

데이터베이스에 대한 고찰

인천전문대학 도서관과 강사

최 승 주

A Study on the Database

Seung Joo Choi

Library Science, Incheon Junior College

We must newly recognize the librarian's professionalism in the post-industrial society. Therefore, this paper is to investigate the development, an outline, system and the industry of database.

I. 서 언

인류 사회의 발전과정은 사회를 근본적으로 변혁시킨 혁신기술군인 사회기술에 의해 변화, 발전되어 왔다. 우리가 오늘날 직면하고 있는 정보사회는 정보력과 기술력이라고 하는 사회기술에 의해 빠른 속도로 변화해 감으로써 물질의 가치에 비해 상대적으로 정보의 가치가 커지고 있으며, 정보가 중요한 자원이 되고, 정보의 생산, 저장 및 분배가 사회의 핵심활동이 되고 있다. 이러한 정보사회의 특징중의 하나인 정보의 양적인 팽창은 사회의 여러 분야에서 필수적이고 기본적인 전제조건으로서 정보처리 과정의 기계화 즉 컴퓨터에 의한 정보처리 및 정보관리를 필요로 하게 되었고, 오늘날 정보의 전달매체로서 점차 그 비중이 커지고 있는 데이터베이스가 이러한 정보사회의 요청에 의해 나타나게 되었다. 정보사회의 전제조건으로서 나타난 데이터베이스는 정보사회가 발달시킨 기술력에 의해서 빠른 속도로 성장, 발달해 가고 있으며, 전통적인 정보 전달매체인 인쇄매체를 대체하는 혁신적인 미래의 정보 전달매체로서 인식되고 있다.

도서관이 보다 적극적이고 능동적으로 정보사회라고 하는 환경에 대처해 나가기 위해서는 사회와 산업의

정보화에 따른 데이터베이스의 성장, 발달에 대한 이해와 데이터베이스의 효율적인 이용을 위한 사서들의 전문적인 지식을 필요로 한다. 그러므로 본 고에서는 데이터베이스의 출현과 발전과정, 데이터베이스의 개요, 데이터베이스 시스템과 산업의 구조등에 대해 살펴봄으로써 데이터베이스에 대한 전반적인 이해와 정보사회가 요구하는 사서직의 전문성을 새롭게 인식하고자 한다.

II. 데이터베이스의 역사적 발전

1. 데이터베이스 개념의 출현과 발전

데이터베이스의 개념이 나타나기 시작한 것은 1950년대 후반부터이다. 산업사회가 고도로 발달하면서 종래의 파일구조 방식으로는 다양한 업무의 효율성을 기할 수 없었기 때문에 데이터베이스 개념이 출현하였다. 즉 업무가 다양화 하면 할수록 발생하는 방대한 자료를 효율적으로 처리하고 여러 업무에서 공통으로 자료를 이용하는 방안을 모색하게 되었다. 업무가 다양화하면 할수록 그에 따라서 데이터 파일이 여러개 생겨나고, 이를 유지하고 관리하는데 많은 노력과 경비가 소요되기 때문에 모든 자료를 가급적 중복을 피하여 한곳에 모아 놓고, 이를 모든 이용자가 동시에

사용할 수 있게 하는 데이터베이스가 필요하게 되었다.

그러나 데이터베이스란 단순히 모든 업무의 데이터를 한 곳에 모아 놓기만 하면 되는 것이 아니라 데이터 상호간에 어떤 특정한 관계를 맺어주어야만 하고, 또 그 데이터베이스를 이용할 수 있도록 하는 소프트웨어가 있어야 한다. 1950년대에 들어서면서 이에 대한 연구가 정부, 기업, 컴퓨터 관련 업체 등 여러 곳에서 활발히 진행되었고 그에 따라서 데이터베이스 개념이 출현하게 되었다.

오늘날의 데이터베이스를 가능하게 한 기술에는 여러가지가 있다. 데이터베이스 내용을 저장하기 위한 기억매체 기술, 데이터베이스의 구축, 검색 및 관리에 사용되는 DBMS, 원거리 온라인 서비스를 위한 통신망 기술, 데이터베이스가 널리 보급되면서 시작된 표준화 문제등이다.

첫째, 데이터베이스 내용을 저장하기 위한 기억매체로서 랜덤 액세스(Random Access)가 가능한 대용량 기억장치인 디스크가 개발되었다. 대용량 기억장치로서 자기테이프 밖에 이용할 수 없었던 시대에는 데이터의 검색 및 기록이 반드시 앞에서부터 순차적으로 처리되어야 했기 때문에 데이터베이스를 생각할 수 없었다. 그러나 랜덤 액세스가 가능한 대용량 기억장치인 디스크가 출현하면서 데이터베이스가 가능했고 또한 발달할 수 있었다. 최근에는 디스크 매체를 증가하는 광매체가 개발되어서 미래의 매체로서 주목되고 있다. 이 광디스크매체는 대기억용량, 고기억밀도, 랜덤

액세스, 기억내용의 안전성등의 특징을 가지고 있으며, 또한 광매체의 대용량은 정보량이 많은 음성이나 화상등을 처리 할 수 있어서 멀티미디어 데이터베이스의 발전이 기대되고 있다.

