

# 이용 기반 데이터베이스 구축 방안에 관한 연구

## A Study on STI Database Construction on Demand

조현양(Hyun-Yang Cho)\*

### 목 차

1 서 론	4 데이터 수집 및 분석
1.1 연구의 목적	4.1 자료 유형별 원문서비스 현황
1.2 연구의 내용 및 방법	4.2 Journal Citation Report
2 표준 데이터 양식	4.3 학술지별 이용 현황
3 대상자료의 선정	5 연구의 제한점
3.1 구축대상 학술지 선정 기준	6 결 론
3.2 입력 대상으로서의 학술지 평가	

### 초 록

본 연구에서는 효율적인 서지정보 데이터베이스를 구축하기 위한 여러 가지 방안들이 제시되었다. 대상 자료의 선정으로부터 입력 양식의 표준화 및 입력시스템의 활용 등에 이르기까지 다양한 요소들이 검토되었으며, 특히 대상 자료의 선정시에 고려하여야 할 부분을 중점적으로 다루었다. 입력 대상 자료의 우선 순위를 결정하기 위하여 이용자 정보요구의 파악이 비교적 용이한 원문복사 서비스 현황을 비교, 분석하였다. 이용현황에 대한 분석 결과 다양한 형태의 자료 가운데 가장 우선적으로 입력하여야 할 대상은 조사결과 Proceeding인 것으로 나타났으며, 학술지는 우선 순위에서 그 다음을 차지하였다. 또한 학술지의 경우 학술지별 원문복사 신청 현황을 분석하여 SCI의 전체 인용빈도와 영향력 지수(Impact Factor)에 의한 순위와 비교·분석하였다.

### ABSTRACT

In this research, several ways of creating effective STI(Scientific & Technological Information) databases were suggested. We put emphasis on the selection of input data, while on the other was handled, such factors as standardization for data entry, data entry system, etc.. In order to decide priority of target data, the status of document delivery service was analyzed. The result shows that conference proceedings were given priority to academic journals. In case of journals, ranking in the number of documents requested at KORDIC (Korea R&D Information Center) and 16 Specialized Information Centers was compared with the ranking in citation frequency and impact factor, appeared at SCI.

키워드: 원문복사서비스, 서지데이터베이스, 서지정보, 이용기반

\* 연구개발정보센터 문헌정보사업실 실장

■ 논문 접수일 : 2000년 5월 18일

## I 서 론

### 1.1 연구의 목적

서지정보는 2차 자료로서 문헌이나 문헌에 수록된 논문단위의 정보에 대한 안내 역할을 수행한다. 이러한 2차 자료를 이용자가 좀 더 쉽게 접근할 수 있도록 컴퓨터를 이용하여 축적하고, 정보검색시스템을 통하여 신속하고 효율적으로 검색할 수 있도록 서지데이터를 모아 놓은 것이 서지정보 데이터베이스이다. 서지정보 데이터베이스 구축의 궁극적인 목적은 이용자와 그가 원하는 원시자료(Source data)를 연결시켜 주는 매개체로서의 역할을 원활하게 수행하기 위한 것이다. 이러한 서지정보 데이터베이스의 구축 목적을 고려해 볼 때 간과해서는 안될 부분이 구축대상 자료의 선정과 입력을 위한 표준화된 데이터 양식의 활용이다.

본 연구의 목적은 연구개발정보센터가 과학기술계 16개 정부 출연 연구소 및 대학과 공동으로 지난 10년간 축적해 온 경험을 토대로 도서관별 자료에 대한 서지정보 데이터베이스의 구축시 고려되어야 할 사항들을 밝히는 데 있다. 서지정보 데이터베이스의 구축과 관련하여 우선적으로 고려할 사항은 크게 3가지로 구분된다. 첫 번째 요소는 향후 기관간 데이터의 상호교환 및 공동활용이 가능하도록 입력 및 데이터 양식을 표준화하는 것이며, 둘째는 이용자의 요구에 부합되는 자료를 선정하는 것이다. 마지막으로 고려할 사항은 데이터의 효율적 입력을 위한 입력시스템의 활용이다. 본 연구에서는 상기의 목적을 달성하기 위하여 표준 데이터 양식을 제안할 것이다. 또한 이용자의 요구에 부합되는 자료를 선정하기 위하여 연구개발

정보센터의 과학기술문헌정보 데이터베이스(이하 SATURN)에 수록된 데이터별 활용빈도와 ISI의 JCR(Journal Citation Report)에 나타난 학술지별 순위를 비교하여 제시함으로써 도서관에서 학술지의 선정과 수집된 자료에 대한 데이터베이스화의 우선 순위 부여에 도움이 되 고자 한다.

### 1.2 연구의 내용 및 방법

본 연구에서는 서지정보 데이터베이스의 효율적 구축을 위하여 OCLC MARC을 기반으로 한 표준 데이터 양식을 제시하였으며, 연구개발정보센터의 SATURN을 이용하여 신청된 원문복사 현황을 분석하여 자료 유형별(국내외 학술지, 국내외 Proceeding) 우선 순위를 부여하였다. 분석 대상 자료로는 1998년(1998년 3월부터 1999년 2월)과 2000년 1/4분기 동안에 16개 과학기술계 정부 출연연구소 및 대학에 신청된 원문복사 통계 데이터가 이용되었다. 또한 JCR에서 제공하는 인용빈도와 영향력 지수에 의한 학술지의 순위와 실제 연구개발정보센터에 신청된 원문 복사 요청 건수를 학술지별로 구분, 비교하였으며, 원문복사 신청 현황에 따른 상위 60위까지의 학술지 목록을 제공하였다.

본 연구에서 언급된 이용자는 두 집단으로 나누어진다. 데이터베이스의 구축을 위한 데이터 표준 양식의 이용자는 사서나 혹은 데이터베이스 구축 담당자이며, 원문복사 이용자는 데이터베이스에 포함된 서지정보를 통하여 직접 원문을 이용하는 최종 이용자를 의미한다.

## 2 표준 데이터 양식

데이터베이스 구축에 있어서 표준화된 양식의 활용은 향후 기관간 데이터의 상호 교환 및 공동활용을 위하여 필수적인 것이다. 표준화된 양식의 사용은 무엇보다 구축된 데이터에 대한 일관성을 유지할 수 있으며, 입력자에게 있어서는 입력상 발생할 수 있는 항목간의 모호함을 방지할 수도 있다. 또한 이용자 측면에서는 검색된 데이터가 일정한 형태를 유지하게 되어 결과물에 대한 친밀감 및 이해도를 높일 수 있다.

