에 관한 업무흐름도는 〈그림2〉와 같다.

〈그림2〉 통합작업 흐름도



통합된 데이타베이스를 생성하기 위한 통합 과정에서 필요한 기준은 데이타포맷, 중복 점검키, 데이타의 소스 MARC 표시 등으로 구분하며 동일한 데이타의 중복 점검 및 통합 기준은 다음과 같다.

1) 데이타 포맷은 OCLC MARC과 LC MARC 데이타를 공유하고 있기 때문에 고유 레코드인 경우에는 두가지 포맷을 공존시키고 OCLC 제어번호, LCCN, ISBN 등의 서지인식번호가 같은 레코드로 인하여 중복 데이타가 발생하는 경우에는 LC MARC 데이타를 선택하여 처리하였다.
2) 중복 점검키는 태그 001의 OCLC 제어번호, 태그 010의 LCCN, 태그 020의 ISBN 으로 하였다.
3) 소스 MARC 표시는 선택키에 의해 리더 부분의 8자수 위치에서 LC MARC은 'l'로 OCLC MARC은 'o'로 하였으며, 이것은 시스템에 의해 자동으로 표시된다.

3.4.1 로칼별 목록데이타 통합

과제수행기관별로 소급변환이 완료된 데이타는 화일별로 저장되어 있기 때문에 우선 각 기관별 로칼 목록데이타 통합작업을 수행해야 한다. 로칼 데이타베이스 생성작업은 두단계를 거치는데 그 첫단계는 화일별 목록데이타를 하나의 로칼 데이타베이스로 생성하는 단계로써 통합대상화일명과 데이타의 사용 MARC을 선택하면 모든 데이타가 통합되며 태그 001은 OCLC MARC인 경우에는 추후 활용을 위해 태그 035로 이동하고 LC MARC인 경우에는 삭제되면서 로칼 서지제어번호가 자동적으로 부여된다. 두번째 단계로는 각 화일의 통합작업 단계가 완료된 통합 데이타 중에서 중복되는 등록번호와 서지인식번호인 ISBN, ISSN, LCCN을 체크하여 중복된 데이타는 에러로 분류하여 〈그림3〉과 같은 형태

〈그림3〉 에러 데이타 수록 화일

```
---------------------------------------------------------------------------
===> LCCN 중복 : 75026615

  현재의 데이타 : 제어번호 = KAIST10026489
  QA613.6
  H57       Differential Topology / Morris W. Hirsch. -- New York :
              Springer-Verlag, 1976.
              x, 221 p. : ill. ; 25 cm. -- (Graduate texts in mathematics ; 33)
  1009425
              LCCN 75-26615//r77

  기존 데이타 : 제어번호 = KAIS10022632
  M
  QA613.6
  H669      Differential Topology / Morris W. Hirsch. -- New York :
              Springer-Verlag, 1976.
              x, 221 p. : ill. ; 25 cm. -- (Graduate texts in mathematics ; 33)

  0031750     LCCN 75-26615//r77
---------------------------------------------------------------------------
```

로 ORCLOG.DAT 라는 화일에 수록되어 에러체크 작업용 화일로 생성된다.

로칼 데이타베이스 생성 작업과정을 수행하여 얻은 결과는 〈표6〉과 같다.

아래의 〈표6〉에서 보는 바와 같이 15개 기관의 소장목록데이타 총213,197건 중 소급변환이 완료되어 MARC 통신형태로 다운로드된 데이타 총 203,331건을 로칼별 데이타베이스로 통합한 결과 146,333건의 서지이며, 고유레코드 건수로는 총 106,245 레코드가 생성되었다.

3.4.2 미변환데이타 처리

당초 사업추진위원회에서는 사업수행중 발생할 수 있는 모든 미변환(no-hit) 데이타는 OCLC나 LC에 위탁하여 일괄 소급변환처리(Retrospective conversion)하도록 하고 일괄 소급변환으로도 해결되지 않는 데이타는 차후에 협의하여 처리하기로 하였으나 no-hit 되는 데이타가 예상외로 많았기 때문에 OCLC나 LC에 의뢰하는 방법은 보류하고 일단 사업주관기관에서 일괄 수집하여 재처리를 시도한 후 그래도 hit 되지 않는 데이타는 추후 오리지날 카탈로깅 작업을 하기로 결정했다.

미변환데이타의 수집대상기관은 12개 기관이었으나 한국표준연구소 1,400건을 포함한 20,300건은 자료송부지연으로 보류시키고 수집된 데이타 19,154건만 주관기관인 한국과학

〈표6〉 로칼 데이타베이스 레코드 현황

기 관 명	목록데이타 통합	
	서지건수	레코드건수
국방과학연구소	18,878 건	12,309 레코드
산업기술정보원	9,583 건	9,382 레코드
통신개발연구원	1,797 건	1,797 레코드
한국국방연구원	1,929 건	1,705 레코드
한국과학기술연구원	14,192 건	9,570 레코드
한국과학기술원	51,701 건	36,670 레코드
한국동력자원연구소	8,310 건	5,572 레코드
한국원자력연구소	16,614 건	11,373 레코드
한국전력공사기술연구원	1,180 건	971 레코드
한국전자통신연구소	7,844 건	5,250 레코드
한국통신연구개발단	3,101 건	2,740 레코드
한국표준연구소	6,672 건	5,291 레코드
한국해사기술연구소	1,444건	1,225 레코드
한국해양연구소	2,112 건	1,826 레코드
한국화학연구소	976 건	664 레코드
계(15개 기관)	146,333 건	106,245 레코드

기술원에서 재작업을 시도하여 총5,345건을 소급변환할 수 있었으며 재작업 보류된 데이타는 계속적으로 수집하여 변환하기로 하였다. 이상의 재작업 과정까지 소급변환을 완료한 데이타를 과제수행기관별, 처리대상 데이타별, 처리내역별로 정리하면 〈표7〉과 같다.

3.4.3 로칼 데이타베이스 통합

통합 데이타베이스 구축을 위해서는 여러기관의 로칼 데이타베이스를 하나의 데이타베이스로 생성하는 단계와 생성된 데이타베이스를 통털어 중복 레코드를 점검하는 단계를 수행해야 하는데 우선 로칼 데이타베이스를 Union Database로 통합하기 위해서 로칼 데이타베이스 내의 데이타는 수행기관에서 다운로드 받을 때 수록된 태그 049, 090, 092, 098, 099중에서 로칼의 소장기관코드를 입력한 태그 049의 서브필드 a만을 남겨 두고 태그 049와 태그 090, 092, 098, 099에 수록된 모든 사항을 삭제한 후 Union Database에 통합시킨다. 이때 LC 청구기호 필드인 태그 050이 존재하지 않

〈표7〉 데이타 변환작업 현황

(단위 : 건)

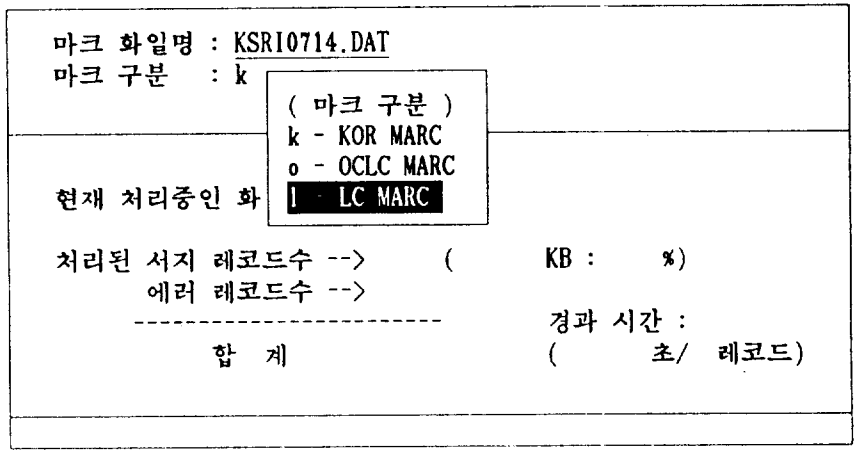
과제수행 기관명	처리대상 데이타			처리내역					
	소장기관명	데이타수	소계	처리		변환		미변환(No-hit)	
				데이타수	소계	데이타수	소계	데이타수	소계
한국과학기술연구원	KIST	25,000	25,000	25,000	25,000	14,192	14,192	10,808	10,808
한국과학기술원	KAIST	60,000	62,800	60,000	62,800	51,701	53,857	8,299	8,943
	KEPCO	1,777		1,777		1,180		597	
	KRICT	1,023		1,023		976		47	
한국동력자원연구소	KIER	11,500	20,000	11,500	20,000	8,310	16,154	3,190	3,846
	ETRI	8,500		8,500		7,844		656	
한국원자력연구소	KAERI	20,000	20,000	20,000	20,000	16,614	16,614	3,386	3,386
한국통신연구개발단	KTRC	5,000	33,531	5,000	33,531	3,101	16,410	1,899	17,121
	KINITI	17,168		17,168		9,583		7,585	
	KISDI	2,938		2,938		1,797		1,141	
	KIDA	8,425		8,425		1,929		6,496	
한국표준연구소	KSRI	8,000	33,000	8,000	28,000	6,672	25,550	1,328	2,450
	ADD	25,000		20,000		15,000		10,000	
한국해사기술연구소	KRISO	8,000	18,866	8,000	14,000	1,444	3,556	6,556	10,444
	KORDI	10,866		6,000		2,112		3,888	
계	15개 기관	213,197		203,331		146,333		56,998	

