

온라인 업로드 기능을 이용한 목록정보 소급 입력

: 도쿄대학 부속 도서관에 있어서의 시험 가동

竹内比呂也, 石井保廣, 谷澤滋生 述

〈서울대 도서관〉 曹順英 譯

도서관의 전산화가 활발해지면서 각 도서관에서는 기정리본 목록의 소급 입력 문제에 큰 관심을 갖게 되었다. 이에 日本의 『大學 圖書館 研究』 第36輯 (1990年 7月)에 게재된 『オンライン アップロード 機能を用いた 目録 情報 遡及 入力: 東京 大學 附屬 圖書館における 試行』 (竹内比呂也, 石井 保廣, 谷澤 滋生 述)을 번역하여 효율적인 DB구축에 다소나마 도움이 되고자 한다.

I. 머리말

토쿄대학 부속 도서관에서는 1986년부터 3년간에 걸쳐 과학 연구비 보조금 『연구성과 공개 촉진비』를 받아 『학술 문헌 소급 정보 검색 데이터베이스 (Scientific Literature Information Database)』라 일컫는 도서목록 정보 대상의 소급 변환/입력 시험 가동 프로젝트 (이하 [SLIDE]라고 한다)를 학술 정보 센터와 공동으로 실시하였다.

이 프로젝트에 앞서 토쿄대학에서는 도서관 업무 전산화 시스템 (Local Integrated Library Information System Project, University of Tokyo 이하 『LILIPUT』이라고 한다)의 구축을 추진하고 있었고, 1986년 3월에 학술정보 센터와 네트워크를 매개로 접속하여, 同센터의 온라인 목록 소재 정보 서비스 (이하 『NACSIS-CAT』라고 한다)와 연동한 목록 시스템을 同年 5월말부터 운용하기 시작하였다.

그후 LILIPUT 시스템 개발의 진전, 목록 단말기의 증설 및 목록업무의

학내에서의 전개에 따라 점차 신규로 입력되는 목록 데이터가 증가하고, 온라인 이용자용 목록(이하 『OPAC』라 한다)에 의한 목록 정보의 제공도 시작하게 되었다. 목록 정보의 입력 및 그 제공 수단이 충실하고 안정됨에 따라 OPAC의 급속한 충실화, 즉 카드 목록으로 제공되고 있는 전산화 이전 목록정보의 OPAC에 의한 제공이 강력히 요청되기에 이르렀다.

이와 같은 학내의 사정에 따라 NACSIS 종합목록 데이터 베이스의 축적도 충분치 않고, 신규 입력의 비율이 높은 상황에서 소급 사업의 시행 가능성을 타진하여 착수하게 된 것이 SLIDE이다. 본고에서는 이 프로젝트 중에서 검토된 소급 변환/입력의 방법에 대해 개관하고, 특히 본 프로젝트에 있어서 최초로 시도되었다고 생각되는 온라인 업로드 방식의 소급 입력에 의한, 목록 데이터 베이스 형식의 시행에 관해 보고하고자 한다.

II. 소급 변환/입력

전산화한 목록시스템에 있어서 목록 정보의 작성, 데이터 베이스의 형식 방법으로는 다음의 세가지 유형을 들 수 있다. 즉

- (1) 온라인 입력
- (2) 일괄 (batch) 변환/일괄 입력
- (3) 일괄 변환/온라인 입력이다

또한 본고에서는 소급을 위한 『변환』과 『입력』이라는 단어에 대해 전자는 『수작업으로 작성한 파일과 기계가 읽지 못하는 파일의 기록을 기계가독형으로의 변환』을 의미하고, 후자는 『정보의, 기존의 데이터 베이스에의 입력(통합)』을 의미하고 있다. 이하 각각의 방법에 대해 특징을 서술하고자 한다.

1. 온라인 입력

목록 담당자는 목록 시스템이 제공하는 입력 화면에, 목록정보를 키보드

등으로 입력해서, 데이터 베이스에 한 건씩 등록하여 종합목록을 형성하는 방법으로는 NACSIS-CAT가 채용하고 있는 것과 같이 가장 일반적인 방법이다.

이 방법은 ①특히 신규 시스템 개발이 필요 없는 점 ②이미 데이터 베이스중에 존재하고 있는 데이터와의 정합성을 유지할 수 있는 점 ③이미 데이터 베이스중에 존재 하고 있는 데이터의 활용이 가능하다는 점 (예컨대 목록의 복사(copy cataloging)등의 장점이 있는 반면 ①작업환경 (단말기대수, 응답시간, 입력시스템의 서비스 시간등)의 영향을 받기 쉬운점 ②직접 키보드로 입력 하기 때문에 단순한 입력 미스가 발생하기 쉽고, 그 때문에 체크 기능을 두어야 하는 점등의 단점을 갖고 있다.

