

온라인 목록 검색 행태에 관한 연구

-LINNET 시스템의 Transaction log 분석을 중심으로-

윤 구 호* 심 병 규**

< 목 차 >

- | | |
|--------------------------------|-----------------------|
| I. 서 론 | Ⅲ. Transaction log 분석 |
| Ⅱ. 온라인목록의 Transaction log 분석 | 1. 검색시스템 이용빈도 |
| 1. Transaction log를 이용한 검색형태개요 | 2. 검색실패율과 검색키 이용빈도 |
| 2. LINNET의 Transaction log 구성 | 3. 검색결과와의 조희 |
| | 4. 절단, 제한 및 연산자 이용형태 |
| | 5. 검색실패의 원인 |
| | Ⅳ. 결 론 |
| | 참 고 문 헌 |

I. 서 론

온라인목록이란 온라인열람목록(OPAC : Online Public Access Catalog) 또는 온라인이용자접근목록(Online Patron Access Catalog)을 지칭하며, 컴퓨터 터미널을 통하여 도서관 소장자료의 데이터베이스를 저자명, 서명, 주제명 등을 검색키로 사용하여 검색할 수 있게 하는 시스템이다.

최근 정보화사회의 대두와 함께 컴퓨터 기술의 향상, 하드웨어 비용의 저하, 지역 및 광역 네트워크의 확대, 기계가독형목록(MARC : Machine Readable Cataloging)의 포맷개발 및 OCLC와 같은 서지도구들의 부상으로 세련된 온라인

* 계명대학교 문헌정보학과 교수

** 포항공업대학교 중앙도서관 사서장.

목록시스템들을 가능케 했고, 또한 도서관 내부에서뿐만 아니라 외부에서의 온라인 목록의 이용을 보편화시키게 되었다.¹⁾ 또한 도서관운영의 기본 방향이 사서중심에서 이용자중심으로 전환되면서 목록의 이용은 이용자와 온라인 목록간의 “커뮤니케이션 프로세스”로 이해하게 하였고, 효율적인 서비스를 제공하기 위한 여러가지 방안을 모색하게 하였다²⁾.

온라인목록의 효율적인 운영을 위해서는 두가지 조건이 필요하다고 볼 수 있다. 첫째는 온라인목록은 사용자들의 요구에 부합할 수 있도록 충분히 유연성이 있어야 한다는 것이고, 둘째는 온라인 목록의 이용내용에 대한 피드백이 반드시 있어야 한다는 것이다.³⁾ 최근 개발되고 있는 대부분의 온라인목록에는 transaction log, 즉 이용자와 온라인 시스템간의 의사소통(dialog)에 대한 자동기록이 생성되고 있는데, 이 transaction log 파일의 활용은 사용자들의 검색행태 연구에 새로운 전기를 마련해 주었다.

Norden과 Lawrence⁴⁾, 도서관진흥재단(CLR : Council on Library Resources)⁵⁾, Dickson⁶⁾, Peters⁷⁾, Millsap과 Ferl⁸⁾ 등은 transaction log를 이용하여 온라인 목록 검색의 성공률과 실패율, 검색키별 이용빈도, 검색가능별 이용률, 검색오류의 유형 및 검색실패의 원인분석 등을 통하여 이용자의 온라인

1) Ferl, Terry Ellen and Millsap, Larry, "Remote use of the University of California MELVYL library system : an online survey", *Information Technology and Libraries*, Vol.11(1992), p. 285.

2) 노문자, "온라인 목록 이용에 관한 연구조사", *승의여전도서관학연구지*, 제15집(1990), p. 7.

3) Larson, Ray R. and Graham, Vicki, "Monitoring and evaluation MELVYL", *Information Technology and Libraries*, Vol.1(1983), p. 93.

4) Buckland, Michael K., "Agenda for online catalog designers", *Information Technology and Libraries*, Vol.11(1992), pp.157-163.

5) Ferguson, Douglas, "The CLR public online catalog study : an overview", *Information Technology and Libraries*, Vol.1(1982), pp.84-97.

6) Dwyer, Catherine M., Gossen, Eleanor A. and Martin, Lynne M., "Known-item search failure in and OPAC", *RQ*, Vol.31(1991), p. 228-236.

7) Alzofon, Sammy R and Van Pulis, Noelle, "Patterns of searching and success rates in an online public access catalog", *College and Research Libraries*, Vol.45(1984), pp.110-115.

8) Millsap, Larry and Ferl, Terry Ellen, "Search patterns of remote users : an analysis of OPAC transactions logs", *Information Technology and Libraries*, Vol.12(1993), pp.321-343.

목록 검색 형태분석을 시도한 바 있다.

포항공과대학교 도서관이 1991년 3월부터 개발을 시작하여 32개월만에 완성한 LINNET (Library Information Network System)은 중국어 및 일본어 검색은 물론, 불리언 논리연산자 및 제한자의 사용등을 이용한 다양한 검색기능을 갖추고 있으며, 동시에 이용자의 검색과정을 완전히 재구성할 수 있는 검색 transaction log⁹⁾를 생성, 보관하고 있다.

본 논문의 목적은 LINNET의 온라인목록 transaction log를 분석하여 도서관 이용자들의 목록검색 행태를 조사하고, 동시에 온라인목록 검색기능을 평가하기 위한 것이다. 이러한 분석과 평가의 결과는 도서관 관리자, 시스템 설계 및 개발자들에게 도서관 이용자들의 검색행태를 파악하는 기초자료로 활용될 것이며, 동시에 아직 시작단계에 있는 한국 도서관의 온라인목록 검색시스템 개발시 여러가지 시행착오를 최소화하는 데에도 기여할 것이다.

II. 온라인 목록의 transaction log 분석

1. Transaction log를 이용한 검색 행태 개요

Transaction log는 “컴퓨터 시스템과 그것을 이용하는 사람간에 발생하는 여러 가지 상호작용을 담은 컴퓨터디스크 또는 테이프상의 기록”으로써¹⁰⁾, 도서관 및 정보학분야에서의 transaction log 분석은 ‘온라인 정보검색시스템과 그 시스템에서 정보를 구하려는 이용자들간의 전자적으로 기록된 상호작용에 관한 연구’라고

9) 이 파일에는 검색일자, 검색의 시작시간 및 종료시간, 검색시도 횟수, 이용한 검색기능의 이름, 검색내용, 이용자의 소속 및 username, hit건수, 상세서지정보 열람건수 등이 수록되어 있다.

10) Peters, Thomas A., "An Introduction to the special section on transaction log analysis", *Library HI TECH*, Vol. 11 (1993), p. 38.

정의할 수 있다.

1960년대 중반부터 시작된 transaction log분석은 다음과 같이 세단계로 발전되어 왔다. 첫번째 단계인 1960년대 중반부터 1970년대 후반에서는 이용자의 행태나 작업내용보다는 주로 시스템 성능 평가에 중점이 주어졌다. 두번째 단계인 1970년대 말부터 1980년대 중반에 transaction log분석이 처음으로 온라인목록 시스템의 연구에 적용되었는데 시스템 이용방법, 이용자들의 검색행태(검색소요시간, 검색오류형태) 등을 주요 연구주제로 삼았다. 세번째 단계인 1980년대 중반이후는 다양화로 특징지을 수 있다. 특정 검색모드(주제명검색 등), 이용자집단(dial access 이용자 등), 정보시스템(CD-ROM, workstation 등), 정보 데이터베이스 등을 주 연구대상으로 선택하였다. 또 동일 연구프로젝트내에서 시스템 분석과 이용자 행태분석을 동시에 실시하기도 하였다¹¹⁾.

이와 같은 발전단계를 통하여 알 수 있듯이 대다수 transaction log 분석은 정보 검색시스템의 실제 이용에 중점을 두어 왔으며, 그 기본요소를 들면 다음과 같다.

- 1) 이용자의 검색명령어
- 2) 검색단어
- 3) 검색일자 및 시간
- 4) 터미널 id
- 5) 검색된 파일
- 6) 검색된 자료수
- 7) 디스플레이된 스크린의 형태와 수¹²⁾

Transaction log 데이터를 수집, 분석 및 해석하는 데에는 네가지 제한점이 있는데 그 내용은 다음과 같다¹³⁾.

11) Peters, Thomas A., "The History and development of transaction log analysis", *Library HI TECH*, Vol. 11(1993), p. 42.
 12) Flaherty, Patricia, "Transaction logging systems : a descriptive summary", *Library HI TECH*, Vol. 11(1993), p. 68.
 13) Kurth, Martin, "The Limit and limitation of transaction log analysis", *Library HI TECH*, Vol. 11(1993), p. 98.