는 경우에는 로칼 LC 청구기호 입력필드인 태그 090은 삭제하지 않는다. 통합작업 중 기존 데이타와 LCCN, ISBN 등 서지인식번호가 중복될 경우에는 작업중인 데이타의 태그 049에 있는 소장기관명만 기존의 통합된 데이타베이스에 포함시키는 과정이 선행되어야 한다.

예를들면 KAIST, KIST, KSRI 세 기관에서 소장하고 있는 도서인 경우에 "049▼a KAIS▼a KIST▼a KSRI"로 수록한다. 또한 Union Database에 있는 데이타와 중복되지만 LCCN, ISBN 등의 서지인식번호가 일부 다른 경우에는 태그 049의 소장기관코드만 등록시키고 중복으로 점검된 신규 데이타에만 들어 있는 서지인식번호를 기존 데이타에 추가한다. 각 로칼 데이타베이스 도입단계를 수행하려면 〈그림4〉와 같이 각 기관 로칼데이타베이스명을 입력하고 MARC 형식을 선택(OCLC MARC은 'o', LC MARC은 'l') 하면 로칼데이타베이스가 통합되면서 LCCN, ISBN, OCLC 제어번호를 점검키로 하여 데이타 중복 점검 및 통합 작업을 수행한다.

데이타 도입과정에 따라 데이타 통합이 완료되면 Union Database에는 서지마스터화일, 미완료데이타화일, 서지인식번호화일 등 3개의 화일에 데이타가 추가되는 단계를 거치게

〈그림4〉 MARC 데이타 일괄입력 화면 예



되고 서지마스터화일에서 여러가지 검색용 정보를 추출하는 과정과 LCCN, ISBN, OCLC 제어번호를 키로 해서 중복 데이타를 에러로 처리하여 화일에 담아주는 기능을 하는 "미완료 데이타 처리" 단계를 수행하게 된다. 이와같은 단계 〈그림5〉를 거쳐 최종적으로 67,207건의 서지레코드를 수록한 Union Database를 구축하였다.

3.4.4 책자형 종합목록 출력

사업추진기관과 수행기관 그리고 데이타 협력기관이 모두 온라인으로 연결되어서 상호 입출력은 물론 검색활용이 가능하도록 하는 것이 본 사업의 궁극적인 목표이지만 현실적으로 불가능한 상황에서는 책자형의 목록이 요구된다. 〈그림6〉과 같은 흐름도를 통해 알 수 있듯이 책자형 종합목록을 출력하기 위해서는 3단계 작업과정이 필요하다. 첫째과정으로 Union Database 마스터화일에서 출력대상 서지레코드를 추출하여 본목록, 서명 및 저자명

〈그림5〉 미완료데이타처리 화면 예

```
[과 학 단 지 U I O N  D B  편 목 / 검 색 ]

1. 청 구 기 호                          [Z8820.2.L53 1984    ]
2. 입 력 일 자                          [91.07.07]
3. 서지제어번호
4. 서지인식번호
5. 저자 키워드
6. 서명 키워드
7. 주제 키워드

[ 1 단계 . 서지 기본정보 추출 ]
처리건수  총     2건중     1건 ( 50% )        경과시간 00:00:01
처리중인  제어번호  CSL00037522
          서    명  Unsteady motions of compressible media w
```

미완료데이타 처리가 끝나면 다음과 같이 처리결과를 보여주는 화면이 나옵니다.

```
1단계 시작: 총건수 2, 시작시간 13:25:24
1단계 종료: 처리건수 2, 종료시간 13:25:26
```

<Enter> 키를 치세요...

색인화일을 생성하는 단계이며, 두번째 과정으로는 색인화일 생성범위를 제한해 주어서 실제 출력에 필요한 색인화일을 생성하는 단계이며, 세번째과정으로는 종합목록색인을 실제 출력하는 단계 등이다. 본 목록의 수록범위는 1972년부터 1990년 상반기까지 입수된 서양서 단행본 총66,808종 중에서 청구기호가 없는 번호가 부여된 레코드 또는 청구기호로 간주할 수 없는 레코드는 제외한 나머지 60,316종을 제한적으로 추출하여 본목록 11권, 색인 4권으로 구성된 총15권 1세트로서 "과학단지장서종합목록" 이라는 제목으로 출간되었다.

