

# 문헌정보학의 지식 구조에 관한 연구

## A Study on Intellectual Structure of Library and Information Science in Korea \*

유 영 준(Yoo Yeongjun)\*\*

### 초록

이 연구는 색인어가 특정 주제 영역의 지식 구조를 표현할 수 있다는 것을 전제로 한다. 여기에서는 문헌정보학 관련 학술지인 정보관리학회지, 한국도서관정보학회지, 한국문헌정보학회지 등에 수록된 논문을 대상으로 국회도서관이 배정한 색인어를 클러스터링하여 문헌정보학의 지식 구조를 파악하였다.

그 과정에서, 색인어간의 연관도 및 동시 출현 빈도를 이용하여 색인어 군집을 생성하였고, 초출색인어와 시기 구분에 의한 시계열 분석을 수행함으로써 문헌정보학의 발전 과정과 그 동향을 밝혔다. 또한 색인어 군집에 의해 도출된 지식 구조와 기존의 전통적인 분류체계의 지식 구조를 비교하여 두 지식 구조간의 차이를 분석하였다.

### ABSTRACT

This study was conducted upon the premise that index terms display the intellectual structure of a specific subject field. In this study, an attempt was made to grasp the intellectual structure of Library and Information Science by clustering the index terms of the journals of the related academic societies at the Library of National Assembly - such as the Journal of the Korean Society for Information Management, the Journal of the Korean Library and Information Science Society, and the Journal of the Korean Society for Library and Information Science. Through the course of the study, index term clusters

\* 이 논문은 2003년 8월 연세대학교 문헌정보학과 박사학위논문을 발표한 것임.

\*\* 연세대학교 문헌정보학과 강사 (arompooh@netian.com)

were generated based on the linkage of the index terms and the frequency of co-occurrence, and moreover, time periods analysis was conducted along with studies on first-appearing terms, in order to clarify the trend and development process of the Library and Information Science. This study also analysed the difference between two intellectual structure by comparing the structure generated by index term clusters with the existing structure of traditional classification systems.

---

키워드 : 지적 구조, 지식 구조, 색인, 단어 동시출현, 클러스터링, 초출색인어, 분류체계, intellectual structure, knowledge structure, index, term co-occurrence, clustering, first-appearing index terms, classification system

## 1 서론

### 1.1 연구의 목적

색인어를 이용해서 지식 구조를 설명하는 방법은 특정 주제 영역의 최신의 색인어를 분석함으로써 시의성을 가질 수 있고, 그 결과로 해당 주제 영역의 변화를 보다 세밀하게 밝혀낼 수 있을 것이다. 또한 해당 주제 영역의 지식 구조의 역동성을 미시적인 관점에서 설명할 수 있을 뿐만 아니라 장기간에 걸친 주제 영역의 발전 과정을 거시적인 관점에서 규명할 수 있는 장점도 갖는다고 할 수 있다.

이런 관점을 견지하면서 색인 과정의 결과로 생성되는 색인어가 문헌의 내용을 표현한다는 점을 고려한다면, 색인어를 이용하여 문헌의 내용을 구조화하거나 문헌 집단으로 대표되는 특정 주제 영역의 지식 구조를 파악하는 것이 다른 방법보다 더 유용하다고 할 수 있다. 더불어 색인어를 이용하여 문헌정보학의 지식 구조를 파악함으로써 기존의 분류체계를 개선할 때 참조로 이용하거나, 시소러스를 구축할 때 용어간의 관계를 보다 정확하게 설정하기 위한 도구로 활용할 수 있고, 그리고 급격하게 변화하는 정보 환경에 맞는 문헌정보학의 교과과정을 개정하기 위한 하나의 판단 기준을 제시할 수 있는 기여도 할 수 있을 것이라고 생각한다. 또한 문헌정보학 이외의 다른 학문 영역의 지식 구조를 파악하고자 할 때에도 이

연구를 참조해서 활용할 수 있을 것이다.

위에 제시한 내용을 근거로 이 연구는 문헌정보학 분야의 학술지에 수록된 논문을 대상으로 작성한 색인이가 표현하는 문헌정보학의 지식 구조를 규명하는 데 그 목적이 있으며, 문헌정보학 영역에서 특정 연도에 최초로 등장하는 연구를 연도별로 파악할 수 있는 초출색인어와 시기 구분에 의한 시계열 분석을 수행함으로써 문헌정보학의 발전 과정과 그 동향을 밝혀 문헌정보학의 지식 구조를 설명하려고 한다. 더불어 색인어 군집에 의해 도출된 지식 구조와 기존의 전통적인 분류체계의 지식 구조를 비교함으로써 기존의 분류체계와 클러스터링 기법을 적용하여 생성한 색인어 군집에 의한 지식 구조간의 차이를 밝히려고 한다. 또한 색인어 군집에 의한 지식 구조가 용어간의 계층관계를 표현하는 데 있어서 어떤 한계가 있는가, 그리고 도서관의 자료를 분류하는 것이 목적인 기존의 분류체계가 지식을 구조화하는 수단으로서 갖는 제한점, 즉 문헌정보학의 학문 발전에 상응하는 유연성을 가지는가 또는 문헌정보학의 학문적 역동성을 수용할 수 있는가 하는 두 관점에서 살펴보는 것도 이 연구가 갖는 목적의 하나라고 할 수 있다.

## 1.2 연구의 방법

문헌정보학의 대표적인 학술지인 **정보관리학회지**, **한국도서관정보학회지**, **한국문헌정보학회지**에 실린 1,162편의 국내 기사에 대해서, 국회도서관이 수작업으로 배정한 색인어 3,933개를 제어과정을 통해 표준화하고 중복되는 색인어를 제외한 2,307 종의 색인어를 분석 대상으로 하였다.

그리고 클러스터링을 이용한 군집 생성의 구체적인 과정은 다음과 같다. 대상이 된 색인어 쌍의 동시 출현 빈도를 이용해 단어간의 연관도를 구하고, 색인어간의 연관도나 동시 출현 빈도를 이용하여 문헌정보학의 지식 구조를 파악한다. 그리고 피어슨 상관계수(Pearson correlation coefficient)를 사용하여 동시출현빈도를 산출하고 클러스터링기법을 적용한 색인어 군집을 생성하였으며, 군집의 개수를 미리 상정하고 통제하여 클러스터링 기법을 적용하였다. 그 이유는 기준값(threshold)에 의해 클러스터링 기법을 적용하게 되면 클러스터의 수를 예측할 수 없으며, 군집 결과를 용이하게 해석하려는 목적으로 이와 같은 방법을 사용하였다. 그리고 클러스터링을 위한 통계처리는 SPSS 10을 사용하였다.

## 2 이론적 배경

특정 주제 영역의 지식 구조를 파악하기 위해서는 해당 영역의 문헌이 담고 있는 주제 개념을 표현한 색인어를 분류 체계를 이용하여 분석하는 것이 한 방법이 될 수 있다. 그러나 이 연구에서는 기존의 분류체계를 이용하지 않고 문헌정보학 분야의 학술지를 대상으로 작성한 색인어를 클러스터링 함으로써 문헌정보학의 지식 구조를 파악하였다.