2. 일괄 변환/일괄 입력

기존의 목록 정보를 근거로 필요한 데이터를 일괄하여 기계가독화 한 후, 당해 데이터를 데이터 베이스에 일괄 입력하는 방법이다. 이 방법에서는 ①비교적 단시간에 대량의 데이터를 기계가독화하고, 데이터 베이스로써 이용자에게 제공할 수 있는 점 ②온라인 입력에 비해 작업환경에 좌우되는 일이 적고, 또한 온라인 입력에 해당하는 작업이 기계적으로 처리되기 때문에 목록 담당자의 부담도 경감된다는 점이 장점인 반면 ①데이터 베이스 가운데 기존 데이터가 있는 경우 신규 입력 데이터와의 정합성을 유지하는 수단이 확립 되어있지 않은점 ②일괄 변환한 기계가독 데이터에 관해서도 데이터간의 정합성을 갖는일이 어려운 점등의 단점을 지니고 있다. 이 방법이 현재 유효하다고 할 수 있는 것은 완전하게 정합성을 유지하고 있는 목록정보를 완결한 것으로서 기계가독화한 후에 신규로 데이터 베이스에 (기존 데이터가 데이터 베이스에 존재하지 않는 상대하에서) 입력하는 경우에 한한다.

3. 일괄 변환/온라인 입력 (온라인 업로드)

이 방법은 입력대상 데이터를 미리 일괄해서 기계가독화한 후 온라인

입력 작업시에 그 작업전개에 따라 필요한 데이터를 온라인 입력화면에 호출해서 (업로드해서) 이용하는 것이다. 이것은 이미 서술한 두가지 방법의 단점을 보완한 것이다. 즉 온라인 입력에 의해, 일괄 입력의 단점이었던 데이터 베이스중 기존 데이터와의 정합성을 유지할 수 있는 한편 미리 일괄해서 목록정보를 기계 가독화함으로써 입력작업의 효율화 및 사전에 데이터 체크기 가능하여 데이터의 신뢰성 향상을 꾀할 수 있게된다. 그러나 이 방법이 가능하기 위해서는 온라인 화면상에 기계 가독화된 데이터를 호출하는 기능(업로드 기능)을 갖고 있어야 한다.

4. NACSIS-CAT/LILIPUT 환경에 있어서 소급변환/입력의 방법

전술한 일반적 소급 변환/입력의 방법 및 그 특징을 기초로, NACSIS-CAT/LILIPUT이라는 환경 아래서의 소급 변환/입력의 방법을 고찰하기에 앞서, 그 전제가 되는 LILIPUT의 목록 시스템에 대해서도 약간의 설명을 하고자 한다. (자세한 것은 이미 발표된 문헌을 참조하기 바람)

LILIPUT은 독자적인 목록작성 시스템을 갖지 않고, NACSIS-CAT와 로컬·데이터 베이스인 토쿄대학 종합목록 데이터 베이스 (이하 『UTLDB』라고 한다) 및 OPAC를 연동하는 것으로 목록시스템을 구성하고 있다. 결국 기본적으로는 NACSIS-CAT 상에서 목록 데이터의 작성, 수정, 추가, 삭제 등이 모두 행해지는 것이다.

* 직접 형성된 데이터 베이스는 NACSIS 종합목록 데이터 베이스이지만, 다운로드에 의해 UTLDB가 형성되고 또 UTLDB를 근거로 OPAC가 작성된다.

따라서 소급 변환/입력의 방법으로서 NACSIS-CAT에 의한 온라인 입력이라든가 NACSIS, UTLDB, OPAC 각 데이터 베이스에로의 일괄 입력이 고려될 수 있다. 그러나 NACSIS 종합목록 데이터 베이스로 참가관에서의 일괄 입력은 허용되지 않기 때문에 다음의 4가지 방법이 검토 대상으로 남아 있다.

- ① 일괄 변환한 소급 데이터를 OPAC로 일괄 격납한다.
 - ② 일괄 변환한 소급 데이터를 UTLDB로 일괄 격납한후, OPAC를 갱신하는 것으로 일괄 격납 데이터를 OPAC에 반영한다.
 - ③ NACSIS-CAT에 의해 온라인 입력하고, 다운로드에 의해 UTLDB, OPAC을 갱신한다.
 - ④ 온라인·업로드 기능을 이용해 NACSIS-CAT로부터 온라인 입력하고, 다운로드에 의해 UTLDB, OPAC을 갱신한다.
- 각 방법의 특징을 表1에 표시하고 있다.