둘째, 데이터베이스를 구축, 검색하고 관리하는 소프트웨어인 DBMS(Database Management System)가 개발되었다. 이 DBMS는 데이터의 중복을 방지하여 컴퓨터의 메모리 공간을 절약시켜 주고, 또한 다양한 이용자에게 다양한 형태로 정보를 제공해주는 장점을 가지고 있다. DBMS는 몇 단계를 거치면서 발전해 왔다. 1970년대 이전의 시기는 태동기로서 1950년대에 개발된 자기디스크로 인해 랜덤액세스에 의한 즉각적이고 효율적인 데이터 액세스는 가능해졌으나 데이터당 축적비용이 상승함에 따라서 DBMS의 필요성이 인식되기 시작한 시기이다.

1970년에서 1980년 까지의 시기는 DBMS의 성장기로서 학문으로서 데이터베이스가 확립되고 정보기술의 기초로서 DBMS의 성장기로서 학문으로서 데이터베이스가 확립되고 정보기술의 기초로서 DBMS가 성장한 시기이다. 1980년대 이후는 상용 데이터베이스가 개발되었고, 퍼스널 컴퓨터의 비약적인 발전과 보급에 따라서 퍼스컴용 DBMS가 개발되는 등의 DBMS 성숙기이다. DBMS를 채택한 데이터베이스 시스템의 구조는 그림 1과 같다.

셋째, 데이터베이스의 원거리 온라인 서비스를 가능하게 하는 정보통신 기술이 발달하였다. 데이터베이스의 상용화에 있어서 가장 핵심적인 것은 정보를 주고

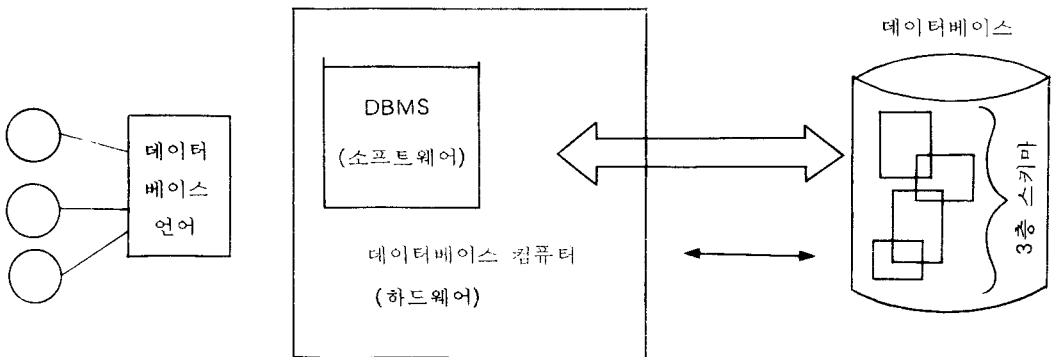


그림 1. 데이터베이스 시스템의 구성.

받는 통로인 통신망이다. 이 통신망 기술의 발달은 데이터베이스 서비스의 질적 양적 향상에 매우 큰 영향을 주었다.

1960년대에 컴퓨터 통신이 처음 시작되었을 때 전용 데이터망이 아닌 PSTN(Public Switched Telephone Network)이 통신의 기본이었다. PSTN은 전용 통신망이 아닌 공중 전화망이었기 때문에 데이터베이스 서비스에 기여하지 못했다. 그 후 전용 데이터망의 하나인 PSDN(Packet Switched Data Network)이 개발됨에 따라서 온라인 정보검색 서비스가 가능하게 되었다. PSDN이 등장하기 전에는 특정한 이용자들만이 전용 회선을 통한 실시간(Real Time) 검색을 할 수 있었으나 PSDN이 개발됨에 따라 정보서비스가 일반화 되었으며 데이터베이스의 상용화가 이루어지게 되었다. 최근의 정보통신망의 동향은 음성 데이터, 화상정보가 통합된 형태로 송·수신될 수 있는 통합디지털 방식의 네트워크 시스템인 ISDN(Integrated Service Digital Network)이 개발되고 있으며, 미래의 바람직한 이상적인 통신망으로 기대되고 있다.

그 밖에 데이터베이스 프로그래밍 언어의 표준화, 시스템 기능의 표준화, 데이터베이스의 표기, 작성의 표준화 등의 표준화 활동과 데이터베이스의 법적 권리 문제인 저작권의 문제 등이 데이터베이스 관련 기술과 환경으로서 데이터베이스의 발달에 기여했다.

