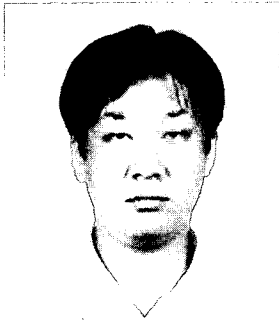


DB제작을 위한 계획수립과 추진방법 ②

— 문헌 DB를 중심으로 —



이 창 한
(산업기술정보원
DB개발실 책임연구원)

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

II. 개념설계와 추진체계

1. 데이터베이스 시스템의 구성

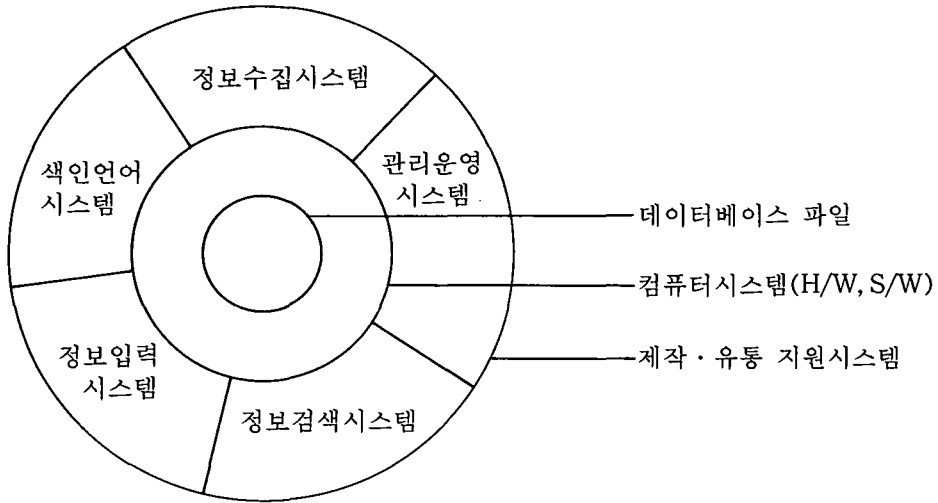
데이터베이스 시스템이란 컴퓨터를 활용하여 데이터를 축적하여 데이터베이스 파일을 생성하고, 생성된 데이터베이스를 이용자에게 전달할 때까지의 일련의 시스템을 가리킨다. 따라서 데이터베이스 시스템은 <도표 II-1>과 같이 데이터베이스 파일, 컴퓨터 시스템 및 제작·유통 지원시스템 등 크게 세가지 부분으로 구분할 수 있다.

데이터베이스 파일이란 이용자에게 전달하고자 하는 데이터를 전산매체에 축적된 형태를 말한다. 데이터베이스 파일은 여러 개의 레코드로 구성되는데 이 레코드가 정보전달의 기본 단위이다. 다시 이 레코드는 필드(항목)들로 이루어져 있으며, 이 필드들은 속성(Attribute)에 의해 그 특징이 구분된다. 좁은 의미의 데이터베이스 설계라면 레코드를 구성하기 위한 필드들을 선정하며, 선정된 필드들에게 속성을 정의하는 것이다.

컴퓨터 시스템은 하드웨어와 소프트웨어로 구성된다. 하드웨어는 호스트 컴퓨터, 워크스테이션, 통신설비, 단말기, 프린터, 보조 기억장치 등을 말하며, 소프트웨어는

<도표 II-1>

데이터베이스 시스템의 구성도



하드웨어를 효율적으로 운영하며, 제반 정보 처리작업을 원활하게 수행하기 위한 운영시스템, 프로그래밍 언어, 특정 목적의 소프트웨어 패키지 등을 말한다.

제작·유통지원 시스템이란 말 그대로 데이터베이스를 효율적으로 제작하고, 제작된 데이터베이스를 효율적으로 유통시키기 위하여 지원하는 시스템이다. 제작 시스템으로는 색인언어 시스템과 입력 시스템이 있으며, 유통시스템으로는 정보수집 시스템, 검색시스템, 관리운영시스템 등이 있다. 이 서브 시스템들은 데이터베이스 파일이나 컴퓨터 시스템과는 달리 비전산화 부분이 개입되는데, 즉 인력, 조직, 업무 체계 등과 컴퓨터 시스템이 혼용되어 구성된다.