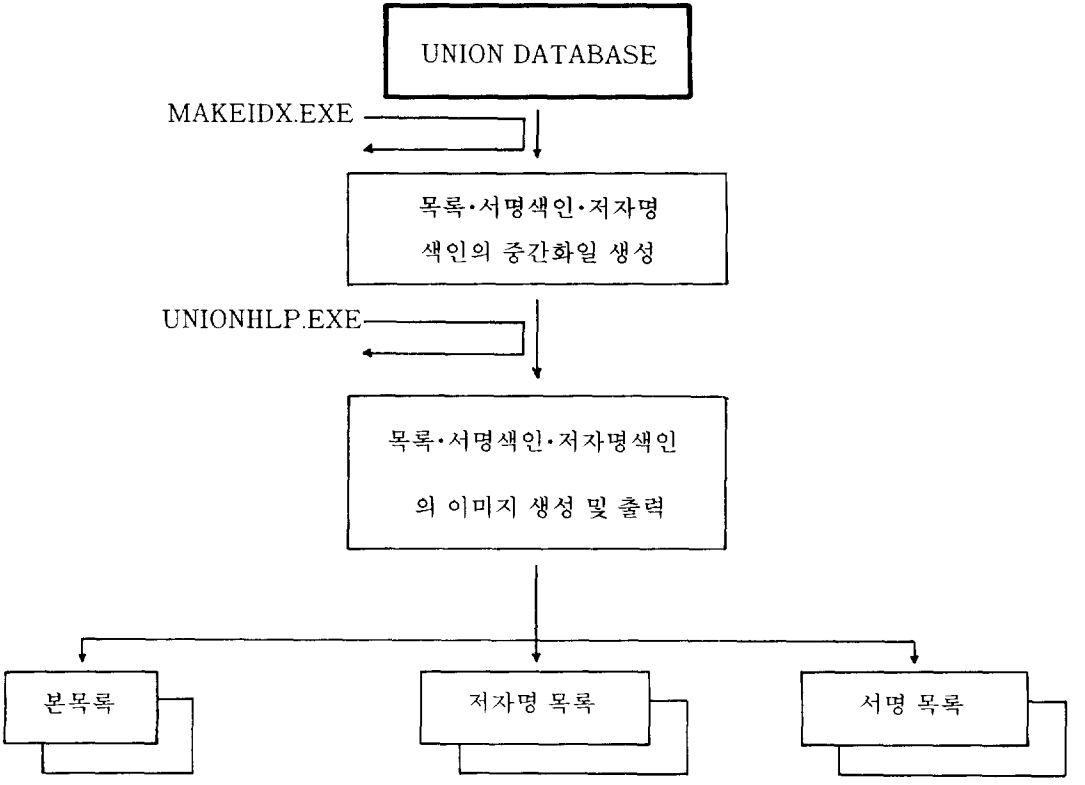
Ⅲ. 과학기술도서종합목록 데이타베이스

1. 사업의 개요

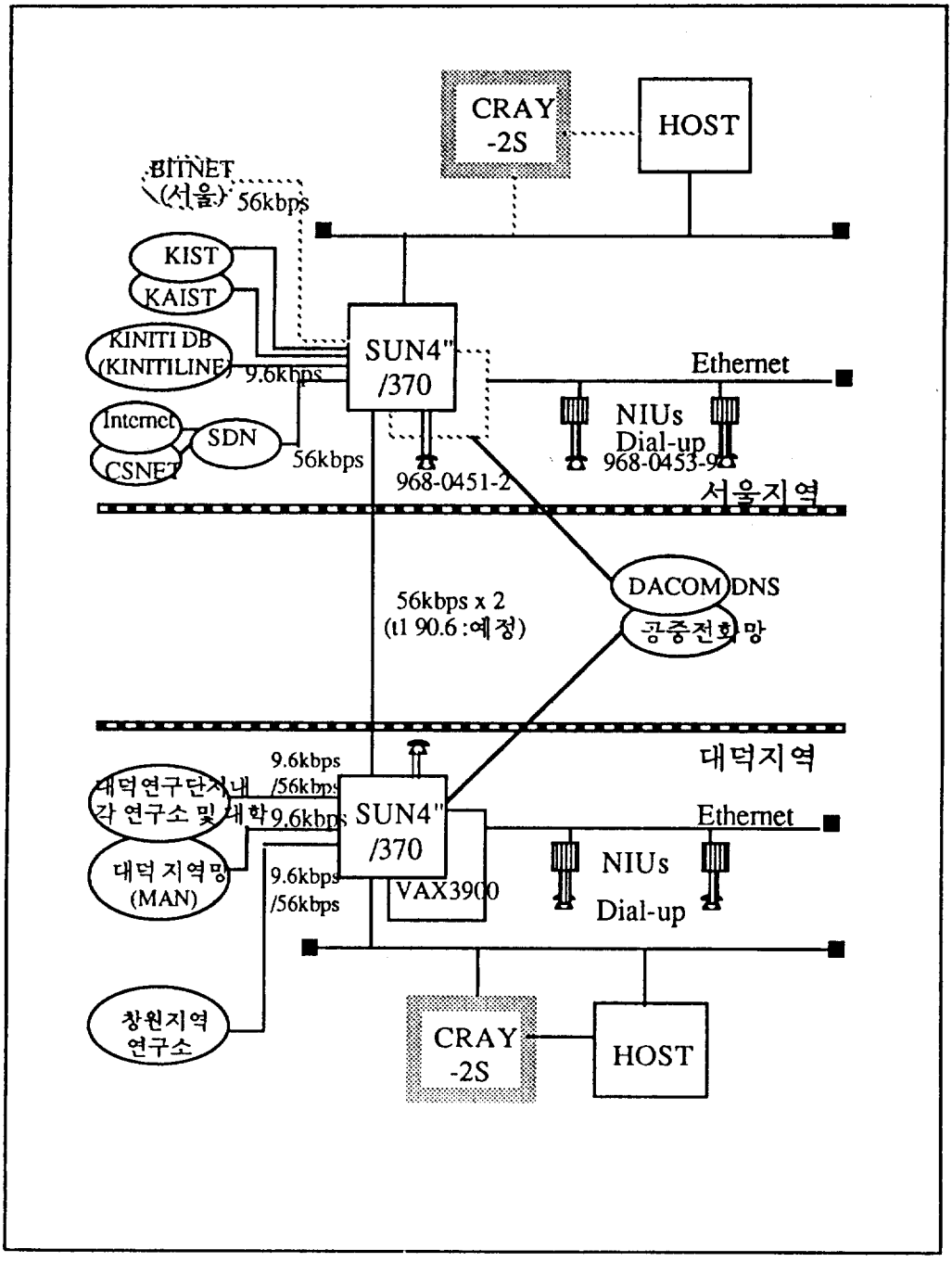
1990년 11월 과학기술처는 시스템공학연구소를 주축으로 하는 과학기술정보유통체제구축 세부시행계획(안)[11]을 마련하고 1) 전문분야별 정보의 수집 처리와 데이타베이스의 구축을 해당 전문 분야 출연연구기관이 분산·수

11) 과학기술정보유통추진위원회, 과학기술정보유통체제 구축시행안, 1990. 11. 과학기술처.

〈그림6〉 책자형 종합목록 출력 흐름도



〈그림7〉 연구전산망(KREONet) 구성도



행하되 사업의 총괄과 관련기술의 표준화을 위한 중앙조정기관으로서 시스템공학연구소를 지정하여 중앙정보센터 기능을 부여하고 2)구축된 데이타베이스는 연구전산망 (〈그림7〉 참조), 공중통신망, 해외망과 연결하여 산업계·학계·연구계에 서비스하며 3)데이타베이스 구축에 필요한 전산자원은 현재 보유하고 있는 컴퓨터 등을 최대한 활용하되 민간기업의 적극적인 참여를 유도한다는 내용을 주요골자로 한 과학기술정보유통체제구축 세부시행계획(안)을 과학기술정보유통추진위원회에 상정하여 1990년 12월 의결한 바 있다.

본 시행계획에 근거하여 과학기술처는 추가 갱정예산에서 총15억원의 예산을 배정받아서 해당 출연기관과 특정연구개발과제사업 명목으로 전문데이타베이스 구축 용역계약을 체결하게 되었으며, 한국과학기술원은 과학기술도서데이타베이스 구축분야에 대한 사업을 할당받게 되었다. 이에따라 한국과학기술원은 전년도에 수행한 바 있는 '과학단지장서종합목록데이타베이스 구축' 과정에서 축적된 기술을 바탕으로 하고 참여기관의 범위를 연구단지정보관리협의회 소속 32개 기관으로 확대하여 과학기술도서종합목록데이타베이스 구축사업을 수행하게 된 것이다.

1991년 5월부터 동년말까지 총8개월간을 사업기간으로 한 본 사업을 통하여 도서 데이타베이스분야 사업비 총1억원을 배정받은 한국과학기술원 과학도서관은 과학기술도서데이타베이스 구축과 동 사업의 지속적 발전을 위한 환경구축에 역점을 두었으며 특히 향후 계획할 온라인입출력시스템을 위한 기기확보에 주력하였으며 총20개 기관이 참여한 본 사업은 데이타의 수집과 처리과정 모두가 전년도에 수행한 과학단지장서종합목록 데이타베이스 구축사업과 유사하여 다만 이미 구축된 데이타베이스와 신규로 처리된 로칼 데이타베이스와의 통합과정 부분에서 새로운 방식이 도입되었을 정도에 그쳤지만, 이것은 향후 다수 데이타베이스 통합기술로 활용될 수 있는 좋은 경험요인이라고 볼 수 있다.

2. 사업추진내역

2.1 데이타베이스 사업추진전략

종합목록데이타베이스는 도서관 자원의 공동활용을 궁극적인 목표로 하는 도서관 네트웍의 핵심이 되는 요소로서 이와같은 데이타베이스를 구축하기 위해서는 1)데이타 표준포맷설정과 같은 표준화 작업 2)수록자료 항목 결정 3)데이타 수집원 및 수집방법 4)목록데이타 변환작업 5)변환 완료된 목록데이타 통합작업 6)보급계획 및 운영계획과 같은 사항들이 사전 검토되어야 될 필요가 있기 때문에 이와같은 사항을 검토한 결과 다음과 같은 추진전략을 수립하였다.

1) 종합목록데이타베이스 구축 대상자료는 연구단지정보관리협의회 회원기관에서 소장하고 있는 과학기술도서 서양서 단행본으로 한다.
2) 목록데이타의 변환 표준포맷은 LC MARC와 OCLC MARC으로 한다.

3) 원만한 변환작업을 위하여 회원 기관들은 주관기관과 상호 데이타를 교환할 수 있는 시스템 환경을 구축한다.
4) 로칼 데이타의 입력범위와 입력필드에 관한 기준을 설정한다.
5) 환경 구축이 여의치 못한 회원 기관의 목록데이타는 주관기관에서 수집하여 소급 변환한다.
6) 변환된 목록데이타의 검증을 통해 누락, 불일치, 표기오류 등의 문제를 해결하고 처리한다.
7) 로칼별 데이타베이스와 통합 데이타베이스 생성을 위한 중복 점검 및 통합 기준을 설정한다.