- 1) 시스템상의 한계로 복잡한 프로그램, 엄청난 데이터 사이즈, 서지 데이터베이스의 계속적인 수정 등이다. 기존 데이터베이스의 목록 레코드들이 계속적으로 수정, 삭제 및 추가되므로 동일한 검색명령으로도 검색 transaction log의 내용이 달라질 수 있다.
- 2) Transaction log는 검색내용을 상세히 나타내지만, 이용자들의 검색의도나 검색결과에 대한 만족도는 나타내지 않는다. 이와 같은 한계점을 극복하기 위하여, 설문조사, 면접조사, 재검색 등의 기법을 transaction log 분석과 병행하여 조사하기도 한다.
- 3) 검색결과가 0건인 경우를 검색실패로 판단하거나, 혹은 검색결과가 있다고 하더라도 반드시 검색성공으로 판단하는 것이 이용자의 검색행태에 대한 정확한 해석인가 하는 점이다.
- 4) Transaction log에는 이용자의 사적인 검색정보가 기록되어 있으므로, transaction log를 분석할 때 이용자들의 프라이버시를 보호해야 한다. 이러한 프라이버시의 보호는 정보검색시스템 모니터링의 확산을 막는 주요한 장애가 될 수 있다.

2. LINNET의 transaction log 구성

LINNET은 1987년 3월부터 KORMARC format과 USMARC format으로 구축된 온라인목록을 가동하고 있다. 1994년 10월 현재 목록 데이터베이스는 포항공과대학교, 산업과학기술연구소, 포항제철의 포항 및 광양연구소에서 소장하고 있는 92,000여 레코드, 141,000여 권의 장서로 구성되어 있으며, 근거리 통신망(LAN : Local Area Network)이 설치된 포항공과대학교와 산업과학기술연구소 및 기타 외부에서 INTERNET, DNS(Dacom Net Service), Dial-Up modem등을 통하여 언제든지 접근할 수 있다. LINNET의 username별 온라인목록 이용자는 LINNET(LINNET을 사용하는 포항공과대학교 및 각 분관의 사서), PLIS(포항공과대학교와 학술 정보 공동이용협정을 체결한 각 기관의 사서), BOOK(모든

일반이용자) 등 세 종류로 구분된다.

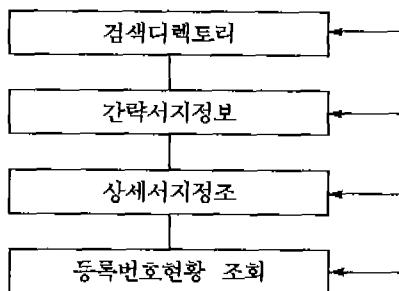
(1) 검색시스템의 구성

LINNET의 검색용 색인은 <표 1>과 같이 각 분관에서 공동으로 사용하는 10개의 공용색인('A'에서 'N'까지)과 분관별로 독립하여 유지, 운영되는 2개의 분관별 색인('R'과 'C') 등 총 12개의 검색키로 나누어져 있다.

<표 1> 검색키의 종류

검색키	내 용	검색키	내 용
A	저장명 색인	X	총서명 색인
AK	저자명 키워드 색인	XK	총서명 키워드 색인
T	서명 색인	P	출판사명 색인
TK	서명키워드 색인	N	서지통정번호 색인
S	주제명 색인	C	분류기호 색인
SK	주제명 키워드 색인	R	등록번호 색인

LINNET의 온라인목록 검색결과를 조회하는 단계는 아래와 같이 4단계로 구성되어 있다. 각 단계에서 이용자는 필요한 데이터를 선택, 조회하거나 프린트할 수 있다. 각 단계별 조회내용은 <그림 1, 2, 3, 4>와 같다.



〈그림 1〉은 이용자가 서명키워드를 이용하여 절단검색과 출판년도 제한을 병행 사용하여 검색하고, 저자명으로 절단검색한 내용 및 그 두가지 검색결과를 and(&) 연산자로 조합한 경우를 보여주고 있다.

T=서명	1. tk=polymer*//1990-	285
X=총서명	2. a=dyson*	13
A=저자명	3. 1&2	1
S=주제명		
P=출판사		
TK=서명 키워드		
XK=총서명 키워드		
AK=저자명 키워드		
SK=주제명 키워드		
N=서지통정번호		
R=등록번호		
C=분류기호		
검색제한 : 전당 300 Records		
연 산 자 : & (and) ! (or)^(not) 제한자 : //(출판년도, 소장기관, 형식부호)		
[명령] 2		
[안내] 검색문, No., C(lear), <-(delete), F1(일어), H(elp), Q(uit)		

〈그림 1〉 단행본검색 디렉토리

〈그림 2〉는 〈그림 1〉의 검색결과중 '[명령]'에서 2번을 선택하여 13건에 대한 간략서지정보를 조회한 화면이다. 저자명의 알파벳 순서로 서명, 출판년도, 참고사항을 보여주며, 참고사항에는 자관의 주제명 또는 타관자료여부를 표기하거나, 수서작업의 진행정도를 표시한다.

1	Dyson, A. E.	Education and democracy	1975	특수교육
2	Dyson, A. E.	Milton : Paradise Lost	1973	영문학
3	Dyson, B. F.	Mechanical testing of engineering	1989	타관자료
4	Dyson, Brain.	Modern academic library : essays	1989	도서관학
5	Dyson, David Jo	Interpretation of electron diffra	1971	타관자료
6	Dyson, Freeman	Symmetry groups in nuclear and pa	1966	물리학
7	Dyson, Freeman	Weapons and hope	1984	군사학
8	Dyson, Geoffrey	육상경기의 역학	1984	오락
9	Dyson, Geoffrey	GEOFFREY DYSON의 육상경기 역학	1985	오락
10	Dyson, N. A.	X-rays in atomic and nuclear phys	1990	물리학
11	Dyson, Norman.	Chromatographic integration metho	1990	화학
12	Dyson, R. W.	Engineering polymers	1990	공업
13	Dyson, R. W.	Specialty polymers	1987	화학

— end (13 records) —

[명령] 12

[안내] No., O(print), S(croll)=No., D(own), H(elp), Q(uit)

〈그림 2〉 단행본검색 간략서지정보

〈그림 2〉에서 확인하고자 하는 번호를 입력하면 해당서명의 상세서지정보 및 소장기관이 나타난다. 〈그림 2〉의 내용중 ‘[명령]’에서 12번을 선택하면 〈그림 3〉과 같이 해당문헌에 대한 상세한 서지정보와 함께 소장기관도 나타내준다.

서명저자 : Engineering polymers/edited by R. W. Dyson

출판사항 : Glasgow : Blackie ;

New York : Chapman and Hall, 1990.

형태사항 : 3 vols : ill. ; 24 cm.

서지주기 : Includes bibliographical references and index.

주 제 명 : Polymers.

Engineering - Materials

ISBN : 0412020815 (Chapman and Hall)

LCCN : 89015811

소장기관 : 1. 포항공과대학교 중앙도서관 2. 산업과학기술 연구정보실

[명령] 1 소장기관의 번호를 입력하십시오.

[안내] No., O(print), U(p), D(own), P(rev.), N(ext), F=NO., S(검색), H(elp), Q(uit)

〈그림 3〉 단행본검색 상세서지정보

<그림 3>에 나타난 상세서지정보 화면에 표기된 소장기관번호 중 '[명령]'에서 1번을 입력하면 <그림 4>와 같이 포항공과대학교 중앙도서관에서 소장하고 있는 해당도서의 등록번호가 권호순으로 나타나며, 각각의 소장위치와 대출, 파본처리, 분실 등 여러가지 현황이나, 반납예정일, 예약등 대출과 관련된 정보가 나타난다.

M0095713	V. 1	도서관 3층	대출중(941010-941020)	8723021	2
M0100428	V. 1, C. 2	도서관 3층	이용가능		
M0095140	V. 2	도서관 3층	분실신고중 (- 941203)	8724002	
M0100392	V. 2, C. 2	도서관 3층	예약보관중 (- 941110)	1102	1
M0100396	V. 3	도서관 3층	파본수리중(941027 -)		
M0100426	V. 3, C. 2	도서관 3층	이용가능		

[명령]
[안내] C(원문복사), B(간략), S(열람), H(elp), U(p), D(own), O(print), Q(uit)

<그림 4> 등록번호현황 조회 화면

(2) 검색방법

LINNET의 검색기능에서는 <표 2>와 같이 서명, 총서명, 저자명, 주제명과 그 각각에 대한 키워드등 총12개의 검색키에 대하여 "검색키=검색어"의 형태로 입력하여 검색한다.