최근에 정보통신산업의 발달에 따라 컴퓨터 시스템이 점차 기존의 비전산화 부분을 대체해 가고 있는 실정이다. 예를 들어 색인언어 시스템에 있어서 컴퓨터에 의한

자동색인이나 자동 초록 등에 대한 연구가 진행되고 있으며, 정보수집 시스템도 과거 카드 기록방식에서 바코드 시스템이나 MARC(Machine Readable Catalogue) 시스템 등으로 대체되고 있다. 또한 컴퓨터 시스템 자체도 데이터베이스의 중요성이 사회 전반적으로 인식됨에 따라 범용의 목적을 탈피하여 데이터베이스 전용의 컴퓨터 시스템으로서 데이터베이스 머신이 연구되고 있다.

2. 제작·유통 지원 시스템의 설계방법

(1) 정보수집시스템

(가) 정보 발생 및 유통경로의 파악

정보원을 파악하기 위해서는 우선 전체 조직도를 보고 주관 핵심부서를 파악하며, 각 부서의 정보관리측면에서의 역할을 명확히 분석하는 것이 중요하다. 그 후 각

부서에서 작성하는 자료의 종류와 분량에 대한 현황을 정보의 형태별·내용별·성격별로 조사할 필요가 있다. 연구소와 같은 조직에서는 연구결과를 보고서로 최종 정리될때까지 어떠한 내용과 형태의 중간보고서와 최종보고서가 연간 몇편 발간되고 있는지 혹은 향후 발간계획에 대하여 명확히 파악하며, 공장이나 개발부서는 대체로 기획조정실을 파악하는 것이 효율적이다.

또한 발생정보에 대하여 어떠한 유통경로에 의하여 필요로 하는 사람에게 배포될 수 있는지를 조사해야 하는데 대체적으로 자료담당자가 자신이 알고 있는 범위에서 필요로 하는 사람에게 배포하는 개인 베이스형태와 배포부서, 보관부서 리스트를 작성하여 활용하는 조직적인 형태로 분류할 수 있다. 이 외에 정보의 보관형태에 대해서도 개인파일, 사업부서단위 파일 및 사업본부단위 파일 등으로 분류하여 조사한다.

(나) 정보요구 파악

정보전담부서에서 데이터베이스로 제작하면 유용할 것으로 판단되는 정보의 종류와 사내정보이용자로서 요구하는 사항과는 반드시 일치하지 않는다. 또한 현실적인 문제로서 타인의 정보는 입수하고자 하면서 자신의 연구결과나 보유정보는 드러내지 않으려는 경우가 많기 때문에 조사결과에 전적으로 의존해서는 안되며, 요구분석시점에서 실질적으로 요구 정도가 높은 정보를 측정하여야 한다.

(다) 타사 자문 실시 및 방법

정보원과 정보요구에 대하여 효율적으로

조사하기 위해서는 이미 정보시스템을 구축 운영하고 있는 타회사에 직접 자문을 받는 것이 효과적이다. 그러나 자사의 업종, 규모, 형태 등에 따라 정보시스템의 차이가 크므로 타사의 실시예를 그 상태 그대로 이용할 수는 없지만 타사의 실시예를 통한 예비지식을 활용한다면 자사 데이터베이스 구축에 기초자료가 될 것이다.

타사 자문실시 방법은 우선 조사하고자 하는 항목을 질문용지에 작성하여 몇개 업체에 대해 실시하되 가능한 업종은 다양화하며 규모면에서는 유사한 업체를 목표로 한다. 또한 유의해야 할 사항은 대부분 형식적인 면은 이미 발표되어 지상으로 충분히 조사가 가능하므로 문서화하기 어려운 구축 과정상의 노하우, 아이디어, 계획과 실제의 차이점, 구축시 지연요소 등을 중심으로 조사하는 것이 바람직하다. 자문실시 조사 항목의 예는 <도표 II-2>와 같다.

(라) 사내 정보원 파악조사·실시

타사의 자문을 통하여 얻은 지식에 기초하여 자사의 정보원을 조사하는데, 조사단위는 어느정도 정보를 정리한 상태에서 결정하되, 조사도중에 부적합한 경우가 발생하면 세분화하거나 통합하여 실제 현상을 충분히 파악할 수 있는 단위로 교체한다. 조사항목으로는 다음과 같다.

