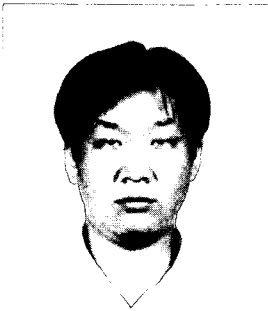


# DB제작을 위한 계획수립과 추진방법

— 문헌 DB를 중심으로 —



이 창 한  
(산업기술정보원  
DB개발실 책임연구원)

“본 글에서는 실제 데이터베이스의 구축단계 중 첫번째인 기획단계에 있어서 데이터베이스시스템을 구성하는 서브시스템들의 기본적인 설계 방향과 이들 시스템을 구축하는데 필요한 조직, 인력, 예산 등을 설정하는 개념설계 방법을 설명하고, 이에 기획안을 효과적으로 표현하기 위한 사업계획서 작성방법에 관하여 주로 인하우스DB에 초점을 맞추어 앞으로 3회에 걸쳐 소개하고자 한다.”

## 차 례

I. DB의 종류와 제작과정
II. 개념설계와 추진체계
1. DB시스템의 구성
2. DB제작 유통지원시스템의 설계방법
3. 추진조직 및 지원
4. DB구축 개념도
III. 사업추진계획서 작성
1. 목적 및 방향
2. 수록항목 및 세부내용
3. 사업계획 승인전략

## I. 데이터베이스의 종류와 제작과정

본장에서는 데이터베이스의 종류를 알아보고, 데이터베이스의 제작부터 활용까지의 전체적인 과정 및 데이터베이스 제작의 전반적인 공정에 대하여 설명하고자 한다.

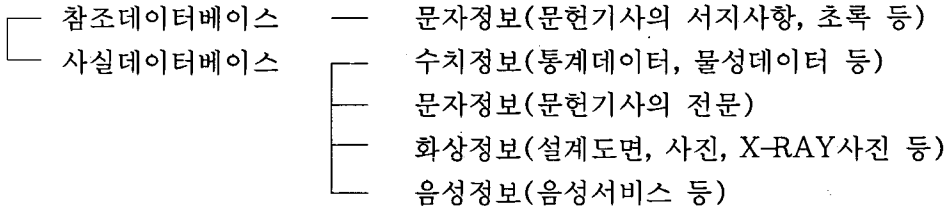
### 1. 데이터베이스의 종류

#### 1.1 형태별 분류

데이터베이스에 수록되어 있는 데이터의

<도표 1-1>

형태별 데이터베이스의 분류



형태에 따라 참조데이터베이스(2차정보 데이터베이스)와 사실데이터베이스(1차 정보 데이터베이스)로 구분할 수 있고, 또한 이들은 문자, 수치, 화상 및 음성과 같이 표현형태에 따라 세분된다.

참조 데이터베이스로는 문헌의 제목, 저자명, 기사수록 잡지명 등의 서지사항과 초록을 수록한 문헌데이터베이스, 기관을 안내하기 위하여 주소, 전화번호, 기관소개 등을 수록한 기관안내 데이터베이스 등이 대표적이고, 사실데이터베이스로는 통계데이터를 수록한 수치데이터베이스, 백과사전의 내용을 수록한 문자정보데이터베이스, 신문 및 사진을 광매체에 수록한 영상

데이터베이스 등이 있다.

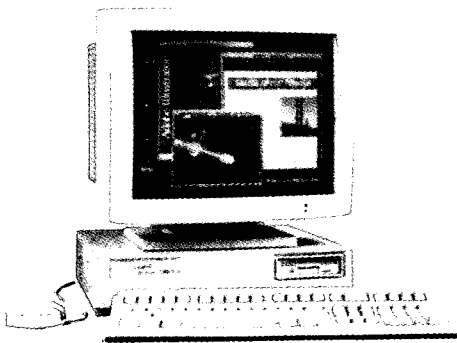
최근에는 데이터 축적매체와 전자출판 기술의 발달에 따라 기사의 전문을 수록한 전문데이터베이스(FULL TEXT DATABASE)가 증가하고, 문자, 수치, 화상, 음성 등이 통합된 멀티미디어가 출현됨에 따라 1차정보와 2차정보를 구분하지 않고 단순히 데이터의 표현형태에 따라 문자데이터베이스, 수치데이터베이스, 화상데이터베이스 및 멀티미디어데이터베이스로 구분하기도 한다.

