

# 이용자 관련 연구논문에 대한 주제어 분석

## A Keyword analysis on the 'user' related research papers : In Library and Information Science

박선미, 숙명여자대학교 문헌정보학과, smpark@sookmyung.ac.kr

오경목, 숙명여자대학교 문헌정보학과, kmoh@sookmyung.ac.kr

Seonmi, Park, Dept. of Library and Information Science, Sookmyung Women's University

Kyung-mook, Oh, Dept. of Library and Information Science, Sookmyung Women's University

본 연구에서는 국내 문헌정보학 분야의 연구 논문 중 이용자 관련 연구 논문 125편을 대상으로 논문에 부여된 주제어간의 연결 관계를 분석 하였다. 사전 작업을 통하여 정리된 226개의 주제어에 대한 연결 관계를 네트워크 분석을 통하여 분석하고 시각화 하였다. 그래프를 통하여 주제어간 연결 강도를 확인하였고, 다른 주제어와 연결성이 높은 상위 20개의 주제어를 제시하였다. 주제어간 근접성이 높은 주제어를 군집화한 결과 14개의 군집으로 정리되었다. 다른 주제어와 연결이 없이 고립된 군집이 8개, 연결된 군집이 6개였다.

### 1. 서론

‘이용자’의 사전적 의미는 어떤 물건이나 서비스 따위를 이용하는 사람을 말한다. 따라서 도서관에서의 이용자란 도서관과 도서관에서 제공하는 서비스를 이용하는 사람을 의미한다고 볼 수 있다. 국내에서 전 학문 분야에 걸쳐 이용자 관련 연구는 1994년 이후 계속 증가 하였는데, 이는 1993년 우리나라에 인터넷 서비스가 처음 시작 되고 이후 인터넷 환경이 급속도로 확산된 것과 무관하지 않다. 과거에 없던 새로운 기술과 서비스의 개발이 있었고, 이에 따른 연구가 함께 진행되었기 때문이다. 문헌정보학 분야에서도 이용자 관련 연구가 증가 하고 있는데, 이는 조재인(2011)의 연구결과에서도 나타난다.

특정 주제에 대한 연구 동향조사를 하는 방법으로는 과거의 연구 내용, 연구 방법, 연구 결과를 조사하여 향후 연구 방향을 제시하는 방식이 대부분이다. 그런데 최근에는 다양한

분야에서 네트워크 분석을 이용하여 개체간의 연관관계를 분석하는 연구가 도입되고 있다. 특히 지식 네트워크는 각종 정보자원을 구성하는 지식 개체의 유형과 패턴을 공간적인 위치나 분포로 보여주는 그래프 또는 네트워크를 말한다. (이수상, 2013)

네트워크 분석을 이용하여 특정 주제 분야의 연구 동향을 파악하고 분석한 연구로 김희정(2011)과 조재인(2011)의 연구가 있다. 김희정(2011)의 연구에서는 웹 아카이빙 주제 영역을 연구 하였으며, 조재인(2011)은 문헌정보학 최근 연구 경향을 분석 하였다.

본 연구에서는 국내에서 발간된 문헌정보학 분야 연구 논문 중 이용자 관련 주제 분야의 연구 논문들에 부여된 주제어를 수집하고, 네트워크 분석을 통하여 주제어 사이의 연관성을 찾아 보고자한다. 특히 본 연구에서는 다음의 결과물을 제시 하고자 한다.

- 주제어간 연결 강도

- 주제어의 연결 중앙성
- 주제어의 군집성

## 2. 연구방법

이용자 관련 연구 논문을 수집하기 위하여 국내 Database인 DBPia에서 키워드명 ‘이용자’로 2013년 6월 19일 검색을 수행하였다. DBPia의 데이터 중 대상을 연구논문으로 한정하기 위하여 자료유형을 ‘전자저널’로 전자저널 유형을 ‘학술저널’로 제한하였고 출판년은 전체로 검색 하였다. 검색 결과 300건의 연구 논문이 검색되었고, 그 중에서 수록 저널별로 정보관리학회지 43건, 한국비블리아학회지 30건, 한국문헌정보학회지 28건, 한국도서관, 정보학회지 24건 총 125건을 본 논문의 분석 대상으로 삼았다. 검색된 논문들은 엑셀로 내려 받아 Microsoft Excel 2007 을 이용하여 기초 자료 분석을 진행 하였으며, Gephi 0.8.2를 이용하여 네트워크 분석을 수행 하였다.