- ① 보유정보의 종류와 양
- ② 파일링 방법
- ③ 파일링 관리부서
- ④ 정보의 보관형태 : 원자료, 복사, 마이크로필름, 기타
- ⑤ 분류형태 : 키워드, 분류표(사내/기준),

<도표 II-2>

타사 자문 조사항목의 예

조사항목		조사내용
축적 정보의 종류와 양	종 류 건 수	연구개발보고서, 생산기술보고서, 자사출원특허, 출장 보고서 등에 관해 구체적으로 기술 각 종류별로 수작업관리 건수와 전산 혹은 DB 관리 건수를 구분하여 기록
파일링 방법	관리 형태 보관장소 보관 형태	분산관리, 집중관리, 병용방식 각 부서, 정보자료실 책자, 마이크로필름, 광디스크 등
색인 방법	통 제 어 자 연 어	시소러스, 주제명 표목, 키워드집 자유키워드, 불용어 처리
정보 관리 시스템	네트워크 조직인력 정보등록 이용과정 결과물	전사적 네트워크도 작성 조직도에서 부서별 정보담당인력 기입 정보등록과정을 차트로 작성 인덱싱 수행자, 담당창구 파악 정보이용과정을 차트로 작성 결과물에 대한 샘플을 첨부
DB 구축상의 문제점	기 간 추진주체 기 타	계획부터 구축까지 소요 기간 시간지연요소 파악 계획안 작성까지 추진주체가 프로젝트팀, 컨설턴트 혹은 병용방식인지 파악 계획과 시스템 가동후 상이점 기타 실정에 합치되지 않는 문제

시계열, 공정별, 기타

- ⑥ 발생정보 공개 유무 : 부서 혹은 전사적으로 월보(계보, 주보, 뉴스레터등) 배포

⑦ 입수 희망정보

- ⑧ 현상의 문제점 : 인원, 공간, 현재관리 체제
이상의 결과를 타사의 사례와 함께 수록

하여 일람가능한 표로 작성한다.

(마) 대상 정보의 결정

어떤 정보가 데이터베이스의 대상이 되는가를 결정하는데 있어서, 정보원 조사에서 수합된 데이터베이스 희망 정보 전체를 데이터베이스화할 수 없으므로 정보원을 계층화할 필요가 있는데 다음과 같은 기준을 고려할 수 있다.

- ① 전사적으로 데이터베이스화를 목적으로 하는 정보
- ② 사업본부 단위의 관리로 이용하는 정보
- ③ 각 부서에서 보유하는 것이 적절한 정보

어떠한 정보부터 데이터베이스화할 것인지 그 우선순위를 검토해야 하는데 다음과 같은 착안점을 참고할 수 있다.

- ① 요구도가 높은 정보
- ② 수집하기 용이한 정보
- ③ 정보관리가 현재 이루어지고 있는 부서부터 출발하여, 그것을 모델케이스로 전체부서에 확대

(바) 자료 등록 및 수집의 체계화

사내에서 발생되거나 외부로부터 수집한 정보를 원활하게 유통시키기 위해서는 우선 사내에 자료등록 체제를 구성하는 것이 가장 중요하다. 이때 대상자료별로 자료의 등록방법, 등록기준, 등록 및 보관 부서에 대하여 기관차원에서 명확히 해야 하며, 정보입력시스템 및 정보관리운영시스템과 긴밀한 연계성을 유지시켜야 한다.

또한 최근에는 국내에도 상용데이터뱅크가 계속 증가되는 추세에 있으며, CD-ROM 등 전산매체가 다양한 분야에 걸쳐

제작되고 있으므로 기관과 관련된 전산자료를 정보수집차원에서 검토할 필요가 있다. 그리고 성공적인 자료 등록 체제를 구축하기 위해서는 각 현업 부서의 보고서 제출의무나 인센티브 제도에 대하여 규정화하도록 한다.