〈表1〉 LILIPUT에 있어서 소급 변환/입력의 비교

	장 점	단 점	
①OPAC로의 직접 격납	· 전체의 작업기간을 단축할 수 있어, OPAC에 의한 목록정보를 신속히 제공할 수 있음.	· OPAC에 대해 직접 데이터를 수정/삭제 할 수 없음.	· 기존 데이터, 타학부 도서관(실) 입력 데이터와의 정합성에 대해 보증할 수 없음.
②UTLDB를 경유해 OPAC로 격납	· 온라인 입력으로 단기간 작업이 이뤄지고 OPAC에 의해 목록정보를 신속히 제공할 수 있음. · UTLDB의 기능으로 약간의 로컬(local)데이터 추가 수정 및 서지/소장의 삭제가 가능하며 그 결과를 OPAC에 반영할 수 있음	· UTLDB 에의 일괄격납은 시스템 부담이 커지고, 다른 업무에 영향이 큼.	· 일괄 격납의 시점에서는 임시로 서지 등록 하고, NC로의 온라인 입력에서 정규의 서지와 치환할 필요가 있음.
③NACSIS-CAT 로의 온라인 입력	· 입력 방법이 통상의 목록작업과 소급 입력으로 일원화 되기 때문에, 기존 데이터, 타학부 도서관(실)등으로 부터의 입력과 경합/중복되는 일을 피할 수 있음.		· 작업 기간이 길어지고, 지속적인 인력의 확보, 비용의 부담이 큼.
④NACSIS-CAT 로의 온라인 입력(업로드 기능 사용)	· 업로드 기능의 작업 지원으로 입력부담을 경감할 수 있음.		· 작업 기간이 길어지고, 지속적인 인력의 확보, 비용의 부담이 큼.

①②는 기계가독화된 데이터를 입력하는 데이터 베이스에 적합하도록 포맷을 조정후 통상 각 데이터 베이스를 갱신하는 경로로 UTLDB 혹은 OPAC에 격납하는 것이다. 이 방법의 최대의 문제점은 데이터 베이스의 기존 데이터와 일괄 격납 데이터와의 정합성이 보증되지 않는다는 점이다. UTLDB, OPAC 모두다 종합목록으로서의 성격을 갖고, 중복서지를 인정하지 않는 데이터 베이스임에도 불구하고, 일괄 입력시 중복 체크 기능은 없다.

또 앞서 서술한 바와 같이 LILIPUT은 데이터의 수정, 추가, 삭제를 NACSIS-CAT상에서 실시하는 기본설계이기 때문에 수정 등을 할 필요가 있을 때는 결국 NACSIS-CAT에서 할 수 밖에 없다. 또 최근에 작성된 목록에 대해서 보증되어 있는 NACSIS 종합목록 데이터 베이스와 로컬 데이터 베이스와의 정합성을 유지 할 수도 없게 된다.

③은 ①, ②에 대한 문제점을 모두 해결할 수 있는 것으로 2.1에서 서술한 문제에 모두 적합하다. SLIDE에 있어서도 1986년 이래 이 방법으로 입력하고 있지만 효율적인 면을 고려해서 기본적으로 학술 정보 센터의 파일 중에 서지 레코드가 존재하는 것에 한정하고 있다.

④는 ①-③에 대한 문제점은 모두 해결 되지만 온라인 입력 화면에 기계가독화된 데이터를 호출하는 기능이 문제가 된다. LILIPUT에 있어서는 이미 입수 처리의 과정에서 생성된 도서 정보(서지 데이터의 일부, 그것을 가공한 검색키 및 소장 데이터)를 NACSIS-CAT에서의 목록 작업시 온라인 화면상에 호출하여(업로드)담당자의 작업 경감에 이용하고 있다. SLIDE에서는 기계가독화한 소급 목록 데이터를 NACSIS-CAT/LILIPUT 양 시스템에 주고 받을때 파일에 상당하는 업로드 테이블에 등록하므로써 온라인 화면에의 업로드를 가능하게 한다.

이상의 고찰에서 최선책이라 생각되는 업로드 기능을 활용한 온라인 입력이 NACSIS-CAT/LILIPUT 이라는 환경에서 가능하며, 1987년부터 시험적으로 채용하기에 이른것이다.