1.2 주제별 분류

수록되는 데이터의 주제분야에 따라 일반, 비지네스, 자연과학·기술, 사회과학·인문과학 등으로 분류될 수 있다. 일반 데이터베이스에는 신문, 법률, 인물, 기관 및 생활정보 등이 있고, 비지네스 데이터베이스에는 기업, 금융, 상품, 시장, 무역정보 등이 있다. 자연과학·기술 데이터베이스에는 의학, 화학, 기계, 금속 등 전문분야의 정보와 산업재산권 정보 등이 있으며, 사회과학·인문과학 데이터베이스에는 인구통계, 교육 등의 정보가 포함된다.

1.3 이용대상별 분류

데이터베이스를 이용하는 대상에 따라서



인하우스(IN-HOUSE) 데이터베이스와 공공(Public) 데이터베이스로 구분될 수 있다. 인하우스 데이터베이스는 개인, 부서, 회사내 등 특정의 제한된 이용자를 대상으로 한 데이터베이스이고, 공공 데이터베이스는 천리안, KINITY-IR, DAIALOG 등 상용 데이터뱅크를 통하여 일반시장에 공개되는 데이터베이스를 말한다.

무형과 검색형으로 구분될 수 있다. 업무형 데이터베이스는 한 조직체에서 효율적인 업무수행을 위하여 주로 사용되며, 대개 다수의 고정필드로 구성되어 있다. 예를 들면 고객관리, 회계관리, 인사관리 데이터베이스 등이 있으며, DBMS(Database Management System)는 관계형 DBMS가 일반적으로 사용되고 있다.

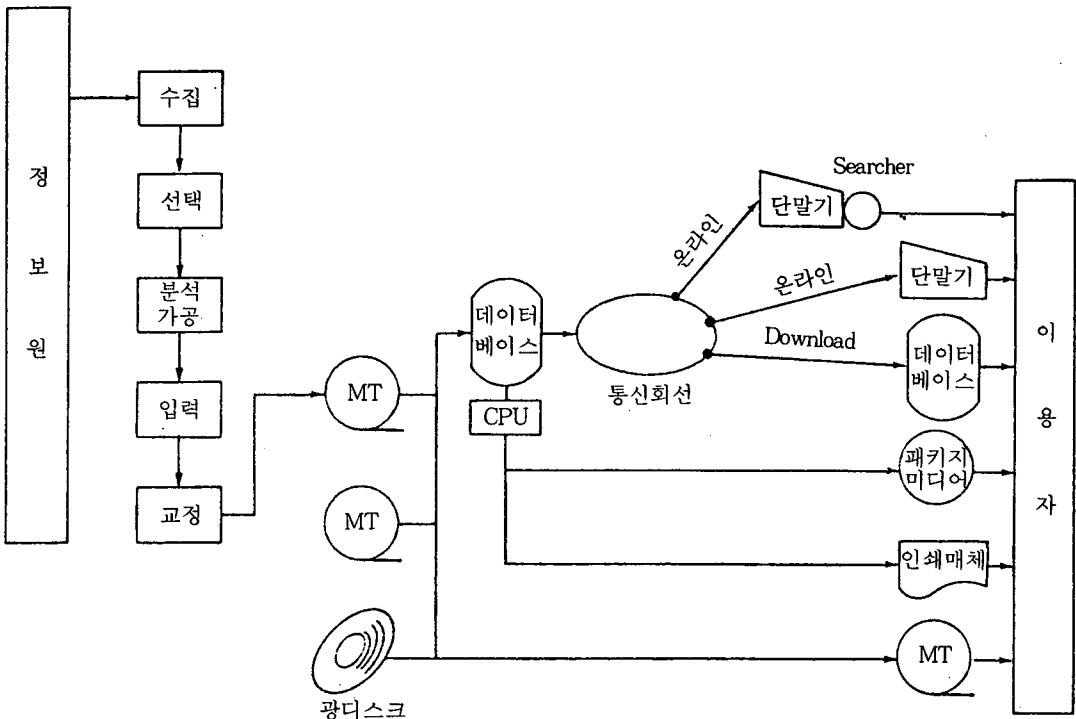
검색형 데이터베이스는 축적된 데이터중에서 검색을 주목적으로 하고 있으며, 다수의 가변길이 필드로 구성되어 있다. 일

#### 1.4 기능별 분류

데이터베이스의 역할과 기능에 따라 업

<도표 1-2> 데이터베이스 시스템의 전체도

정보생산	수집·가공·추적	제 공		이 용	
정보제공자 (Information Provider)	데이터베이스 제작자 (Database Producer)	데이터베이스제공자 (Database Distributer)	통신사업자 (Carrier)	검색대행업자 (Information Broker)	이용자 (User, End User)



반적으로 색인언어와 초록으로 정보내용이 표현되는 문헌정보가 중심이 되고 있으며, DBMS는 정보검색시스템(Information Retrieval System)이 주로 사용되고 있다.