(2) 입력시스템

(가) 개념 항목의 선정

정보수집시스템의 개념설계 과정에서 정보원의 유통경로와 정보 요구에 대한 분석을 통하여 대상 정보에 대한 윤곽이 드러나면, 이를 위하여 어떠한 데이터를 입력하여야 하는지 결정해야 한다. 실제 입력 항목의 선정과 정의는 데이터항목 설계에서 수행하게 되는데 이에 앞서 정보관리기술자와 시스템기술자간 혹은 조직내 관련 부서간의 논의 자료로서 활용할 목적으로 개략적이 항목도출이 필요한데 이것이 개념항목이다.

개념항목을 선정할 때에는 검색과 표시 요구에 주안점을 두어 정리하며, 실제 데이터베이스에서는 복수로 분리될 항목이라도 하나의 개념으로 볼 수 있으면 1개 항목으로 간주한다. 또한 이 개념 항목은 실현 가능성이 아직 명확하지 않으며 다만 이용자의 요구를 충족시킨다는 관점에서 작성하면 된다. 만약 데이터베이스의 정보원으로서 기존 시스템의 전산 부산물을 활용할 경우나 사외로부터 기계가독형의 과일을 구입할 경우라도 개념항목표 작성은 필요한데, 그 이유는 자사에서 불필요한 항목을 삭제하거나 필요한 정보를 추가시키는 경우가 많기 때문이다.

<도표 II -1>

기술보고서의 개념항목표 사례

개념항목	검 색	표 시	반 복	비 고
자료번호		○		작성년도+부문코드+일련번호
작성일자	○	○		
작성자	○	○	○	
표제	○	○		
초록	○	○		코드표에 의한 통제
분류	○	○	○	
키워드	○	○	○	

(나) 데이터 입력방식

등록정보를 DB화 하기 위하여 우선 등록시트의 내용을 전산 입력하여야 하는데 크게 온라인방식과 배치방식이 있다. 온라인 입력방식은 단말기를 통하여 호스트컴퓨터에 입력하는 것으로서 주로 리얼타임 데이터베이스나 갱신주기를 단축시키기 위한 방법이다. 배치방식은 자기테이프나 플로피디스크 등 자기기억매체에 일괄적으로 입력·처리하는 것이다. 입력 대상자료의 성격에 따라 병용방식도 검토할 필요가 있다.

데이터입력은 데이터베이스 구축이 정상화되었을 때, 예산과 비용이 가장 많이 소요되는 과정중의 하나이므로 입력항목, 등록건수로부터 입력량과 입력비용을 추정하며, 입력시간 등을 고려하여 입력방법을 결정한다. 이와 같이 입력시간 및 비용면에서 보았을 때, 사내에서 입력처리하는 것이 여의치 않을 경우가 발생할 수 있으므로 데이터입력전문센터에 입력 외주화를 검토할 필요가 있는데 이때 온라인입력의 경우는 패스워드 및 데이터 보안의 문제

등이 발생할 수 있으므로 배치방식으로의 주입력을 수행하는 것이 바람직하다고 생각된다.

최근에는 국내에도 국·한·영문 광학문자판독기(OCR, Optical Character Reader)가 개발되어 거의 실용화 단계에 있으므로 비교적 정형화 되어 있는 자료에 있어서 그 입력량이 많을 경우에는 광학문자판독기의 도입을 적극적으로 검토할 필요가 있다.

(3) 정보검색시스템

(가) 데이터베이스관리 시스템의 일반적인 특징 검토

자사내 범용컴퓨터를 활용하여 데이터베이스를 구축할 때, 소프트웨어 즉 데이터베이스관리시스템(DBMS, Database Management System)의 선정은 시스템 구축에 있어서 매우 중요한 사항이다. DB화하고자 하는 정보의 특성을 충분히 관리할 수 있으며, 자사의 목적에 부합한 기능을 보유하고 있는 DBMS의 선정을 위해서 충분한 검토가 필요하다. 또한 선정 후보 소프트

웨어에 대하여 사내 호스트컴퓨터에 탑재 가능 여부를 파악해야 하며, 만약 DBMS 성능의 우수성을 우선한다면 자사컴퓨터에 탑재 불가능하더라도 외부기관의 컴퓨터 활용 방안도 검토할 수 있다.