연구개발정보센터의 SATURN은 과학기술분야에 특화된 정보를 중점적으로 소장하고 있는 기관들이 공동으로 참여하여 구축하고 있다. 여러 기관이 분리된 장소에서 하나의 데이터베이스

를 구축하기 위해서 가장 먼저 고려되어야 할 사항이 바로 표준화된 데이터 양식의 도입이었다. 1991년 과학기술계 출연 연구소 기술정보실을 중심으로 데이터베이스의 구축에 필요한 양식의 표준화를 위하여 기술위원회가 구성되었으며, 표준양식은 OCLC의 MARC for Serials를 근간으로 하여 작성되었다. 그러나 당시 작성된 데이터 양식은 가능한 한 OCLC MARC을 전용하는 차원이어서 데이터베이스 구축시 불필요한 항목이 많이 포함되어 있었다. 이러한 문제점을 보완하기 위하여 기술위원회에서 MARC 포맷을 보다 단순화시키는 작업이 매년 계속되어 왔다.

〈표 1〉은 기술위원회에서 작성된 서지정보 데이터베이스 구축을 위한 축약된 표준양식이다. 이 양식은 연속 간행물은 물론 conference

〈표 1〉 태그별 정의

태그	태그명	항목수	크기	서브 필드
001	AN	1	11	제어 번호
008	KD	1	6	기사 입력일자(0 - 5 col)
	DT	1	1	자료 유형(6 col)
	DTC	1	16	자료 유형(Display Only)
	LA	1	3	언어 코드(23 col)
010	LN		64	▼a LCCN
020	IB		64	▼a ISBN(국제표준도서번호) ▼c price(가격)
027	SR	2	64	▼a 표준기술레포트번호
049	LO	5	64	▼a 원문 소장기관 ▼l 소장기관 등록번호
088	RN	3	64	▼a 보고서번호
090	CN (청구기호)	2	64	▼a 분류 기호 ▼b 저자 기호 ▼f 별칭 기호
099	SC	3	64	▼a 주제 코드
100	PNH	1	192	▼a 개인 저자 ▼u 저자 소속 기관
700	PNA	10	128	▼a 공동 저자 ▼u 저자 소속 기관

〈표 1〉 태그별 정의(계속)

태그	태그명	항목수	크기	서브 필드
110	COH	1	192	▼a 기관 저자
710	COA	5	128	▼a 공동 기관 저자
245	TI	1	256	▼a 기사 제목 ▼b 부제 ▼n 권책번호 ▼p 권책명
250	ED	1	64	▼a 판차 사항
260	PB (출판사항)	1	128	▼a 출판지 ▼b 출판사(학위논문은 대학명) ▼c 출판일
300	PH (형태사항)	1	128	▼a 페이지 ▼b 삽도 표시 ▼c 자료 크기
440	SN	3	128	▼n 규격 번호
500	GN	2	256	▼a 주기 사항
502	DN	2	128	▼a 학위 정보(학위, 대학, 학위수여일)
503	SH	2	128	▼a 규격 갱신 기록
510	CI	2	128	▼a 인용 자료명(DB인 경우 DB명) ▼c location within source(수록된 권호, 페이지, 연도, DB인 경우 DB access no.)
513	RT	2	64	▼a 보고서 종류(연구, 경영, 예술, 기술, 기타) ▼b 수록 기간
520	AB	1	2046	▼a 초록
536	FI (연구후원정보)	3	128	▼a 연구 후원 기관 ▼b 연구 계약 번호
590	SA (규격승인사항)	5	64	▼a 규격 승인 국가 ▼b 규격 승인 기관 ▼c 규격 승인 일자 ▼d 규격 종류
599	AD	5	64	▼a 논문 지도 교수
650	SU	20	64	▼a 주제명(자연어, 저자 keywords)
653	ID	20	64	▼a 식별자(identifier, 통제어)
773	HI (원문수록처)	5	512	▼t 서명(title, 기사제목, 보고서제목, 학위논문제목) ▼a 주표목(main entry) ▼b 판차 사항(edition) ▼d 발행일자(출판지, 출판사, 출판년) ▼g 권호 또는 통권(년도, 페이지-페이지) ▼k 시리즈 사항 기술 ▼p 축약 서명(abbreviated title) ▼r 보고서 번호(report no.) ▼S 통일 서명(uniform title) ▼u 표준 보고서 번호(standard rn) ▼x ISSN ▼y 잡지 약어 코드(coden) ▼z ISBN
810				▼a 작성자명 또는 작성자 고유번호
820				▼a 참고 문헌수
830				▼a 원 제목(번역했을시)

proceedings, 학위논문, 기술보고서, 규격 등 다양한 형태의 자료를 데이터베이스화하는데 활용될 수 있다.

여기에서 데이터베이스에 포함되어야 할 필수 항목을 제외한 항목의 선택은 기관의 몫이다. 입력대상 자료 유형별로 차이가 있지만 주어진 33개의 항목 가운데 필수 항목은 접근번호, 소장기관, 자료유형, 사용언어, 저자, 기사제목, ISSN(ISBN 혹은 보고서번호), 초록, 원문수록처, 작성자 명 등 10개 항목이다. 필수 항목은 이용자가 검색된 데이터의 내용 파악에 필요한 최소한의 필드로 구성되어 있으며, 현재 SATURN의 데이터별 평균 항목 수는 13.6개이다

지금까지 데이터 입력작업은 일반 워드 프로세서 상에서 이루어졌으나 항목에 대한 표기가 레코드별로 반복 입력되어야 하는 불편함이 따르고, 반복적이고 불편함에 따른 입력상의 실수가 빈번하게 발생하였다. 이러한 입력상의 실수를 방지하고 입력자의 편의를 도모하기 위하여 연구개발정보센터는 2000년부터 입력시스템을 개발, 보급하고 있다. <표 2>에 나타난 바와 같이 입력시스템의 이용과 더불어 13개 기관의 평균 에러 발생율은 1999년 16.98%에서 2000년 0%가 되

었다. 여기에서 에러란 필수 항목의 누락, 원문 수록처(학술지명)에 대한 표현의 일관성 미비 등이며, 철자에 대한 입력 실수는 감안하지 않았다. 특히 에러율 0%가 가능했던 근본 이유는 2000년부터 온라인으로 입력시 에러의 검증이 자동적으로 시스템상에서 처리되도록 하였기 때문이다.

