

의학관계잡지의 데이터베이스 구축에 관한 기초연구

가톨릭대학교 의과대학 도서관

정 학 수

An Establishment of Database for the Medicial Journal

Jung Hag Su

Medical Library, Catholic University Medical College

I. 서 론

사서는 이용자가 무엇을 얻고 싶어 하는지에 대해 항상 정확하게 파악하려고 하고 알맞은 정보를 제공할 수 있어야 한다. 이를 위해 자료의 축적 방법이라든지 검색방법, 전달되어질 통로 등에 대해 많은 연구와 발전이 있어 왔지만 필요로 하는 정보적 요구에 대해 만족할만한 수준의 봉사는 되고 있지 않다는 것이 현실이다.

이러한 현실이 근래에 들어와서 컴퓨터라는 매체를 통해 도서관과 이용자간의 차이를 조금이나마 해결하고 있다고 볼 수 있다.

다양한 형태로 배출되는 일련의 자료를 컴퓨터의 기억장치에 저장시키고 조직화 시켜 필요로 하는 순간 얻을 수 있다는 것이 기존의 방법보다는 그속도나 효율성에서 분명 한단계 진일보 하였다고 보는 것이 타당하리라 여려진다.

컴퓨터에 어떤 자료를 저장하여야 하는가? 어떻게 저장시켜야 하는가? 어떻게 조직화 시켜야 하는가? 얼마나 효율적인 정보를 얻을 수 있게 하는가? 하는 물음 들에 대한 가장 정확한 해답을 얻기 위하여 우리는 많은 노력과 연구를 하고 있는 것이다.

본고에서는 이러한 일련의 상황을 도서관의 전산화 중 국내에서 발행되는(외국잡지 제외) 의학 관련 학술잡지의 데이터베이스를 구축하는데 필요한 기초작업으

로서 데이터베이스의 구축이 왜 필요하며, 데이터베이스화 하기 까지의 최대한 빠른 방법을 생각해 보며, 과학적인 데이터베이스 구축의 방법에 대해 언급하고자 한다.

II. 본 론

1. 데이터베이스의 정의

데이터베이스란 정보처리를 위한 기본적 시스템구성요소인 데이터의 집합으로 데이터의 종류, 양, 구성 및 그의 색인 구조 등이 여러 목적에 이용될 수 있도록 적합하게 설계, 작성된 데이터 파일의 군으로서 다수의 사용자가 이용가능한 필요한 정보를 얻기 위하여 불필요한 정보를 피하여 서로 연관된 데이터를 체계적으로 저장해 놓은 것으로 정의 할 수 있다.

데이터베이스는 기록되는 정보의 내용과 유형에 따라 참조데이터베이스(Reference Database)와 소스 데이터베이스(Source Database)로 나눌 수 있는 데 그 내용은 다음 표 1과 같다.

2. 데이터베이스화할 내용

데이터 베이스를 구축하는데 있어서 국내에서 발행되는 의학 관련 잡지의 전문을 데이터베이스 하는 것이 가장 바람직 하고도 최선의 방법이 될 수 있지만 국내에서 발행되는 의학 관련 학술잡지에서 치과학, 간호학 및 수의학 잡지를 제외한 113종의 잡지중 분

표 1. 데이터베이스의 종류

종 류	형 식	내 용
참조 데이터베이스		
서지 데이터베이스	문헌 목록	초록, 서지사항
리퍼럴 데이터베이스	디렉토리	인명, 기관명, 잡지소장목록
소스 데이터베이스		
전문 데이터베이스	전문	잡지논문, 법령, 백과사전
수치 데이터베이스	수치	각종통계, 실험, 관측데이터
문자·수치 데이터베이스	문서, 수치	신문, 조사, 분석보고서
화상·영상 데이터베이스	화상, 영상	화상, 영상데이터, 프로그램

식된 96종의 잡지에 1989년에서 1990년까지 2년간 실린 논문 수만도 11,160편으로 그 자료의 양이 방대하여 전문데이터베이스를 만든다는 것은 시간과, 인원, 장비, 경비가 절대 부족한 현실에서 추진하기에는 매우 어렵고도 힘든 작업이 될 것이다.

이에 초록과 서지사항을 포함하는 서지데이터베이스(Bibliographic Database)를 구축하는 것이 바람직 하리라 여겨진다.