2. 데이터베이스의 발전

데이터베이스는 각 시대별로 그 성격을 달리하며 발전해 왔다. 1960년대는 상용 데이터베이스가 실현되기 이전으로서 정부의 강력한 지원과 국가의 프로젝트 수행을 주축으로 한 데이터베이스 산업의 기초가 다져진 시기이다.

- 1960년 McCarthy에 의해서 Time Sharing System이 개발되었다.
- 1961년 CAS (Chemical Abstracts Sefcicel)에서 Chemical Title이라는 정기간행물이 최초로 전산 제작에 의해서 발행되었다.
- 1964년 미국 국립의학도서관(NLM)에서 Batch 방식의 MEDLARS 개발
- 1965년 SDC사에서 ORBIT 개발

- 1966년 DIALOG 개발

1970년대는 데이터베이스가 여러 기술을 중심으로 상용화 되기 시작했으며, 공중 통신망을 통한 데이터베이스의 온라인 서비스가 시작된 데이터베이스 산업이 발전하게 된 시기이다. 데이터베이스의 주 대상분야는 과학, 기술분야였으며, 이용자는 학자, 기술자 등의 전문가였다

- 1971년 MEDLINE, OCLC 공중망 이용 서비스 개시
- 1972년 DIALOG, ORBIT 최초의 온라인 상용 데이터베이스 서비스
- 1973년 LEXIS 개시
- 1975년 VAN service 시작
- 1977년 BRS 서비스 개시
- 1979년 COMPUSERVE 서비스 시작

1980년대는 데이터베이스 시장이 확대되고, 기술이 아니라 수요가 데이터베이스 산업을 주도해 가며, 데이터베이스 대상분야가 비즈니스를 중심으로 확대된 데이터베이스 시장이 확대된 시기이다.

- 1984년 Gateway service 시작
- 1985년 상업용 CD-ROM 개발
- 1988년 DIALOG를 이용해서 화상정보 검색.

1990년대는 퍼스널 컴퓨터의 광범위한 보급으로 인해 데이터베이스 이용자가 일반 이용자로 다양화 되어지고, 대상분야도 취미, 레저 등의 생활정보에 이르기까지 확대되어진 시기이다.

III. 데이터베이스의 개요

1. 데이터베이스의 정의

데이터베이스란 모든 종류의 관련정보를 컴퓨터를 이용하여 효율적으로 검색하기 위하여 체계적으로 정리 축적한 것이라고 할 수 있다. CODASYL(Conference on Data Systems Languages)의 데이터베이스 정의는 다음과 같다¹⁾. 데이터베이스란 어느 한 조직의 다수 응용 시스템이 사용하기 위해 통합 저장된 운영 데이터의 집합이라고 정의할 수 있다. 이 데이터베이스의 정의는 몇가지 의미들을 내포하고 있다.

첫째, 데이터베이스는 통합된 데이터(Integrated

1) 丘龍元. 컴퓨터과학개론, 대림, 1992, pp454-455

data)이다. 이것은 데이터베이스에서 데이터들이 원칙적으로 중복되지 않는다는 것을 의미한다. 데이터의 중복은 일반적으로 복잡한 부작용을 초래한다. 그러나 중복성은 완전히 배제하는 것은 아니다. 경우에 따라서는 효율성 때문에 중복을 불가피하게 허용할 수도 있는데 이것을 최소의 중복(Minimal redundancy) 또는 통제된 중복(controlled redundancy)이라고 한다.

둘째, 데이터베이스는 저장된 데이터(Stored data)이다. 이것은 책상이나 화일 캐비닛에 들어 있는 데이터가 화일 캐비닛에 들어 있는 데이터가 아니라 테이프나 디스크와 같이 컴퓨터가 액세스하여 처리할 수 있는 저장장치에 수록된 데이터를 말하는 것이다.

셋째, 데이터베이스에 있는 데이터는 운영 데이터(operational data)이다. 어느 조직이나 기관들은 그 기능을 수행하는데 반드시 유지해야 할 데이터가 있다. 그러므로 데이터베이스는 단순한 데이터의 집합이 아니라 그 존재목적이 뚜렷하고 유용성을 가지고 있는 없어서는 안될 그러한 데이터의 집합인 것이다. 단순한 입출력 데이터 작업 처리에 필요한 임시 데이터는 운영 데이터가 아니다.

넷째, 데이터베이스는 공용 데이터(Shared data)이다. 이것은 어느 하나의 응용 프로그램이나 응용시스템을 위한 데이터가 아니라 한 조직에 있는 여러 응용시스템들이 공동으로 소유하고 유지하며 이용하는 공용 데이터인 것이다. 즉 상이한 다수 사용자들이 서로 다른 목적으로 데이터 베이스의 같은 데이터를 공유할 수 있다는 것을 의미한다.