Ⅲ 온라인·업로드 방식에 의한 소급변환/입력의 실제

1. SLIDE의 작업개요

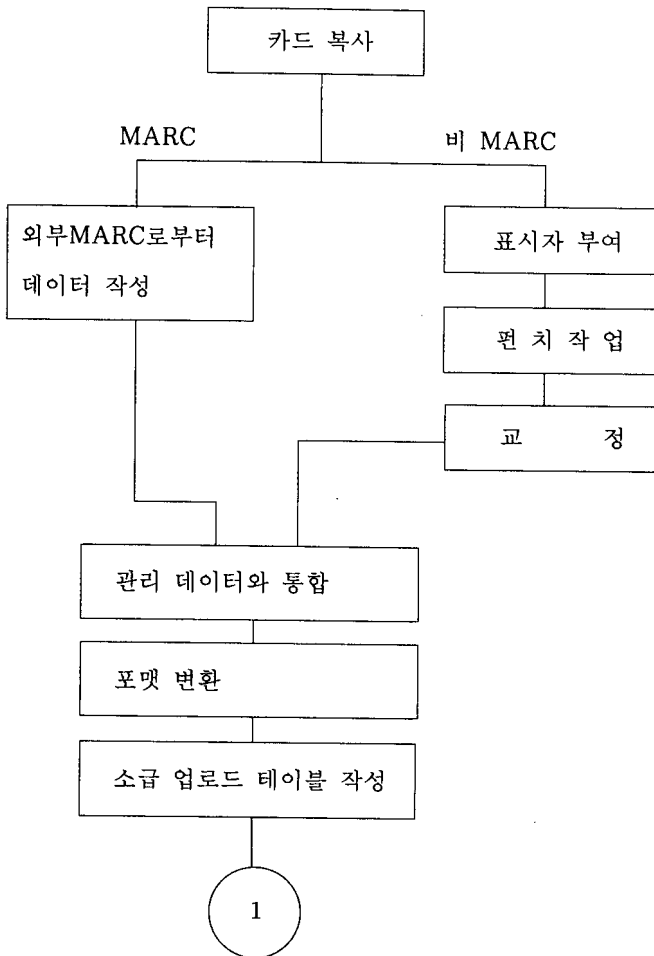
다음의 圖1은 작업의 흐름을 나타낸다. 이 작업은 업로드.테이블 작성까지의 데이터 변환단계와 NACSIS-CAT상에서의 입력작업단계로 대별된다. 데이터 변환 단계에서는 우선 SLIDE에서 대상으로 하는 목록 카드(사무용 서가 카드를 사용)를 이후의 작업에서 사용할 원고로 쓰기 위하여 복사하였다. 그리고 MARC (JPMARC 또는 LCMARC)와의 상호 참조를 행하고 해당 데이터가 MARC로 되어 있는 것과 독자적인 새 데이터의 작성이 필요한 것으로 나누었다. 독자적인 목록의 작성분은 복사한 목록(copy cataloging) 원고에 대해 NACSIS-CAT에서의 온라인 화면의 각 입력 필드에 카드상의 각 기술이 대응하는 형태로 표시자(tag)를 부여 하였다. 표시자 부여가 끝난것은 외부에 위탁하여 기계가독 형식의 데이터로 자기 테이프에 받았다. MARC로 되어있던 데이터 역시 자기 테이프를 받았지만 NACSIS-CAT에 대응하도록 데이터를 조정 하였다. 이와 같이 외부위탁에 의해 준비된 기계가독 데이터를 데이터 체크후 업로드.테이블에 등록하였다.

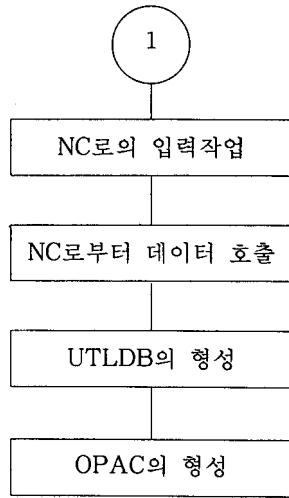
NACSIS-CAT상에서의 입력 작업은 온라인·업로드 기능을 이용해서 온라인 화면상에 업로드 데이터를 업로드 테이블로부터 호출하여 서지, 소장 데이터를 작성했다. 여기서 작성된 목록 데이터는 학술정보 센터가 제공하게되고, 도쿄 대학의 개별판 자기 테이프에 위해, 통상의 목록업무에서 작성한 데이터와 함께 LILIPUT에 합쳐져서 UTLDB, OPAC로 격납 되는 것이다.

2. 데이터 변환

데이터 변환이 최대한 완전해야 됨은 말할 것도 없다. 업로드 기능을 이용한 입력작업에서는 작업의 효율성 및 품질이 업로드하는 데이터에 달려 있다고 할 수 있기 때문에 가능한한 정확한 데이터를 준비하도록 노력하

였다. 즉 사전에 변환 하고있는 데이터의 신뢰성이 높으면 목록작업에서 입력 담당자는 검색, 키, 일서 입력에서 가나 한자 변환, 그 읽기 등 키보드의 입력에서 많은 부분을 업로드 기능으로 대체 할 수 있기 때문에 목록작업시 본래의 목록작업에는 없는 이들 작업에 주의를 기울일 필요가 없게 되고 서지의 동정, 서지계층의 표현등에 의해 집중할 수가 있다. 이 같은 이유로 데이터 베이스의 일관성 유지면에서도 업로드 기능을 이용할 수 있을 것이다. 구체적으로는 다음과 같은 작업을 수행하였다.





〈圖1 SLIDE작업의 흐름도〉

(1) 원고의 준비

소급변환의 정보원으로서 목록 카드를 사용하였다.

목록 카드 중에서는 작업대상 도서와의 대응이 가장 명확하다는 이유에서 사무용 서가카드를 채택하였고, 외주업자에게의 지시, 서지 기술에의 표시자 부여등 써넣을 스페이스 확보와 카드의 점유기간을 최소화하고 통상의 업무에 지장이 없도록 용지에 복사를 해서 작업용 원고로 사용하였다.