3. 국내 의학 관련 학술잡지의 검색도구와 시스템

발표된 논문이나 자료들이 본래의 존재 자체로서 큰 의미를 갖지만 정보원으로서 학문의 발전에 끼치는 영향을 간과해서는 안된다. 즉 이미 발표된 정보는 새로운 가치의 정보를 재생산 하는데 밑거름이 된다는 것이다. 그렇다면 이러한 정보원에 접근하는 방법 또한 매우 중요한 것이다. 어떠한 학자나 연구인이 자기분야에 관련된 정보를 충분히 찾아서 이용할 때 연구활동이 올바르게 진행될 것이기 때문이다. 이런 이유로 탐색수단의 필요성이 대두되었고 등장하게 된 것이다.

이영철 등이 발표한 국내의학논문에 인용된 잡지명 순위에 따른 국내잡지의 분포를 분석하여 보면(표 2) 국내 발표 논문에서도 국내 의학 잡지의 인용빈도가 크게 미흡하다는 것이 나타나고 있는데, 그 이유를 의학 도서관에 근무하는 사서의 입장에서 보면 외국잡지에 비해 검색자료와 시스템이 잘 갖추어 지지 않았다는 데서도 한 원인을 찾을 수 있다.

국내 의학문헌의 검색을 하기 위해서는 정기간행물 기사 색인, 의학 약학 논문 총목록, Index Medicus

표 2. 국내의학논문에 인용되는 잡지의 순위에서 국내잡지의 분포

순 위	잡지수	누 계
1~ 10위	1	1
11~ 20위	1	2
21~ 30위	1	3
31~ 40위	1	4
51~ 100위	9	13
101~ 200위	12	25
201~ 500위	38	63
501~1,000위	38	101
1,001~1,530위	39	140
계	140	140

(Korea), Medical Abstracts Korea, Korean Medical Abstracts, 학술총람, Korean Index Medicus 등의 색인지를 이용하여야 하는데 이는 모두 도서의 형태로 출판되고 있어 정보의 출현에서부터 찾을 수 있는 도구의 정비까지 걸리는 시간적 차이 때문에 정보적 가치가 반감되어 가고 있다.

적시에 정확한 정보를 검색할 수 있는 시스템을 개발하는 것이 무엇보다도 선행되어야 할 과제인 것이다.

4. 데이터베이스화 할때 고려할 사항

데이터베이스를 구축하는데 있어서 다음과 같은 사항은 충분히 고려되고 연구되어야 하는데 그 첫째로는 축적 및 검색되기 위한 데이터의 종류와 그양이며, 둘

넷째 데이터의 저장매체, 셋째 데이터를 수집하는 방법, 넷째 데이터가 축적되고 검색될 기술등이다.

첫째 데이터의 종류와 양에 있어서 서지사항과 초록을 기본적으로 구성요소로 볼 때 문자-데이터가 그 주종을 이루게 될 것이므로 화상과 표가 들어가는 데이터베이스보다는 상대적으로 저장용량을 적게 차지하게 되어 많은 양을 한꺼번에 저장하여 검색하게 되므로 효율성을 높일 수 있을 것이다.

둘째 데이터의 저장매체로는 대용량을 한꺼번에 저장하여야 하고, 데이터에 접근하는 속도가 빨라야 하는데 현재까지의 과학기술로는 CD-ROM이 가장 유력하다. CD-ROM은 디지털 정보를 레이저 빔을 이용하여 새겨넣은 재생전용의 매체로 고밀도의 대용량(최저 540Mbyte에서 최고 680Mbyte)을 갖고 있으며 기록된 데이터를 읽기만 하고 지울 수는 없다. CD-ROM에 기록된 데이터를 검색하는데 소요되는 시간은 평균 0.8초이며 이는 마그네틱 디스크에 비해서는 속도가 느리지만 마그네틱 테이프나 플로피 디스크보다는 빠른 속도이다.

셋째 데이터를 수집하는데 있어서 이미 발행되어진 잡지에 실린 논문을 데이터베이스화 하는데는 시간과, 비용, 인력이 충분하게 갖추어지면 언젠가도 가능하므로 앞으로 발행되어질 잡지에 대해서 살펴보기로 한다.