일본의 저작권법에서는 데이터베이스를 논문, 수치, 도형 및 그 밖의 정보의 집합물로서 이같은 정보를 전자계산기를 이용하여 검색할 수 있도록 체계적으로 구성한 것²⁾이라고 정의하고 있다.

2. 데이터베이스의 특징

데이터베이스는 다음과 같은 특성을 가지고 있어야 그 기능을 다할 수 있다.

첫째, 실시간 접근성(Real Time Accessibility). 데이터베이스를 구축해 놓고 이를 효율적으로 사용하

기 위해서는 실시간 처리가 가능해서 어떠한 문의에 대해서도 즉각적인 응답이 이루어지도록 되어야 한다. 특히 오늘날과 같이 급변하는 상황에서 일괄처리에 의한 처리응답은 데이터베이스 이용자들의 요구를 만족시켜 주기 못할 것이다. 이용자들이 원하는 질의의 처리응답 방식을 일괄처리에 의한 방식보다는 즉각적이고 효율적인 실시가 처리방법일 것이다.

둘째, 지속적인 변화(continuous evolution) 데이터베이스는 정적이 아니고 동적이다. 즉 데이터베이스는 삭제, 갱신을 통해서 가장 최근의 정확한 데이터를 유지하면서 성장한다.

셋째, 동시공유(concurrent sharing). 데이터베이스는 상이한 목적을 가진 응용을 위한 것이기 때문에 동시에 여러 사용자가 접근 이용할 수 있어야 한다. 이는 시차를 두고 여러 사람이 같이 사용한다는 의미보다는 어느 순간에 두사람이상의 사용자에게 의하여 하나의 화일이 사용됨을 의미한다. 즉 상이한 다수의 사용자들이 서로 다른 목적으로 데이터베이스의 같은 데이터를 공유함을 의미한다.

넷째, 내용에 의한 참조(content refernece) 데이터베이스 내에 있는 데이터 레코드들은 주소나 위치에 의해 참조되는 것이 아니라 데이터의 내용 즉 데이터가 가지고 있는 값에 의해서 참조된다.

3. 데이터베이스의 유형

데이터베이스는 그 형태나 용도, 기능에 따라서 다양한 유형으로 분류할 수 있는데, EUSIDIC(European Association of Scientific Information Dissemination Centers)³⁾의 분류, Cuadra⁴⁾의 분류, 일본의 데이터 통신협회⁵⁾의 분류등은 데이터베이스를 1차정보 데이터베이스와 2차정보 데이터베이스로 구분하고 있다.

1) 1차정보 데이터베이스

1차정보 데이터베이스는 fact database, data-

3) EUSIDIC database guide, oxford, Learned Information 1983, p.324

4) J. wanger, R.N. Landau, "Nonbibliographic online database service" JASES. Vol. 31, NO. 3, 1980. p171-180

5) 日本データ通信協會, データベース, ネットワークの採來動向 デルフマによる, 東京, 日本 データ通信協會, 1986, p411

2) データベース振興センタ編, データベース白書, 1990. p. 16.

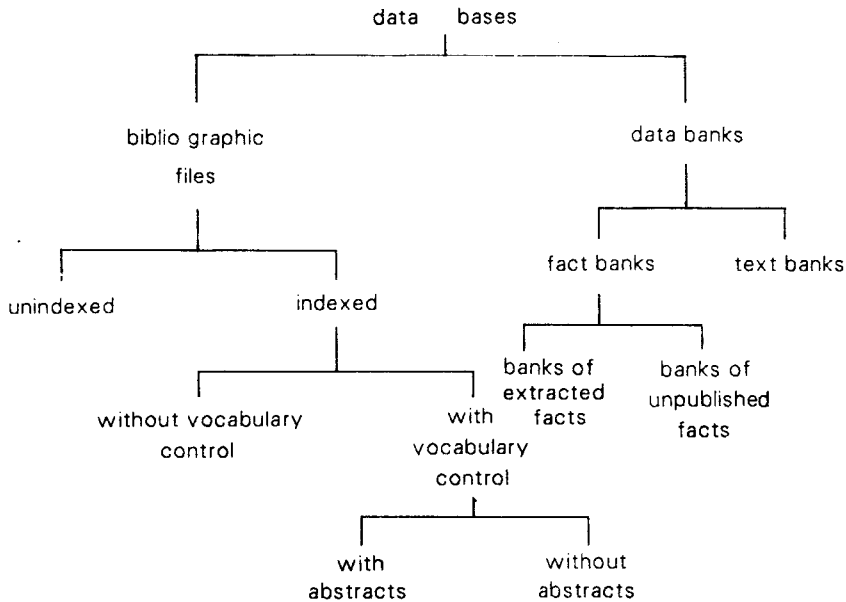


표 2. Cuadra의 분류체계

REFERENCE Databases

- Bibliographic 서지사항 및 초록
- Referral 그 외의 안내정보

SOURCE Databases

- Numeric 수치
- Textual-Numeric 수치와 문장의 조합
- Properties 물리적, 화학적 성질
- Full text 전문(문헌, 논문)

bank, source database 등으로 지칭되며, 필요정보 그 자체의 제공이 목적인 데이터베이스로서 축적정보의 형태는 문자정보, 수치정보, 화상정보, 음성정보 등 다양하다. 이 1차정보 데이터베이스는 전문을 대상으로 하는 full-text database, 수치정보를 대상으로 하는 numeric database, 문자와 수치가 동시에 존재하는 textual database, 문자·화상·음성정보가