(2) 데이터의 작성 및 교정

작업용 원고를 근거로 데이터를 기계가독화하는 작업을 외주했지만, 전술한것과 같이 일단 JPMARC(또는 LCMARC)를 검색하여 해당하는 서지데이터가 MARC상에 채록되어있는 것은 MARC 데이터를 새로 작성하는것과 같은 데이터 형태로 포맷을 변환시킨 후 자기테이프에 받았다. 그때 MARC에서의 추출분이 새로 작성한 데이터와 정합성을 유지하도록 포

맷 변환시 각 데이터 구성요소에 표시자를 부여하였다. 이렇게 하면 새로 작성하는 작업이 필요없게될 뿐만 아니라 원고로 사용하는 카드상에는 별로 기재되어 있지 않은 ISBN, NBN등을 포함하고 있어 질이 좋은 데이터를 얻을 수 있고 이 데이터를 검색키로해서, NACSIS-CAT상에서 효율이 높은 검색이 가능하게 되고 또 MARC 데이터의 일괄 편집에 의한 납기 단축이 가능하게 된다.

또 MARC에 해당 데이터가 없는 것은 새로 데이터를 작성하기 위해 복사원고를 근거로 업자에게 데이터 작성을 의주했는데, 그 준비로서 원고에 표시자 등을 부여했다. 이 작업은 각 서지 데이터의 데이터 작성시에 NACSIS-CAT 온라인 화면의 각 데이터필드와 대응하도록 작성 데이터의 구성 요소에 대해 필드 명칭을 탭(tab)으로 부여함과 동시에 필요한 구분 기호를 넣고, 또 일서에 있어서의 특별한 읽기도 부여하였다. 작업은 직원의 지도아래 아르바이트생들이 하였다. 카드상의 변환대상이 되는 목록 데이터에는 모두 표시자를 부여해서 업자에게 넘겨 준후 기계가독데이터로 변환한 것을 자기 테이프에 받았다. 변환 데이터는 도서관의 전산시스템의 단말기에서 온라인 편집이 가능한 형태로 보내짐과 동시에 교정용 리스트를 작성해 육안으로 체크하고 불완전한 것·내용에 문제가 있는것·틀린 것 등은 적당한 원고 또는 원자료를 참조하여 수정하므로써 보다 완전한 데이터를 만든다.

실제로 데이터 수정은 도서관내에 단말기에서 편집 기능을 이용해 직접하므로써, 기계가독 형식으로 들어간 데이터에 대해서는 다시 업자에게 수정을 의뢰하는 시간의 낭비를 줄였다.

3. 포맷의 변환

LILIPUT시스템 내부에서는 각 업무 모듈(module)간의 연락, 데이터 교환을 위해 학술정보센터가 규정하고 있는 서지교환 포맷(B 포맷)에 준하는 데이터 교환형식을 사용하고 있다. 기계가독화한 목록 데이터는 일단 이 교환 형식으로 변환 해야 LILIPUT 시스템에서의 취급이 가능하다.

LILIPUT에서는 외주 데이터로부터의 포맷 변환 기능을 갖고 있지 않기

때문에 이번 SLIDE에서는 기계가독화한 목록 데이터를 교환형식으로 교환하는 프로그램을 독자적으로 작성했다.

4. 업로드·테이블의 준비

개개의 서지 데이터 구성요소(NACSIS-CAT에서의 데이터·필드) 각각을 가변장 레코드로 하고, 그 상위에 논리 서지 단위를 갖는 교환형식의 포맷에 대해, 업로드 테이블은 도서의 물리적 단위가 1레코드가 되는 가변장 데이터 형식을 취하고 있다. 또 각 레코드는 자료 ID(등록번호)를 우선키(주접근키)로 하는 직접 접근 가능한 파일 형식(VSAM)을 취하고 있다.

업로드·테이블을 작성하는 절차는 검수를 필한 도서의 데이터에 대해서 목록 작업을 할 때 입력 작업을 줄이기 위해 사용하는 업로드·테이블을 준비하는 절차의 일부분을 유용(流用)하고 있다. SLIDE에서는 교환형식의 포맷으로 한 데이터를 이 업로드·테이블 작성 경로로 보냄으로써 업로드·테이블에 등록하였다.

단 소급작업과 통상의 목록 작업에서 같은 업로드·테이블을 사용하는 것은 혼란을 가져올 우려가 있어서 각각 다른 업로드·테이블을 준비하였다.

5. 온라인 입력 작업

NACSIS-CAT상에서의 온라인 입력작업은 파트타이머를 이용하였다. 작업자에 대해서는 약2주간의 강의 및 NACSIS-CAT교육 모드에 의한 교육 훈련을 시켰다. LILIPUT에서는 필요에 따라서 교열 리스트가 작성되기 때문에 이것을 직원이 체크하고 문제가 생기면 작업자에게 피드백을 보낸후 데이터 수정을 하였다.

작업용 자료로서, 복사원고, 외주 데이터로부터 작성한 리스트 및 열람관리용의 초기데이터로 격납되고 있는 UTLDB로부터 출력한 관리 데이터의 리스트를 사용했지만 목록의 질적인 면을 고려하여 필요한 경우에

는 실제 실물을 확인하도록 지도하였다.

실제의 입력 작업은 검색 화면에 있어서 검색키의 업로드로부터 시작되며 그 개요는 다음과 같다.