또한 DBMS를 선정하는데 있어서 고려할 사항은 일반적으로 별도의 구분을 하지 않고 DBMS를 호칭하는데, 사내 데이터베이스를 구축하는데 있어서 용도에 따른 개념 구분을 명확히 해야 한다. DBMS를 용도에 따라 구분하면 주로 문헌데이터베이스용으로 개발된 검색형 DBMS - 정보검색시스템 (IRS, Information Retrieval System이라고도 함) - 와 수치나 사실 데이터 등을 주로 관리하는 업무형 DBMS가 있는데 최근 양자간에 동시에 활용할 수 있는 소프트웨어가 등장하고 있으며, 가까운 장래에 통합될 전망도 있으나 현 상태에서 그 특징을 구분하면 <도표 II-4>와 같다. 기업에서 문헌정보를 위주로 데이터베이스화하여 검색하고자 할 경우는 검색형 DBMS를 선택하며, 사내정보관리 및 여타분야에도 활용하고자 할 경우에는 업무

형 DBMS를 선택하는 것이 바람직하다.

(나) 데이터베이스관리 시스템의 조사

데이터베이스관리 시스템의 구축 방법은 크게 직접 개발하는 방법과 시장에 유통되고 있는 패키지를 도입하여 필요 기능을 추가하는 방법으로 대별된다. 직접 개발하는 방법은 특수한 상황이 아니라면 막대한 예산과 장기간의 소요기간이 필요하므로 가급적 도입하는 방법이 경제적이다 판단된다.

실제로 소프트웨어 산업의 발달로 국내에서 유통되고 있는 상용 데이터베이스관리시스템의 종류가 다양화되고 있으므로 이들중 목적에 가장 부합된 것을 선택하기 위해서는 현재 국내에 유통되는 데이터베이스관리 시스템의 종류에 대하여 조사할 필요가 있다. 국내 유통 데이터베이스관리 시스템은 <도표 II-5>를 참고하기 바란다.

(다) 데이터베이스관리 시스템의 평가

DBMS를 평가할 때 여러 관점에서 평가해야 하는데 평가항목으로는 일반공통사항, 등록·갱신, 검색, 출력 등이며 각 항목별 유의점은 다음과 같다.

① 공통사항

컴퓨터 기종의 제약, 유저프로그램의 이용 가능성, 시소러스 처리기능, 한글처리 등이 중요한 검토사항이다. 그 중에서도 한글처리는 문헌 데이터베이스 제작시 특히 관심을 기울여야 할 부분으로 외국 DBMS 제작사의 제품을 사용할 경우에는 한글처리 능력이 완벽한지 더욱 주의해야 할 필요가 있다.

<도표 II-4> 검색형 DBMS와 업무형 DBMS의 비교

구 분	검색형 DBMS	업무형 DBMS
적 용 분 야	참조형 정보 서비스	업무데이터 일원관리
데이터구조	단 순	복 잡
수록항목수	소량(수십-백)	다량(수천-만)
항 목 값	개념적·지식적	수치적·사실적
갱 신 방 법	일괄추가	동시 갱신

<도표 II-5>

국내 주요 데이터베이스관리 시스템의 종류

구분	명칭	판매·연락처	H/W	O S
검색형	CDS-ISIS	산업기술정보원	IBM PC 호환기	DOS
	KIROS	산업기술정보원	IBM 대형 컴퓨터	MVS
	COHIRES-II	쌍용컴퓨터	IBM 대형, VAX 등	MVS, VMS
	BRS-Search	선경정보시스템	IBM PC·대형, UNIX 등	MVS, DOS, UNIX
	FAIRS	한국후지쯔	FACOM	FACOM OS
업무형	DIMS	삼성데이터시스템	IBM 대형	MVS
	MANTIS	뉴컴퓨터서비스	IBM 대형, VAX	MVS, VMS
	ADABAS	펜타컴퓨터코리아	IBM, DEC, UNIX 등	DOS, MVS, UNIX
	SYBASE	삼보컴퓨터	EBM PC 386 SUN W/S	UNIX
	INFORMIX	다우기술	IBM, UNIX 등	DOS, UNIX
	BASIS+	휴먼웨어	IBM 대형, VAX, CYBER	MVS, UNIX 등
4 GL	FOCUS	지구어시스트	IBM, UNIX 등	DOS, MVS, UNIX 등
	NATURAL	펜타컴퓨터코리아	IBM, DEC, UNIX 등	DOS, MVS, UNIX 등

* 4GL은 제 4세대언어(4 Generation Language)의 약자로서 엄밀히 따지면 DBMS는 아니지만 범용 DBMS로서 활용할 수 있음.