연구 논문에 사용된 주제어는 저자가 직접 입력한 형태의 비통제 주제어이므로 주제어의 형태 또한 다양하다. 때문에 분석을 위하여 사전 작업을 실시하였다. 먼저 한글 주제어와 영어 주제어가 병기 되어 있는 경우에는 한글 주제어를 우선으로 선택 하였다. 그리고 ‘이용자’ 또는 ‘이용자연구’ 자체가 주제어로 부여 되어 있는 경우는 분석에서 제외하였다. 분석 대상 연구논문이 모두 이용자관련 연구이기 때문이다. 다음으로 ‘전거제어’, ‘전거통제’와 같이 동일한 용어의 다양한 표현들 및 유사한 용어들을 하나로 통일 하였으며, ‘서비스 품질’, ‘서비스품질’등과 같이 동일한 용어에 대한 띄어쓰기가 다른 경우 이를 하나로 통일 하였다.

이러한 사전 작업에 의해서 추출된 주제어는 모두 228개로 정리되었다.

## 3. 데이터 분석

### 3.1 기초 자료 분석

228개의 주제어 중 사용 빈도수가 많은 상위 20개의 주제어를 정리하면 <표 1>과 같다.

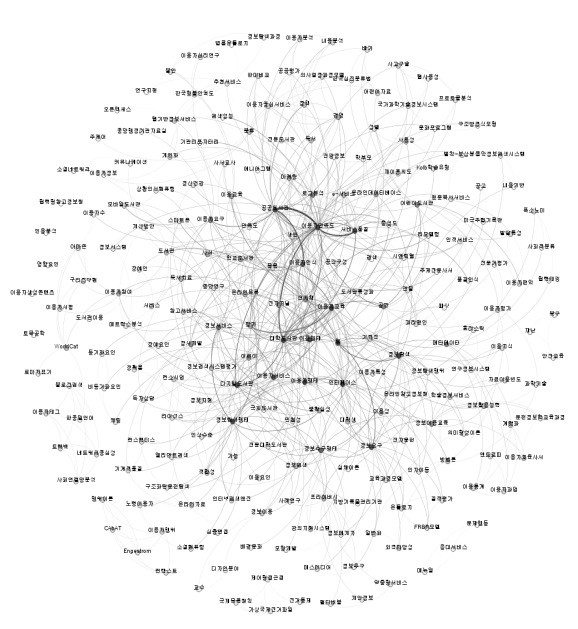
<표 1> 빈도수 상위 20개 주제어

순위	주제어	논문 수
1	이용자만족도	24
2	공공도서관	16
3	대학도서관	16
4	이용자교육	9
5	이용자인식	9
6	인터페이스	9
7	서비스품질	8
8	웹	8
9	이용자행태	7
10	정보탐색	7
11	정보탐색행태	7
12	공간	6
13	디지털도서관	6
14	정보서비스	6
15	정보요구	6
16	정보추구행태	6
17	어린이도서관	5
18	이용자서비스	5
19	이용행태	5
20	평가	5
합계		165

가장 많이 사용된 주제어는 이용자만족도, 공공도서관, 대학도서관, 이용자교육, 이용자 인식, 인터페이스, 서비스품질, 웹의 순이었으며 20위까지의 주제어가 사용된 논문 수는 165건 이었다. 전체 주제어가 사용된 논문의 총 수는 462건으로 상위 20개의 주제어가 사용된 비율은 약 35.7%로 나타났다. 20위권 밖의 주제어들은 사용된 논문수가 4회 이하였고, 1회만 사용된 주제어는 151개로 전체 주제어 228개 중 약 66.2%에 달한다. 소수의 연구가 진행된 주제 분야에 대한 원인 분석 및 해당 주제 분야의 연구 진행을 위한 후속 연구가 필요할 것이다.

### 3.2 주제어간 네트워크 분석

주제어간 네트워크 분석은 Gephi를 사용하여 자료의 분석과 시각화를 실시하였다. 먼저 한 논문에서 사용된 주제어를 서로 연결된 주제어 쌍으로 정리 하였다. 사전 작업을 통하여 추출된 228개의 주제어 중에서 하나의 논문에서 주제어가 단독으로 사용 되어 주제어간 연결 관계 확인의 대상에 해당하지 않는 주제어 2개를 제외하였다. 226개의 주제어를 분석 대상으로 삼아 동일한 논문 내에서 사용된 주제어끼리 주제어 쌍으로 정리 했으며, 그 결과 682개의 주제어 쌍이 정리 되었다. 선정된 주제어 간 한 논문에서 동시에 사용된 횟수를 주제어 간 연결강도로 했을 때 226개의 주제어 간의 연결강도를 그림으로 표시하면 <그림 1>과 같다.



<그림 1> 주제어간 연결 강도

주제어간 연결의 강도는 선의 굵기로 표시했다. <그림 1>의 연결 강도는 각 주제어의 연결 회수를 원형상에 배치한 것으로 대체적인 주제어의 연결 관계를 파악할 수 있다.