색인어를 클러스터링하기 위해서는 먼저 분석 대상이 되는 색인어간의 유사도를 측정하기 위해 동시출현빈도를 이용하면서 특정한 공식 또는 계수를 사용하는 데, 이 때 사용하는 공식을 유사계수라고 한다(이재윤 2002, 11). 유사계수는 단순 동시출현빈도와 상대 동시출현빈도를 사용할 수 있다. 단순 동시출현빈도 유사계수는 아래와 같은데, 단어  $x$ 와  $y$ 의 유사계수를  $sim(x,y)$ 로 단어  $x$ 와  $y$ 의 동시출현빈도를  $f(x,y)$ 로 나타낼 수 있다.

$$sim(x, y) = f(x, y)$$

일반적으로는 단순 동시출현빈도를 이용하는 방식 대신 동시출현빈도와 개별 단어의 빈도, 전체 단어 빈도를 이용하여 단어 사이의 통계적인 연관성을 객관적으로 평가하는 상대 동시출현빈도 방식의 유사계수가 주로 이용된다(이재윤 2002, 11).

이 연구에서는 색인어를 군집하기 위해 상대 동시출현빈도 방식의 피어슨 상관계수(Pearson correlation coefficient) 공식을 사용하였고, 공식은 다음과 같다.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2}} = \frac{n[\sum_{i=1}^n x_i y_i] - [\sum_{i=1}^n x_i][\sum_{i=1}^n y_i]}{\sqrt{n[\sum_{i=1}^n x_i^2] - [\sum_{i=1}^n x_i]^2} \sqrt{n[\sum_{i=1}^n y_i^2] - [\sum_{i=1}^n y_i]^2}}$$

$x_i$  : 단어  $x$ 와 행렬의  $i$ 번째 단어와의 동시출현빈도

$y_i$  : 단어  $y$ 와 행렬의  $i$ 번째 단어와의 동시출현빈도

딩 등(Ding, Chowdhury, and Foo 2001)의 연구에서 사용한 피어슨 상관계수  $r$ 은 두 단어의 유사도를 측정하기 위해 두 단어와 동시출현단어 행렬의 다른 단어들과의 동시출현빈도를 이용하여 구한다. 그리고 피어슨 상관계수를 사용한 이유는 일반적으로 가장 많이 사용되는 코사인 계수보다 군집 결과가 좋게 나타났기 때문이다.

### 3 색인어 클러스터링에 의한 지식 구조

#### 3.1 색인어 선정 및 표준화

색인어 클러스터링의 대상은 정보관리학회지, 한국도서관정보학회지, 한국문헌정보학회지 등으로 한정하였고, 서지학분야의 학회지인 서지학연구는 부여된 색인어가 고유명사 즉 인명, 지역명, 시대명, 도서명 등으로 구성되어 있으며, 서지학분야의 내용을 정확하게 이해하고 색인어를 부여하는 것이 서지학분야 연구자들에게 국한될 수 있다는 판단에 따라 본 연구에서는 제외하였다.

또한 군집의 개수를 미리 한정하여 클러스터링 기법을 적용함으로써 문헌정보학의 지식 구조를 용이하게 제시하였고, 클러스터링의 대상이 된 세 학회지의 색인어 수는 3,933개이었지만 ASIS 시소러스(ASIS Thesaurus of Information Science and Librarianship, 2nd ed., 1998)와 정보학사전을 사용하여 색인어를 제어하고 표준화하였다. 그리고 ‘문헌정보학’이나 ‘도서관’과 같이 너무 일반적이어서 주제어로서의 역할을 할 수 없는 색인어들은 제외하였다.

#### 3.2 정보관리학회지의 색인어 클러스터링

군집을 생성한 과정을 구체적인 예를 들어 설명하면, 그림 1의 40군집 중 ‘공공도서관’, ‘인터넷’, ‘웹사이트’, ‘커뮤니케이션’ 등 네 개의 군집이 인터넷이라는 하나의 군집이 되어 20군집 중의 하나를 형성하였고, 20군집의 ‘분류체계’, ‘인터넷’, ‘정보검색’ 등 세 군집이 ‘인터넷’이라는 10군집 중의 하나를 형성하였고, 동일한 ‘인터넷’ 군집이 최종적으로 5군집을 형성하게 되었다. 그리고 나머지 두 학회지와 세

학회지를 통합한 군집도 위와 동일한 군집 생성 과정을 거쳐 군집 결과를 생성한 것이다.

40군집은 전체적으로 정보 검색 영역 군집과 정보조직 영역의 비율이 높았고, 20군집은 ‘인터넷’, ‘정보검색’, ‘대학도서관’, ‘데이터베이스’, ‘KORMARC’, ‘OPAC’ 등의 순서로 출현빈도가 높게 나타났고, 정보 검색 영역의 군집이 가장 많았다. 또한 10군집의 특징은 ‘정보검색’ 군집이 ‘인터넷’ 군집으로 통합되었다는 점이라고 할 수 있다. 출현 빈도도 두 번째로 높고 일반적인 인식으로도 정보관리학회지의 가장 핵심적인 개념으로 인식되고 있는 ‘정보검색’ 군집이 ‘인터넷’ 군집에 포함된 것은 ‘인터넷’과 ‘정보검색’이 밀접한 관계를 맺고 있다는 사실을 보여 주었다. 마지막으로 최상위 군집인 5군집에는 ‘데이터베이스’, ‘인터넷’, ‘대학도서관’, ‘KORMARC’, ‘OPAC’으로 나타났고, 목록 영역에 속하는 군집이 두 개 포함되어 목록 영역의 주제적 중요성을 보여주었다.

그리고 계층구조적 관점에서 최상위 군집과 하위 군집간의 관계를 살펴보면 다음과 같다. ‘KORMARC’ 군집과 ‘대학도서관’ 군집은 연관성이 매우 높았고, ‘OPAC’ 군집은 20군집부터 독자적인 개별 군집을 이루어 독립성이 강한 군집으로 나타났다. 또한 ‘인터넷’ 군집도 하위 군집간의 연관성이 높게 나타났지만, ‘데이터베이스’ 군집은 타 군집에 비해 하위 군집간의 연관성이 낮은 것으로 볼 수 있다.



그림 1. 정보관리학회지 군집 결과

그리고 계층구조적 관점에서 최상위 군집과 하위 군집간의 관계를 살펴보면 다음과 같다. 'KORMARC' 군집은 '메타데이터'를 하위 군집으로 포함하였고, '대학도서관'은 '디지털도서관' 과 '전산화'를 하위 군집으로 포함하여 연관성이 매우 높았고, 'OPAC' 군집은 20군집부터 독자적인 개별 군집을 이루어 하위 군집을 가지지 않는 독립성이 강한 군집으로 나타났다. 또한 '인터넷' 군집도 하위 군집간의 연관성이 높게 나타났지만, '데이터베이스' 군집은 '동시인용분석', '자동색인', '시소러스', 'SCI' 등을 하위 군집으로 포함하여 타 군집에 비해 하위 군집간의 연관성이 낮은 것으로 볼 수 있다.

### 3.3 한국도서관정보학회지의 색인어 클러스터링

그림 2의 40군집은 도서관 경영 영역의 군집이 가장 많았고, 정보 조직 영역, 정보 검색 영역 등 다양한 주제가 포함되었다. 그리고 20군집은 '대학도서관'과 '공공도서관'의 출현빈도가 다른 색인에 비해 월등하게 높았고, 특정 주제 영역에 속하는 색인어 군집에 치우치지 않으면서 다양한 주제 영역의 색인어 군집으로 분산된 것으로 나타났다. 또한 10군집은 '공공도서관', '학교도서관', '대학도서관' 등 특정 관종의 도서관 군집, '디지털도서관', '인터넷', '정보기술' 등 정보기술관련 군집, 'DDC'와 '목록규칙'과 같은 정보 조직 영역 군집, 그리고 정보기술에 따라 변화하고 있는 '정보서비스'와 도서관 인력의 교육과 관련된 '전문교육' 등의 군집이 포함되었다. 최상위 군집인 5군집은 '공공도서관', '정보서비스', '전문교육', '대학도서관', '목록규칙' 등의 군집으로 구성되었다. 특히 '공공도서관' 군집은 '공공도서관', 'DDC', '디지털도서관', '학교도서관', '인터넷' 등 다섯 개의 하위 군집을 포함하여 가장 많은 하위 군집을 가지면서 다양한 주제 영역을 포괄하는 군집이었는데, 하위 군집간의 연관성이 다른 군집보다 약한 것으로 볼 수 있다. 이와는 반대로 '대학도서관' 군집은 20군집부터 하나의 독립된 군집을 이루어 5군집까지 하나의 군집을 유지하였고, '목록규칙' 군집은 'AACR' 군집을 하위군집으로 포함하는 것으로 나타났다.