이상과 같은 추진전략의 구체적인 수행과정은 (1) 데이타 수집 (2) 표준화 (3) 로칼데이타 처리 (4) 통합 데이타베이스 생성 등으로 크게 4단계로 구분하였다.

〈표8〉 데이타 수집 현황

(단위 : 건)

기 관 명	데 이 타 수	데 이 타 수 집 형 식
국방과학연구소	1,797	인쇄물
국방품질관리소	1,174	인쇄물
산업과학기술연구소	8,200	플로피디스켓(MARC 통신형태)
산업기술정보원	9,944	인쇄물
통신개발연구원	403	M/T(간략 MARC 통신형태)
포항공과대학	60,000	플로피디스켓(MARC 통신형태)
한국과학기술연구원	8,954	인쇄물
한국과학기술원	1,985	인쇄물
한국국방연구원	1,980	인쇄물(카드목록 복사)
한국동력자원연구소	116	인쇄물(카드목록 복사)
한국원자력연구소	641	플로피디스켓(MARC 통신형태)
한국인삼연초연구소	1,654	인쇄물
한국전기연구소	2,312	인쇄물
한국전력공사기술연구원	266	인쇄물(등록원부 복사)
한국전자통신연구소	5,909	인쇄물
한국조폐공사기술연구소	1,979	인쇄물
한국통신연구개발단	1,703	플로피디스켓(MARC 통신형태)
한국표준연구소	142	플로피디스켓(MARC 통신형태)
한국해사기술연구소	104	플로피디스켓(MARC 통신형태)
한양화학(주)중앙연구소	4,437	인쇄물
계(20개 기관)	113,700	

2.2 데이타 수집

수록 대상자료로 선정된 데이타를 수집하는 방법과 수집 데이타의 형식은 인쇄물 또는 플로피 디스켓에 수록하는 등 각 회원 기관의 정보관리 담당부서로 부터 기관별 편의성에 따라 선택할 수 있도록 하였으며 사업주관기관인 한국과학기술원 외 4개 기관(한국과학기술연구원, 한국원자력연구소, 한국해사기술연구소, 한국표준연구소)에서는 이미 상호 동일한 편목모듈을 도입하여 표준 MARC을 사용하고 있기 때문에 MARC 통신형태로 변환된 데이타를 수집하를 수 있었으며, 그 데이타 수집 현황은 〈표8〉과 같다.

2.3 데이타 소급변환

수집된 데이타는 사업주관기관에서 전체 처리를 원칙으로 하여 주관기관과 동일한 목록모듈을 사용하고 있는 4개 기관에서는 표준 MARC에 의한 소급변환 처리결과를 MARC 통신형태로 변환시킨 상태로 접수되었기 때문에 수정이나 추가 편집과정없이 그대로 수용했으며 간략 MARC 통신형태로 데이타를 송부한 통신개발연구원의 데이타는 추가 편집과정을 거친 후 로칼데이타베이스로서 활용하였고 포항공대와 산업과학기술연구소에서 송부한 데이타는 MARC 통신형태이기는 하지만 표준 MARC 통신형태와는 일부 항목에서 편차가 있었기 때문에 송부된 데이타 전체를 텍스트 화일로 환원시킨 후 일정 절차에 따라 일괄작업으로 표준화해서 로칼데이타베이스로 활용하였다.

2.4 소급변환 완료 데이타

20개 기관으로부터 수집한 목록데이타는 총 113,700건에 달했으며 이를 처리한 결과 즉, 소급변환을 완료하고 MARC 통신형태로 다운로드한 데이타는 총95,450건으로서 로칼별 데이타베이스 생성을 위한 기초자료 준비를 할 수 있게 되었다.

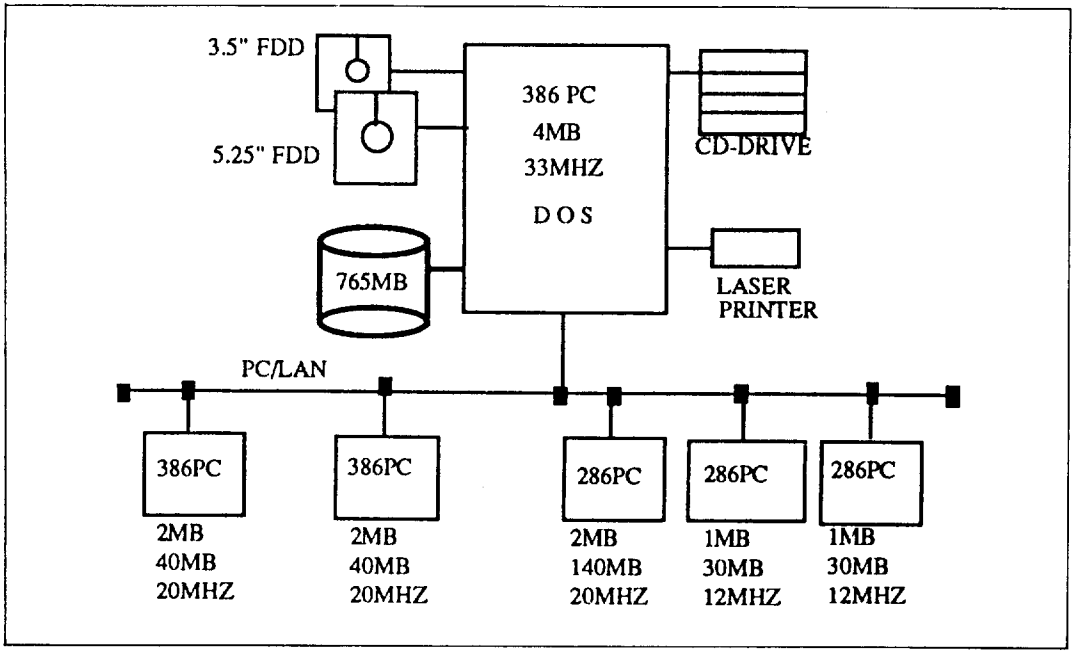
3. 데이타베이스 구축기술

3.1 시스템 환경

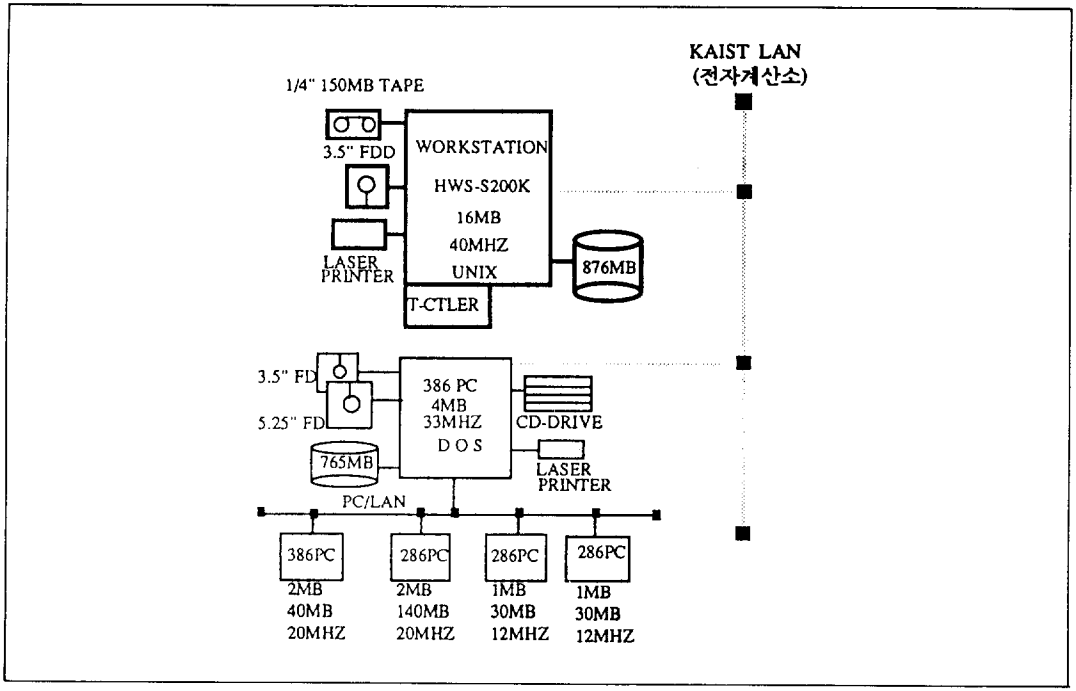
과학기술도서종합목록데이타베이스 구축을 위한 기본장비는 전년도에 수행한 과학단지장서종합목록데이타베이스 구축때 사용한 장비와 소프트웨어를 그대로 활용할 수 있었는데 그것은 전년도의 사업규모와 금년도 사업규모가 거의 같은 수준이었기 때문에 가능했던 것으로 평가된다. 다만 전체 데이타 용량의 증가로 인하여 부족한 하드디스크만 700메가급으로 보충하는 정도로 가능했으며, 일부 확충된 시스템 환경은 〈그림8〉과 같다.