### 3 대상 자료의 선정

#### 3.1 구축 대상 학술지 선정 기준

SATURN의 입력 대상 자료 유형은 크게 국내 외 학술지, Proceeding, 기술보고서, 학위논문, 규격 등의 5가지로 구분되며, 1991년 이후 지금까지 구축된 데이터베이스의 총량은 300만 건에 이르고 있다. 데이터베이스는 구축된 데이터의 양이 증가할수록 이에 대한 활용도 또한 증가하는 것이 일반적이다. 그러나 도서관의 입장에서 소장된 모든 자료를 한번에 데이터베이스화하는 것은 불가능하며, 따라서 입력대상 자료에 대한 우선 순위를 결정할 필요가 있다. 이러한 결정은 이용자의 정보 요구에 기반한 것이어야 한다.

<표 2> 기관별 1일 평균 구축건수 및 에러 발생율

기관명	입력기 이용입력 건수/일 및 에러율 (2000년)	에디터 이용입력 건수/일 및 에러율 (1999년)	비 고
기초과학지원연구소	150건, 0%	80건, 14.2%	
생명공학연구소	120건, 0%	80건, 14.7%	
원자력연구소	130건, 0%	90건, 20.0%	
해양연구소	120건, 0%	89건, 21.5%	
표준과학연구소	140건, 0%	92건, 15.0%	
평균건수 및 에러율	135건, 0%	87건, 16.98%	

그러나 구축 단계에서 이용자의 요구를 조사하고 이를 반영하는 것은 현실적으로 불가능하다. 이러한 상황에서 가장 쉽게 접근할 수 있는 방법이 대상 자료에 대한 인용빈도를 분석하여 이용하거나 기 구축된 데이터베이스의 이용도를 활용하는 것이다. 첫 번째의 인용빈도에 따른 대상자료의 선정은 SCI를 활용하거나 직접 인용 문헌을 분석하여 활용하는 것이다. 두 번째는 기존 데이터베이스의 이용도를 활용하는 방법으로 데이터베이스를 통하여 검색된 문헌에 대한 원문복사 제공 건수를 분석하여 이용자가 선호하는 대상자료를 선정하는 것이다.

오늘날 과학기술의 급진적인 발전으로 인하여 과학기술정보의 유효주기는 상대적으로 짧아지고 있으며, 새로운 정보에 대한 이용자의 요구 변화가 타 분야에 비하여 현저한 것이 사실이다. Gupta(1990)는 물리학 분야의 논문의 유효주기에 대한 연구를 수행하였으며, 타 과학기술분야와 마찬가지로 물리학 분야의 논문도 유효주기가 짧은 것으로 나타났다. Brown(1979)도 화학 분야에 있어서 인용형태에 대한 분석 실험을 수행하였으며, Gupta와 유사한 결론을 내리고 있다. 즉 과학기술정보의 특성상 유효주기가 짧기 때문에 과학기술계에 종사하고 있는 학자나 연구자들은 보다 최신성이 보장되는 컨퍼런스 논문들을 선호하고 있는 것처럼 보인다. 이러한 현상에 대한 분석은 본 논문의 "4. 데이터 수집 및 분석"에서 다루어 질 것이다.

### 3.2 입력 대상으로서의 학술지 평가

일반적으로 연구나 학술활동을 위하여 연구자는 학술지에 대한 의존도가 높다. 학술지는 연구성과의 전달 및 기록 수단으로, 연구업적에

대한 인증 수단으로 역할을 한다. 그러나 학술지별로 중요도에 대한 차이가 있고 이를 평가하기 위한 방법의 하나로 인용분석과 같은 계량서지학적 방법이 활용되고 있다.

학술지의 평가는 일반적으로 이용빈도와 인용빈도를 이용하여 이루어지고 있다. 이용빈도에 대한 측정은 대부분 도서관의 경우 학술지에 대한 대출을 금하고 있어서 대출통계에 의존하기 보다는 이용현황을 직접 관찰하는 것이 보다 확실한 방법일 수 있다. 또 다른 방법은 도서관 이용자를 대상으로 설문조사를 수행하는 것이다. 그러나 도서관에서 이용현황을 직접 관찰하거나 설문을 통하여 파악한다는 것은 현실적으로 어려움이 따른다. 이용자에 대한 설문조사는 이용자가 본인이 이용하는 학술지의 구독 증지를 우려하여 실제보다 이용빈도를 과장되게 표현할 수도 있다. 또한, 이용현황을 지켜보면서 이를 카운트한다는 것 또한 어렵다. 이러한 현실을 고려해 볼 때 이용빈도의 측정을 위한 가장 용이한 방법은 학술지에 대한 복사횟수를 체크하거나 복사 신청건수를 카운트하는 것이다.

다음으로 학술지를 평가하기 위하여 널리 활용되는 방법은 학술지별, 논문별 인용빈도를 분석하여 학술지의 순위를 부여하는 것이다. 논문의 수준을 평가하는 직접적인 요소 가운데 하나는 발표된 논문이 타 논문에 미친 영향력을 평가할 수 있는 피인용회수(Number of Cited)이다. 즉 타 논문에 의해 많이 인용될수록 그 논문이 많은 연구자에 의해 읽혀진 것이며, 논문의 질이 높음을 의미하는 것이다. 과학기술분야에 발표되는 학술지에 포함된 논문에 대한 질적 평가를 위한 지표로써 ISI에서는 Science Citation Index를 매년 발표하고 있다. SCI는 기초과학을 비롯하여 의학, 생물학, 공학 등 응용과학 분야

의 학술지 및 Proceeding에 수록된 단위 논문에 대한 인용정보를 다양하게 포함하고 있으며, 학술지에 대한 중요한 평가지표로 활용되고 있다.