따라서 클러스터링기법도 주제 개념간의 계층구조를 표현할 수 있음을 보여주었다. 결과적으로 '공공도서관' 군집을 제외한 군집 결과는 클러스터링에 의한 군집 분석이 의미가 있는 것으로 나타났지만, '공공도서관' 군집에 속하는 하위군집

의 결과는 클러스터링에 의한 군집 분석 즉 출현빈도에 근거한 통계적 분석 기법의 한계를 드러낸 것으로 해석할 수 있을 것이다.

### 3.4 한국문헌정보학회지의 색인어 클러스터링

그림 3의 40군집은 세 학회지의 군집 결과 중에서 주제 영역의 분포가 가장 높은 결과를 보였으며, 특히 다른 학회지에 비해 상대적으로 참고 서비스 영역의 색인어가 많이 포함된 것으로 나타났다. 그리고 하나의 색인어가 독립된 군집을 이룬 예들을 살펴보면, ‘공공도서관’, ‘학교도서관’, ‘디지털도서관’ 등 출현 빈도가 높은 색인어들과 출현 빈도가 낮은 ‘도서관조직’, ‘연속간행물’, ‘분류체계’ 등이 함께 포함된 것으로 나타났다. 그리고 20군집은 ‘분류’, ‘MARC’, ‘DDC’, ‘AACR’, ‘OPAC’ 등 정보 조직 영역의 군집들이 가장 많았고, ‘공공도서관’은 출현 빈도가 두 번째로 높았지만 ‘대학도서관’ 군집을 구성하는 하위 색인어로 포함되어 20군집의 개별 군집에 포함되지 못했다. 또한 ‘교과과정’, ‘상호대차’, ‘사서’, ‘정보봉사’, ‘참고봉사’ 등의 군집이 포함되었고, ‘웹’, ‘디지털도서관’, ‘데이터베이스’, ‘인터넷’ 등의 정보기술관련 군집들도 포함되었다. 또한 10군집은 ‘교과과정’, ‘학교도서관’, ‘참고봉사’ 등 다른 학회지의 군집 결과에서는 크게 부각되지 않았던 색인어들이 포함되었고, 나머지 군집들 중 ‘대학도서관’을 제외하면 ‘AACR’, ‘DDC’, ‘OPAC’ 군집 등 정보 조직 영역이 가장 많은 군집을 이루었으며, ‘정보검색’, ‘웹’, ‘디지털도서관’ 등 정보기술과 관련 있는 군집들이 포함되었다.

최상위 군집인 5군집의 구성을 보면, ‘정보검색’ 군집이 가장 많은 하위 군집을 포함하여 다양한 주제개념을 대표하는 군집으로 나타났지만, 하위 군집간의 관련성이 낮았다. 그러나 ‘대학도서관’ 군집은 ‘사서’를 하위 군집으로 포함하였고, ‘참고봉사’ 군집은 ‘연구자’와 ‘정보봉사’를 하위 군집으로 포함하여 하위 군집간의 관련성이 높았다. 또한 ‘AACR’ 군집은 ‘전자저널’을 하위 군집으로 포함하여 관련성이 낮았지만, ‘전자저널’ 군집을 구성하는 색인어들이 ‘KCR’과 ‘KORMARC’인 점을 고려한다면, ‘AACR’ 군집은 목록규칙 또는 자동화목록형식과 깊은 관련성을 갖는 군집이라고 할 수 있다. 그리고 ‘OPAC’ 군집은 이용자와 관련된 하위군집을 포함한 것으로 나타났는데 이러한 결과는 세 학회지를 통합한 군집 결과와 동일하다.