<표 2> 검색 키

검색 키	검색문 사용예	검색 결과
T=서명	T=CHEMISTRY	서명이 chemistry인 도서
X=총서명	X=PHYSICS SERIES	총서명이 physics series인 도서
A=저자명	A=BOYCE, BETTY L.	저자명이 boyce, betty l. 인 도서
S=주제명	S=HYDRAULIC FLUIDS	주제명이 hydraulic fluids인 도서
P=출판사	P=PERGAMON	출판사명이 pergamon인 도서
TK=서명키워드	TK=CHEMISTRY	서명키워드가 chemistry인 도서

XK=총서명키워드	XK=PHYSICS	총서명키워드가 physics인 도서
AK=저자명키워드	AK=BOYCE	저자명키워드가 boyce인 도서
SK=주제명키워드	SK=HYDRAULIC	주제명키워드가 hydraulic인 도서
N=서지통정번호	N=0003-0213	ISBN이 0003-0213인 도서
R=등록번호	R=M0123456	등록번호가 M0123456인 도서
C=분류기호	C=HF1441	L, C 분류기호가 HF1441인 도서

저자명검색에서는 '성, 이름' 순으로 입력하며, 서지통정번호는 ISBN, ISSN, CODEN등의 통제번호로 구성된 데이터를 검색할 경우 이용한다.

모든 검색키에 "*"를 사용하여 우측절단을 할 수 있다. 즉, 'TK=PHYSICS*'로 입력하면 서명키워드가 PHYSICS인 데이터 및 'PHYSICS'로 시작하는 모든 데이터가 검색된다.

또한 모든 검색키에 "/"를 사용하여 <표 3>과 같이 여러가지 제한검색을 할 수 있다. 즉, 'TK=CHEM*//1992-'로 입력하면 서명중의 한 단어가 'CHEM'으로 시작하는 데이터 중 1992년 이후에 출판된 자료가 검색된다.

<표 3> 제한검색 유형 및 사용예

제한 유형	제한검색문 사용예	검 색 결 과
년도 제한	TK=CHEM*//1992 TK=CHEM*//1992- TK=CHEM*//-1992 TK=CHEM*//1990-1993	1992년에 출판된 도서검색 1992년 이후에 출판된 도서검색 1992년 이전에 출판된 도서검색 1990과 1993년 사이에 출판된 도서검색
소장 기관 제한	TK=CHEM*//POST	LINNET을 공동사용하는 분관중 분관코드가 'POST'인 기관(포항공과대학교)에서 소장하고 있는 도서만 검색
내용 형식 제한	TK=CHEM*//A	USMARC, KORMARC등의 TAG 008에서 사용되고 있는 내용형식부호중 "A"(초록)가 입력된 도서 검색

먼저 이루어진 검색결과를 <표 4>에서 보듯이 출판년도, 소장기관, 내용형식부호를 이용하여 다시 제한할 수도 있는데, 그 예는 다음과 같다.

<표 4> 검색결과와 제한

사용예	검 색 결 과
1//1992	"1"번 검색결과중 출판년도가 1992년인 도서만 다시 검색
1//RIST	"1"번 검색결과중 'RIST' (산업과학기술연구소)에서 소장중인 도서만 다시 검색
1//B	"1"번 검색결과중 내용형식부호가 'B' (서지자료)인 도서만 다시 검색

제한검색에서 사용되는 내용형식부호는 USMARC과 KORMARC의 Tag 008의 내용형식부호를 그대로 사용하였는데, 그 유형은 <표 5>와 같다.

<표 5> 제한검색의 내용형식부호

형식부호	내 용	형식부호	내 용
A	초록 (Abstracts/summaries)	P	프로그램 텍스트
B	서지 (Bibliographies)		(Programmed texts)
C	책지목록 (Catalogs)	Q	필름해설서
D	사전 (Dictionaries)		(Filmographies)
E	백과사전 (Encyclopedias)	R	명감 (Directories)
F	편람 (Handbooks)	S	통계 (Statistics)
G	법률기사 (Legal articles)	T	기술보고서
I	색인 (Indexes)		(Technical reports)
J	족보	V	판례 및 판례집
K	음반해설서 (Discographies)		(Legal cases)
L	법령집 (Legislation)	W	법률보고서
N	문헌분석서 (Surveys of the literature)		(Law reports and digests)
O	평론 (Reviews)	Y	연감 (Yearbooks)

논리연산은 & (AND), ! (OR), ^ (NOT) 등과 같은 논리연산자를 사용하여 검색할 수 있게 하였으며, 또 이 연산자를 이용하여 검색결과도 다시 연산할 수 있다. '&', '!', '^'의 기호를 이용하여 논리연산을 할 때에는 기호만 사용해도 되지만, 'AND', 'OR', 'NOT'의 문자를 사용할 때는 반드시 앞뒤에 1 바이트 이상의 공란이 있어야 한다. 논리연산의 사용에는 <표 6>과 같다.

〈표 6〉 논리연산자 사용 예

검색문 사용예	검 색 결 과
T=PHYS* ! A=TOM	서명이 PHYS로 시작하거나, 혹은 저자가 TOM인 모든 자료를 검색
T=PHYS* & SK=P*	서명이 PHYS로 시작하고, 동시에 주제명키워드가 P로 시작하는 자료를 검색
T=PHYS* ^ SK=P*	서명이 PHYS로 시작하는 자료중, 주제명키워드가 P로 시작하지 않는 자료를 검색

위의 검색결과를 각각 1, 2, 3으로 간주하고, 그 결과를 논리연산하는 예를 보면 <표 7> 과 같다.

〈표 7〉 검색결과 논리연산

1 AND 2	1번의 검색결과와 2번의 검색결과 중 공통자료를 검색
1 OR 2	1번이나 2번의 검색결과를 모두 검색
1 ^ 2	1번의 검색결과중 2번의 검색결과를 제외한 자료를 검색
1 ! 2 ^ 3	1번과 2번 검색결과를 합한 내용중 3번의 검색결과를 제외한 자료를 검색

(3) 검색 transaction log의 구성

LINNET에서는 시스템 관리자가 각 이용자들의 검색내용과 시스템의 성능을 상세히 분석할 수 있도록 transaction log를 유지하고 있다. 모든 이용자의 검색활동을 재구성할 수 있도록 충분한 정보를 제공하는 이 파일의 raw data는 아래와 같이 복잡한 문자와 숫자 및 기호로 구성되어 있다. 각 필드 사이에 있는 '@@'는 필드를 구분하는 식별자로 사용된다.

〈raw data의 예〉

```
^OAH(940820, 1. 03) = OM1@@tk=encyclo*@@55955, 31441@@POST@@USER
@@250@@25@@15@@55955, 31470@@1
```

Transaction log의 파일명은 '^OAH'이고, 각각의 필드는 가변장(Variable field)으로 구성되어 있다. 1차 구분자로 이용자가 검색기능에 로그인한 일자(예 : 940820은 1994년 8월 20일에 로그인 했음을 의미함)를 사용하고, 2차 구분자로 로그인 차수 및 차수별 검색시도 횟수를 보관하고 있다. (예 : 1. 03은 해당일에 처음으로 검색기능에 로그인 하였고, 세번째 생성된 transaction log를 의미) 나머지 필드를 데이터를 저장하고 있는데 데이터의 내용을 구체적으로 분석하면 <표 8>과 같다.

〈표 8〉 검색 transaction log의 구성

필드 구분	내용
이용한 기능명	이용한 기능명. 예) OM1 (일반이용자용의 검색 기능)
검색내용	이용자의 입력내용. 예) TK=encyclo*
검색시작 일시	검색시작시 컴퓨터에서 부여하는 년월일시분초정보. 예) 55955, 31441
소속기관 코드	각 분관별로 부여한 소장기관 코드 예) POST (포항공과대학교 중앙도서관)

username	LINNET 로그인시 사용한 username구분자. 예) user
검색건수	검색결과 건수. 예) 250
상세서지정보 열람건수	검색결과내에서 상세서지 정보 열람건수. 예) 25
등록번호 조회 횟수	상세서지정보 열람후 등록번호 현황 열람건수 예) 15
검색완료 일시	검색완료시 컴퓨터에서 부여하는 년월일시분초 정보. 예) 55955, 31470
검색자 신분코드	검색에 로그인한 이용자의 신분. 예) 1 : 학부생, 2 : 대학원생 등

III. Transaction log 분석

Transaction log의 분석대상 기간은 이용자들이 비교적 검색기능을 활발하게 이용한 학기초인 1994년 8월 20일부터 중간고사 이전인 동년 10월 10일까지로 하였다. 이 52일 동안 생성된 총 40,627건의 데이터를 대상으로 결과분석에 필요한 여러가지 프로그램을 별도로 작성하여 통계처리하였다.