Garfield(1972)는 인용분석 자료는 도서관에서 학술지의 선택을 위한 참고자료로써, 장서로서의 학술지의 보관에 대한 판단요소로써 또한 인용건수 대비 학술지의 가격을 고려한 소장장서의 평가자료로써 그 외에도 많은 분야에서 활용될 수 있음을 밝히고 있다. 학자들에게 있어서 인용 분석자료는 그 분야의 학문의 발전 동향을 파악하기 위하여 필수적으로 점검해야 할 안내 자료로도 활용되고 있다. 그러나 학술지를 평가하는데 있어서 인용분석에만 의존하는 것은 또 다른 오류를 불러올 여지가 있다. 특히 논문 성격을 띄지 않고 새로운 기술에 대한 소개나 동향을 파악할 수 있는 글을 주로 게재하는 동향지 성격의 Science 지나 Scientific America 등과 같이 학자들에게 정기적으로 검토되고 읽혀지지만 실제 논문 작성시 그리 많이 인용되지 않는 경우도 있기 때문이다.

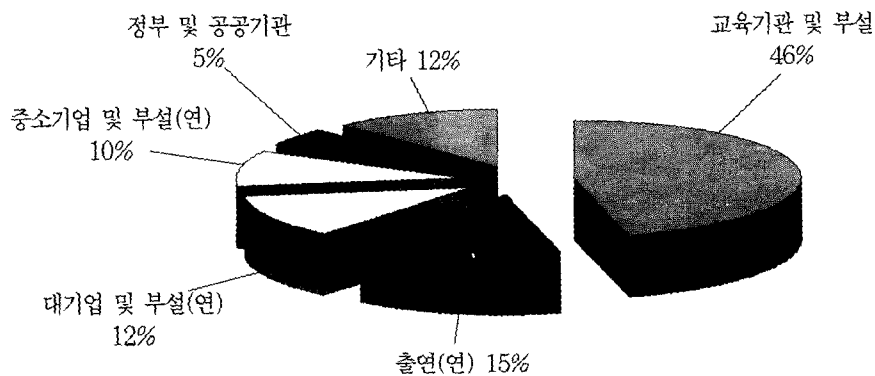
당 연구소가 보유하고 있는 SATURN의 참여 기관은 구축대상 자료의 선정을 위하여 기관 내에

서의 이용도에 따라 입력의 우선 순위를 부여하고, 데이터를 입력한다. 이용자들은 구축된 데이터베이스를 검색하여 원문을 입수할 수 있으며, 이에 따라 학술연구 활동에 이용하는 일반 이용자의 학술지별 선호도를 파악하는 것이 가능하다.

## 4 데이터 수집 및 분석

### 4.1 자료 유형별 원문 서비스 현황

본 논문에서는 1998년 3월부터 1년간 그리고 2000년 1월부터 3월까지 3개월간 연구개발 정보센터를 통하여 신청된 원문복사 44,200건(1998년 38,600건, 2000년 1분기 5,600건)을 이용하여 신청된 논문이 포함된 자료 유형별, 국내외 간행물별, 각 단위 학술지별로 통계를 분석하였다. 분석된 통계수치에 따라 논문이 포함된 국내외 학술지별 우선 순위를 부여하였다. 또한, 학술지의 경우 이를 실제 SCI에서 제시한 인용횟수별, 영향력 요소별 학술지의 순위와 비교, 분석하는 작업을 수행하였다.



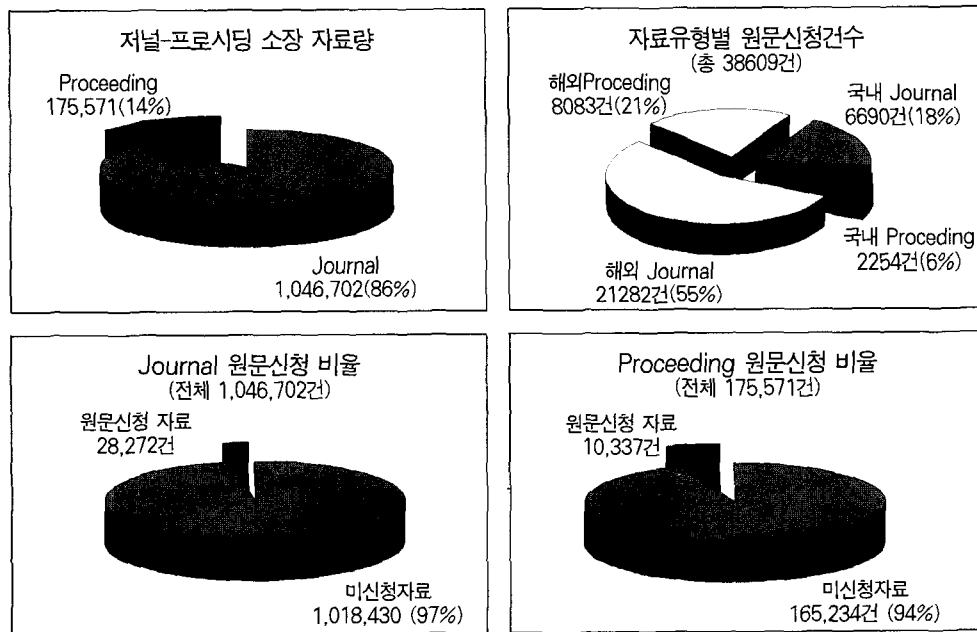
〈그림 1〉 직업별 이용자 구성 현황

본 연구에서의 원문복사 서비스대상은 인터넷을 통하여 연구개발정보센터의 서지정보 데이터베이스인 SATURN에 접속하여 필요한 원문을 신청하는 불특정 다수의 이용자이다. 원문복사를 신청한 직업별 분포는 교수, 석박사, 학생, 연구원, 기타로 다양하며 그 분포도는 <그림 1>과 같다.

1998년 실험기간 당시까지 구축된 데이터의 양은 총 117만 건이며 2000년 3월까지 구축된 데이터는 300만 건에 이르고 있다. 데이터의 구성은 학술지나 Proceeding 등에 수록된 단위 논문에 대한 서지정보로 이루어져 있다. 학술지와 Proceeding의 구성비율은 98년 86:14 (학술지: 1,046,702, Proceeding: 175,571)과 2000년 87:13(학술지: 2,621천건, Proceeding: 381천건)으로 나타났으며, 자료 유형별 구축량에 대한 국내의 구분은 하지 않았다. 대상 기간

동안에 신청된 원문복사 건수는 총 44,209건 (98년: 38,609, 2000년: 5,600건)이었으며, 이 가운데 원문복사 신청 건수는 해외 학술지가 24,050건 (98년: 21,282건, 2000년: 2,768건)으로 가장 많았다. 그 다음으로 해외 Proceeding이 9,604건(98: 8,083건, 2000년: 1,521건)으로 2위를 차지하였으며, 8,004건 (98년: 6,990건, 2000년: 1,014건)의 국내 학술지, 2,551건(98년: 2,254건, 2000년: 297건)의 국내 Proceeding의 순으로 나타났다.