## 2. 데이터베이스 유통의 전체과정

데이터베이스를 구축하는 데에는 잡지, 신문기사, 연구보고서 등의 문자정보로부터 수치, 도형, 화상 등 다양한 형태의 정보가 인쇄매체, 자기테이프, CD-ROM 등 정보매체로 수집된다. 이러한 정보를 생산하여 제공하는 저자나 출판사가 데이터베이스 구축측면에서 보면 정보제공자(IP, Information Provider)라고 부를 수 있다.

데이터베이스 생산자(DP, Database Producer)는 정보를 수집, 평가하고 데이터베이스에 수록할 정보를 일정기준에 따라 선택한다. 또한 정보의 내용을 분석하고 검색을 위한 색인언어 즉 키워드 혹은 분류를 부여하거나 입력 데이터를 추출하는 작업을 하며, 이를 퍼스널컴퓨터나 단말기를 이용하여 입력, 교정 등을 반복하여 마그네틱테이프나 광디스크 등의 기록매체에 수록한다. 상용 데이터베이스라면 산업기술 정보 데이터베이스를 제작하는 산업기술정보원, 국내 기업정보 데이터베이스를 제작하는 한국무역시장정보(주), 국내외 무역 관련 데이터베이스를 제작하는 무역협회 등을 들 수 있다. 국내의 경우 주요 데이터베이스 생산자는 자체 정보통신망을 통하여 데이터베이스 제공자의 역할도 겸하고 있는 것이 대부분이다.

데이터베이스 제공자(Database Distributor)는 생산된 데이터베이스 파일을 독자

적인 정보검색시스템에 탑재하고 통신회선을 통하여 이용자에게 제공한다. 국내의 주요데이터베이스 제공자로서는 HiTEL의 한국PC통신, 천리안의 데이콤, KINITI-IR의 산업기술정보원, KRISTAL의 연구개발정보센터 등이 있다.

통신회선으로는 공중전화망, HiNET-P나 DNS 등과 같은 공중 패킷교환망 및 VAN 등을 활용할 수 있다. 여기에는 통신회선을 제공하는 통신회선업자가 가입된다. 또한 해외에서는 별도의 데이터베이스 제공 시스템으로부터 서비스되고 있는 데이터베이스들을 이용자가 마치 동일 시스템에서 사용하는 것처럼 보이는 게이트웨이 업체들도 출현하고 있다. 예를 들면 EasyNet와 NIFTY-Serve 등이다.

이용자는 데이터베이스 제공자로부터 통신업자의 통신회선을 통하여 퍼스널컴퓨터와 단말기를 이용하여 온라인 검색으로 이용할 수가 있으며, 이외에 CD-ROM이나 인쇄매체 등을 통하여 데이터베이스를 이용할 수 있다. 또한 이용자를 대신하여 검색을 대행하는 검색대행업자(Information Broker)가 있다.

## 3. 데이터베이스의 제작 과정

### 3.1 기능적 측면에서의 결정사항

데이터베이스의 제작에 있어서 기본방향은 취급하는 데이터의 형태, 주제, 이용대상의 결정에 앞서 데이터베이스 제작의 목적을 설정하고 그 목적을 달성하기 위하여 어떠한 기능을 부여해야 하는가에 따라 결정된다. 즉 앞서 소개한 데이터베이스의 분류 방법중에서 기능별 분류에서 업무형

데이터베이스로 할 것인지 혹은 검색형 데이터베이스로 할 것인지를 선택하여야 한다.

기업의 정보시스템중 판매관리를 예로 들면, 과거의 업무형 데이터베이스에서는 판매장부의 집계 및 처리를 컴퓨터가 사람을 대신하여 업무의 효율을 높히는데 만족하였으나, 향후에는 판매예측 혹은 마케팅 전략 등에 관한 정보를 획득할 수 있는데 필요한 데이터를 여러가지 형태로 수집하여 분석하여야 하므로 적절한 정보를 찾을 수 있는 검색기능을 주체로 한 데이터베이스가 요구될 것이다. 즉 기존의 기업 정보시스템의 목적이 단순히 기업의 생산성 향상이나 경영자의 의사결정 지원에서 경쟁력 우위나 사업전략의 지원으로 변화하고 있다.