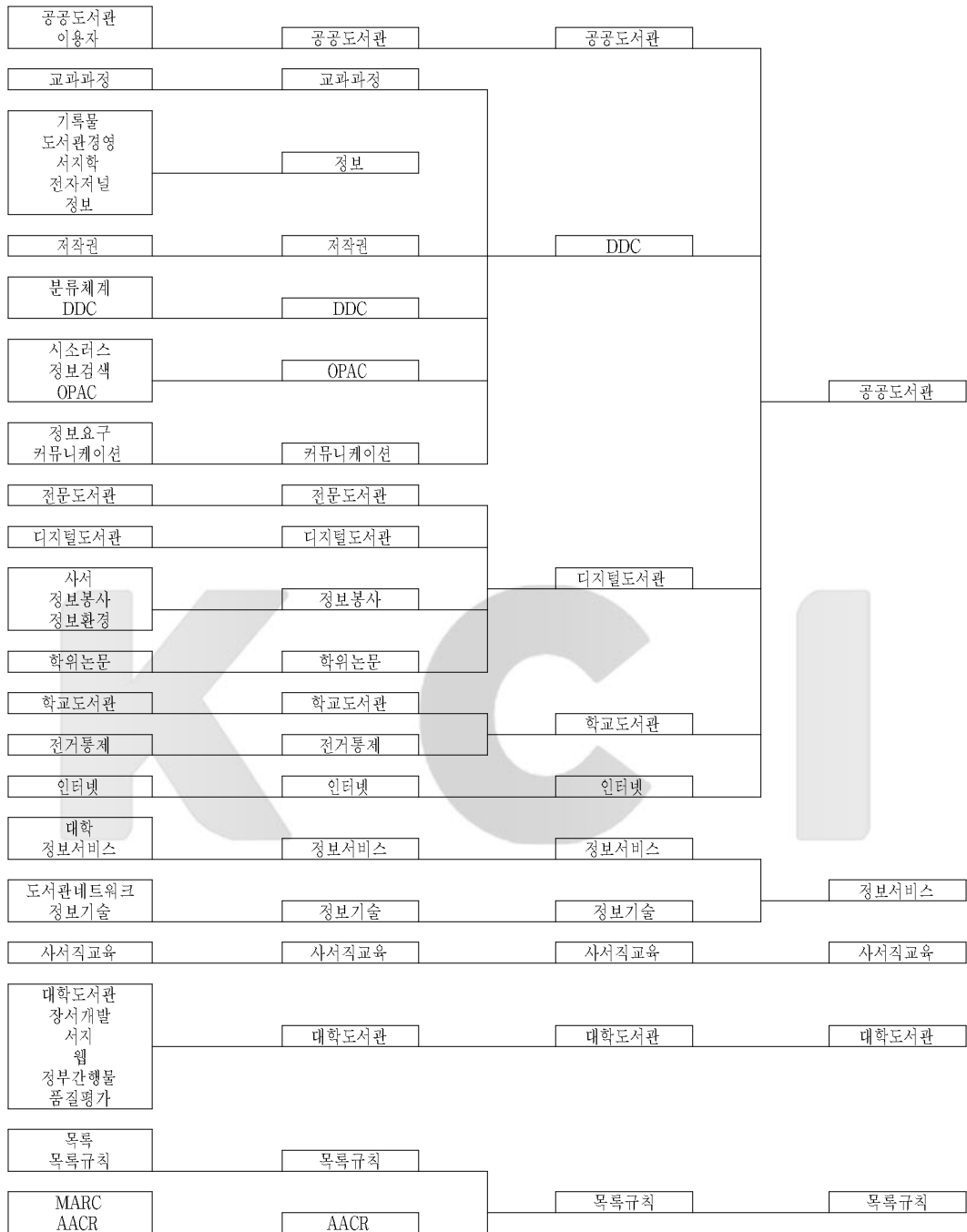


그림 2. 한국도서관정보학회지 군집 결과



그림 3. 한국문헌정보학회지 군집 결과

정보 조직의 목록 영역인 ‘AACR’ 군집과 ‘OPAC’ 군집이 최상위 군집에 포함 되어 목록 영역의 중요성을 확인할 수 있었고, 특이한 사항은 ‘참고봉사’ 군집이 포함된 것인데 다른 학회지의 군집 결과와는 다른 특징으로 볼 수 있다. 그리고 ‘대학도서관’ 군집은 세 학회지를 통합한 군집 결과와 개별 학회지의 최상위 군집에 모두 포함된 것으로 나타남으로써, 특정 학회지의 특성과 상관없이 다양한 연구가 이루어지고 있음을 알 수 있었다.

### 3.5 세 학회지를 통합한 색인어 클러스터링

그림 4의 40군집은 전체적으로 목록 영역을 포함하는 정보 조직 영역의 군집이 가장 많았고, 정보 검색 영역이 두 번째로 많은 군집을 이루었으며, 인터넷과 관련된 영역의 군집도 높은 비율을 나타내었다. 그리고 ‘목록’, ‘분류’, ‘주제’가 하나의 개별 군집을, ‘지식관리’, ‘컨텐츠’, ‘컨텐츠관리’가 하나의 개별 군집을 이루었는데, 이 두 군집은 명확하게 군집을 대표할 수 있는 색인어를 선택하는데 어려움이 있었다. 그 이유는 색인어간의 출현빈도도 비슷하였고, 특정한 하나의 색인어가 명확하게 군집을 대표하지 못하는 군집이었기 때문이었다. 또한 ‘저작권’ 군집은 17개의 색인어들이 하나의 개별 군집을 형성하여 다양한 주제 개념을 포함하는 군집이었다. 그리고 20군집은 ‘MARC’, ‘KORMARC’, ‘OPAC’ 등 목록관련 군집과 ‘시소러스’, ‘정보검색시스템평가’, ‘정보검색’, ‘메타데이터’ 등 정보검색관련 군집이 개별화되었고, ‘대학도서관’, ‘이용자연구’, ‘교과과정’, ‘기록물’, ‘컨텐츠관리’ 등도 군집을 이루었다. 또한 ‘데이터베이스’, ‘DDC’, ‘정보기술’, ‘디지털도서관’, ‘정보시스템’, ‘공공도서관’, ‘학술정보’, ‘정보서비스’ 등도 20군집에 포함되었다. 또한 10군집의 결과는 목록 영역인 ‘KORMARC’과 ‘OPAC’ 군집으로 클러스터링되었고, 정보 검색 영역은 ‘정보검색’과 ‘메타데이터’ 군집으로 개별화되었다. 그리고 20군집을 이루었던 ‘이용자연구’, ‘교과과정’, ‘기록물’, ‘컨텐츠관리’ 등은 10군집에서도 군집을 유지하였고, 다양한 주제개념으로 구성된 ‘공공도서관’ 군집이 가장 많은 색인어를 포함하였고 출현빈도가 가장 높은 ‘대학도서관’ 군집도 10군집에 포함되었다. 최상위 5군집은 ‘KORMARC’, ‘정보검색’, ‘공공도서관’, ‘OPAC’, ‘대학도서관’ 등이 포함되었으며, 특징적인 것은 목록 영역의 군집이 2개가 포함됨으로써 문헌정보학에서 목록 영역이 차지하는 중요성을 잘 보여주는 결과라고 할 수 있다.

5군집에 포함된 군집들을 계층적인 관점에서 분석해보면 다음과 같다. 'KORMARC' 군집은 'MARC', '목록규칙', 'KORMARC', '전자저널' 등의 하위 군집을 포함하였고, 하위 군집인 '전자저널'이 주제적인 연관성이 낮은 것으로 볼 수도 있지만, '전자저널' 군집을 구성하는 색인어가 '목록작성', '전자저널', '접근점' 등 목록 영역의 색인어를 포함함으로써 목록과 관련된 군집으로 해석할 수 있다.

'정보검색' 군집은 '시소러스', '정보검색시스템평가', '인터넷', '정보검색', '메타데이터' 등의 하위 군집을 포함하였고, '정보검색'과 '인터넷'이 가장 중요한 색인어로 나타나 '인터넷상에서의 정보검색'이 핵심적인 주제를 구성하는 것으로 나타났고, '시소러스'와 '메타데이터'는 정보 검색 영역을 구성하는 중요한 하위 주제로 분석되었다.

'OPAC' 군집은 '이용자'와 'OPAC' 등을 하위 군집으로 포함함으로써, 이용자나 이용 행태 등의 주제와 깊은 관련성을 가지는 것으로 해석할 수 있다. 마지막으로 '대학도서관' 군집의 하위 군집은 '대학도서관'과 '사서' 등을 하위 군집으로 포함하여, 대학도서관과 사서의 두 주제가 관련성이 있음을 확인할 수 있었다. 또한 '대학도서관' 군집에 속한 색인어 중에는 '참고봉사', '장서개발' 등이 출현빈도가 높은 것으로 나타났다.

전체적으로 군집 결과를 살펴보면, 클러스터링 기법은 전통적인 분류체계의 계층구조 상에서 동일 범주 내의 연관성은 잘 표현하는 것으로 나타났지만, 군집의 크기가 큰 경우에는 하위 군집간의 연관성이 떨어지는 결과를 보였다. 이러한 결과는 군집의 크기가 작은 군집의 하위 군집간의 연관성이 오히려 높았음을 반증하는 것이라고 할 수 있다. 결론적으로 클러스터링 기법은 주제적 특성이 유사한 문헌정보학 내의 주제나 개념들을 특정한 군집으로 묶어주는 결과를 보여주었을 뿐만 아니라, 군집 내의 계층 구조도 잘 반영하는 것으로 해석할 수 있다.

#### 4 초출색인어와 시기 구분에 의한 시계열 분석

특정 연도에 출현한 색인어 중 그 해에 처음으로 색인된 초출색인어를 분석하는 방법은 문헌정보학의 역사적인 발전 과정과 연구 경향의 변화를 일목요연하게 파악하는데 유용하다.