이용자가 검색기능에 로그인할 때마다 각자의 신분에 해당하는 번호를 선택하게 하여 신분을 구분하였다. 이용자의 신분은 그 소속기관에 관계없이 학부생, 대학원생, 교수, 연구원, 직원, 사서 및 기타의 일곱가지 신분으로 나누었다.

정확한 이용자의 검색시스템 이용행태를 파악하기 위하여 12종류의 검색키에 대해 사용한 절단 및 비절단 검색, 제한자 3종, 논리연산자 3종 등 모든 유형을 대상으로 분석하여, 검색시스템에 로그인하는 빈도와 로그인후의 검색시도 횟수등을 조사하였다.

절단검색, 연산자, 제한자는 온라인 목록검색에서 독립적으로 그 역할을 수행하므로 검색키(색인)와는 별도로 개별적인 변수로 취급하여 분석하였다.

검색결과에 대한 만족도를 조사하기 위하여 이용자의 검색결과에 대한 상세서지정보 및 등록번호별 현황 조회횟수를 조사하였다. 간략서지정보에서 상세서지정보로 상세서지정보에서 등록번호 현황 조회로 디스플레이할수록 이용자들은 검색결과에 대한 만족도가 높아지는 것으로 추정하였다. 검색결과가 0건인 경우를 검색실패로, 1건 이상의 경우를 검색성공으로 간주하고 성공률과 실패율을 조사한 후, 검색실패의 원인을 분석하였다. 또한 검색횟수 및 hit건수를 중심으로 검색기능의 효율성을 판단하였으며, 제한자 및 연산자의 사용빈도별 hit건수를 비교, 분석하여 일반이용자의 검색기능 이해정도를 조사하였다.

1. 검색시스템 이용빈도

(1) 검색기능 로그인 횟수

조사대상기간인 총 52일에 대하여 이용자들의 시스템 로그인 횟수를 조사하여 보았다. <표 9>에서 보듯이 조사기간중 시스템에 로그인한 횟수는 총 11,218회로, 하루 평균 로그인 횟수는 216회였다.

<표 9> 검색기능 로그인 횟수

Username	LINNET	PLIS	BOOK	계
로그인 횟수	465	59	10694	11218

Username별로는 <표 9>에서와 같이 BOOK(일반이용자용)이 10,694회(95.33%), LINNET(사서용)이 465회(4.15%), PLIS(협정기관용)가 59회(0.52%) 로그인하여, 주로 BOOK이 많이 이용하는 것으로 나타났다.

이용자들의 시간대별 이용빈도는 오후 3시에서 오후 4시 사이가 1,192회(10.63%)로 가장 많이 이용하였으며, 오후 2에서 오후 3시 사이가 1,165회(10.39%), 오후 4시에서 오후 5시 사이에는 1,138회(10.14%) 이용하며, 결국 하루중 오후 2시부터 오후 5시 사이에 시스템을 가장 많이 이용한 것으로 나타났다. 또한 포함

공과대학교 도서관 이용자들의 주 활동시간대인 오전 9시부터 오후 6시 사이의 로그인인 총 8,418회(75.04%)로 나타났으며, 나머지 시간대에는 2,800회(24.96%)로 나타났는데, 이는 각 username별로도 동일한 경향을 보이고 있다. 특히 일반이용자용 username인 BOOK은 0시에서 7시 사이에도 계속 이용한 것으로 나타났는데, 이는 포항공과대학교의 기숙사에서 검색이 가능하도록 Network이 구축되어 있는 까닭으로 보인다.

LINNET 이용자들의 로그인 활동을 신분별로 구분하면, 학부생이 6,018회(53.65%), 대학원생이 2,950회(26.30%), 연구원 565회(5.04%) 순으로 나타났다. 각 신분별 로그인 시간대가 조금씩 다르게 나타났는데, 학부생과 대학원생은 오후 1시 부터 5시 사이에 가장 많이 로그인 하였고, 또한 대학원생들이 소수이지만 꾸준히 0시에서 8시사이에도 계속하여 로그인 하는 것으로 나타났다. 특히 학부생과 대학원생은 새벽 0시부터 포항공과대학교 도서관이 문을 닫는 새벽 2시까지도 많이 이용하는 것으로 나타나 학생들의 생활행태를 어느정도 파악할 수 있게 하였다.

(2) 요일별 검색기능 로그인 횟수

요일별 로그인 횟수를 <표 10>과 같이 신분별로 구분하여 조사하였다.

(표 10) 요일별 검색기능 로그인 횟수

요일	학부	대학원생	교수	연구원	직원	사서	기타	계
일	609	145	13	6	9	2	17	802
월	953	570	44	100	62	81	82	1893
화	909	467	30	97	54	81	88	1726
수	934	476	36	76	74	92	98	1786
목	1062	514	36	95	81	99	81	1968
금	908	502	30	133	69	87	116	1845
토	643	276	41	58	44	53	83	1198
계	6018	2950	230	565	393	49	565	11218

모든 신분의 로그인 횟수를 종합하여 보면 월요일에서 금요일 사이는 큰 차이없이 (1,726회~1,968회) 로그인 하는 반면, 토요일은 1,198회로 횟수가 떨어졌다. 그러나 일요일에도 802회나 로그인하여 평일의 절반정도를 기록하였다. 특히 학부생의 경우는 토요일과 일요일의 로그인 횟수가 평일의 65% 수준으로 검색시스템을 상당히 꾸준하게 이용하는 것으로 나타났다.

(3) 신분별 검색 횟수

조사대상 기간동안 총 40,627회의 검색이 이루어졌는데, 이는 매일 평균 781회씩 검색한 셈이다($40,627 \div 52 \text{일} = 781 \text{회 로그인/1일}$).

〈표 11〉 신분별 검색 횟수

신분	학부	대학원생	교수	연구원	직원	사서	기타	계
계	20986	10981	849	2221	1803	1563	2216	40627

이 기간동안의 총 로그인 횟수가 11,218회(〈표 10〉 참조)였으므로, 1회 로그인 시 평균 3.62회 검색한 결과가 된다($40,627 \div 11,218 \text{회 로그인} = 3.62 \text{회 검색/1회 로그인}$). 신분별로는 학부생 20,986회(51.66%), 대학원생 10,981회(27.03%), 연구원 2,221회(5.47%) 순이었고, 시간대별 검색활동을 살펴보면 로그인 횟수가 많아지는 시간대와 마찬가지로 오후 1시에서 5시 사이의 검색이 15,466회(38.07%)로 가장 많이 검색하는 것으로 밝혀졌다.

(4) 신분별 1회 로그인 시의 평균검색 횟수

검색시스템 이용자가 한번 로그인하여 몇번 검색을 시도하고 로그아웃하는가를 신분별로 조사해 보았다.

〈표 12〉 신분별 1회 로그인시의 평균검색 횟수

구 분	학부생	대학생	교 수	연구원	직 원	사 서	기 타	계
로그인횟수	6018	2950	230	565	393	495	565	11218
검색 횟 수	20986	10981	849	2221	1803	1563	2216	40627
로그인시 평균검색수	3.49	3.72	3.69	3.93	4.59	3.16	3.92	3.62

〈표 12〉에서 보는 바와 같이 대학내 일반직원의 평균 검색회수가 4.59회로 가장 높게 나타났으며, 반대로 사서의 평균검색회수가 3.16회로 가장 낮게 나타났다. 이 수치는 검색시스템에 익숙한 이용자들이 비교적 적은 검색회수로도 쉽게 소기의 목적을 달성하고 있음을 보여주고 있다. 전반적으로는 큰 차이없이 전체 평균인 3.62회에 근사한 검색을 시도하는 것으로 나타났다.