그러나 1998년의 경우 구축된 데이터 량과 원문복사 건수를 비교하여 볼 때 학술지는 전체 구축량의 약 3%(28,272건/1,046,702건)에 해당되는 건수가 신청되었으나 Proceeding의 경우 전체 구축량 대비 약 6%(10,337건/165,234건)가 실제 원문복사 요청이 들어온 것으로 나타났다. 또한 2000년 1/4분기의 경



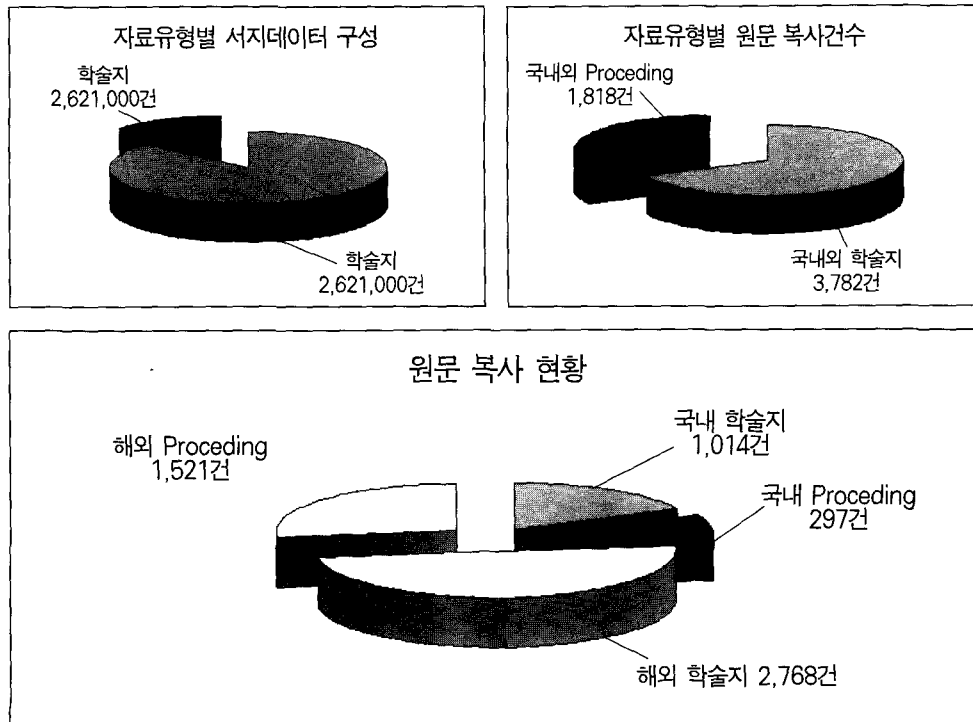
<그림 2> 1998년도 원문신청 분석 결과



우 구축된 데이터 량과 원문복사 건수를 비교하여 볼 때 학술지는 전체 구축량의 약 0.14%(3,782건/2,621천건)에 해당되는 건수가 신청되었으나 Proceeding의 경우 전체 구축량 대비 약 0.48%(1,818건/381천건)가 실제 원문복사 요청이 들어와 구축건수 대비 원문복사 이용률이 학술지에 비해 현저하게 높은 것으로 나타났다. 그 가운데 구축량 대비 원문복사 신청률은 국내 Proceedings가 가장 높게 나타났다. <그림 2>는 자료유형별 원문신청 건수 및 백분율, 학술지의 원문 신청비율, Proceeding의 원문 신청비율, 그리고 자료유형별 원문복사 신청 현황을 파이 도표로 나타낸 것이다.

분석된 자료에서 나타나듯이 실험 대상기간

(98년과 2000년 1분기)동안 자료유형별 실제 구축된 데이터 양 대비 실제 원문복사 신청 비율은 학술지(98년: 3%, 2000년: 0.14%)와 Proceeding(98년: 6%, 2000년: 0.48%) 모두 감소하였다. 데이터 양 대비 원문복사 신청 비율의 급격한 감소는 1999년 1년 동안 정보의 디지털화에 집중적으로 투자하여 일시에 많은 양의 데이터베이스를 구축하였기 때문이다. 그러나 총 구축 건수 대비 원문복사 신청 비율은 1998년의 경우 Proceeding이 학술지보다 약 2배정도 높은 비율을, 그리고 2000년에는 약 3.4배가 높은 것으로 나타났다. 즉 1998년과 2000년을 비교하여 볼 때 연구자들이 연구개발활동에 필요한 정보의 입수원으로써 학술지보다는 Proceeding에 더 의존하는 것을 알 수



<그림 3> 2000년 1분기 서지DB 데이터량 및 원문신청 건수

있다. 그러므로 과학기술분야의 데이터베이스의 구축시 이용자의 정보 요구에 근거하여 자료 유형별 우선 순위를 결정한다면 학술지보다는 Proceeding에 우선 순위를 부여하는 것이 합당할 것이다.