① 검색화면에 있어서 'UL(ULOAD)' 명령후 자료 ID를 입력하고 송신하면 업로드 테이블로부터 검색용 데이터가 호출되어 화면에 표시된다(圖2-①). 이 키를 기본으로 검색을 하고 NC 파일, 참조 MARC 파일에 해당서지가 있으면 그 데이터를 유용하기 때문에 업로드.테이블중의 서지 데이터는 이용되지 않는다.

② 해당서지가 없는 경우 'Create' 명령으로 서지 신규 작성 화면으로 보낸후, 다시 UL 명령과 자료 ID를 입력함으로써 서지 데이터가 업로드-테이블로부터 화면상에 호출된다(圖2-②)

그 다음은 서지를 이용할 때와 같이 불필요한 데이터의 삭제, 부족한 데이터의 추가, 링크 부여 등의 작업을 하고, 서지의 입력이 종료한 후 'Register' 명령의 실행에 의해 소장입력 화면으로 이행한다.

③ 소장화면에 있어서도 마찬가지로 'UL' 명령과 자료 ID 를 입력해서 업로드·테이블로부터 해당 데이터의 소장 정보를 호출하여 화면상에 표시한다(圖2-③)

이 소장 정보의 업로드는 입력의 수고를 줄여 줄 뿐만 아니라 코드, 기호등 긴스트링입력시 발생하는 오타 (typing miss)를 크게 줄여 줄 수 있다.

和圖書書誌檢索. 簡略標示 (教育用)

> : ULOAD 0000520163

TITLE =

AUTH =

AKEY : 日の著 ISBN : NBN : NOLCN :

PUB = YEAR :

PLACE = CNTRY : LANG :

SH =

WORDS =

ID : PID :

FILE :

↓ ① 명령어 입력

> :
 <BN01827332
 GMD : SMD : YEAR : CNTRY : ja TILL : TXTL : ORCL : REPRO :
 VOL : ISBN : PRICE : XISBN :
 ISSN : NBN : NDLCN : GPON :
 OTHN :
 TR : 日本の名著 : 近代の思想 / 桑原武夫 編 || ニホンノメイチョ : キン
 ダイノシソウ
 ED :
 PUB : 東京 : 中央公論社, 1962
 PHYS : 3,310p; 18cm.
 VT :
 CW :
 NOTE :
 PTBL : 中公新書 チュウユウ シンショ<>1
 AL : 桑原武夫 || クワバラ, タケオ
 UTL :
 CLS :
 SH :

↓ ② 호출 화면 (서지)

和圖書所藏新規入力 (教育用) (BN01827332)
 > : ULOAD 0000520163
 <BN01827332> 日本の名著 : 近代の思想 / 桑原武夫 編.--
 (中公新書 ; 1)
 <FA001787>
 <CD0003388187>
 LOC : 開架
 VOL : CLN : A10;858 RGIN : 000520163 CPYR :
 LDF : QFF=RC87G
 LTR :

③ 호출 화면 (소장)

<圖 2> 온라인·업로드의 흐름

IV. 결과 및 그 평가

1. 입력결과

이번 시험가동 결과 얻어진 실적을 表2에 종합하였다. 전체적으로 9,404건의 서지와 11,953건의 소장 데이터가 입력되었다. 이것은 파트타이머 3인의 약 5개월에 걸친성과이다. (트레이닝기간 포함)

또 11월 및 12월은 NACSIS-CAT로 종합목록 데이터베이스 또는 참조 MARC에 해당 서지가 있는 것을 작업대상으로 하고 실제로 업로드 기능을 이용한 새로운 서지의 입력을 시작한 것은 1월부터이다.

表2의 실적에 관해서 간단한 분석을 한 것이 表3이다. 이것은 실제 작업에 소요된 일수를 구해 입력 실적으로부터 서지, 소장 각각의 1일당 입력 건수를 나타낸 것이다. 이 표로부터 다음과 같은 점이 명확해졌다.

① 작업자의 목록작업, 시스템에 대한 숙련 등으로 잠시 입력건수가 증가 경향에 있는데 그 중에서도 특히 업로드 기능의 이용이 가능해지고부터 눈에 띄게 입력 건수가 증가함을 알 수 있다. 이것은 업로드 기능이 예상대로 효과를 거둔 것이라고 할 수 있을 것이다.

② 2월의 실적을 보면 전입력 건수에 대해 3% 정도의 신규 입력을 시험적으로 했는데 다른 달에 비해 서지, 소장 모두 입력 건 수가 떨어짐을 볼 수 있다. 작업자의 신규입력에 대한 숙련도와도 상관이 있겠지만 일반적으로 입력 작업에 있어서 신규 입력이 작업자에게 주는 부담은 상당히 큰 것이라고 생각된다.

③ 3월은 서지 데이터의 업로드 비율이 낮음에도 불구하고 서지, 소장 입력 모두 많은 것은 서지 데이터만이 아니라 검색키 및 소장 데이터의 업로드가 유효하게 가동되었기 때문이라고 생각된다.

또 업로드 기능에는 앞서 말한 바와 같이 직접적인 효과 외에 아래와 같은 효과도 지적할 수 있다.