② 등록·갱신

문헌 데이터베이스를 제작하는데 있어서 DBMS 기능중 자칫하면 검색기능만이 강조되기 쉬운데 실제 데이터베이스의 등록 갱신 기능도 매우 중요한 평가항목으로서 특히 데이터베이스 갱신이나 키워드 자동 추출 기능에 대하여 검토한다.

③ 검색

검색개시부터 완료까지 용어절단(Truncation), 마스터파일 검색, 스트링서치(String Search) 등과 같은 기능적인 측면 외에 조작성, 대화 진행방법 등 운영적 측면의 평가도 필요하다.

또한 검색 질문식 누적 표시 기능, 검색

결과의 보관기능 보유여부를 검토해야 하며 검색결과와 소팅, 그래프 표시, 수치 데이터 해석 등 단순 검색 이외에 분석 기능도 포함될 수 있다. 실제로 DIALOG의 PTS(Predicasts Terminal System) 파일에서는 산업데이터, 경제데이터의 통계처리도 가능하며 전용명령어가 준비되어 있다.

④ 출력

검색결과와 소팅(Sorting), 머징(Merging) 혹은 출력형식의 자유도를 평가해야 한다.

⑤ 기타

기타 기능으로서 SDI(Selective Dissemination of Information), KWIC(KeyWord In

<도표 II-6> DBMS 평가분석 사례

평가항목		KIROS	CDS/ISIS
공 통	DB 형태 H/W 유저프로그램 단말기·네트워크 호스트 언어 링크 언어 한글 처리 유사 시스템	검색형 DBMS IBM 30XX, 43XX EXIT 루틴 활용 IBM PC 호환기 ASSEMBLER, PL/I ASSEMBLER, PL/I 가능 STAIRS, BRS	검색형 DBMS IBM PC ISISPAS 루틴 활용 LAN 가능 PASCAL PASCAL 가능
갱 신	데이터 입력 데이터 갱신 키워드 추출	온라인, 배치 온라인, 배치 가능	온라인, 배치 온라인, 배치 가능
검 색	마스터 검색 링크 검색 폴텍스트 검색 대소관계 트렁케이션 논리연산 HISTORY 동의어 보존기능 통계처리	가능 가능 가능 포매티드 필드 전방일치 AND OR NOT XOR 가능 불가 질문식·검색결과 불가	가능 불가 가능 SORT WORKSHEET 전방일치 AND OR NOT 가능 ANY TERMS 활용 검색결과 불가
출 력	출력형식 SDI	자유도가 적음 가능	자유도가 많음 가능
기 타	KWIC 사용편의성 확장성	불가 정형적 업무에 편리, 운 영메뉴다양 관련 운영 틀이 다양, STAIRS와 연계 확장성 좋음	불가 메뉴방식 편의성 탁월 중형, 대형 컴퓨터용 한 글화 연구 진행중

Context) 인덱스 작성, 장애대책, 복구방법 등의 부대기능과 DBMS를 개발·유통시키는 제작 기관의 장래 개발 방침도 중요한 평가 항목이 된다.

이러한 평가 항목에 기초하여 검색형 DBMS인 CDS/ISIS와 KIROS의 평가한 사례는 <도표 II-6>과 같다. 이러한 평가 분석은 초기에는 자료나 매뉴얼을 통하여 실시하며 최종적으로 실제 후보 DBMS를 컴퓨터에 탑재한 후 각 항목에 대하여 벤치마크 테스트(Benchmark Test)를 수행하여야 한다. 또한 최근 전반적인 소프트웨어의 라이프사이클이 매우 짧기 때문에 DBMS 개발 및 유통 동향을 지속적으로 조사하여 내부 시스템의 적용 가능성을 파악하므로써 새로운 환경에 적응해 나가야 한다.