(5) 시간대별 로그인 횟수와 검색 횟수

검색시스템에 로그인하여 검색을 시도한 횟수를 시간대별로 조사하여 보았다. 이 조사는 〈표 12〉의 내용과 유사하지만 어느 시간대에 많이 검색을 시도하는지를 알아보고자 하는 의도였다. 조사결과, 가장 많이 이용한 시간대인 오전 9시부터 오후 9시까지의 로그인횟수(10,080회)와 검색횟수(36,702회)를 대비해 보았더니 3.64회로 평균(3.62회)과 큰 차이가 없는 것으로 나타났다. 다만 새벽 3시부터 6시사이 즉, 도서관이 문을 닫은 이후의 시간에 검색시스템에 로그인 한 이용자는 평균 4.58회로 낮 시간대에 비해 비교적 여러번 검색을 시도한 것으로 나타났다.

2. 검색실패율과 검색키 이용빈도

(1) 검색실패율

신분별로 0건 hit 발생횟수(검색실패율)를 조사해 보았다.

〈표 13〉 신분별 0건 hit 발생 횟수

구 분	학부생	대학원생	교수	연구원	직원	사서	기타	계
검색횟수	20986	10981	849	2221	1803	1563	2216	40627
0건hit횟수 (%)	10108 (52.91)	4981 (26.07)	369 (1.93)	1108 (5.80)	795 (4.16)	517 (2.71)	1225 (6.42)	19103 (100)

〈표 13〉에서 보듯이 총 검색횟수 40,627회 중 검색결과가 0건 hit로 나온 경우는 19,103회로 평균 47.02%의 검색실패율을 기록하였다. 신분별로는 학부생이 52.91%의 실패율을 보여 가장 많이 검색에 실패하는 것으로 나타난 반면, 사서가 2.71%, 교수가 1.93%로 낮은 실패율을 보였다. 이는 LINNET의 검색시스템에 익숙한 정도가 성공률과 실패율을 결정하는 중요한 요인인 것으로 여겨진다.

(2) 신분별 검색키 사용횟수

이용자의 신분별로 어떤 검색키(색인)를 많이 이용하는가를 조사하였다. 검색결과와의 조합 또는 연산인 경우(예 : 1 & 2 또는 2//1990- 등의 유형임)는 검색키를 입력하지 않으므로 조사대상에서 제외하고, 검색키를 입력한 경우만을 조사대상으로 하였다.

〈표 14〉 신분별 검색키 사용 횟수

검색키	학부	대학원생	교수	연구원	직원	사서	기타	계
A	1367	726	92	120	145	40	200	2690
AK	1090	989	59	70	30	14	49	2302
C	54	38	14	1	10	63	2	182
N	18	8	0	7	1	11	0	45
P	254	176	21	33	33	35	66	618
R	69	33	6	10	20	110	12	260
S	1044	336	42	112	67	9	133	1743
SK	1470	759	89	267	69	35	144	2833
T	4653	2052	213	388	661	695	488	9150

TK	10305	5391	302	1136	676	508	1038	19360
X	205	147	14	36	25	24	22	473
XK	300	224	8	43	15	10	20	620
계	20829	10879	860	2223	1752	1554	2174	40276

각 검색키의 총 사용횟수를 살펴보면 <표 14>에서 보듯이 서명키워드가 19,360회(48.07%)로 가장 많이 사용되었고, 서지통정번호가 45회(0.11%)로 가장 적게 사용되었다.

검색키의 사용빈도를 각 신분별로 살펴보면 다음과 같다.

학 부 생 : TK > T > SK > A > S > AK > XK > P > X > R > C > N

대학원생 : TK > T > AK > SK > A > S > XK > P > X > C > R > N

교 수 : TK > T > A > SK > AK > S > P > C = X > XK > R > N

연 구 원 : TK > T > SK > A > S > AK > XK > X > P > R > N > C

직 원 : T > TK > A > SK > S > P > AK > X > R > XK > C > N

사 서 : T > TK > R > C > A > P = SK > X > AK > N > XK > S

기 타 : TK > T > A > SK > S > P > AK > X > XK > R > C > N

직원과 사서를 제외한 모든 이용자들은 서명키워드를 가장 많이 이용하였고, 등록번호(R), 서지통정번호(N), 분류기호색인(C)은 매우 저조한 이용률을 보이고 있다. 또 교수와 사서들은 다른 신분에서는 저조한 이용률을 보인 분류기호색인(C)을 상대적으로 많이 이용하는 것으로 나타났다. 특히 서명과 서명키워드의 사용횟수는 28,510회로 전체 이용횟수의 70.79%를 차지한 반면, 그 다음으로 많이 이용된 저자명과 저자명키워드가 4,992회(12.39%) 및 주제명과 주제명키워드가 4,576회(11.36%)의 이용률을 보이고 있어, 도서의 제목과 관련된 정보를 이용하여 대부분의 검색이 이루어지는 것으로 나타났다. 또한 서명, 저자명, 주제명의 세가지의 정보를 이용한 검색이 전체검색의 94.54%를 차지하고 있는 것으로 나타나 서지데이터에서 추출한 출판사, 서지통정번호 및 총서명과 관련된 검색키는 일

만이용자들에게 있어서는 큰 의미가 없는 것으로 판단된다.

(3) 검색키별 검색실패율과 성공률

어느 검색키가 검색률이 높고 낮은가 즉, 각 검색키별 검색성공률과 실패율을 조사하였다.

〈표 15〉 검색키별 검색실패율과 성공률

검색키	0건 hit	1건이상 hit	계	실패율(%)	성공률(%)
A	1388	1233	2621	52.96	47.04
AK	635	1607	2242	28.32	71.68
C	117	55	172	68.02	31.98
N	45	0	45	100.00	0.00
P	197	380	577	34.14	65.86
R	68	192	260	26.15	73.85
S	1246	449	1695	73.51	26.49
SK	1607	966	2573	62.46	37.54
T	3797	4776	8573	44.29	55.71
TK	7346	10739	18085	40.62	59.38
X	377	71	448	84.15	15.85
XK	469	126	595	78.82	21.18
계	17292	20594	37886	45.64	54.36

연산자를 사용한 경우(예 : tk=physics and ak=adams or sk=nuclear)는 2개이상의 검색키를 사용하므로 어느 검색키로 검색에 성공 또는 실패하였는지 판단하기가 곤란하여 조사대상에서 제외하였고, 제한자를 사용한 경우도 다양한 제한자의 사용으로 성공과 실패율의 조사에 영향을 미칠 것으로 판단되어 조사대상에서 제외하였으며, 그리고 검색결과를 이용하여 제한 또는 연산한 경우도 같은 이유로 제외하였다. 본 조사는 연산자 및 제한자를 사용하지 않은 단순검색 형태만(예 : tk

=physics 또는 ak=adams등)을 대상으로 하였다.

조사대상 총 37,886건을 분석해 본 결과 <표 15>에서와 같이 0건 hit된 경우가 17,292회였고, 1건 이상 hit된 경우가 20,594회로 나타나 검색성공률이 54.36%, 검색실패율이 45.64%를 기록하므로써 검색실패율이 상당히 높은 것으로 나타났다. 이를 각 검색키별로 보면 검색성공률은 등록번호가 73.85%, 저자명 키워드가 71.68%, 출판사명이 65.86% 순으로 나타났고, 검색실패율은 서지통정번호가 100.00%, 총서명이 84.15%, 총서명키워드가 78.82%, 주제명이 73.51%, 청구번호가 68.02% 순으로 높게 나타났다. <표 14>의 검색키별 사용횟수에서는 서명키워드, 서명, 주제명키워드 등의 순으로 이용빈도가 높게 나타났으나, <표 15>의 검색키별 검색성공률에서는 서명키워드, 서명 및 주제명키워드 등의 성공률이 그리 높지 않은 것으로 나타나 결국, 이용빈도가 높은 검색키가 반드시 성공률이 높은 것은 아님을 나타내 주었다.

(4) 단어 사용 횟수

이용자들이 검색시 1개의 단어를 입력하여 단순한 형태의 검색을 시도하는지 또는 여러개의 단어를 동시에 입력하여 복잡한 형태의 검색을 시도하는지를 파악하고자 검색문 입력시 사용한 단어의 갯수를 공란(blank)을 기준으로 조사하여 보았다.

<표 16> 단어 사용 횟수

구 분	1개	2개	3개	4개	계
발생횟수	37745	1013	40	11	38809
hit 건수	1199709	8798	494	399	1209400
단어 1개당 평균 hit수	31.78	4.34	4.12	9.07	30.28

정확한 조사를 위하여 검색결과와 조합 또는 연산에 해당되는 검색문(예 : 1 & 2 또는 2//1990-1992) 및 검색키만 입력하고 검색단어 입력부분에 공란 또는 특수 문자를 입력한 경우(예 : 'tk= '의 경우임)는 조사대상에서 제외하였다.