① 단순한 오타(타이프 미스)가 거의 없게 되고, NACSIS-CAT에의 입력 후 직원에 의한 데이터 체크 및 입력 작업자에 의한 재입력(수정)이라

는 부담이 현저하게 감소했다.

② 입력 작업자가 단순한 키 입력 작업에만 주의하는 것이 아니라 본래의 목록 작업에 있어서의 고찰, 판단, 서지 동정, 계층 표현 등을 집중적으로 할 수 있게 되어 입력 데이터에 대한 신뢰성 향상을 기대할 수 있게 되었다.

③ 작업 공정을 데이터의 준비, 기계 가동화, 교정, 업로드 테이블 작성, NACSIS-CAT에 의한 입력 등 각 단계마다 분할 할 수 있기 때문에 작업 일정, 요원 배치 등의 계획성과 생산성이 향상되었다.

〈表2〉 업로드 기능을 이용한 소급입력의 실적

(작업담당자 : 3人/日)

년월	서지입력건수					소장입력 건 수	특기사항
	NC	참조 MARC	업로드	신규 입력	계		
'87. 11	623	529			1,152	1,510	11/9 작업 개 시
'87. 12	825	887			1,712	2,232	
'88. 1	449	1,030	1,155	1	2,635	3,441	당월부터 업 로드
'88. 2	324	220	2,116	82	2,742	3,120	
'88. 3	480	634	249		1,363	1,650	3/12 작업 종 료
합계	2,701	3,300	3,520	83	9,604	11,953	

※ NC : 전국 종합목록 데이터 베이스의 서지 데이터 이용

참조MARC : 참조MARC로부터 서지데이터이용

〈表3〉 소급입력실적의 효율분석

年 月	월간작업일수			서적데이터업 로드비율(%)	서지 * 1 작업일수	소장 * 2 작업일수
	전일	반일	계			
'87. 11	12	6	15.0	—	76.8	100.6
'87. 12	15	7	18.5	—	92.5	120.6
'88. 1	14	8	18.0	43.8	146.4	191.2
'88. 2	16	7	19.5	77.2	140.6	160.0
'88. 3	6	2	7.0	18.3	194.7	235.7

주) * 1 : 월간서지 입력건수 ÷ 월간작업일수

* 2 : 월간소장 입력건수 ÷ 월간작업일수

2. 문제점

온라인 업로드 기능을 이용한 SLIDE작업에서 다음과 같은 문제점이 드러났다.

(1) 변환 정보원의 문제

재래의 목록 카드를 소급 작업의 원고로 했지만, 카드 목록에서 변환시킨 데이터에 있어서도 대개 NACSIS-CAT에서 얻을 수 있는 필수항목을 채우고 있다. 그러나 내용적으로는 목록 카드에 적용되고 있는 목록 규칙과 NACSIS-CAT에서 입력 기준 (및 기초가 되는 목록 규칙)과의 차이 때문에 발생하는 문제가 있다.

구체적으로 출판년의 취급방법을 예를 들면 NACSIS-CAT에서는 동일판에 있어서는 최초의 인쇄 년도를 채용하고 있는 반면 목록 카드에서는 소장 자료의 출판 년도를 채용하고 있다. 이같은 문제 때문에 NACSIS-CAT 입력 후 데이터 체크용으로 출력된 교열 리스트를 보고 직원이 작업자에게 실물 대조를 지시하여 필요한 경우에 데이터 수정을 하도록 하였다.

또 계층 관계를 갖는 도서는 NACSIS-CAT에서는 어미(親)서지를 독립된 서지로 다룰 필요가 있겠지만 카드에는 어미서지의 목록 정보가 충분히 기술되어 있지 않기 때문에 신규로 계층 관계를 갖는 목록을 작성할 경우 자주 원자료를 참고할 필요가 생긴 것이다.

(2) 변환시의 문제

기계가독화로 변환할 때 카드를 복사한 원고에 표시자를 부여했는데 그 작업시간이 상당히 많이 걸려 아직 개선의 여지가 남아 있다고 할 수 있다. 또 기계가독화데이터에 대한 교정 작업도 데이터의 질을 유지하기 위해서는 불가결한 것으로 이것도 시간을 요하는 작업이었다.

(3) 업로드 기능의 문제

업로드 기능은 NACSIS-CAT가 제공하는 각 작업 화면 중 초기 상태를 상정해 프로그램이 준비되어 있기 때문에 실제 업로드에서는 초기 화면에 준비되어 있지 않은 데이터 항목의 반복 (복수 저자의 경우 두번째 이하의 AL 필드등)은 한 건만 유효하다. 또 각 필드에 정해진 문자 수를 초과

한 데이터에서는 초과분이 업로드 되지 않는다. 이같은 데이터에서는 작업자가 다시 입력해야만 한다. 단 이들 반복되는 항목, 초기 필드 길이를 넘는 데이터에 관해서도 업로드 테이블에서는 정규 데이터로서 유지할 수 있기 때문에 업로드 기능의 확장으로 그 문제가 해결 되리라 생각된다.