합쳐진 multimedia database 화합물의 구조식이나 물성을 포함하고 있는 properties database 등으로 구분할 수 있다.

2) 2차정보 데이터베이스

2차정보 데이터베이스는 referral database, reference database, bibliographic database 등으로 지칭되며, 필요정보의 안내를 목적으로 하는 데이터베이스로서 축적정보의 형태는 문자정보가 주류를 이룬다. 이 2차정보 데이터베이스는 문헌의 서지사항이나 초록에 대한 문헌 안내정보 데이터베이스인 bibliographic database, 인명, 단체, 기관, 잡지소장목록 등의 디렉토리 데이터베이스인 referral database로 구분할 수 있다.

데이터베이스의 제공형태에 의한 유형은 온라인과 오프라인으로 구분되어 지며, 데이터베이스를 구성하는 데이터의 시간경과의 특성에 의해서는 주식정보와 같이 실시간으로 변하는 데이터로 즉각적으로 제공하는 즉답형 데이터베이스(Immediate database)와

표 3. 일본데이터통신협회의 분류체계

정보의 목적	데이터베이스의 종류	정보의 주요소	예
안 내	문헌 데이터베이스	문자정보	
	그 이외의 안내 데이터베이스	문자정보	소장목록, 기관안내
사실의 제시	사실 데이터베이스	문자정보	사람이나 사물에 관한 리스트, 디렉토리, 辭典, 事典 전문정보(법령, 의사족)
		수치정보	사회활동의 제시표 및 그의 통계
			실험, 관측으로 얻어진 물질의 구조식
화상정보	지도, 설계도		

문헌정보와 같이 데이터의 축적이 중심이 되는 기록형 데이터베이스(historical database)로 구분될 수 있다.

IV. 데이터베이스 시스템 및 모형

1. 데이터베이스 시스템

데이터베이스는 데이터, 하드웨어, 소프트웨어, 사용자로 이루어지며 데이터베이스 관리 시스템은 중심으로 구성된다. 데이터베이스 관리 시스템(DBMS; Database Management System)이란 여러 운영 프로그램들이 공유할 수 있도록 불필요한 중복을 피해 데이터를 저장하고, 그 데이터를 필요로 하는 응용프로그램이나 사용자들에게 보다 효율적으로 축적된 데이터를 제공하기 위한 시스템이다. 그림 2는 DBMS를 채택한 데이터베이스 시스템의 구조이다.

1) 3층 스키마

데이터베이스에서 관리하는 모든 데이터들은 일정한 규칙에 의해 구성되어지며 이 구성원칙은 데이터를 저장하고 관리하는 방법과 데이터에 접근하는 방법의 기본원칙이 된다. 이러한 전반적인 데이터베이스 기술을 스키마(schema)라 하며 관점에 따라 여러개의 서브 스키마로 나누어진다.

(1) 외부 스키마(external schema): 데이터베이스 사용자가 처리할 업무에 따라 필요로 하는 개념 스키마의 일부분을 의미한다. 즉 응용프로그램이 데이터를 바라보는 관점을 나타내며, 이것을 서브스키

마 또는 뷰(view)라고도 한다.

(2) 개념 스키마(conceptual schema): 기관, 조직체의 입장에서 본 데이터베이스의 전체구조를 말하며 이것은 각 응용 프로그램들이 필요로 하는 데이터를 중복없이 종합한 조직전체의 데이터베이스 구조이며 하나만 존재한다. 개념 스키마는 논리적인 데이터베이스 전체의 구조를 의미한다. 즉 데이터베이스 화일에 저장되어 있는 데이터의 형태를 나타내며, 레코드와 데이터 항목의 이름을 부여하고 그들 사이의 관계를 명시해 준다.

(3) 내부 스키마(internal schema): 물리적 저장 장치의 측면에서 본 전체 데이터베이스의 구조를 의미하며, 외부 스키마가 여러개 있을 수 있는데 비해 내부 스키마는 개념 스키마와 마찬가지로 하나만 존재한다. 물리적인 데이터 구조인 내부 스키마는 기억장치 내에 데이터가 저장되어 있는 데이터의 물리적 설계도이다. 즉 데이터베이스가 영속적인 기억장치 위에서 저장되고 접근되는 방법들을 설명하고 기술해 놓은 것이다.