그 결과 1회 검색시 사용된 단어는 <표 16>에 나타난 바와 같이 최대 4개였고, hit건수는 단어를 1개 사용할 경우에는 평균 31.78건, 2개를 사용했을 경우에는 평균 4.34건으로써, 단어수가 많아질수록 hit건수는 줄어드는 것으로 나타났다. 이는 2개이상의 단어를 사용하는 대부분의 이용자가 or연산자 보다는 and연산자를 사용한 것으로 추정된다. 검색단어를 4개 사용한 이용자는 전체평균과 매우 다른 내용을 보이고 있어 특이하다. 특히 단어를 1개만 사용한 경우가 37,745건으로 전체 발생횟수(38,809회)의 97.26%를 차지하여 대부분의 이용자가 검색키를 선택한 후 단어 1개를 입력하는 단순한 검색방법을 사용하는 것으로 판단된다.

(5) 검색문의 길이

LINNET시스템은 이용자가 검색문을 입력할 때 최대 40 byte까지 입력할 수 있도록 설계되어 있다. 이러한 환경에서 이용자가 입력한 검색문의 길이를 5byte 간격으로 구분하여 각각의 발생빈도 및 입력유형을 파악하였다. 조사대상에서 제외하는 검색문의 유형은 <표 16>의 방법을 적용하였다.

<표 17> 검색문의 길이

구 분	발 생 빈 도	발 생 률 (%)
1 - 5 바이트	12039	31.02
6 - 10 바이트	15728	40.53
11 - 15 바이트	6224	16.04
16 - 20 바이트	2905	7.49
21 - 25 바이트	1124	2.89
26 - 30 바이트	615	1.58
31 - 35 바이트	167	0.43
36 - 40 바이트	7	0.02
계	38809	100.00

그 결과 6 내지 10 바이트를 사용한 경우가 40.53%로 가장 많이 발생하였고 그 다음이 1 내지 5 바이트(31.02%), 11 내지 15 바이트(16.04%)로 나타나, 총검색 38,809회중 87.59%가 1 내지 15 바이트 이내에서 이루어진 것으로 나타났다. 또한 이용자들이 검색문의 길이를 30 byte 이내로 입력한 경우가 전체의 99.55%(38,635건)를 차지하였으므로, 검색시스템의 개발담당자가 검색용 색인을 만들거나, 이용자의 검색문 입력허용 범위를 제한할 때 30 byte로 설계해도 사용에 지장이 없을 것으로 판단된다.

3. 검색결과와의 조회

(1) 상세화면 조회율

〈표 18〉에 나타난 바와 같이 총 hit건수 1,219,089회중 상세화면 조회건수는 24,566회로, 검색결과건수에 대하여 평균 2.01%의 상세화면 조회율을 나타냈다.

〈표 18〉 상세화면 조회율

hit건수 (a)	상세화면 조회 건수 (b)	상세화면조회비율(%) ($b \div a \times 100$)
1219089	24566	2.01

(2) 등록번호화면 조회율

〈표 18〉에서 설명한 상세화면을 조회한 후 등록번호현황 화면을 몇회 조회하였는가를 조사하였다. 이용자의 등록번호현황 조회목적은 상세화면을 조회하고 이용자에게 필요한 도서로 판단하여, 각 도서관(분관)별 소장위치 및 청구기호를 확인하고 열람 또는 대출가능성을 확인하기 위한 것으로 추정된다. 등록번호화면 조회율을 확인함으로써 이용자의 검색결과에 대한 최종 만족도를 판단할 수 있을 것이다.

〈표 19〉 등록번호화면 조회율

상세화면조회건수 (a)	등록번호화면 조회 건수 (b)	등록번호조회비율(%) (b ÷ a × 100)
24566	15789	64.27

조사대상기간중 이용자들은 총 24,566회 상세화면을 조회한 후, 15,789회 등록번호화면을 조회하여 평균 64.27%의 조회율을 나타내었다.

(3) 검색키별 상세화면 조회율

이용자가 검색문 입력시 사용한 각 검색키별 상세화면 조회율을 조사하여, 어느 검색키를 사용한 경우에 가장 만족스러운 검색결과를 얻는지 조사하였다. 〈표 15〉에서 설명한 것과 같이 정확한 조사를 위하여 연산자나 제한자 검색을 한 경우는 포함시키지 않았다.

〈표 20〉 검색키별 상세화면 조회율

검색키	hit 건 수	상세화면 조회건수	상세화면 조회율(%)
A	25392	1183	4.66
AK	47974	1592	3.32
C	3722	26	0.69
N	0	0	0.00
P	43293	357	0.82
R	192	191	99.48
S	24698	493	2.00
SK	80727	1520	1.88
T	108220	5662	5.23
TK	850162	12500	1.47
X	4465	155	3.47
XK	10168	49	0.48
계	1199013	23728	1.98

〈표 20〉과 같이 총 1,199,013건을 검색하고 23,728건의 상세화면을 조회하여 평균 1.98%의 조회율로 나타났고, 각 검색키별로는 등록번호(R)가 99.48%로 가장 높은 조회율을 보인 반면, 서지통정번호(N)는 0.00%로 나타나, 이용자가 서지통정번호(예: ISBN, LCCN, ISSN, CODEN 등)에 익숙하지 않기 때문인 것으로 밝혀졌다. 서명(5.23%)과 저자명(4.66)을 이용한 검색결과의 상세화면 조회율이 타 검색키를 이용한 조회율보다 높게 나타났다. 특히 서명(T), 저자명(A), 주제명(S), 총서명(X)으로 검색하는 경우에는 각각의 키워드색인을 이용한 검색에 비하여 hit건수는 적었으나, 상세화면 조회율은 상대적으로 높은 비율을 보여, 키워드색인을 이용한 검색보다 일치색인(Exact Index)을 이용한 검색에서 높은 만족도를 보이는 것으로 나타났다.

4. 절단, 제한 및 연산자 이용행태

(1) 검색키별 절단검색 이용빈도

〈표 14〉와 동일한 방법을 적용하여 특정 검색문의 유형은 제외하고, 연산자를 사용하지 않은 단순검색의 경우를 대상으로 검색키별 절단검색의 이용빈도를 조사하였다.

〈표 21〉 검색키별 절단검색 이용빈도

검색키	절단검색횟수	비절단검색횟수	계	절단검색률(%)
A	723	1967	2690	26.88
AK	299	2003	2302	12.99
C	78	104	182	42.86
N	2	43	45	4.44
P	107	511	618	17.31
R	0	260	260	0.00
S	384	1359	1743	22.03
SK	500	2333	2833	17.65
T	3803	5347	9150	41.56

TK	3001	16359	19360	15.50
X	122	351	473	25.79
XK	113	507	620	18.23
계	9132	31144	40276	22.67

이용자들의 총 40,276회 검색 횟수중 절단검색 횟수는 9,132회(22.67%)로 나타났다. 검색키별 절단률을 보면, 청구번호 및 서명을 이용하여 검색한 경우가 각각 42.86% 및 41.56%로 타 검색키에 비하여 절단기능을 많이 사용하였다. 특히 서명(T), 저자명(A), 주제명(S), 종서명(X)으로 검색하는 경우에 다른 키워드색인을 이용한 경우보다 절단검색을 많이 사용한 것으로 나타났다.

(2) 검색키별 제한자 이용빈도

출판년도, 소장기관, 형식부호등의 제한자 사용빈도를 각 검색키별로 구분하여 어느 검색키에서 제한자를 가장 많이 사용하는가 조사하였다.

〈표 22〉 검색키별 제한자 이용빈도

검색키	제한자사용	제한자비사용	계	제한자사용률(%)
A	12	2678	2690	0.45
AK	16	2286	2302	0.70
C	44	138	182	24.18
N	1	44	45	2.22
P	36	582	618	5.83
R	0	260	260	0.00
S	6	1737	1743	0.34
SK	41	2792	2833	1.45
T	18	9132	9150	0.20
TK	104	19256	19360	0.54
X	2	471	473	0.42
XK	8	612	620	1.29
계	288	39988	40276	0.71

이용자들은 전체 검색건수 중 0.71%만 제한자를 이용한 것으로 나타나, 절단검색의 22.67%에 비하여 이용률이 매우 낮았다. 검색키별 제한자 이용률을 보면, 분류기호(C)의 24.18%, 출판사(P)의 5.83%를 제외한 대다수의 검색기에서 낮은 이용률을 보이고 있다.