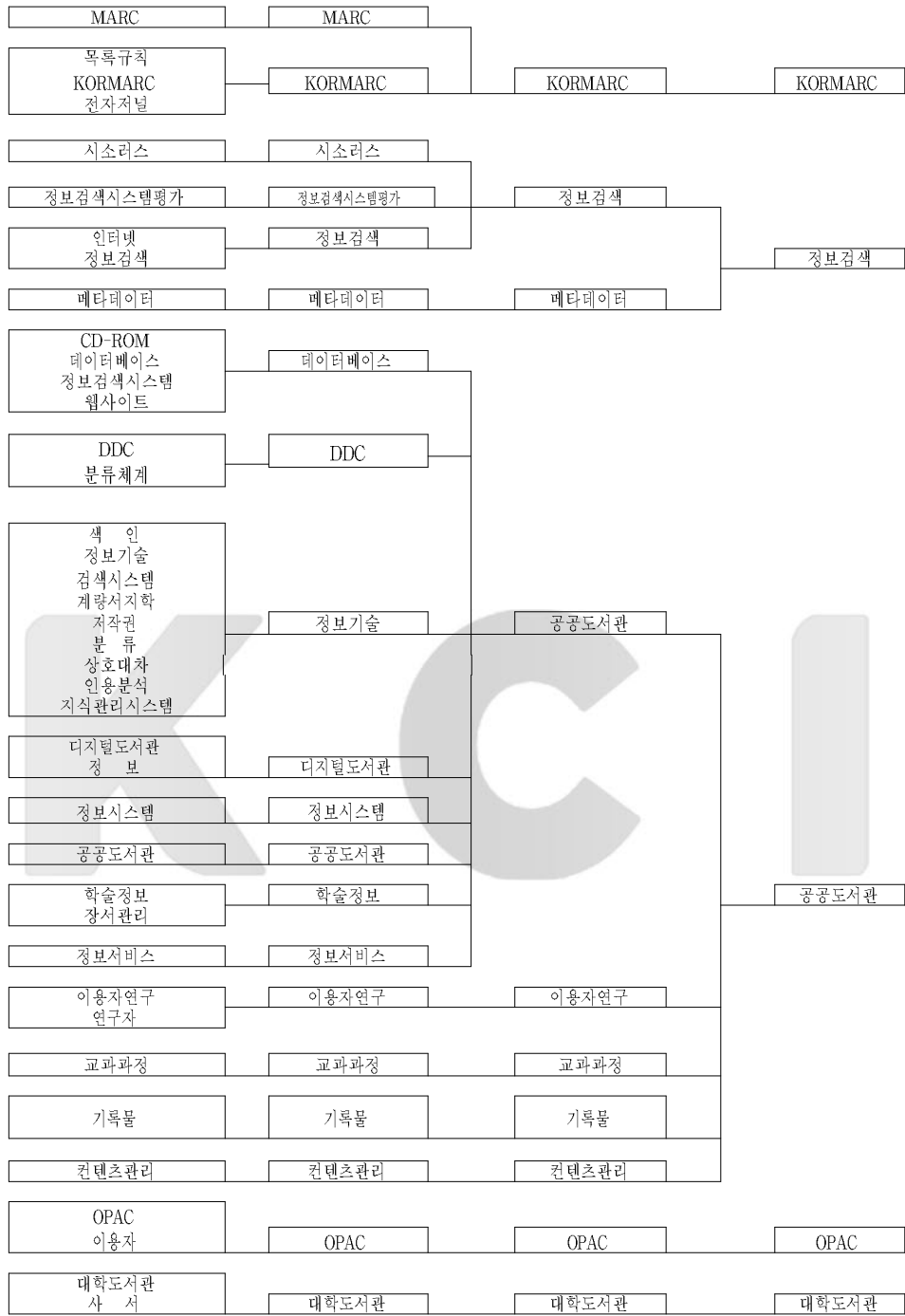


그림 4. 세 학회지를 통합한 군집 결과

#### 4.1 초출색인어를 이용한 시계열 분석

표 1은 연도별로 초출색인어를 조사하여 초출색인어의 특징을 파악함으로써 문헌정보학의 발전 과정과 연구 경향을 살펴보기 위한 것인데, 연도별로 초출색인어의 횟수 평균, 최다 횟수 초출색인어와 그 횟수, 상위 5개 초출색인어의 횟수 평균 그리고 초출색인어 중 횟수가 상위 5위안에 든 색인어 목록을 제시하였다.

그리고 색인어 횟수는 정보 검색의 문헌 빈도(document frequency), 즉 학술지에 실린 기사에 출현한 색인어의 횟수를 의미한다. 그리고 연도별로 최다 횟수 초출색인어와 상위 5위안에 드는 색인어를 추출함으로써, 문헌정보학 영역의 발전 과정을 일목요연하게 파악할 수 있었다. 또한 상위 5개 초출색인어의 횟수 평균값이 10이상인 연도를 중심으로 연구의 흐름을 살펴보았다.

표 1의 내용을 구체적으로 살펴보면, 먼저 최다 횟수 초출색인어와 초출색인어 중 횟수 상위어 5개를 중심으로 문헌정보학의 발전 과정을 살펴보면 다음과 같다. 1978년은 ‘공공도서관’이 64회로 가장 높은 색인 횟수를 기록하였고, ‘AACR’과 같은 목록규칙에 관한 연구도 포함되었는데, 국회도서관이 색인을 배정한 첫해에 공공도서관 분야가 가장 활발하게 연구가 이루어진 것을 알 수 있다.

도서관 관종별 연구의 발전 과정은 위에서 지적하였듯이 공공도서관에 관한 연구가 처음이었고, 1980년에 학교도서관, 1981년에 대학도서관, 전문도서관에 관한 연구는 1985년에 처음 수행된 것으로 나타났다. 이어서 1979년은 ‘데이터베이스’와 ‘정보검색시스템’ 등과 관련된 영역이 가장 활발하게 연구가 이루어졌고, 이러한 연구의 흐름은 ‘AACR’에 관한 연구와 함께 1981년의 ‘KORMARC’을 이용한 대학도서관의 전산화를 위한 기초 연구로서의 의미를 가지면서, 순차적인 발전 과정을 보였음을 확인할 수 있었다.

1984년은 ‘이용자’에 관한 연구가, 1985년은 ‘정보검색’과 ‘정보검색모형’ 등 정보검색에 관한 본격적인 연구가 시작되었고, 뒤이어 1986년은 정보검색의 하위 영역인 ‘클러스터분석’이나 ‘계량서지학’에 관한 연구가 시작되었다. 그리고 1989년은 ‘정보서비스’나 ‘상호대차’에 관한 연구와 더불어 ‘OPAC’에 관한 연구가 가장 활발한 것으로 나타났고, ‘분류체계’, ‘하이퍼텍스트’, ‘브라우저’, ‘자동초록’ 등의 연구가 시작되었음을 알 수 있다. 또한 1978년에 사서직에 대한 연구가 시작된 이래로 1990년이 되어서야 보다 구체적으로 사서라는 사람에 대한 연구로 발전한 것을

확인할 수 있고, 'CD-ROM' 과 같은 새로운 개념과 함께 'MEDLINE' 과 같은 특정 주제 영역의 데이터베이스를 다루는 연구와 문헌정보학의 발전에 따른 문헌정보학교육에 대한 검토가 이루어졌음을 알 수 있다.

1990년 이후로는 주목할 만한 변화가 나타나지 않았지만, '선거통제', '정보시스템', '전문검색시스템', '서지데이터베이스' 에 관한 연구들이 수행되었다. 특히 1995년은 한국 문헌정보학의 획시거적인 구분을 짓는 분기점이라고 할 수 있을 정도로 확연한 특징을 나타낸 연도라고 할 수 있다. 1995년의 상위 5개 초출색인어의 횟수 평균값이 전후 10년 사이에 가장 높았고, '인터넷' 을 포함하여 '디지털도서관', '웹', '전자도서관' 등 기존의 발전 과정이나 문헌정보학의 연구 환경을 급격하게 변화시킨 연구들이 최초로 등장한 시점이라고 할 수 있다.

그리고 1996년 이후의 연구는 1995년에 나타난 연구의 논리적인 결과로서 네트워크를 이용한 '원문제공서비스' 나 인터넷 상에서의 정보조직 수단인 '메타데이터', 인터넷 상에서의 '검색엔진' 등 인터넷이나 네트워크와 밀접한 관련을 가지는 연구가 순차적으로 등장하였고, 2000년과 2001년은 '기록보존' 이나 '지식관리' 등 최근의 연구 동향을 보여주고 있지만 1995년에 시작된 연구 경향이 차츰 새로운 변화를 맞이하고 있음을 보여 주고 있다.