2) DBMS

DBMS란 응용프로그램이 데이터에 대한 모든 액세스가 가능하도록 데이터베이스를 관리하는 소프트웨어이다. 그 기능은 데이터를 조직하고, 데이터를 조작하며 데이터를 제어하는 것이다. 즉 응용프로그램들이 데이터베이스를 공유할 수 있게 관리해 주는 소프트웨어 시스템이다. 사용자가 프로그램을 통해 데이터를 요구하게 되면 DBMS는 사용자의 요구를 수행하기

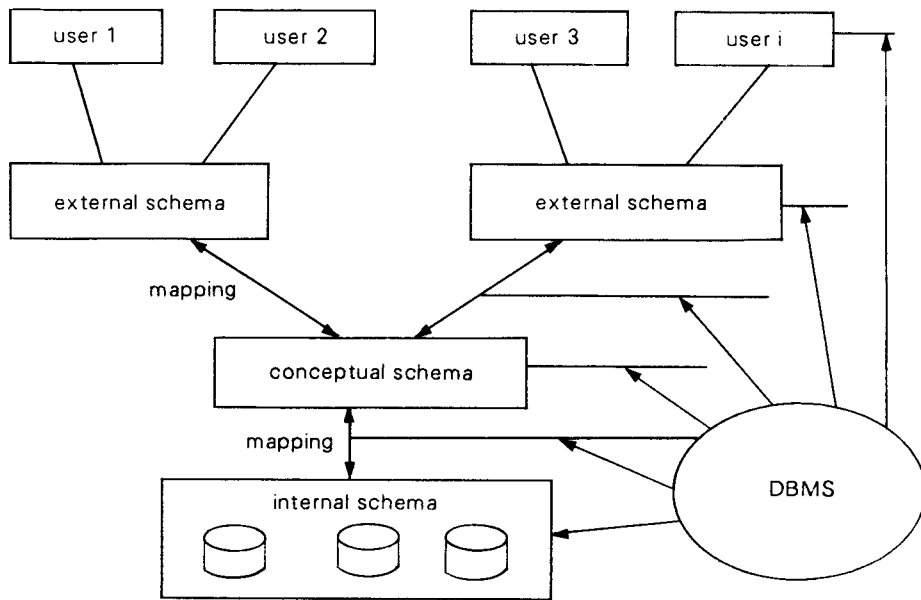


그림 2. DBMS의 논리적 구조.

위해 요구사항을 접수하여 번역하고, 데이터베이스의 구조에 따라 번역된 내용을 가지고 외부 스키마를 탐색한다. 이어서 개념 스키마와 내부 스키마의 대응관계를 찾아 기억장치내의 물리적인 데이터베이스를 탐색하고, 탐색된 결과는 프로그램의 요구대로 이용자에게 전달된다.

DBMS를 통해 얻을 수 있는 장점은 데이터의 중복을 최소화할 수 있고, 데이터를 공유할 수 있으며, 데이터의 일관성을 유지할 수 있고, 데이터의 무결성(integrity)을 유지할 수 있고, 데이터의 보안을 보장할 수 있으며, 데이터의 표준화를 실현할 수 있는 것 등이다.

3) 데이터베이스 언어

데이터베이스 언어란 데이터베이스를 구성, 조작, 검색하는데 필요한 언어들을 총칭하며 데이터 정의어, 데이터 조작어, 데이터 질의어가 여기에 포함된다.

데이터 정의어(data description/definition language)는 데이터베이스의 구조를 기술하는 언어로서 이 언어에 의해서 데이터베이스의 외부 스키마, 개념 스키마, 내부 스키마가 정의되어 진다.

데이터 조작어(data manipulation language)는 응용 프로그램이 데이터베이스에 원하는 내용을 표현하기 위해 제공하는 언어이다. 이 조작어는 데이터베이스 내용의 검색, 삽입, 수정등을 담당한다.

데이터 질의어는 데이터베이스 화일 구성과 범용 프로그래밍 언어를 정확히 알지 못하는 단말 사용자들이 이용하는 언어이다.

4) 데이터베이스 사용자

데이터베이스를 액세스하는 사람은 모두 사용자이며, 단말사용자, 응용 프로그래머, 데이터베이스 관리자등으로 구분할 수 있다.

단말사용자(end user)는 단말기에서 질의어를 사용하여 데이터베이스를 액세스하는 일반 사용자이다. 이들은 대개 데이터베이스에 대한 지식과 범용 프로그래밍 언어에 관한 전문지식 없이 데이터베이스를 이용하는 사람들이다.

응용 프로그래머(application programmer)는 데이터 조작어를 이용하여 데이터베이스를 응용하는 사용자로서 호스트 언어를 사용하여 화일을 처리하는 능력을 가지고 있는 사용자이다.