#### 4.2 Journal Citation Report

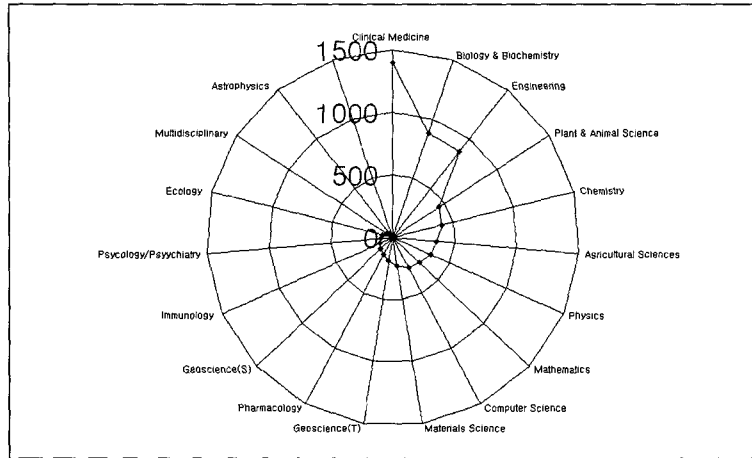
JCR은 ISI가 분야별 중요 학술지를 선정하여 연도별 총 인용수(Total Cites), 영향력 지수, Immediacy Index, 연간 총 수록 논문수, Cited Half-Life 등 학술지를 평가할 수 있는

자료를 제공해 주는 데이터베이스이다. JCR은 5300 여 종의 학술지에 대한 분석자료를 제공해 주는데, 18개의 중분류로 구분된다. 그런데 하나의 학술지가 둘 이상의 세부분류에 포함될 수 있으므로, 분류 기준에 의한 총 학술지 수는 6348이 되고, 각각의 중 분류에 속하는 학술지의 수는 <표 3>과 같다.

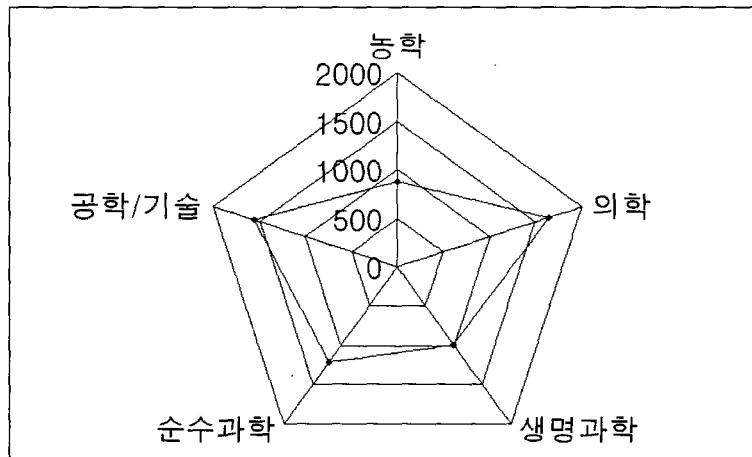
분야별 학술지 분포를 그림으로 나타내면 <그림 4>와 같다. 그림에서 보는 바와 같이 분야간 색인되는 학술지의 편차가 현저하며, 따라서 인용횟수 및 논문 게재수에 의한 분야간 비교는 신빙성이 결여될 수도 있다. 다만 분야별로 출

<표 3> 분야별 JCR 색인 학술지 수

중분류 명	대분류명	학술지 수
Agricultural Sciences	A	352
Astrophysics	S	38
Biology & Biochemistry	P	875
Chemistry	S	411
Clinical Medicine	C	1397
Computer Science	T	277
Pharmacology	C	157
Psychology/Psychiatry	C	98
Ecology	A	86
Engineering	T	870
Geoscience(T)	T	184
Geoscience(S)	S	137
Immunology	P	117
Materials Science	T	225
Mathematics	S	292
Multidisciplinary	T, S, P	56
Physics	S	334
Plant & Animal Science	A	442



〈그림 4〉 분야별 JCR 색인 학술지 분포



〈그림 5〉 대분류 분야별 학술지 수

판되는 학술지의 수에 있어서도 차이가 있기 때문에 포함 학술지의 차이는 부분적으로 인정되어야 할 것이다. 하나의 예로 의학 분야의 경우 1397종의 학술지가 색인되어 있고, 가장 적은 천체물리학의 경우 38종에 지나지 않는다. SCI는 논문에 대한 인용빈도 및 영향력 지수에 근거하여 학술지의 중요도를 평가한다. 대부분의 경우 동일 분야 혹은 유사 분야의 연구결과를

참조하게 되며, 결과적으로 많은 학술지가 포함된 주제분야일수록 학술지의 중요도가 높게 평가된다는 것이 SCI가 안고 있는 문제점이다.

#### 4.3 학술지별 이용현황 비교

본 절에서는 실제 도서관에서 수집된 자료의 활용도와 SCI의 인용빈도간의 상관관계를 파

악하였다. 자료의 활용도는 통계작성이 비교적 용이한 원문복사 이용 현황을 분석하여 부여된 학술지별 순위를 결정하고 이를 SCI의 인용빈도 순위와 비교하였다. SCI 인용빈도는 도서관에서 자료선정을 위하여 혹은 학술지에 대한 서지정보 데이터베이스 구축시 우선 순위의 결정을 위한 가이드 라인으로 활용되고 있다. 4.2에서 언급된 문제점에도 불구하고 SCI가 가이드 라인으로써 활용될 수 있는 이유는 색인 대상 학술지 선정에 있어서 다음과 같은 엄격한 기준을 적용하기 때문이다.

- 동일분야 연구자들의 평가를 거쳐 논문을 수록하는가?
- 학술지 기고자에게 논문 게재료를 부과하는가?
- 저자와 편집인들이 지역적 대표성을 가지는가?
- 학술지가 출판 기준을 지키는가?
- 학술지가 출판 시기를 일정하게 유지하는가?
- 학술지의 논문이 다른 학술지에 의해 많이 인용되는가? (Impact Factor)
- 출판된 후 얼마나 빨리 인용되는가? (Immediacy Index)
- 얼마나 오랫동안 인용되고 있는가? (Cited Half-life)

〈표 4〉는 SATURN의 검색을 통하여 신청한 원문 복사 건수와 SCI의 학술지별 인용 순위 및 영향력 지수를 나타내고 있다. SATURN에 포함된 2,842 종 가운데 1998년 한해 동안 1회 이상 원문복사 요청이 있었던 학술지는 1,940 종에 이르고 있다.