또한 2001년까지는 뚜렷하고 주도적인 주제 영역이 대두된 것은 아니고 다양한 주제 영역들이 혼재하고 있어 향후 문헌정보학의 연구 동향을 예측하기에는 미흡하며, 2002년 이후의 초출색인어를 분석해 보아야 정확한 연구 경향을 파악할 수 있을 것이다.

#### 4.2.2 시기 구분을 이용한 시계열 분석

전체 색인어를 대상으로 평면적인 시기 구분에 의한 시계열 분석을 하지 않고, 3장의 세 학회지를 통합한 색인어 클러스터링의 군집 결과 중 10군집 결과를 문헌정보학의 주요한 10개 연구 영역으로 상정하여 시기에 따른 주제 영역의 변화를 분석하였으며, 각 군집에 포함된 색인어 목록은 표 2와 같다.

표 1. 연도별 초출 색인어의 특징

연도	횟수 평균	최다 횟수 초출 색인어	최다 횟수	상위 5개 초출 색인어의 횟수 평균	초출 색인어 중 횟수 상위어				
					1위	2위	3위	4위	5위
1978	4.30	공공도서관	64	21.4	공공도서관	정보봉사	교과과정	ACR	사서직
1979	3.78	데이터베이스	26	17.6	데이터베이스	시소러스	정보검색시스템	장서개발	참고봉사
1980	7.00	학교도서관	25	9.4	학교도서관	정보기술	사서교사	공공도서관법	비서성
1981	6.18	대학도서관	105	31.4	대학도서관	KORMARC	검색효율	MARC	전산화
1982	2.00	자동색인	14	4.4	자동색인	고서	IFLA	목록이용	관외대출
1983	1.47	교육과정	8	3.4	교육과정	평생교육	경부	도서관학과	도서관학교육
1984	3.17	이용자	26	13.2	이용자	목록규칙	저작권	정보	비도서자료
1985	2.51	정보검색	32	11.6	정보검색	전문도서관	학위논문	연속간행물	정보검색모형
1986	3.13	정보서비스	27	14.4	정보서비스	DDC	상호대차	클러스터분석	제량서지학
1987	1.97	KDC	10	6.2	KDC	목록	전문가시스템	도서관경영	독서교육
1988	1.88	온라인	11	6.4	온라인	검색시스템	교육	정보원	사회과학자
1989	2.68	OPAC	29	11.6	OPAC	분류체계	하이퍼텍스트	브라우저	자동초록
1990	2.16	사서	26	10.0	사서	CD-ROM	정보관리	MEDLINE	문헌정보학교육
1991	1.40	전자통제	7	4.0	전자통제	서지학	과학기술정보	구문	도서관정보학
1992	1.54	정보시스템	10	5.6	정보시스템	인터페이스	정보유통	초록	정보접근
1993	1.59	전문검색시스템	7	6.6	전문검색시스템	정보사회	SCI	전문교육	대학교육
1994	1.34	서지데이터베이스	6	3.8	서지DB	정보탐색	문헌	문헌분류	불논리검색
1995	2.07	인터넷	38	20.2	인터넷	디지털도서관	웹	전자도서관	인용분석
1996	1.48	원문제공서비스	9	6.6	원문제공서비스	연구자	품질평가	KCR	홈페이지
1997	1.44	메타데이터	13	7.6	메타데이터	전자저널	대학	이용자인터페이스	Z39.50
1998	1.33	검색엔진	10	6.2	검색엔진	서비스	XML	학술지	도서관서비스
1999	1.30	더블링크어	6	5.4	더블링크어	웹사이트	평가지표	기록물	지식관리시스템
2000	1.15	기록보존소 외 권	3	3.0	기록보존소	대학기록보존소	디지털정보	인터넷자원	조직문화
2001	1.11	지식관리 외 5건	3	3.0	지식관리	정보정책	불평행동	컨텐츠관리	컴퓨터파일

표 3은 3장의 세 학회지를 통합한 색인어 군집 결과 중 10군집 결과를 네 시기로 구분한 시계열 분석 결과인데, 시기 구분을 통해 각 시기별로 문헌정보학내에서 어떤 연구가 이루어졌는지를 파악할 수 있다.

10군집 결과의 시기별 변화를 살펴보면 다음과 같다. ‘공공도서관’이 모든 시기에 색인어 빈도가 가장 높았고 꾸준히 색인어 빈도 수가 증가하였고, ‘대학도서관’은 1999-2001년 시기를 제외하고 나머지 세 시기의 색인어 빈도가 두 번째로 높았으며, 색인어 빈도가 계속 증가하였다. 또한 ‘정보검색’은 모든 시기의 색인어 빈도가 지속적으로 증가하면서 1999-2001년 시기의 색인어 빈도가 ‘대학도서관’을 추월하여 두 번째로 올라선 것으로 나타났으며, ‘KORMARC’ 과 ‘OPAC’ 도 모든

시기에서 색인어 빈도가 증가하였지만, ‘이용자연구’와 ‘교과과정’은 1999-2001년 시기에서 색인어 빈도가 줄어들었다. 그리고 ‘메타데이터’는 1994-1998년 시기에 처음 색인어 빈도가 나타나면서 증가 추세에 있으며, ‘기록물’과 ‘지식관리’는 1999-2001년 시기에 처음 색인되었다.

표 2. 10군집에 포함된 색인어 목록

군집	소속 색인어 (색인 빈도 순)
1	KORMARC, MARC, 목록규칙, AACR, 전자저널, 비도서자료, 전거통제, KCR, XML, 서지기술, 컴퓨터파일, 편목, 목록법, 전자정보원, 상호협력, 표목, 목록작성, 접근점
2	인터넷, 정보검색, 웹, 시소러스, 검색효율, 자동색인, 검색엔진, 클러스터분석, 전문검색시스템, 하이퍼텍스트, 자동분류, 정보자원, CC, 브라우저, 용어, 한글문헌, 검색성능, 적합성평가, 정보검색시스템평가, 평가기준, 문헌, 키워드탐색
3	공공도서관, 디지털도서관, 정보서비스, 데이터베이스, 학교도서관, DDC, 정보검색시스템, 분류체계, 정보기술, 정보봉사, 색인, 상호대차, CD-ROM, KDC, 저작권, 정보, 정보시스템, 커뮤니케이션, 학술정보, 원문제공서비스, 전자도서관, 전문도서관, SCI, 정보관리, 정보사회, 학위논문, 사서직, 전산화, 주제전문사서, 수서, 품질평가, 웹사이트, 평가지표, 네트워크, 검색시스템, 도서관봉사, 서지학, 목록, 분류, 서지데이터베이스, 인용분석, 정보검색모형, 교육, 학술지, 홈페이지 Z39.50, 도서관네트워크, 검색, 인용문헌, 계량서지학, 동시인용분석, 도서관경영, 대학, 정보원, 사서교사, 국립대학교서관, 멀티미디어, 정보전문직, 국립중앙도서관, DIALOG, 온라인정보검색, MEDLINE, 정보이용행태, 정보통신, 초록, 한글, 시각장애인, 서지, 장서평가, 전문데이터베이스, 지식베이스, 주제, 지식관리시스템, 정보유통, 독서교육, 마케팅, 시스템, 정보전문가, 장서관리, 백과사전, 한국문헌목록정보, ETLARS, 교수, 의학도서관, 한국의학논문, 평가모형, 문헌분류, 보조표, 인터넷자원, SGML, 자동초록, 자원공유, 고서, 참고문헌, 지적구조, 과학, 도서관조직, 도서관호법, 독서요법, 불논리검색, 우리말, 정보자원관리, 정보학교육, 서지정보, 연구동향, 학술잡지, 이용행태, 지식경영, TQM, 전자출판, 디지털정보, 유통, 지적자유, 고문헌, 과학기술정보, 정보정책, 교육정보화, 마케팅전략, 정보교육, 평생교육, 문헌정보학교육, 정보화사회, 자료선택, 산업분야, 정보분석, 정보사서, 정보접근, 효율성
4	OPAC, 이용자, 온라인, 인터페이스, 서비스, 이용자인터페이스, 도서관서비스, 정보탐색, 열람목록, 주제탐색, 불평행동
5	메타데이터, 더블린코어, RDF, 검색도구
6	대학교서관, 사서, 장서개발, 참고봉사, 연속간행물, 정부간행물, 부산대학교, 참고사서, 직무만족, 계속교육, 전자참고서비스, 조직, 조직문화, 참고질문, 표준화, 대학교서관사서, 도서관자동화
7	이용자연구, 정보환경, 연구자, 정보요구, 과학기술, 전자공학, 사회과학, 사회과학자, 정보이용, 정보추구, 정보탐색행태
8	교과과정, 교육과정, 대학교육, 전문가시스템, 전문교육, 학부제, 인공지능
9	기록물, 기록보존소, 대학기록보존소
10	지식관리, 콘텐츠, 콘텐츠관리