시스템공학연구소로부터 지원받은 사업비는 총1억원이었으며, 본 사업비는 여타 전문연구소에 할당된 연구비와 유사한 성질을 띄고 있는데 그것은 데이타베이스 구축 자체에도 중요성을 부여하고 있으나 향후 계속될 데이타베이스 구축을 위한 시스템 환경 구축에 더욱 더 큰 비중을 둔 것이라고 볼 수 있다. 이에따

〈그림8〉 시스템환경 A



〈그림9〉 시스템환경 B



라 주관기관인 한국과학기술원 과학도서관에서는 향후 계속적으로 종합목록데이타베이스 구축의 중심기관으로서의 위상 정립을 위한 목표를 설정하고 전체 예산액의 60% 이상을 환경구축 경비로 사용하였으며 그 중 50%는 하드웨어 구입비로, 나머지 50%는 소프트웨어 구입비로 각각 집행한 바 있다. 이에따라 구축된 시스템 환경은 〈그림9〉와 같다.

3.2 데이타처리용 표준 MARC의 선택

신뢰성 높은 데이타베이스 구축을 위하고 각 도서관의 장서구성 성격, 관종별 특성, 현재 운영 또는 계획 중인 자동화시스템에서 사용하는 MARC과의 호환성은 물론 서지레코드의 유통을 통한 도서관간의 상호협력을 원활하게 하기 위해서는 신뢰성과 활용성이 공히 갖추어진 MARC을 채택해야 한다. 본 사업수행을 위해서 주관기관에서는 서양서 분야에서 세계적으로 인정되고 있는 표준포맷인 US MARC을 채택하고 US MARC을 기본으로 한 LC MARC 과 OCLC MARC 레코드를 대용량 축적매체인 CD-ROM에 수록하여 제공하고 있는 서지데이타베이스 BiblioFile과 CAT CD450을 목록데이타 변환작업의 표준도구로 하되 각 회원기관에서 보유한 형편에 따라서 선택할 수 있도록 하였다.

US MARC을 근간으로 하여 주관기관에서 사용하고 있는 목록모듈로 다운로드된 MARC 포맷 서지레코드는 〈그림10〉과 같다.

〈그림10〉MARC 포맷 서지레코드 예

```
제어번호: CSL00011311   상태: 수정   갱신일: 91.05.15   구분: OCLC MARC

유형: a   서지형식: m   입력수준:    기술형식:      입력일: 84.09.23
삽도: a   연도유형: s   발행년도: 1984,          발행국: dcu  언어: eng
색인: 1   지식수준:    자료형태:    내용형식: b   목록전거:    기판:
전기:     회의간행: 0   기념논문:    중복기입:     문학형식:    수정:

LCCN          >>010     ▼a 83-18614
지역시스템    >>035     ▼a ocm10020299
목록전거      >>040     ▼a DLC ▼c DLC ▼d m.c.
ISBN          >>020     ▼a 0070503281 : ▼c $35.00 (est.)
소장기관표시  >>049     ▼a KAER ▼a KIST ▼a KIER ▼a KAIS
LC분류기호    >>050 00  ▼a QA901 ▼b .A53 1984
기본개인명    >>100 1   ▼a Anderson, Dale A. ▼g (Dale Arden), ▼d 1936-
서명저자사항  >>245 04  ▼a Computational fluid mechanics and heat
              transfer / ▼c Dale A. Anderson.
출판사항      >>260     ▼a Washington : ▼b Hemisphere Pub. Corp. ;
              ▼a New York : ▼b McGraw-Hill, ▼c c1984.
대조사항      >>300     ▼a xii, 599 p. : ▼b ill. ; ▼c 25 cm.
서지적주기    >>504     ▼a Bibliography: p. 571-590.
일반주기      >>500     ▼a Includes index.
주제부출주기  >>650  0  ▼a Fluid mechanics.

[F1 = 도움말] 고정필드 F  가변필드 V  저장종료 S  취소종료 X ESC
```

3.3 로칼 데이타베이스

3.3.1 로칼데이타 처리

로칼데이타 처리를 위한 목록레코드의 선정과 편집기준, 로칼데이타의 입력범위 (청구기호의 입력필드, 소장기관코드의 입력필드와 입력방법, 소장사항의 입력필드와 입력방법, 등록번호의 입력필드와 입력방법) 그리고 로칼 데이타베이스 생성 등의 모든 절차와 방법은 과학단지장서종합목록 데이타베이스의 그것과 동일하다. 다만 참여기관 수의 증가로 인하여 코드 수가 확장되었으며 〈표9〉의 소장기관코드 중 산업기술정보원, 한국통신연구개발단, 한국동력자원연구소 등과 같이 기관명의 변경과 기관조직개편 등의 상황변동이 있었던 경우는 여러 개의 코드를 갖고 있기 때문에 상황이 끝나는 시점에서 일괄 조정하여 단일화해야 한다.

〈표9〉 기관명, 기관약호, 소장기관코드

기 관 명	기 관 약 호	소장기관 코드
국방과학연구소	ADD	ADDI
국방품질관리소	DQAA	DQAA
산업과학기술연구소	RIST	RIST
산업기술정보원	KINITI	KIET, KINI
통신개발연구원	KISDI	KISD
포항공과대학	POSTECH	POST
한국과학기술연구원	KIST	KIST
한국과학기술원	KAIST	KAIS
한국국방연구원	KIDA	KIDA
한국동력자원연구소	KIER	KIER, KIRE
한국원자력연구소	KAERI	KAER
한국인삼연초연구소	KGTRI	KGTR
한국전기연구소	KERI	KERI
한국전력공사기술연구원	KEPCO	KEPC
한국전자통신연구소	ETRI	ETRI
한국조폐공사기술연구소	KSPM	KSPM
한국통신연구개발단	KTRC	KTRC, KTA, KTAR
한국표준연구소	KSRI	KSRI
한국해사기술연구소	KRISO	KRIS
한국해양연구소	KORDI	KORD
한국화학연구소	KRICT	KRIC
한양화학(주)중앙연구소	HYCC	HYCC

3.3.2 로칼 데이타베이스 생성

전년도 수행한 과학단지장서종합목록 데이타베이스 구축시 수행한 로칼데이타 처리방식과 동일하게 처리된 데이타는 통합 데이타베이스 생성을 위한 사전작업으로써 로칼 데이타베이스로 구축되어야 한다. 각 기관별로 입수된 목록데이타를 소급변환작업을 거친 결과 생성된 로칼 데이타베이스 생성현황은 〈표10〉과 같다.