한권의 잡지는 학회, 협회, 학교, 병원, 정부, 군기관, 연구소, 제약회사, 출판사, 신문사 등 여러 기관에서 발행하는데 각 발행처에서 필요로 하는 논문을 원저자로 부터 수집하여 자체 심사기준에 의해 심사한 다음 출판사로 넘겨 도서의 형태로 일반에 알려지게 된다. 이러한 과정에서 데이터베이스 구축에 필요로 하는 서지사항과 초록을 얻을 수 있어야 하는데, 얻는 시점은 다음과 같은 전제조건에 충실하여야 한다.

전제조건으로는 논문의 인쇄시점과 비슷하게 데이터베이스가 구축되어 논문의 발표와 동시에 검색가능하게 되어야 할 것이며, 정확한 원자료를 얻을 수 있는 시점이어야 한다.

즉 잡지의 발행처에서 검색시스템에 대한 충분한 가치의 인정으로 원자료를 출판사로 보내는 시점에서 또 다른 한본의 원논문이 검색시스템 유지기관에 도착하게 된다면 효율적인 데이터베이스 구축이 이루어지게 될 것이다(그림 1).

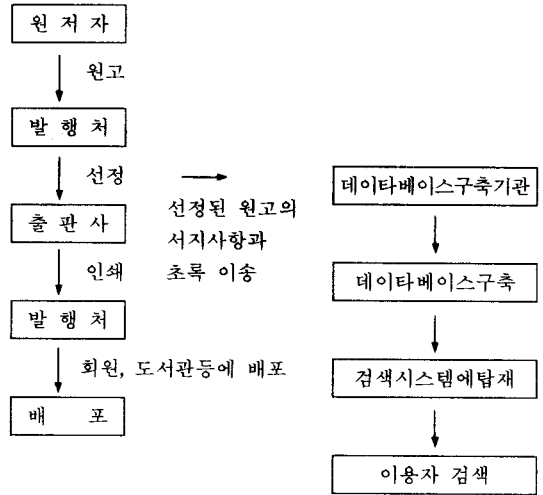


그림 1. 데이터베이스 구축까지의 과정도.

넷째 데이터가 축적되고 검색될 기술에는 축적매체와 검색시스템의 유기적인 연관관계에 의해서 최대의 능력이 발휘될 수 있도록 설계되어야 하겠으며, 일대일의 검색 시스템을 만들 것인가, 일대 다의 검색시스템이 되어야 하는가의 문제, 구축된 데이터베이스를 이용자가 온라인에 의해 검색되도록 할 것인가, 오프라인에 의해 검색되어야 할 것인가를 결정하고, 검색용어의 선택에 있어서도 자연어에 의한 검색 방법을 취할 것인가, 주제어에 의해 검색될 것인가, 두가지 모두를 혼용한 방법으로 설계되어야 할 것인가 하는 여러 측면들이 앞으로 고려되어야 할 것이다.

III. 결 론

이제 우리는 검색시스템의 획기적인 방향 전환에 한 발을 들여 놓은 상태이다. 의학 관계 잡지의 데이터베이스 구축을 하는데 앞으로 여러분야에 걸쳐 다양한 논의가 거듭되어야 한다.

앞으로는 키워드가 하나하나의 문자를 타이핑하여 저장하던 시대에서 이미지파일로 받아들여 문자화하는 기술이 발달하여 대용량의 정보도 빠른 시간내에 데이터베이스를 구축할 수 있게 됨에 따라 짧은 시간과 적은 비용을 만족할만한 검색시스템을 유지할 수 있으리라 본다.

참 고 문 헌

- 1) 윤화진: 입법활동지원을 위한 자료의 데이터베이스화에 관하여. 국회도서관보, 22권 1호, 1985, pp87-96.
- 2) 이승하: 한국의학관계잡지 통계분석. 한국의학도서관, 15권 1-2호, 1988, pp. 1-4.
- 3) 김호배, 윤봉자: 우리나라 의학정보 색인사업. 한국의학도서관, 15권 1-2호, 1988, pp 15-23.
- 4) 강미혜: 도서관정보매체로서 CD-ROM의 유용성과 그 이용. 국회도서관보, 29권 4호, 1992, pp5-25.
- 5) 이영철, 윤봉자, 오은숙, 김호배: 효율적인 상호대치를 위한 국내의학잡지의 인용문헌 분석. 한국의학도서관, 18권 1-2호, 1991, pp1-48.