단순한 업무의 효율화를 위한 데이터베이스라면 대상이 되는 데이터에 표현된 문자나 수치를 입력, 처리하여 여러가지 형식으로 출력할 수 있으면 된다. 그러나 판매전략이나 시장예측을 수립하기 위해서는 이용자가 요구하는 목적에 맞추어 대상이 되는 데이터의 내용을 평가하거나 분석된 데이터를 입력하여 검색할 필요가 있다.

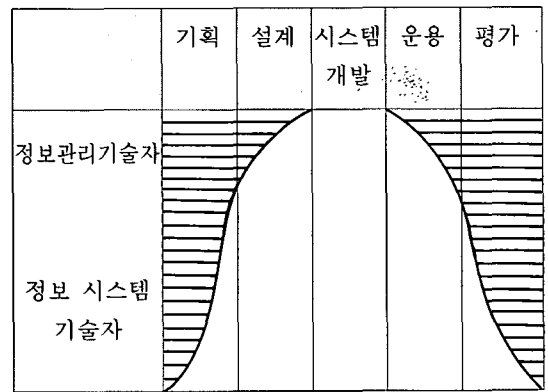
본고에서는 이용자의 목적과 요구에 맞추어 정보를 수집하고, 정보가 갖는 의미내용을 분석, 가공, 축적 및 검색할 수 있는 검색형 데이터베이스의 구축 방법에 대하여 살펴보고자 한다.

### 3.2 제작공정

데이터베이스의 제작공정은 일반적으로 기획, 설계, 개발, 운용, 평가로 구분할 수 있다. 또한 데이터베이스의 제작과 관련된

인력으로는 정보제공자, 이용자, 정보관리기술자, 정보시스템기술자, 경영자 등이 있으며, 데이터베이스 구축에 있어서 중심역할을 수행하는 것은 정보관리기술자와 정보시스템기술자이다. 특히 상기 다섯가지 공정의 업무중 정보관리기술자가 담당하는 부분은 <도표 1-3>과 같다.

<도표 1-3> 정보관리기술자의 업무영역



정보관리기술자가 담당하는 구축 공정으로서 기획, 설계, 운용, 평가에 대하여 설명하면 다음과 같다.

기획은 정보부문의 입장에서 뿐만아니라 조직전체의 경영방침, 중장기계획에 맞추어 데이터베이스의 위치를 설정하고, 데이터베이스 시스템을 조직 전체의 공통 시스템으로 구축하고자 하는 것이다. 데이터베이스의 기획에는 구축목적, 조직 내외의 환경, 대상 정보, 일정, 추진인력, 추진방법, 소요예산 등을 결정하여야 하며, 향후의 정보통신기술의 변화와 저작권 등의 법적 문제도 포함하도록 한다.

설계는 이용자의 정보요구에 근거하여 색인언어, 입력, 검색, 운영, 관리 등 다각



적인 관점에서 검토하고, 어떠한 데이터를 어떻게 표현하는가 즉 데이터 항목의 선정과 정의를 하는 것이다. 또한 데이터의 입출력방법의 설계, 검색시스템의 설계, 색인언어의 설계, 정보수집시스템의 설계 등이 포함된다. 실제 동 과정에서 데이터베이스시스템의 기능이 결정되므로 가장 중요한 공정이라 할 수 있다.

운용은 실제로 데이터를 수집, 가공, 입력하는 데이터 축적과정과 데이터베이스의 이용과정을 의미한다. 데이터의 축적에 있어서 변화하는 정보와 이용자의 요구에 대하여 데이터의 통일성과 일관성을 유지하는 것이 중요하다. 즉 데이터의 품질관리이다. 한편, 데이터베이스의 이용에는 이용자를 위한 교육, 이용 매뉴얼의 작성, 이용

촉진과 검색기술의 향상이 포함된다.

평가는 구축한 데이터베이스를 이용자의 입장에서 객관적으로 평가하고, 다음 과정에서 부족한 점을 갱신시켜 나가는 것을 의미한다. 검색시스템의 이용자 편의성, 데이터 내용의 구축 시기, 원문서비스 제공 시기 등이 포함된다.

정보관리기술자는 이용자의 목적과 정보요구를 분석 예측하고, 대상 정보를 선정하여 필요한 데이터 항목을 세분화하여 정의한다. 또한 검색 목적에 맞는 색인 시스템과 유저 인터페이스를 설계하며, 정보의 가공, 축적 및 제공에 관여한다. 한편 정보시스템기술자는 정의된 데이터 항목과 필요한 기능을 주어진 하드웨어와 소프트웨어를 활용하여 개발하는 것이다. [B]