연결 중앙성을 분석한 결과 상위 20개에 해당하는 주제어들은 <표 2> 와 같다.

<표 2> 연결 중앙성 상위 20개 주제어

순위	주제어	연결 중앙성	논문수 순위
1	이용자만족도	46	1
2	대학도서관	42	3
3	공공도서관	29	2
4	인터페이스	27	6
5	이용자교육	27	4
6	이용자행태	20	9
7	디지털도서관	19	13
8	정보탐색행태	18	11
9	이용자서비스	17	18
10	서비스품질	17	7
11	이용행태	17	19
12	정보요구	17	15
13	정보탐색	17	10
14	웹	17	8
15	이용자인식	16	5
16	정보추구행태	15	16
17	정보서비스	14	14
18	로그분석	13	20+
19	어린이도서관	12	17
20	메타데이터	12	20+

<표 1>의 주제어가 사용된 논문 수 상위 20개 주제어들과 비교해 보면 대부분 유사 하지만 순위가 다르게 나타났다. 연결 중앙성 상위 5개 주제어는 이용자만족도, 대학도서관, 공공도서관, 인터페이스, 이용자교육 이다. 로그분석, 메타데이터와 같은 경우는 <표 1>에 나타나지 않은 주제어 즉, 주제어의 사용 빈도 순위에서 20위 밖의 주제어였는데 <표 2>에서는 20위 안에 포함되었다. 이러한 주제어는 많은 수의 논문에서 사용된 주제어는 아니지만, 논문 내에서 함께 사용된 다른 주제어가 많은 경우에 해당된다.

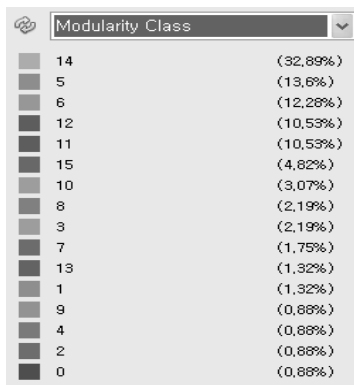
마지막으로 주제어간의 군집성을 알아보기 위해 각 주제어간의 긴밀도를 그래프로 시각화한 결과가 <그림 2> 이다.



<그림 2> 주제어의 군집성

226개의 주제어는 총 14개의 군집을 이루고 있다. 6개의 군집은 군집간 연결 관계가 형성되어 있으며, 8개의 군집은 다른 군집과 연결 관계가 없는 고립된 군집인 것을 확인할 수 있다.

<그림 3>은 <그림 2>의 군집별 주제어 비율을 보여 준다. 14번 군집에 포함된 주제어는 32.89%이며, 5번 군집에 포함된 주제어가 13.6%, 6번 군집에 포함된 주제어가 12.28%임을 보여준다.



<그림 3> 군집별 주제어 비율

상위 3개 군집에 포함되는 주제어는 총 134개 이다. 전체 분석대상 주제어의 약 60%에 해당하는 주제어가 논문 내에서 상호 연관 관계를 가지고 연구가 되었다는 것을 알 수 있다.

#### 4. 결론

본 연구의 목적은 문헌정보학 분야의 이용자 관련 연구 논문에서 사용된 주제어간의 연관 관계를 네트워크 분석 방법에 의해서 분석하는 것이다. 연구 논문 125편을 대상으로 주제어를 분석 하였으며, 사전 작업을 통하여 226개의 주제어를 분석 대상으로 정리 하였다. 분석 도구로는 Gephi를 이용하여 관계를 분석하고 시각화 하였다. 그래프를 통하여 주제어간 연결 강도를 확인해 보았고, 중앙 연결성을 분석하여 다른 주제어와 연결성이 높은 상위 20개의 주제어를 제시하였다. 226개의 주제어간에 근접성이 높은 주제어를 군집화한 결과 14개의 군집으로 정리되었다. 다른 주제어와 연결이 없이 고립된 군집이 8개, 연결된 군집이 6개였다.

#### 참고문헌

DBpia, [Cited 2013. 6. 19].  
 <<http://www.dbpia.com>>.  
 Gephi, [Cited 2013. 6. 20].  
 <<https://gephi.org>>.  
 김희정. 2011. 네트워크 분석을 기반으로 한 웹 아카이빙 주제영역 연구. 『한국비블리아학회지』. 22(2): 235-248.  
 이수상. 2013. 『네트워크 분석 방법론』. 서울 : 논형.  
 조재인. 2011. 네트워크 텍스트 분석을 통한 문헌정보학 최근 연구 경향 분석. 『정보관리학회지』. 28(4): 65-83.