〈표 4〉에 나타난 바와 같이 1998년도의 경우 원문복사 요청 순위 상위 30위 이내에 랭크

된 문헌들이 실제 SCI의 인용빈도 및 영향력 지수에서는 대부분 상위에 랭크되지 않았다. 이는 원문복사 이용자의 인용빈도 순위가 SCI에서 매년 발표되는 학술지의 순위와는 상당한 차이를 보이고 있음을 보여준다. 원문복사 이용도 상위 30위 이내에 랭크된 학술지(Proceeding 포함) 가운데 17개 (약 60.1%)의 학술지가 SCI에 등재되지 않았다. 물론 국내 학술지 및 Proceeding의 경우 SCI에서 인용빈도 및 영향력 지수에 대한 평가가 이루어지지 않았기 때문에 직접적인 관련이 없다고 볼 수도 있다. 그러나 많은 국내 과학기술계 학자나 연구자들이 국내 학술지 및 Proceeding에 발표되는 논문들을 참고하고 있음이 나타났다. 특히 일본 기계학회에서 발간되는 영문학회지인 “Transactions of Japan Society of Mechanical Engineers” 경우 원문복사 요구는 1998년과 2000년 공히 1등으로 랭크되어 있으나, SCI 인용 순위 및 영향력 지수에서는 하위권에 머물러 있다. “Physical Review A” 만이 유일하게 SCI 인용 및 원문복사 요청 순위에서 10위 내에 나타나고 있다

〈표 5〉에서 나타난 바와 같이 2000년 1분기 동안 이용된 학술지(Proceeding 제외)의 경우 역시 1998년도의 상황과 유사하게 나타났다. 755 종의 학술지가 실험기간 동안에 1회 이상 원문복사 서비스가 이루어졌으며, 상위 60위에 랭크된 학술지 가운데 29개 (48.3%) 학술지가 SCI에 등재되지 아니한 것으로 나타났다. 그러나 31개 학술지의 경우에 있어서 영향력 지수나 인용횟수에 따른 SCI 순위와는 많은 차이를 보이고 있다. 본 논문에서는 원문복사 신청 현황을 분석하여 부여된 학술지별 순위는 첨부하지 않았다.

결론적으로 우리의 경우 도서관에서 자료의 선정과 소장자료의 데이터베이스화를 위한 우선 순위를 결정하는 데 있어서 SCI는 유용한 자료 이긴 하지만 이를 기준으로 삼는 것은 문제가 있다. 왜냐하면 SCI에 포함된 학술지 및 Proceeding가 엄정한 기준에 의하여 선정되기는 하지만 구미 중심이며, 특히 영어권 중심이기 때

문으로 해석할 수 있다. 반면, 원문복사 현황 분석을 통한 소장 자료의 활용도 파악은 최종 이용자의 정보요구가 보다 적극적으로 반영된 것으로 판단된다. 궁극적으로 자료의 수집과 이의 활용을 촉진시키기 위한 데이터베이스의 구축이라면 활용도 분석에 근거하여 학술지별 우선 순위를 부여하는 것이 바람직하다고 하겠다.

〈표 4〉 1999년도 SCI 인용순위와 1998년도 원문복사 신청건수와의 비교표

자 료 명	원문복사횟수	SCI 순위	Impact Factor
Transactions of Japan Society of Mechanical Engineers	1396	2592	0.126
Journal of Fluid Mechanics	940	131	1.609
Physical Review A	449	7	6.14
Japanese Journal of Applied Physics	420	108	1.261
Communications of ACM	315	623	1.293
AIAA Journal	292	458	0.671
Computers and Structures	292	942	0.376
Transactions of ASME: Journal of Fluids Engineering	290	1976	0.315
Journal of Korean Institute of Communication Science	289	575	0.681
Journal of Applied Polymer Science	254	210	0.841
Transactions of Korean Institute of Electrical Engineers	246		
Journal of Applied Polymer Science	230		
기술현대	203		
Journal of Aircraft	202		
Mitsubishi Juko Giho	188		
화학공학(Journal of Korean Institute of Chemical Engineers)	187		
새물리	182		
Journal of Korean Institute of Telematics and Electronics	177		
자동차공학회지	173		
Polymer	169		
대한기계학회논문집	162		
Journal of Society of Naval Architects of Japan	162		

〈표 5〉 1999년도 SCI 인용순위와 2000년도 원문복사 신청건수와의 비교표

제 목	신청횟수	인용순위	순위
Transactions of the Japan Society of Mechanical Engineers	129		
대한국계학회논문집	99		
일본기계학회논문집 A(Transactions of the Japan Society of Mechanical Engineers A)	94		
AIAA Journal	66	6,183	2718
Sensors and Actuators: A Physical	39	1782	2566
전기학회지(일본)	38	50	5294
Journal of Aircraft	35	860	4185
Journal of Applied Polymer Science	35	12709	2233
Surface and Coatings Technology	35	3824	2202
Journal of Materials Science	33	12103	2900
Journal of Sound and Vibration	31	5137	2694
International Journal for Numerical Methods in Engineering	30	4901	1939
Ceramic Engineering and Science Proceedings	28		
기술현대	27		
Computers and Structures	27	2029	4824
소음,진동	26		
Japanese Journal of Applied Physics	26	21527	1554
Journal of Fluid Mechanics	26	18312	1089
Mitsubishi Juko Gihou	26		
자동차공학회지	25		
한국요업학회지	25		
Transactions of the ASME : Journal of Heat Transfer	25	2582	2945
Nuclear Engineering and Design	24	833	4941
기계와 재료	22		
대한국계학회지 = Journal of the Korean Society Mechanical Engineers	22		
Biomaterials	22	4160	954
한국항공우주학회지	21		
Materials Science and Engineering A: Structural Materials: Properties, Microstructures and Processing	21	5385	2549
Thin Solid Films	21	11898	1961
일본조선학회논문집	20		
Journal of Computational Physics	20	6928	1425
한국폐기물학회지	19		
Perkin Transactions 1	19		
Transactions of the ASME : Journal of Fluids Engineering	19	1023	3425
마이크로소프트웨어	18		
전기학회논문지	17		
Computer Methods in Applied Mechanics and Engineering	17	3834	2130
Journal of the Ceramic Society of Japan	17	764	3654
Macromolecules	17	40831	334
Nuclear Instruments & Methods in Physics Research, Section A, Accelerators, Spectrometers, Detecto	17	10453	2210
관서조선학회지	16		
한국정밀공학회지	16		
American Nuclear Society Transactions	16		
철도총연보고(RTRI Report)(일)	15		
Electrochemical Society, Journal	15	21638	759
대한조선학회지	14		
약학회지	14		
자동차기술 = Journal of Society of Automotive Engineers of Japan	14		
한국박용기관학회지 = Journal of the Korea Society of Marine Engineers	14		
Canadian Journal of Fisheries and Aquatic Sciences	14	9823	1025
Nuclear Technology	14	773	4383
Physical Review A	14	43099	499
Transactions of the ASME : Journal of Mechanical Design	14	501	4158
공기청정기술 = Air Cleaning Technology	13		
대한용접학회지 = Journal of the Korean Welding Society	13		
대한조선학회논문집	13		
화학공학	13		
Bulletin of Chemical Society of Japan	13	12948	1447
Earthquake Engineering & Structural Dynamics	13	791	3368
IEEE Transactions on Magnetics	13	372	2680