〈표10〉에서 보는 바와 같이 20개 기관의 목록데이타 총113,700건을 처리한 결과 95,450건의 데이타가 소급변환 완료되고 MARC 통신형태로 다운로드되었으며, 이를 다시 로칼별 데이타베이스 생성을 위해 중복 및 에러 점검과정을 거쳤을 때 총 66,928개의 서지레코

〈표10〉 로칼 데이타베이스 현황

기 관 명	목 록 데 이 타 변 환 작 업			로칼 DB 레 코 드 수
	대상서지건수	변환 (HIT)	미변환 (No-hit)	
국방과학연구소	1,797	875	922	559
국방품질관리소	1,174	1,001	173	971
산업과학기술연구소	8,200	8,200	8,200	8,200
산업기술정보원	9,944	2,821	7,123	2,485
통신개발연구원	403	360	43	309
포항공과대학	60,000	60,000	—	39,720
한국과학기술연구원	8,954	6,815	2,139	3,719
한국과학기술원	1,958	1,448	537	1,409
한국국방연구원	1,980	1,860	120	1,760
한국동력자원연구소	116	89	27	67
한국원자력연구소	641	641	—	641
한국인삼연초연구소	1,654	1,515	139	949
한국전기연구소	2,312	1,186	1,126	949
한국전력공사기술연구원	266	199	67	152
한국전자통신연구소	5,909	2,695	3,214	2,249
한국조폐공사기술연구소	1,979	1,334	645	628
한국통신연구개발단	1,703	996	707	924
한국표준연구소	142	142	—	142
한국해사기술연구소	104	104	—	102
한향화학(주)중앙연구소	4,437	3,169	1,268	993
계(20개 기관)	113,700	95,450	18,250	66,928

드를 생성해 낼 수 있었다. 표준 서지 데이타베이스인 BiblioFile 또는 CAT CD450를 이용하여 각 기관별 로칼데이타를 처리한 결과 3가지 유형의 데이타 변환 비율을 발견할 수 있었다. 즉, (1) MARC 형태로 수집한 데이타는 로칼에서 중복 데이타 점검을 완료한 데이타이기 때문에 주관기관에서는 별도의 작업없이 로칼 데이타베이스 생성작업을 할 수 있었으며 중복 데이타 비율 또한 낮으며 (2) 신규로 수집한 데이타를 주관기관에서 소급변환작업한 목록데이타 변환(Hit)율은 약 80-90% 정도로 나타났으며 (3) 소급변환 비율이 높지 않게 나타난 기관의 경우는 이미 구축한 통합 데이타베이스 생성작업시 발생한 미변환(no-hit) 데이타를 포함하여 작업했기 때문에 목표한 비율보다는 저조한 결과가 나타난 것이다.

3.4.2 통합 데이타베이스 생성

통합 데이타베이스를 생성하는 기준은 로칼 데이타베이스 생성기준과 동일하며 시스템은 통합 작업과정에서 태그 001에 통합 데이타베이스의 서지제어번호를 자동으로 생성하여 기록한다. 소재정보를 제공하는 입력필드인 태그 049의 경우에는 서브필드 a만 통합 데이타베이스에 포함시키고 태그 049의 기타 서브필드와 태그 09X 필드의 데이타는 모두 제외시키되 LC 분류기호 정보필드인 태그 050이 존재하지 않으면 태그 090을 분류기호정보로 제공한다. 통합작업 중 기존 데이타와 LCCN, ISBN 등 서지인식번호가 중복될 경우에는 작업중인 데이타의 태그 049에 있는 소장기관명만 기존의 통합 데이타베이스에 포함시킨다. 이와같은 통합 데이타베이스 생성을 위해서는 여러 기관의 로칼별 데이타베이스를 하나의 데이타베이스로 생성하는 단계와 생성된 데이타베이스를 통해 중복 레코드를 점검하는 단계를 수행해야 하는데 각 로칼별 데이타베이스의 도입단계를 수행하기 위해서는 〈그림4〉와 같이 각 기관 로칼 데이타베이스명을 입력하고 MARC 형식을 선택하면 로칼 데이타베이스가 통합되면서 LCCN, ISBN, OCLC 제어번호를 점검키로 하여 데이타 중복점검과 통합 작업이 시작된다.

데이타 도입과정의 실행에 따라 데이타가 통합되면서 통합 데이타베이스에는 서지마스터화일, 미완료데이타화일, 서지인식번호화일이 생성된다. 또한 〈그림5〉와 같은 미완료데이타 처리과정에 따라 데이타가 추가되는 단계를 거치면서 서지마스터화일에서 필요한 색인화일이 작성되고 LCCN, ISBN, OCLC 제어번호 등으로 중복 데이타를 점검하여 발생된 데이타는 에러로 처리하며 화일에 수록한 후 에러 점검작업시 활용하게 된다.

위와같은 중복 데이타 처리기준에 따라 통합작업을 수행한 결과 소재정보가 수록된 총 53,563건의 고유서지(목록)레코드 목록 데이타베이스가 생성되었다.

위의 통합작업과 동일한 방법으로 전년도에 구축한 과학단지장서종합목록데이타베이스의 서지레코드 67,207건과 신규로 구축한 53,563건의 서지레코드를 통합하는 작업을 수행하였던 결과 총99,176건의 표준 MARC 형태의 서

지레코드와 173,173건의 소장기관정보를 수록한 과학기술도서종합목록데이타베이스가 구축되었다.

Ⅳ. 데이타베이스 구축사업 결과 분석

1. 사업 성과

국가연구개발활동의 중심지로 성장해 온 과학단지에 소장되어 있는 과학기술문헌자료를 공동활용함으로써 부족한 정보자원에 대한 갈증을 다소나마 해소하기 위하여 관련기관끼리 소장문헌의 목록데이타에 대한 표준화를 통해 하나의 데이타베이스로 구축하고 구축된 표준 MARC 서지데이타베이스를 활용하여 도서관 온라인공동편목체제를 구축한다는 목표아래 수행된 과학단지장서종합목록데이타베이스 구축사업은 연구단지정보관리협의회 회원기관 중 15개 기관으로부터 서양서 단행본 213,197건의 목록데이타를 수집하여 한국과학기술원 외 6개 기관이 분담처리한 결과 146,333건의 서지데이타를 표준 MARC 포맷으로 소급변환하고 1차로 15개기관 총106,245건의 서지레코드를 갖춘 로칼 데이타베이스를 구축하였으며, 2차로 구축된 로칼 데이타베이스를 통합한 결과 67,207건의 고유서지레코드를 수록한 통합 데이타베이스 구축을 완료함으로써 당초 목표로 한 공동편목에 대한 기본적 틀을 갖추는데 성공하였다.

1차 사업의 성공적 사례와 사업수행기간 중 축적된 기술을 바탕으로 하는 2차 사업은 참여기관의 확대로 총20개 기관으로부터 데이타를 수집하고 수집된 데이타 중 MARC 통신형태로 자료를 송부한 4개 기관을 제외한 전 기관의 데이타는 주관기관인 한국과학기술원에서 일괄 소급변환하는 절차를 거쳤으며 총113,700건의 수집된 데이타를 소급변환한 결과 95,450건의 서지데이타를 로칼 데이타베이스로 구축하고 본 데이타를 통합한 결과 53,563건의 고유서지레코드를 수록하는 종합목록 데이타베이스 구축을 완료하였으며 이 데이타베이스를 전년도에 이미 구축한 67,207건의 과학단지장서종합목록 데이타베이스와 통합하여 최종적으로 총99,176건의 표준 MARC 형태의 고유서지레코드와 173,173건의 소장기관 정보

〈표11〉 데이타베이스 입력 및 통합 데이타 통계

구 분	데이타 수집	변환 데이타	로칼 DB	Union DB
1차 사업	213,197건	146,333건	106,245건	67,207건
2차 사업	113,700건	95,450건	66,928건	53,563건
Union DB	326,897건	241,783건	173,173건	99,176건

* Union DB의 레코드수가 1차, 2차의 합이 아닌 것은 1차와 2차의 통합에 의한 중복 데이타 21,594건이 포함되었기 때문임.

를 수록한 과학기술도서종합목록 데이타베이스를 구축하게 된 것이다. 〈표11〉

2. 기대효과

분담목록을 목표로 한 데이타베이스 구축사업에 관한 한 선진 외국에서 이미 상당 수준에 도달해 있는 반면에 우리나라에서는 늦었기는 하지만 그마나 최초로 시도된 사업인 만큼 사업의 결과에 따라 활용할 수 있는 여지가 많은 것은 분명하다.

2차에 걸친 과학기술도서종합목록데이타베이스구축사업을 수행한 결과를 바탕으로 우리는 다음과 같은 기대효과를 예상할 수 있다. 즉, (1)서지데이타를 표준 MARC 포맷으로 통일함으로써 서지정보의 신뢰도를 향상시킬 수 있으므로 정확한 자료의 검색과 활용이 기대되며 (2)표준화된 서지데이타를 공동활용함으로써 편목작업시 소요되는 인력 및 시간이 최소화되어 자료의 신속한 이용을 기할 수 있으며 (3)각 기관에 분산 소장되어 있는 소재정보를 수록하므로써 필요한 정보자료의 소장위치를 파악하여 자료의 공동활용 및 장서구성의 질적 향상을 극대화할 수 있으며 (4)연구전산망을 기본으로 구축되는 과학기술정보유통체제를 통한 전국민 온라인서비스를 위한 공헌도 또한 클 것으로 기대된다.