그림 5의 내용을 구체적으로 살펴보면, ‘공공도서관’과 ‘대학교서관’은 전체 색인어에서 차지하는 비율이 갈수록 낮아짐으로써 문헌정보학 내에서 갖는 두 영역의 연구 비중이 줄어들고 있음을 보여주었다. 그리고 ‘정보검색’과 ‘KORMARC’

은 대체로 비율이 높아지는 경향을 보임으로써 연도가 계속될수록 문헌정보학 내의 연구 비중이 늘어나고 있으며, 'OPAC'은 연구 비중이 낮아지고 있고, '이용자 연구'는 연구의 비중이 높아짐과 낮아짐을 반복하는 경향을 보이고 있다.

표 3. 시기별 색인어 빈도(4기로 구분)

기간 색인어	1978-1986	1987-1993	1994-1998	1999-2001
KORMARC	10	14	27	65
정보검색	17	21	76	92
공공도서관	89	126	283	313
OPAC	1	7	59	34
메타데이터	0	0	5	20
대학도서관	25	27	81	78
이용자연구	5	5	24	20
교과과정	7	9	17	12
기록물	0	0	0	11
지식관리	0	0	0	9
누계	154	209	572	654

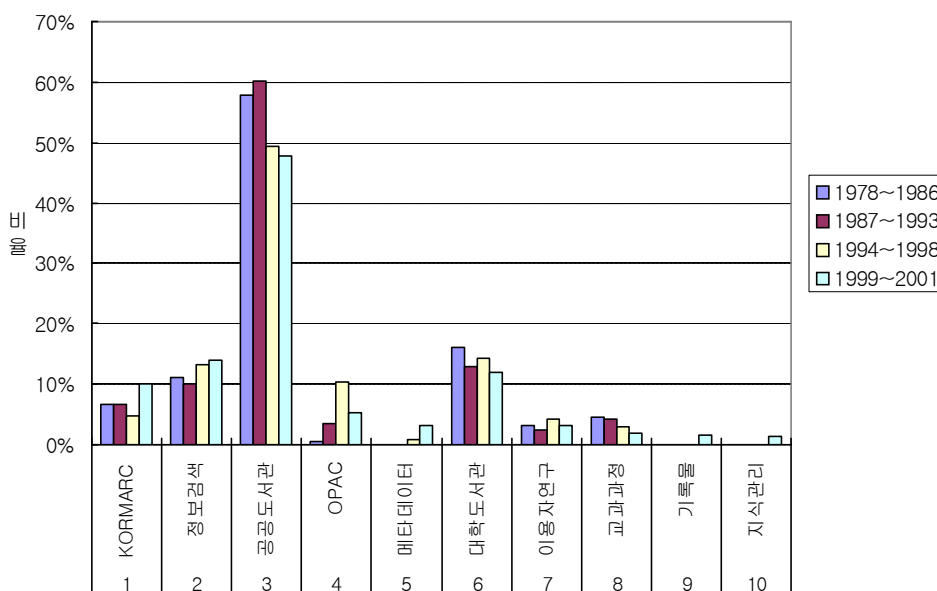


그림 5. 10군집 색인어 시기별 분포(4기로 구분)

## 5 색인어 군집과 분류체계에 의한 지식 구조의 비교

### 5.1 듀이십진분류체계와 색인어 군집의 지식 구조 비교

기존의 전통적인 분류체계와 색인어 군집 결과로 생성한 지식 구조를 비교함으로써 각 지식 체계의 특성을 살펴보고, 보다 나은 지식 구조의 가능성을 검토해보려고 한다. 듀이십진분류체계에 따라 클러스터링의 결과로 생성된 40 군집의 색인어 군집을 구조화한 결과는 다음과 같다.

020 정보, 학술정보, 기록물

.02 기술적 과정

.04 정보검색시스템평가, 정보검색, CD-ROM, 데이터베이스, 정보검색시스템, 검색시스템, 계량서지학, 인용분석, 지식관리시스템, 디지털도서관, 정보시스템, 콘텐츠관리

.21 장서관리

.092 사서

025 .21 장서관리

025 .3 색인

.316 MARC, KORMARC, 메타데이터, OPAC

.32 목록규칙

025. 4 분류

.43 분류체계

.431 DDC

.49 시소러스

025 .52 정보서비스

.58 이용자연구, 이용자

025 .62 상호대차

027 .4 공공도서관

027 .7 대학도서관

---

004 정보기술  
.678 인터넷, 웹사이트  
050 전자저널  
346.0482 저작권

-----  
미분류 색인어 : 연구자, 교과과정

위의 구조화 결과 중 020에 분류한 ‘정보’, ‘학술정보’, ‘기록물’ 등 세 색인어는 듀이십진분류의 문헌정보학 하위 주제에 분류할 수 있는 해당 항목이 없어 최상위 주제인 문헌정보학(020)에 분류하였다. 또한 ‘사서’는 듀이십진분류 본표(Schedules)의 문헌정보학(020)과 특정 직종이나 영역에서 활동하는 인물을 범주로 제시한 보조표 7(Table 7)의 문헌정보학과 관련된 영역의 직업을 가진 사람(092)의 번호를 합성한 결과이고, 나머지 결과는 모두 본표에 제시된 번호에 의해 분류되었는데, ‘정보기술’(004)과 ‘인터넷’(004.678) 그리고 ‘웹사이트’(004.678)는 데이터처리 및 컴퓨터과학 영역에 분류되었고, ‘저작권’(346.0482)은 법학 영역 중 민법에 분류되었다.

그리고 색인어 군집 결과를 구조화한 결과, 정보소장 및 검색시스템(025.04) 등 정보 검색 영역이 가장 많은 수를 차지하였는데, 듀이십진분류의 주제 전개는 ‘도서관 운영(library oprations)’에 초점을 맞추어서 분류체계를 구성하였기 때문에 정보 검색 영역의 하위 주제를 전개하지 않은 것으로 보인다.

## 5.2 한국십진분류체계와 색인어 군집의 지식 구조 비교

한국십진분류체계는 듀이십진분류체계의 영향을 받았지만 오랜 기간 동안 사용되어 오면서 우리 환경에 맞게 수정된 분류체계이고, 우리 문헌정보학의 지식 구조를 규명함에 있어 우리의 유일한 분류체계인 **한국십진분류법**을 이용하는 것이 의미가 있다고 판단하여, 한국십진분류체계의 지식 구조에 의해 국회도서관이 문헌정보학관련 학술지에 실린 기사를 대상으로 배정한 색인어 군집의 지식구조를 서로 비교하였다.