## 5 연구의 제한점

본 연구의 결과는 제한된 예산 및 시간 동안에 소장 자료에 대한 데이터베이스를 보다 효율적으로 구축하고자 하는 기관에게 있어서 대상자료의 선정과 또한 도서관에서 장서구성의 우선 순위 결정에도 도움이 될 것이다. 그러나 본 연구는 과학기술 분야에 속하는 특정 집단을 대상으로 수행하였기 때문에 연구의 결과를 실제 상황에 그대로 활용하기에는 몇 가지 제한점이 따른다.

본 연구의 수행에 있어서 예상되는 한계점은 다음과 같다. 이러한 제한점은 연구의 결과에 중대한 영향력을 미칠 수도 있다. 그러나 상기에 언급된 본 연구의 목적에 나타나듯이 양질의 데이터베이스를 구축하기 위한 지침을 제공하는 것이기 때문에 대상자료에 대한 정확한 순위를 부여하는 것은 본 연구의 주요 목적이 아니다.

1) 개별 원문복사에 대한 신청자의 수준이나 직업 분포는 고려하지 않았다.

2) 동일한 자료 유형 내에서의 구축된 데이터의 건수에 대한 고려도 부족하였다. 예를 들어 A 학술지의 경우 5,000건의 논문이 포함되어 있으나, B 학술지의 경우 단 1,000건의 데이터만이 포함되어 있을 수도 있다.

다만 구축된 연도는 가능하면 동일한 범위로 한정시켰다.

3) 원문 복사 서비스의 성격상 자관 소장 자료의 경우 원문복사서비스 기관을 이용하지 않고 자관에서 직접 복사하여 사용하게 되며, 이러한 경우는 고려하지 않았다.

4) 원문복사서비스를 위한 데이터베이스의 포함 주제범위의 한계(과학기술분야) 또한 고려하지 않았다.

## 6 결 론

본 연구를 통하여 서지정보 데이터베이스를 보다 효율적으로 구축하기 위한 방안들이 제시되었다. 기관간 정보의 상호 교환 및 공동활용을 위한 MARC 기반의 표준화된 데이터 양식을 제안하였으며, 여러 최소화를 위한 입력시스템의 활용이 검토되었다. 특히 대상 자료의 선정 시에 고려하여야 할 부분을 중점적으로 다루었다. 본 연구의 분석결과 서지정보 데이터베이스의 구축을 위한 입력시스템의 사용은 업무의 능률 향상, 에러의 감소 측면에서 월등히 효과적인 것으로 나타났다. 다음으로 다양한 형태의 자료 가운데 과학기술분야에서 가장 우선적으로 입력하여야 할 대상은 Proceeding인 것으로 나타났으며, 학술지는 우선 순위에서 그 다음을 차지하였다. 현재 구축중인 과학기술 문헌정보 데이터베이스에 포함된 데이터 가운데 학술지에 수록된 논문이 많은 부분을 차지하고 있지만 우선 순위에 의한 입력을 고려한다면 Proceeding을 먼저 입력하는 것이 구축된 데이터베이스의 활용도를 높일 수 있는 하나의 방안이 될 것이다.

국내외 학술지 및 Proceeding의 이용 현황은 해외 학술지, 해외 Proceeding, 국내 학술지, 그리고 국내 Proceeding의 순으로 나타났다. 그러나 전체 소장 데이터 대비 이용률은 해외 Proceeding, 국내학술지, 해외 학술지, 국내 Proceeding 순인 것으로 나타났다. 또한 학술지별 이용자의 선호도는 상당한 차이가 있는 것으로 나타났다.

마지막으로 원문복사 현황과 SCI의 인용빈도 및 영향력 지수에 대한 순위를 비교해 본 결과 국내 이용자들이 빈번히 이용하는 학술지들

이 SCI에서의 순위와는 많은 차이가 있는 것으로 나타났다. 도서관에서 학술지의 선정과 데이터베이스 구축을 위한 우선 순위를 정하는 데 있어서 영어권 이용자 중심의 SCI에 의존하기

보다는 원문복사 현황 분석자료나 실제 국내 학회지에 수록된 논문의 인용색인을 분석하여 활용하는 것이 이용도에 기반한 방안이 될 것이다.

## 참 고 문 헌

- 사공철, 구자영, 김석영. 1989. 『과학기술문헌정보론』. 서울: 구미무역 출판부.
- 시스템공학연구소. 1991. 『과학기술정보 유통체제 구축 사업 (I)』. [대전]: 동연연구소.
- 심병규, 김기영. 1997. 최근 5년간 국내 과학기술자들의 연구활동에 관한 고찰. 『도서관학논집』, 27: 137-169.
- 연구개발정보센터. 1999. 『과학기술문헌정보데이터베이스 구축사업 (IX)』. [대전]: 동연연구소.
- 연구개발정보센터. 1999. 『SCI DB 분석을 통한 기초과학수준 평가체제 수립에 관한 연구』. [대전]: 동연연구소.
- Brown, Pauline. 1980. "The Half-Life of the Chemical Literature." *Journal of the American Society for Information Science*, 31(1): 61-63.
- Chen, Ye-Sho, Chong, P.P. and Tong, M.Y. 1995. "Dynamic behavior of Bradford's Law." *Journal of the American Society for Information Science*, 46(5): 370-383.
- Esler, S.L. and Nelson, M. L. 1998. "Evolution of Scientific and Technical Information Distribution." *Journal of the American Society for Information Science*, 49(1): 8-91.
- Garfield, Eugene. 1972. "Citation Analysis as a Tool in Journal Evaluation: Journals can be ranked by frequency and impact of citations for science policy studies." *Science*, 178: 471-479.
- Gupta, U. 1990. "Obsolescence of Physics Literature: Exponential Decrease of the Density of Citations to Physical Review Articles with Age." *Journal of the American Society for Information Science*, 41(4): 282-287.