(3) 연산자 이용빈도

LINNET의 검색시스템에서 제공하는 and(&), or(!), not(^)의 연산자중 이용자들이 어떤 연산자를 가장 많이 사용하였는가를 조사하였다.

〈표 23〉 연산자 이용빈도

연산자유형	발생횟수(%)	hit건수	평균 hit 건수
and (&)	1680 (91.35)	7923	4.72
or (!)	140 (7.61)	8685	62.04
not (^)	19 (1.03)	1042	54.84
계	1839 (100.00)	17650	9.60

조사기간동안 이루어진 총 40,627건의 검색 중 1,839회에 걸쳐 연산자를 사용(4.52%), 매 22회 검색시마다 평균 1회 연산자를 사용한 것으로 나타났다. 연산자의 유형별 사용횟수는 and연산자가 1,680회로 가장 많이 사용(91.35%)되었고, 그 다음이 or연산자(140회, 7.61%), not연산자(19회, 1.03%)의 순이었다. and 연산자의 사용률이 or연산자의 사용률보다 아주 높은 까닭은 이용자들이 포괄적인 검색보다는 원하는 정보를 정확하게 검색하고자 하는 의도를 나타낸 것으로 여겨진다. 또한 〈표 22〉에서 설명한 바와 같이 제한자 사용률이 0.71%인 점과 〈표 23〉의 연산자 이용률이 4.52%에 불과한 점을 고려할 때 이용자는 아주 단순한 검색을 선호하는 것으로 판단된다.

연산자를 사용할때 어떤 검색키를 이용하여 상호연산검색을 하는가를 살펴보면, 2개의 검색키를 연산자로 조합한 경우는 총 1,009건이 발생하였다. 서명키워드씩

인(TK)을 and(&) 연산자로 조합한 경우가 508회로 가장 많았으며, 서명색인(T)을 and(&) 연산자로 조합한 경우가 239회, 주제명키워드색인(SK)을 and(&) 연산자로 조합한 경우가 83회의 순으로 나타났다. 이와는 대조적으로 서지통정번호색인(N)과 등록번호색인(R)은 전혀 조합이 이루어지지 않은 것으로 나타났다. <표 23>에서 설명한 바와 같이 이용자들이 and연산자 중심으로 검색을 하였고, 특이한 점은 서로 다른 검색키를 연산한 경우(예 : tk=physics & ak=adams)는 아주 적게 나타나, 이용자는 1회 검색시 2종류 이상의 상이한 검색키를 사용하기 보다는 같은 종류의 검색키를 이용하는 검색(예 : 서명)을 선호하는 것으로 나타났다.

5. 검색실패의 원인

검색실패의 원인을 살펴보면 <표 24>와 같이 데이터베이스에 없는 자료를 검색한 경우가 15,764회(79.42%)로 가장 많은 부분을 차지했고, 그 다음이 검색키 지정 오류(18.95%), 명령어 구축 실패(1.63%)의 순으로 나타났다.

<표 24> 검색실패의 원인

원 인	건 수 (%)
데이터베이스에 없는 자료	15,764 (79.42)
검색키 지정오류	3,761 (18.95)
명령어 구축 실패	324 (1.63)
계	19,849 (100.00)

‘데이터베이스에 없는 자료’는 검색문은 정상으로 입력하였으나 hit건수가 없는 경우이었다. 예를 들어 우측절단을 이용하여 ‘T=한국의 역*’을 입력하면 서명이 ‘한국의 역’으로 시작하는 ‘한국의 역사’, ‘한국의 역할과 위상’등 여러가지 문헌들이 hit되지만, 일치검색인 ‘T=한국의 역’을 입력하여 hit건수가 0건인 경우가 이에 해당된다.

‘검색키 지정오류’는 시스템에서 제공하지 않는 검색키를 입력한 경우로, 두가지 유형으로 구분할 수 있다. 첫째는 이용자가 시스템에서 제공하지 않는 검색키를 입력하는 경우이고(예, j=한국의 나방), 두번째는 한/영문 모드의 선택의 오류로 인하여 영문으로 구성된 검색키를 한글코드값으로 입력하게 되는 경우이다. 그 대표적인 사례를 보면 <표 25>와 같다.

<표 25> 한, 영문 지정 전환 오류의 예

입 력 내 용	실 제 내 용
□ = 이청준	A = 이청준
ㅅ = 이어도	T = 이어도
ㅌ = 한국대표문학전집	X = 한국대표문학전집

‘명령어 구축 실패’는 의미없는 특정 코드값을 입력하여 검색에 전혀 사용되지 않는 경우(예 : \$#^^^* & ^\$#ㄴㅇㄹ)였다.

IV. 결 론

본 연구는 포항공과대학교 중앙도서관의 전산시스템인 LINNET의 온라인목록 transaction log를 분석하여 이용자들의 검색시스템 이용행태를 조사, 분석하고, 그 결과를 토대로 온라인목록 검색시스템의 개발시 참고자료로 활용하며, 동시에 LINNET시스템의 피이드백을 위한 기초자료로 이용하기 위해서이다. 본고에서 조사된 내용을 요약하면 다음과 같다.

첫째, LINNET 이용자들은 1994년 8월 20일부터 동년 10월 10일까지 총 11, 218회 LINNET의 온라인 목록검색기능에 로그인하여 총 40,627건의 검색 transaction log데이터를 남겼다. 이는 하루 평균 216회 로그인을 하여 781회의

검색을 한 것으로써, 한번 로그인할 때마다 평균 3.62회의 검색을 한 것이다. 이용자들은 오후 2시에서 오후 5시 사이에 시스템에 가장 많이 이용하였으며, 검색방법을 숙지한 이용자들이 비교적 적은 횟수의 검색으로도 만족할만한 결과를 얻는 것으로 판단되었다.

둘째, 이용자들은 LINNET의 12개 검색키중 서명키워드를 가장 많이 이용하며, 등록번호, 서지통정번호 및 분류기호는 거의 이용하지 않는 것으로 나타났다. 특히 서명, 서명키워드, 저자명, 저자명키워드, 주제명 및 주제명키워드를 이용한 검색이 전체의 94.54%를 차지하였다.

셋째, LINNET의 온라인 목록 검색결과 중 1건 이상이 검색되는 검색성공률은 52.98%였다. 12종의 검색키중 검색성공률이 가장 높은 것은 등록번호, 저자명키워드였으며, 가장 낮은 것은 서지통정번호였다. 검색성공률은 등록번호, 저자명키워드, 출판사명의 순으로 나타나 이용빈도가 높은 검색키가 반드시 검색성공률이 높은 것은 아니었다. 이용자들은 한번 검색시스템에 로그인한 후 평균 4회 검색을 시도하고, 1회당 약 30건의 데이터를 검색하여, 한번 로그인할 때마다 약 120건의 검색결과를 얻는 것으로 나타났다. 이용자의 검색문중 30 byte이내로 입력된 경우가 99.55%를 차지한 것으로 나타나 검색시스템의 색인을 만들거나 검색문 입력허용 범위를 결정할 때 30byte를 기준으로 설정해도 무난할 것으로 판단되었다.

넷째, 이용자들이 검색후 상세서지정보 화면을 조회하는 비율은 평균 2.01%로 매우 낮은 반면, 상세서지정보를 확인한 후 등록번호현황을 조회하는 비율은 64.27%로 상당히 높은 것으로 나타났다.

다섯째, LINNET의 온라인 목록 이용자들은 전체 검색횟수 중 22.67%가 우측 절단기능을 이용한 것으로 나타났고, 제한자 및 불리언 연산자는 거의 이용하지 않는 것으로 나타났다. 서명을 이용하여 검색할 때 우측절단 이용빈도가 가장 높았으며, 연산자를 이용할 경우에는 대부분 'and(&)'연산자를 이용한 것으로 나타났다.

여섯째, 검색실패의 원인을 살펴보면 데이터베이스에 없는 자료를 검색한 경우가 15,764회(79.42%)로 가장 많은 부분을 차지했고, 그 다음이 검색키 지정오류(18.95%), 명령어 구축 실패(1.63%)의 순으로 나타났다.

이상과 같이 LINNET을 중심으로 이용자의 검색행태를 조사하여 분석한 결과에 따르면 이용자의 다양한 입력내용을 시스템에서 조정하면, 검색성공률도 높이고 이용자의 만족도도 높일 수 있을 것으로 판단되며, 그 주요 방법을 살펴보면 대략 다음과 같다.