V. 발전방향

1. 미변환데이타 처리

본 사업의 수행으로 구축한 과학기술도서종합목록 데이타베이스는 일과성이 될 수 없다. 데이타베이스는 항상 최신성과 신뢰성이 보장되어야 하는데 최신성 보장을 위해서는 계속 입수되는 정보자료를 추가 입력 해야 하며, 신뢰성 보장을 위해서는 사업진행 중 소급변환에 실패한 데이타 즉, 미변환데이타를 완전히 소화해서 입력함으로써 명실상부한 표준 서지데이타베이스로 유지될 수 있는 것이다. 미변환데이타 처리 방안으로서는 첫째, 1차적으로는 사업개시 당시 거론했던 방식 즉, 외국의 전문적인 서지데이타 소급변환기관에 의뢰해서 변환처리 하되 나머지 처리불능 통보 데이타는 자체에서 오리지날 카탈로깅으로 처리하는 방법이 있으며, 두번째 대안으로서는 모든 데이타를 오리지날 카탈로깅으로 처리하되 입력 데이타에 대한 철저한 품질관리로 데이타베이스의 신뢰성을 높이는 방식이다.

위에 제시된 두가지 방법 중 한가지를 택할 때는 방식 자체의 편의성보다는 신뢰성 제고에 기준을 두는 것이 마땅하겠지만 할당되는 예산이 여의치 못할 때는 그 형편이 고려된 합당한 방향이 선택되어야 할 것이다.

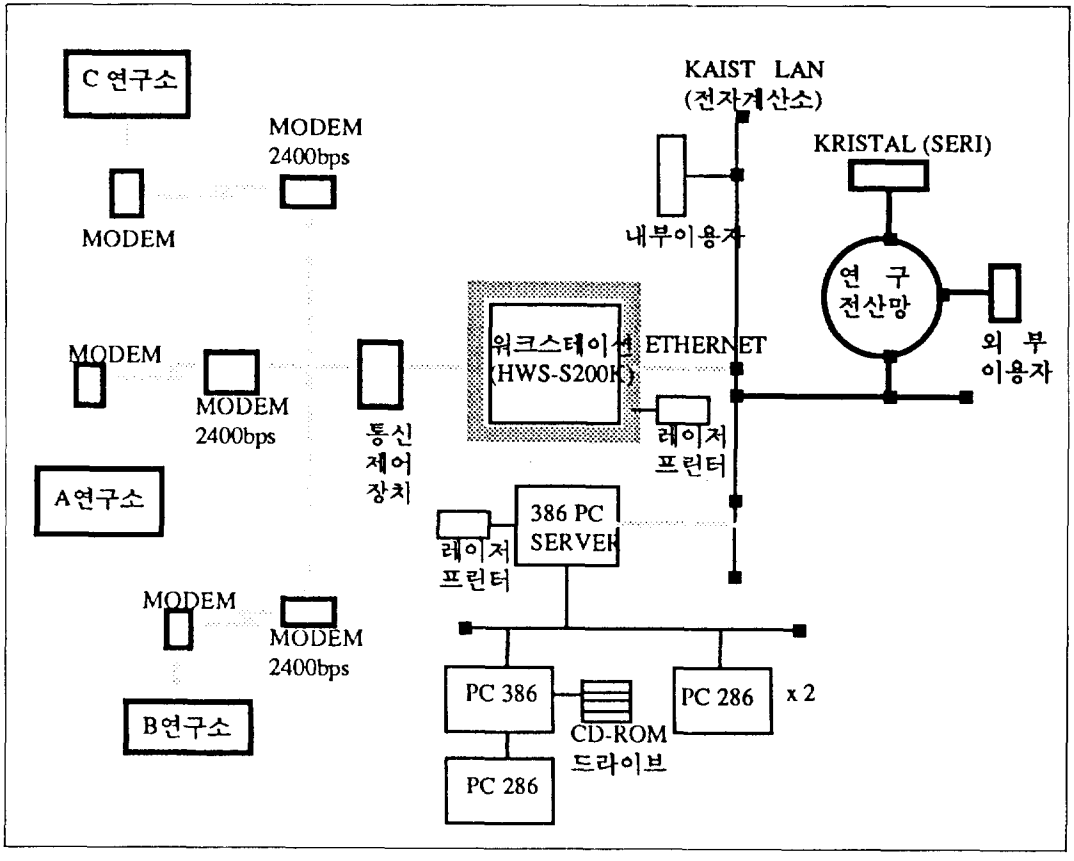
2. 온라인공동편목

공동편목이란 개별기관 단위로 입수되는 문헌정보를 편목할 때 중복 정보때문에 발생될 수 있는 회원기관의 인적, 시간적 낭비요인을 제거하고, 경제적인 방법을 찾자는데 그 목적이 있다. 즉, 목록담당자는 기존의 통합 데이타베이스를 우선 검색하고 목록하고자 하는

데이타가 존재하면 다운로드해서 로칼 데이타베이스로 업로드하여 그대로 활용하고 필요한 데이타가 존재하지 않는 경우에는 회원기관으로서 책임있는 오리지날 카탈로깅을 실행하기 위해서 존재한다 해도 과언이 아닐 것이다.

이와같은 공동의 편의를 위하여 활용할 수 있는 데이타베이스를 구축한 이상 이를 효과적으로 활용할 수 있는 방안이 적극적으로 모색되어야 한다. 2차에 걸친 사업을 수행하는 과정에서는 주관기관은 물론 각 회원기관의 시스템 환경이 매우 열악했던 관계상 인쇄물 형태나 디스켓에 수록하는 정도로 데이타를 수집해서 일괄작업(Batch) 할 수 밖에 없었으나 제2차 사업수행기간 중 확보된 장비를 중심으로 온라인으로 입출력할 수 있는 시스템 구성을 할 수 있었기 때문에 (〈그림11〉 참조) 신년도에 수행될 계속사업기간중에는 온라인 입출력이 가능해 질 것으로 판단된다. 이와같은 온라인 입출력시스템 구축이 성공적으로 되기 위해서는 온라인 입출력시스템을 운용할 수 있는 소프트웨어 개발이 병행되어야 하는 것은 물론이지만 전체 회원기관의 적극적인

〈그림11〉 온라인시스템 구성도



참여도 반드시 요구되는 요인이므로 이를 간과하지 않을 경우 국내 최초의 온라인공동편목시스템이 될 것으로 확신한다.

〈그림11〉에서 보는 바와 같이 중앙통제기관인 한국과학기술원에서는 대용량의 워크스테이션을 호스트로 하고 각 회원기관에서는 2400BPS 정도의 모뎀을 사용하여 데이타를 송부하고 송부된 데이타는 적절한 검증과정과 통합과정을 거쳐서 연구전산망을 통하여 시스템공학연구소 국가과학기술정보유통사업단으로 직송하게 됨으로써 직송된 데이타는 기존 데이타베이스와 함께 전국의 과학기술정보 이용자에게 서비스될 수 있을 것이다.