데이터베이스 관리자(DBA; database administrator)는 데이터 정의를 사용하여 데이터베이스를 DBMS에 표현하고 관리하기 위한 목적으로 데이터베이스에 액세스하는 사용자로서, 이 DBA는 데이터베이스에서 정보를 추출하는 것이 아니라 정보의 검색, 저장, 관리를 위해 지원해 주는 사람이다.

2. 데이터베이스의 모형

데이터베이스의 모형은 사용자 측면에서 본 데이터 구조의 종류이며, 관계모형, 계층모형, 네트워크모형 등이 있다.

관계모형(**relational model**)은 데이터 구조를 이용한 것이 아니고 수학적 이론에 기초를 두고 있는 모형으로서 세가지 모형 중에서 가장 추상적인 모형이다.

계층모형(**hierachical model**)은 데이터베이스를 구축하는 각 레코드가 계층구조 또는 트리구조 형식으로 구성된 모형이다. 즉 각 레코드형은 몇개의 자식(**child**) 레코드를 가지고 있으며, 하나의 부모(**Parent**)레코드를 갖는다. 이 계층모형은 데이터 항목 사이에 자연적인 계층관계를 가진 데이터 구조의 응용에 적합하다.

네트워크 모형(**network model**)은 계층모형의 확실적인 트리구조를 탈피한 모형으로서 서로 관련이 있는 세그먼트들이 그물처럼 얽혀서 전체구조가 하나의 네트워크(망)처럼 되어있다. 계층구조에서 자식레코드는 오직 하나의 부모레코드를 가질 수 있지만, 네트워크 구조에서는 자식레코드가 부모레코드를 갖지 않아도 좋으며, 여러개를 가질 수도 있다는 것이 두 모형의 중요한 차이점이다.

V. 데이터베이스 산업의 구조

데이터베이스 산업은 일반 제조업과 마찬가지로 데이터베이스 제작업과 유통업의 2가지 부분으로 크게 나눌 수 있다. 데이터베이스 제작자(DP; **database producer**, IP; **information provider**)란 자기가 보유하고 있는 정보원이나 정보 생산업자로부터 제공받은 자료, 데이터를 분석하고 처리해서 데이터베이스를 구축하는 사람들을 일컫는다.

데이터베이스 유통업자(Distributer, vendor)란

표 4. Growth in number of Vendors, Producers and Database.

year	Vendor	Producer	Database
1975	105	200	301
1979	263	316	528
1982	311	422	773
1985	614	1210	3010
1988	750	1733	4200
1989	770	1950	5578
1990	850	2224	6750
1991	933	2372	7637
1992	1438	3007	7907
1993	1629	2744	8261

데이터베이스 제작자가 구축한 데이터베이스를 통신회선에 실어서 데이터베이스 이용자에게 제공하는 서비스업자이다. **Gale Directory of Databases**⁶⁾에 의하면 1993년 말 전세계적으로 유통되고 있는 데이터베이스의 수는 8261개이며, Vendor수는 1629, IP 수는 2744이다.

1. 온라인 정보제공 기관

IP가 구축한 데이터베이스를 통신회선에 실어서 이용자들에게 제공하는 Vendor로서 DIALOG는 현재 유통되고 있는 Vendor중에서 세계 최대의 Vendor이다. DIALOG는 1966년 Lockheed 사가 미항공우주국(NASA)이 소유하고 있던 자료를 기초로 개발되었으며, 1972년 부터는 전세계의 이용자를 대상으로 온라인 상용 데이터베이스 서비스를 실시하고 있다. 또한 1988년 부터는 화상정보 서비스도 지원하고 있다.

ORBIT는 System Development Corporation 이 1965년 미공군과 미의학도서관(NLM)과의 유대관계 속에서 개발되었으며, 1972년부터 온라인 상용 데이터베이스 서비스를 실시하고 있다.

BRS(Bibliographic Retrieval Services)는 DIALOG, ORBIT와 더불어 세계 3대 Vendor중의 하나

6) K. Y. Marcaccio ed., Gale directory of Database, Gale Research, 1994, p.xx.

표 5. 데이터베이스 유형별 Vendor

데이터베이스 유형	Vendor
Abstracting and Indexing Database	<ul style="list-style-type: none"> • DIALOG • Information service • Infopro Technologies • NLM (US National Library of Medicine) • Questel • STN international • Data-star
Full-text Database	<ul style="list-style-type: none"> • DIALOG • Mead Data Central • West publishing company • NewsNet Inc • Data Times corporation
Numeric Business Database	<ul style="list-style-type: none"> • Reuters Information Services Ltd • DRI/McGraw-Hill
Numeric Scientific Databases	<ul style="list-style-type: none"> • CAN/SND • STN international • FIZ Karlsruhe • INKADAT
consumer-oriented Database	<ul style="list-style-type: none"> • compuServe Information service • America online • PRODIGY

표 6. 데이터베이스 벤더와 데이터베이스 수⁷⁾

벤더명	국·지역	데이터베이스 수
DIALOG	미국	173
BRS	미국	67
ORBIT	미국	63
ESA-IRS	이탈리아	38
QUESTEL	프랑스	38
INKA	독일	36
DIMDI	독일	32
DATA-STAR	영국	22
NLM	미국	17
Pergamon- Infoline	영국	17

7) 1983년의 통계이므로 데이터베이스 수에서 현재와 차이를 보일 것이다.