V. 맺음말

NACSIS-CAT/LILIPUT이라는 환경에 있어서 온라인 업로드 방식에 의한 소급 입력을 시험가동 했는데 이로인해 온라인 업로드 방식을 통상의 온라인 입력 방식과 비교할 때 온라인 입력 방식이 월등히 효율이 좋은 방식이라는 것이 명확해 졌다.

또 NACSIS-CAT로부터 토쿄대 로컬 데이터 베이스라는 최근에 작성된 목록 데이터와 완전히 같은 경로로 입력되기 때문에 각 데이터베이스에 있어서도 정합성이 유지된다.

데이터 베이스의 정합성을 유지하면서 일괄 입력하는 방법이 없는 상황 아래서는 제한된 수의 단말기를 사용해서 온라인 입력하는 방법 뿐이 없기 때문에, 이같은 조건 아래서 신속하게 소급 입력을 하려면 1 레코드당 입력 부담은 (load) 최소화, 효율은 최대화해야 할 것이다.

따라서 일괄 변환/온라인 입력 방식은 업로드 기능을 갖는 시스템이면 NACSIS/LILIPUT 이외의 환경에서도 응용이 가능한 방법이다. 또 실제 다른 목록 시스템에 의해 작성된 기계가독 데이터를 NACSIS-CAT에서 작성되고 있는 최신의 로컬 데이터 베이스에 통합하기 위해서 NACSIS-CAT로 온라인 업로드하고 있는 사례도 있다. 이것은 단순히 개별 도서관의 소급이라는 의미 뿐 아니라, NACSIS-CAT에서의 전국 종합목록 형성이라는 점에서도 큰 의미를 갖고 있다고 생각된다.

본 시험가동에서는 데이터 변환시 부담의 크기가 문제가 되었다. 당초 NACSIS 종합 목록 데이터 베이스, 참조 MARC에서의 히트(Hit)율이 상당히 낮을 것이라고 예상 되었기 때문에 모든 데이터를 변환 시킨 후 입력한다는 순서를 밟았는데 현재는 양파일에서 모두 히트율이 향상하고 있

고 별도의 방법 (예를 들면 확실하게 히트하는 것에 대해서는 검색 키가 되는 부분과 소장 데이터 만을 변환하는 방법)을 생각 할 수도 있다.

더우기 앞으로 NACSIS종합 목록 데이터 베이스가 그 내용이 충실해지고, 특수한 자료이외에는 서지 데이터의 업로드에 대한 필요성이 낮아질 것이지만 입력결과의 항목에서 이미 언급한 바와 같이 그 대부분이 참가관 개별 방법이 되는 소장관련 데이터의 업로드 기능이 입력 효율에 미치는 영향도 무시할 수 없는 것이어서 앞으로의 작업에 참고가 될 것이라고 생각된다.

끝으로 우리에게 귀중한 작업의 기회를 주신, 본 과학 연구대표 山崎弘郎 東京大學 공학 부교수를 비롯해 연구 분담자 여러분 및 田中久文 圖書館 情報大學 사무국장 (소화 63. 3. 31까지 부속 도서관 사무부장), 또 연구실시에 있어서 계획 지도에 참가했던 永田治樹 北海道大學 연구 협력 과장 (소화 63. 3. 31까지 부속 도서관 정리과 도서관 전문원) 및 田村潤二 岡山大學 부속 도서관 정보서비스 과장 (평성 원년 3. 31까지 정보관리과 도서관 전문원), 宮本明雄 滋賀 醫科大學 도서관장 (평성 2. 3. 31까지 정보관리과 도서관 전문원) 그 외에 협력 지도를 아끼지 않은 増田元氏의 많은 관계 직원 및 화콤-하이테크 여러분에게 감사의 뜻을 전한다.

참고문헌

- 1) 東京大學附屬圖書館 『東京大學附屬圖書館 システム LILIPUT』 (1988)
- 2) Yong, Heartsill [編] 丸山昭二郎 (감역) 『ALA 圖書館情報學辭典』 東京, 丸善, 1988.
- 3) 永田治樹, 増田 元, 大西直樹 “アップロードとダウンロード: 目録・所在情報 ネットワークとローカルシステム” 『情報の科學と技術』 V. 37, No. 5, 1987, p. 175-184.
- 4) 遠 昭二, 吉竹 忍, 川端美明 “北大圖書館における 圖書データ大量 遡及 入力 の 試み : 學情センターファイルを利用しての目録カードオンライン 入力” 大學圖書館研究, No. 34, 1989, p. 49-61.

- 5) 文部省 科學研究費. 試驗研究 (1) 課題番號 62880006, 研究代表者: 黒田 晴雄 『大量文獻情報遡及變換入力システムの高度化に 關する研究』 1990. 2
- 6) 杉町 宏, 北 克一 “學術情報センター-目錄システムとその柔構造密結合方式による目錄作成” 『情報の科學と技術』 V. 40, No. 2, 1990, p. 123-128.