3. 데이타베이스 수록범위

구축된 과학기술정보관련 데이타베이스를 연구전산망의 전국 네트워을 통해서 전국에 산재되어 있는 과학기술정보 이용자에게 서비스하는 것을 목표로 설립된 시스템공학연구소 과학기술정보유통사업단은 한국과학기술원에서 주관한 과학기술도서종합목록 데이타베이스외에도 한국과학기술연구원(KIST) 외 8개 연구소를 대상으로 총 4억2천5백만원의 예산을 지원하여 〈표12〉와 같은 전문정보 데이타베이스 구축사업을 수행토록 한 바 있다. 그러나 9개 연구소의 전문정보 데이타베이스는 각 연구소에서 수행하는 연구과제에 필요한, 연구소 특성에 맞는 전문연구정보만이 수록되어 있을 뿐 소위 포괄적 의미의 데이타베이스는 아니다. 따라서 과학기술도서종합목록 데이타

〈표12〉 전문정보 데이타베이스 구축현황

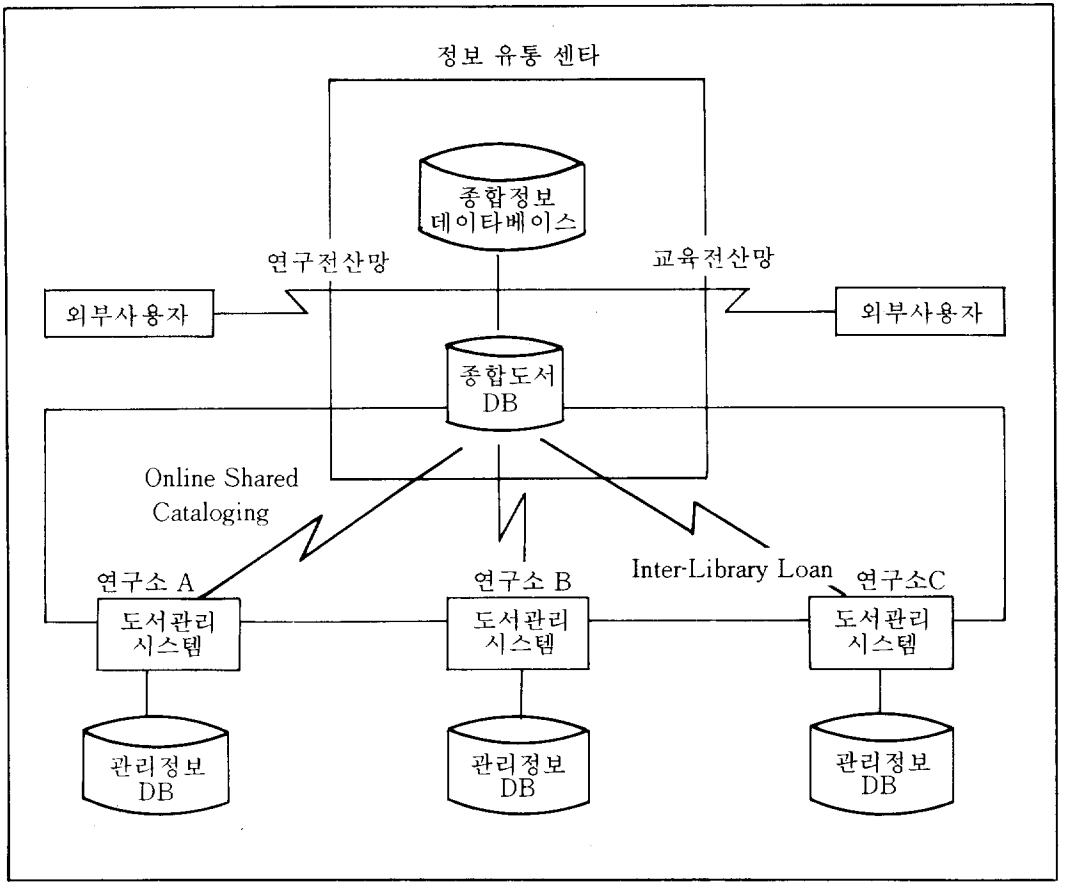
과제수행기관명	과 제 명	예 산
한국과학기술연구원	신소재분야 심층 데이타베이스 구축	4천만원
한국기계연구소	기계전문기술정보 데이타베이스 구축	4천만원
한국동력자원연구소	에너지 분야 전문정보 데이타베이스 구축	7천만원
한국원자력연구소	원자력분야 전문정보 데이타베이스 구축 사업	4천만원
한국전기연구소	전력·전기분야 전문정보 데이타베이스 구축 및 VAN서비스 환경조성	3천만원
한국표준연구소	신소재 열물성, 열역학 및 기계적 특성 데이타베이스 구축	5천5백만원
한국해사기술연구소	해사기술정보 데이타베이스 구축 및 VAN 구축 사업	5천5백만원
한국해양연구소	해양자료 정보서비스 체제 구축	6천5백만원
한국화학연구소	화학공업 데이타베이스 시스템 설계 및 조성에 관한 연구	3천만원
계		4억2천5백만원

베이스와 같이 각 기관에 분산되어 있는 과학기술정보를 효율적으로 활용하기 위한 데이타베이스가 더 많이 구축되어야 하는데 우선적으로 본 데이타베이스가 수록범위를 서양서 단행본에 한정할 수 밖에 없었던 상황을 탈피하여 동양서 단행본도 역시 데이타의 품질관리에 철저를 기하면서 입력되어야 할 것이다. 그런 다음 도서의 형식을 갖추지 못한 과학기술정보자료 즉, 연속간행물, 논문, 기술보고서, 지도, 시청각자료, 마이크로폼 자료 등 모든 형태의 정보자료를 연구단지정보관리협의회 회원기관의 벽을 넘어서 전국적으로 산재해 있는 과학기술연구기관 및 관련기관, 제조업체 등으로 확대해야 할 것이다.

4. 정보서비스

본 데이타베이스에는 정보자료의 소재정보도 수록되어 있기 때문에 불특정 다수의 이용자가 자신에게 필요한 정보의 소재를 확인한

〈그림12〉 종합목록데이타베이스 유통 개념도



후 실제 그 정보를 활용할 수 있는 시스템을 구성해야 한다. 그렇게 하기위해서 우선적으로 구성되어야 되는 시스템은 도서관 네트웍이며 구성된 네트웍을 통하여 이용자가 정보소재지까지 일일이 가지 않고도 필요한 정보를 이용할 수 있는 도서관 상호대차제도의 구축과 활성화가 요구된다. 또한 이용자의 다양하고 전문화된 정보요구에 신속하며 정확하게 대응할 수 있도록 하기 위하여 선택적 정보제공(SDI), 최신정보주지서비스(Current Awareness Service) 등과 같은 향상된 정보서비스 방안의 모색과 추진이 정보유통업무를 담당할 국가과학기술정보유통사업단에서 적극적으로 개발되어야 할 것이다.

위와 같은 목표를 달성하기 위해서 시스템공학연구소에서는 〈그림12〉와 같은 정보유통 기본시안을 마련하고 그에 대한 단계적 실행계획을 추진 중에 있으므로 수준높은 컴퓨터 지식과 문헌정보학 지식이 추진과정에서 잘 접목된다면 기대이상의 성과도 거둘 수 있을 것으로 전망한다.

Ⅵ. 결론 및 제언

우리나라에서 최초로 수행된 도서종합목록 데이타베이스 구축사업은 궁극적으로 본 데이타베이스를 활용한 공동편목(Shared Cataloging)을 목표로 하고 구축된 정보자료의 공동활용을 목적으로 하여 수행되었던 것인 만큼 여기에 거는 기대감이 매우 컸던 것이 사실이다. 선진 외국에서는 이미 이에 대한 기술축적은 물론 실제 활용면에서는 후발국의 정보시장까지 넘보고 있는 싯점에서 국내 기술개발과 기술축적의 측면에서 본다면 매우 고무적이라고 볼 수 있다. 우리의 정보를 남에게 의존하게 된다면 영원히 정보예속국의 처지를 면하기는 어려울 것이기 때문에 이를 극복하는 관점에서 보아준다면 더할나위 없이 바람직한 사업이었다고 평가될 수 있을 것이다.

이와같은 긍정적인 여러 요인들을 생각할 때 다음과 같은 사항이 강조됨으로써 더욱 훌륭한, 더욱 발전적인 모습의 데이타베이스로써 활용될 수 있을 것이다.

첫째, 데이타베이스 수록범위의 대폭 확대를 통하여 진정한 의미의 과학기술정보통합데이타베이스로써 성장시켜야 할 것이다. 둘째, 구축된 데이타베이스를 기반으로 한 정보서비스 체제를 더욱 다양하게 구성함으로써 정보이용자들에게 필수불가결한 정보원으로써 활용되어야 할 것이다. 셋째, 종합목록데이타베이스는 여타 전문정보 데이타베이스와는 달리 서지정보 데이타베이스이기 때문에 참여기관 모두가 당사자 이해관계를 떠나서 자기의 일로 생각하는 풍토가 정착되어야 할 것이다.

이상과 같은 요인을 생각하면서 참여자 모두가 힘을 한데 모아서 추진할 때 진정한 정보화 사회의 밑거름으로서 국가발전에 기여할 수 있으리라고 본다.