첫째, 이용자들이 찾고자 하는 자료가 데이터베이스에 없거나, 검색키를 잘못 지정하여 0건 검색이 된 경우는 이용자 교육 및 도움말 강화와 함께 검색프로그램을 보완하면 실패율을 낮출 수 있을 것으로 판단된다. 특히 검색문은 정확히 입력되었으나 데이터베이스에 없는 자료로 인식되어 0건 검색된 경우에는 검색문에 절단자 또는 연산자를 추가하여 검색이 되도록 시스템상의 지원이 필요할 것으로 판단된다.

둘째, 시스템에서 이용하지 않는 검색키를 입력하므로써 검색에 실패한 경우는 이용자 교육 등을 통하여 개선할 수 있을 것이다. 그러나 이용자들이 한/영문모드 변환에 미숙하여 검색키를 잘못 지정함으로써 검색에 실패한 경우에는 시스템에서 자동으로 검색키를 교정해 줄 필요가 있는 것으로 나타났다. 예를 들면, 이용자들이 영문으로 표기해야 하는 검색키를 영문키보드에 해당하는 한글코드로 입력할 경우, 시스템에서 그에 해당하는 영문표기로 전환해 주는 것이 그 방법이 될 것이다.

셋째, 시스템에서 이용자의 검색문에 대해 우측절단이 가능하도록 지원 하는 방안도 고려할 수 있다. 모든 검색문에 대해 우측절단을 자동으로 지원하면 검색성공률이 현저히 향상될 것으로 예상된다. 그러나 이 방법을 채택할 경우 검색건수가 지나치게 많아지거나, 시스템의 과부하 발생 등의 문제가 예상되므로 신중하게 결정해야 할 것으로 여겨진다.

넷째, 키워드 검색에서 연산자를 자동으로 지원하는 방안을 들 수 있다. LINNET은 서명, 저자명, 주제명, 총서명에 대하여 공란을 기준으로 키워드를 추출, 자연어 키워드 검색기능을 제공하고 있다. 키워드가 1개의 단어로 구성되어 있다는 것을 모르는 이용자가 검색키로 키워드를 지정하고 2개 이상의 단어를 입력한 경우 (예 : 'tk=한국의 민속학', 2개 이상의 키워드를 동시에 입력한 경우임), 시스템에서는 '한국의민속학'을 검색하여 0전이 검색된다. 이와 같이 2개이상의 단어를 입력할 경우에 and(&) 또는 or(!) 연산자를 자동으로 지원하는 것이(예 : 'tk=한국의 & 민속학' 또는 'tk=한국의 ! 민속학') 효과적인 검색방법이 될 수 있을 것이다.

참 고 문 헌

- 김재일 외. "KTALIS도서자료 검색시스템의 이용기능 분포도 및 응답시간을 중심으로 한 시스템성능에 대한 고찰.". 정보관리학회지. 5권 (1988). pp. 145-156.
- 노문자. "온라인 목록 이용에 관한 연구조사.". 송의여전 도서관학연구지. 제15집 (1990). pp. 7-17.
- 랑카스터, F. 윌프리트 저;윤구호;김태승 공역. 정보검색시스템. 서울: 구미무역출판부. 1985.
- 사공복희. "온라인 목록과 그 설계에 관한 연구.". 도서관. 44(1989). pp. 3-21.
- 이두영. 도서관네트워크구조론. 서울: 구미무역출판부. 1987.
- 李愼裁. "국회도서관 온라인 시스템에서의 정보검색과 키워드에 관한 연구.". 국회도서관보. 28권(1991). pp. 10-17.
- 정영미. 도서관정보전산화론. 서울: 구미무역출판부. 1989.

車培根. 社會科學研究方法. 서울:세영사. 1981.

포항공과대학 중앙도서관. Network을 이용한 도서관정보전산화: LINNET을 중심으로. 포항: 포항공과대학교 중앙도서관, 1993.

Lancaster, F. W. 저;장혜란 역. 도서관서비스평가론. 서울: 구미무역 출판부, 1990.

Akeroyd, John. "Information seeking in online catalogues.". *Journal of documentation*. Vol. 46, No. 1(1990). pp. 33-52.

Alzofon, Sammy R. and Van Pulis, Noelle. "Patterns of searching and success rates in an online public access catalog.". *College and Research Libraries*. Vol. 45, No. 4(1984). pp. 110-115.

Barreau, Deborah K. "Using performance measures to implement an online catalog.". *Library Resources & Technical Services*. Vol. 32, No. 4 (1988). pp. 312-322.

Berger, Michael G. "The MELVYL system, : the next five years and beyond.". *Information Technology and Libraries*. Vol. 11, No. 2(1992). pp. 146-157.

Buckland, Michael K. "Agenda for online catalog designers.". *Information Technology and Libraries*. Vol. 11, No. 2(1992). pp. 157-163.

Cherry, Joan M. "Improving subject access in OPACs : an exploratory study of conversion of user's queries.". *Journal of Academic Librarianship*. Vol. 18, No. 2(1992), pp. 95-99.

Connell, Tschera Harkness. "User acceptance of library catalog results : an exploratory study.". *Library Resources & Technical Services*. Vol. 35, No. 2(1991). pp. 191-201.

Cooper, Michael D. "Usage patterns of an online search system."

- Journal of the American Society for Information Science*. Vol. 34, No. 5 (1983). pp. 343-349.
- Dwyer, Catherine M., Gossen, Eleanor A. and Martin, Lynne M. "Known-item search failure in an OPAC.". *RQ*. Vol. 31, No. 2 (1991). pp. 228-236.
- Ensor, Pat. "User characteristics of keyword searching in OPAC.". *College and Research Libraries*. Vol. 53, No. 1 (1992). pp. 72-80.
- Ferl, Terry Ellen and Millsap, Larry. "Remote use of the University of California MELVYL library system : an online survey.". *Information Technology and Libraries*. Vol. 11, No. 3 (1992). pp. 285-303.
- Ferguson, Douglas. "The CLR public online catalog study : an overview.". *Information Technology and Libraries*. Vol. 1 (1982). pp. 84-97.
- Flaherty, Patricia. "Transaction logging systems : a descriptive summary.". *Library HI TECH*. Vol. 11, No. 2 (1993). pp. 67-78.
- Kalin, Sally W. "Support services for remote users of online public access catalogs.". *RQ*. Vol. 31, No. 2 (1991). pp. 197-213.
- Kalin, Sally W. "The Invisible users of online catalogs : a public services perspective.". *Library Trends* Vol. 35, No. 4 (1987). pp. 587-595.

The Search Pattern of Online Public Access Catalog - The Analysis of LINNET Transaction log -

Koo-ho Yoon *

Byung-kyoo Shim**

The purpose of this study is about the search pattern of LINNET (Library Information Network System) OPAC users by transaction log, maintained by POSTECH (Pohang University of Science and Technology) Central Library, to provide feedback information of OPAC system design.

The results of this study are as follows.

First, for the period of this analysis, there were totally 11,218 log-ins, 40,627 transaction logs and 3.62 retrievals per a log-in. Title keyword was the most frequently used, but accession number, bibliographic control number or call number was very infrequently used.

Second, 47.02% of OPAC searches resulted in zero retrievals. Bibliographic control number was the least successful search. User displayed 2.01% full information and 64.27% local information per full information.

Third, special or advanced retrieval features are very infrequently used. Only 22.67% of the searches used right truncation and 0.71% used the qualifier. Only 1 boolean operator was used in every 22 retrievals.

* Professor, Department of Library and Information Science, Keimyung University.

** Associate Director, Center Library, Pohang University of Science and Technology

The most frequently used operator is 'and(&)' with title keywords. But bibliographical control number(N) and accessionnumber(R) are not used at all with any operators.

The causes of search failure are as follows.

1. The item was not used in the database.
(15,764 times : 79.42%).
2. The wrong search key was used.
(3,761 times : 18.95%)
3. The senseless string(garbage) was entered.
(324 times : 1.63%)

On the basis of these results, some recommendations are suggested to improve the search success rate as follows.

First, appropriate user education and online help function let users retrieve LINNET OPAC more efficiently.

Second, several corrections of retrieval software will decrease the search failure rate.

Third, system offers right truncation by default to every search term. This methods will increase success rate but should considered carefully.

By applying this method, the number of hit can be overnumbered, and system overhead can be occurred.

Fourth, system offers special boolean operator by default to every keyword retrieval when user enters more than two words at a time.

Fifth, system assists searchers to overcome the wrong typing of selecting key by automatic korean/english mode change.