표 7. 정부관련 IP와 데이터베이스

정부관련 IP	데이터베이스
미국국립의학도서관(NLM)	<ul style="list-style-type: none"> • MEDLARS • MEDLINE • TOXLINE • Health planning and Administration • HISTLINE • CATLINE • AVLINE • SERLINE
미교육성	• ERIC
미상무성	• NETIS
미의회도서관	<ul style="list-style-type: none"> • LC MARC • RE MARC
일본과학기술정보센터	• JOIS
대영도서관	• LISA

표 8. 전문학회 및 조직체 IP와 데이터베이스

전문학회 및 조직체 IP	데이터베이스
미국화학회	<ul style="list-style-type: none"> • CA SEARCH • CASSI • CHEMICAL INDUSTRY NOTES
미금속학회	• METADEX
미국심리학회	• PsycInfo
Institution of electrical engineering	• INSPEC

나이이며, 1976년 부터 상용서비스를 실시하고 있다.

전세계적으로 유통되고 있는 1629개의 Vendor중에서 데이터베이스 유형별 대표적인 Vendor는 표 5와 같다. 또한 벤더별 데이터베이스 수는 표 6과 같다.

2. 데이터베이스 작성기관

데이터베이스 작성기관(IP, DP)이란 정보를 수집하

표 9. 민간단체 IP와 데이터베이스

민간단체 IP	데이터베이스
Bioscience Information service	• BIOSIS Review
Engineering Information Inc	• COMPENDEX
Predicasts Inc	• PTS Prompt
Institute for scientific inf	• SCISEARCHN
Dun & Bradstreet corporation	• D&B DUN's MARKET IDENTIFIERS • D&B DUN's • FINANCIAL RECORDS PLUS • D & B DONNELLEY DEMOGRAPHICS

고 분석해서 데이터베이스를 구축하는 조직체를 말하며 작성기관의 성격에 따라서 정부관련 조직, 전문학회, 민간단체등으로 구분되어 질 수 있다. 표 7, 표 8, 표 9는 2700여개의 IP중에서 IP 유형별 대표적인 IP와 데이터베이스 이다.

VI. 결 언

우리가 오늘날 직면하고 있는 사회는 정보의 생산, 저장 및 분배가 사회의 중요한 활동이 되고, 정보의 양적인 팽창이 심화되고 있으며, 정보의 가치가 점점 더 커지고 있는 정보사회이다. 이러한 정보사회의 정보 전달매체로서 출현한 데이터베이스는 사회와 산업

의 정보화에 따라 그 중요성이 점점 더 증가되고 있다.

도서관과 관련된 정보환경의 급격한 변화 발전은 단순히 자료를 안내하는 소극적이고 수동적인 사서의 역할을 하는 것이 아니라 정보사회에 적극적이고 능동적으로 대처해 전문성을 발휘할 수 있는 전문가로서의 사서의 역할을 요구하고 있다. 그러므로 컴퓨터 및 네트워크에 관련된 지식, 데이터베이스에 대한 전반적인 지식, 데이터베이스를 이용한 정보검색 및 검색결과 분석, 가공에 관련된 지식등이 정보환경에 적극적으로 대처하기 위한 전문가로서 사서들이 갖추어야 하는 지식 및 조건일 것이다.

참 고 문 헌

- 丘龍完, 컴퓨터과학개론, 대림, 1992.
徐泰雪, 崔明圭, 데이터베이스 産業의 動向과 政策, 産業技術情報院, 1992.
박준식, 김정현, 뉴미디어와 도서관, 계명대학교출판부, 1992
사공철 등저, 최신정보검색론, 구미무역, 1990
세계데이터베이스 91. 産業. 技術情報院. 1991
신양호, 석상기, 데이터베이스, 상조사, 1983
이경호 역, 정보시스템론, 대구대학교 출판부, 1986
黃漢雄, 金幸源, 編著, 데이터베이스 활용기법 매일경제신문사, 1994.
Marcaccio, K. Y. et al: *Gale Directory of Databases, Gale Research, 1994*
김선호, 온라인 사서에 대한 고찰, 도서관 Vol. 46, No. 3, 1991. pp24-55.
趙載浩 譯, 常用데이터베이스: 要點과 活用(1), 情報管理研究, Vo. 24, No. 2, 1993. pp36-53