다음에 제시한 결과는 한국십진분류체계에 따라 클러스터링의 결과로 생성된

40군집의 색인어 군집 각각을 구조화한 내용이다.

- 020 정보, 학술정보, 기록물
    - .71 교과과정
  - 023 .2 사서
  - 024 .31 OPAC
    - .316 MARC, KORMARC, 메타데이터
    - .32 목록규칙
  - 024 .4 분류체계, 분류, 색인
    - .41 DDC
    - .47 시소리스
  - 024 .6 장서관리
  - 025 .168 상호대차
  - 025 .2 정보서비스
  - 025 .4 이용자연구, 이용자
  - 026 .3 공공도서관
  - 027 .6 대학도서관
- 
- 004.2 정보검색시스템평가
  - 005 정보기술, 인터넷, 정보검색, CD-ROM, 데이터베이스, 정보검색시스템, 웹사이트, 검색시스템, 계량서지학, 인용분석, 지식관리시스템, 정보시스템, 콘텐츠관리
  - 011 저작권
  - 050 전자저널
- 미분류 색인어 : 디지털도서관, 연구자

위의 구조화 결과 중 ‘정보’, ‘학술정보’, ‘기록물’ 등의 색인어는 한국십진분류법 중 문헌정보학의 하위 주제에 해당하는 적절한 번호가 없어 최상위 주제이며, 하위 주제를 포괄할 수 있는 문헌정보학(020)에 분류하였다. 그리고 색인어 군집 결과 중 가장 많은 수를 차지하는 정보 검색 영역은 모두 종류의 프로그래밍, 프로그램, 데이터(005)의 하위 개념으로 포함되어 있고, ‘디지털도서관’ 같은 색인어는

분류하기에 적절한 주제 영역과 수준을 찾기가 어려웠다.

한국십진분류법이 도서관의 자료를 분류하는 기능을 수행하는 것이 근본적인 존재 이유이므로 문헌정보학 자체의 지식을 구조화하기에 부적절한 체계를 가지고 있는 것으로 보인다. 따라서 문헌정보학의 지식 구조를 보다 명확하게 구조화하기 위해서는 우선 문헌정보학의 주제 영역을 새롭게 분석하여 지식을 구조화할 수 있는 분류체계로서의 가능성을 고려할 필요가 있으며, 구체적으로는 정보 검색 영역과 정보 조직 영역 그리고 정보기술 영역 등을 문헌정보학 내에서 합리적으로 구조화할 수 있어야 하며, 새롭게 등장하는 지식을 신속하게 수용할 수 있는 체계상의 유연성을 가지는 것이 바람직 할 것으로 보인다.

## 6 결 론

정보관리학회지, 한국도서관정보학회지, 한국문헌정보학회지 색인어 군집을 생성하여 문헌정보학의 주요한 주제 개념들을 밝혀내었으며, 세 학회지를 통합한 색인어 군집에 의한 분석 결과를 통해, 목록 영역을 포함하는 정보 조직 영역 군집 정보 검색 영역, 인터넷과 관련된 영역이 순서로 중요도가 높음을 알 수 있었다.

또한 초출색인어를 이용한 시계열 분석의 결과에 의해 ‘인터넷’, ‘디지털도서관’, ‘웹’, ‘전자도서관’ 등의 초출색인어가 등장한 1995 년부터 문헌정보학 발전과정의 분기점이었음을 확인할 수 있었다.

그리고 색인어에 의해 문헌정보학 영역의 지식 구조를 설명하는 방법은, 색인어의 계층관계를 완벽하게 구조화하지는 못했으며, 전통적인 분류체계에 의한 지식 구조가 도서관의 자료를 분류하는 것이 목적이라는 사실을 확인하였다.

위의 결론을 토대로 이 연구와 관련된 향후의 연구 과제는 다음과 같다. 용어간의 계층관계를 구조화함에 있어 미흡함을 보인 클러스터링에 의한 용어 군집 모형과 도서관의 자료를 분류하는 기능을 수행하는 기존의 분류체계와는 별도로 문헌정보학 학문 발전의 역동성을 수용할 수 있는 지식 구조화 수단의 가능성을, 시소러스와 같은 지식 표현 체계를 개발하려는 시도에서 찾아야 할 것이다.

또한 색인어 군집 결과에서 확인할 수 있었듯이, 정보 조직 목록 영역에 비해 상대적으로 발전이 지체되고 있는 정보 조직 분류 영역의 연구 활성화를 위해 전통적인 듀이십진분류와 한국십진분류법과 같은 도서관의 자료분류표 자체에 대한

연구와 함께 지식을 구조화할 수 있는 새로운 분류체계에 관한 연구도 이루어져야 할 것이다.

## 참 고 문 헌

- 김태수. 1999. 『목록의 이해』. 서울: 한국도서관협회.
- 김태수. 2000. 『분류의 이해』. 서울: 문헌정보처리연구회.
- 이미경. 2002. 『동시출현 단어 분석을 통한 지식구조 파악에 관한 연구: 인공지능 분야를 대상으로』. 석사학위논문. 연세대학교.
- 이재윤. 2002. 『단어 동시출현 기반 질의확장의 성능 최적화에 관한 연구』. 박사학위논문, 연세대학교.
- 사공철, 김태수, 정영미, 최석두 공편 『정보학사전』. 2001. 서울: 문헌정보처리연구회.
- Aitchison, Jean, Alan Gilchrist, and David Bawden. 2000. *Thesaurus Construction and Use: A Practical Manual*. London: Aslib.
- ANSI Z39.4. 1968. *Basic Criteria for Indexes*. ANSI Z39.4. New York: the Institute.
- Callon, M., John Law, and A. Rip. 1986. *Mapping of the dynamics of science and technology*. London: Macmillan. 재인용: 이미경. 『동시출현 단어 분석을 통한 지식구조 파악에 관한 연구: 인공지능 분야를 대상으로』. 석사학위논문. 연세대학교
- Coulter, Neal, Ira Monarch, and Suresh Konda. 1998. "Software engineering as seen through its research literature: a study in co-word analysis." *Journal of the American Society for Information Science*, 49(13): 1206-1223.
- Ding, Y., G. Chowdhury, and S. Foo. 2001. "Bibliometric cartography of information retrieval research by using co-word analysis." *Information Processing and Management*, 37: 817-842
- Fugmann, Robert. 1993. *Subject Analysis and Indexing: Theoretical Foundation*

- and Practical Advice*. Textbooks for Knowledge Organization, vol. 1. Frankfurt/Main: Indeks Verlag.
- He, Q. 1999. "Knowledge discovery through co-word analysis." *Library Trends*, 48(1): 133-159.
- ISO-2788. 1986. 단일언어 시소러스 제정 및 개발지침. 정동열 역. 『시소러스 개발지침』. 서울: 문헌정보처리연구회, 1994: 3-93.
- Iyer, Hemalata. 1995. *Classificatory Structures: Concepts, Relations and Representation*. Textbooks for Knowledge Organization, vol. 2. Frankfurt/Main: Indeks Verlag.
- Milstead, Jessica L. 1998. *ASIS Thesaurus of Information Science and Librarianship. 2nd ed.* Medford: Information Today, Inc.
- Noyons, E.C.M. and A.F.J. van Raan. 1998. "Monitoring scientific developments from a dynamic perspectives: self-organized structuring to map neural network research." *Journal of American Society for Information Science*, 49(1):68-81.