(4) 색인언어 시스템

(가) 색인언어의 기본방향 설정

데이터베이스에서는 정보내용의 핵심을 색인언어로 표현하므로 색인 방법의 결정은 시스템 구축상 매우 중요한 부분이다. 색인언어의 종류로는 통제어 방식과 자연어 방식이 있는데 각각 장단점이 있으므로 자사의 업종·분야·규모에 따라 결정해야 한다. <도표 II-6>에서 보는 바와 같이 통제어 방식은 분류, 시소러스 및 용어집 등을 이용할 수 있으며, 자연어 방식에는 자유 키워드 부여와 자동 키워드 추출 방법이 있다.

이들 여러가지 색인언어중에서 채택할 수 있는 기준으로서 예를 들면 자사의 사업분야가 명확하게 한정되어 있을 경우에

는 통제어 방식의 시소러스를 활용할 수 있으며, 시시각각으로 사업 대상이 변화할 경우에는 자연어 방식인 자유키워드로 색인하는 것이 효율적이다. 또는 예산과 인력이 충분하다면 통제어와 자연어를 병용하는 것도 고려할 수 있으며, 선진국의 주요 데이터베이스는 통제어와 자연어의 장점을 모두 수용하여 검색효율을 극대화하기 위하여 대부분 병용 방식을 채택하고 있다.

(나) 기존 통제어 색인방법의 활용방안 검토

시소러스 방식을 채택할 경우에는 기존 시소러스를 그대로 이용하는 방법, 기존 시소러스의 해당분야를 선택하여 사내용으로 재편성하는 방법과 자사의 고유 시소러스를 작성하는 방법 등을 고려할 수 있다.

현재 국내에서 발간된 시소러스로서는 1992년 시스템공학연구소가 제작한 40,000여 용어 규모의 과학기술용어 시소러스와 1993년 한국언론연구원에서 제작한 10,000여 용어 규모의 신문기사 종합시소러스가 있으며, 통제 용어집으로서는 산업기술정보원과 시스템공학연구소가 공동으로 제작한 40,000여 용어 규모의 과학기술 한·일·영 용어 번역집이 있다.

자사 고유의 시소러스제작은 극히 소규모가 아니라면 많은 인력과 예산이 필요하므로 사전에 충분한 검토가 필요하다. 이외에 통제어 방식으로는 전술한 과학기술 용어번역집 혹은 핸드북 등에서 필요한 용어를 추출하여 사내의 색인작업용 키워드 집으로 제작, 색인하는 키워드집 활용방식이 있다.

<도표 II-7>

통제어와 자연어의 비교표

구 분		통 제 어			자 연 어	
		분류 코드	시소러스	용 어 집	자유키워드	자동 추출어
통제어와 자연어의	장 점	검색효율(재현율, 적합율) 제고			색인부하가 적음 전문용어·신조어 색인가능	
	단 점	검색시 참조용 자료 필요, 색인 부하가 높 음, 개념 변화 추가기능 저조			검색 효율 저조	
검색어수		수십-백	수천-만	수백-만	무 한	무 한
색인 부하		소	대	중	소	무
기억 용량		소	중			대
적용 대상		큰 개념을 어느 정도 한정할 때	세부 개념, 구조적인 검색이 필 요할 때	병렬 개념, 구조적인 검색 불필 요할 때	보조 키워 드로서 유 효	긴 문장항 목이 포함 되어 있을 때

(다) 자연어 방식의 검토

자유키워드 방법은 기사의 사용 용어를 그 상태대로 사용하거나 색인자 임의로 키워드를 부여하는 것인데 이때에는 여러명의 색인자가 가능한 일정한 형태의 색인어를 추출할 수 있는 색인 매뉴얼의 제작이 필수적이다. 자동 추출은 컴퓨터에 의한 자동색인을 의미하는데 최근 자연어처리 소프트웨어를 활용하면 더욱 효율적이라고 생각되나 한글 데이터에 대해서는 아직 완벽하게 적용할 수 있는 단계는 아니라고

판단된다.

(5) 정보관리운영시스템

(가) 데이터 구축·이용 실적 통계

데이터베이스는 한번 구축하였다고 당해에 종료될 수 있는 성질의 사업이 아니며, 지속적으로 추가, 수정 등의 갱신이 수반되어야 가치가 있으므로 일정 기간마다 분야별·작성자별 등의 기준에 의한 구축 실적을 파악 분석할 수 있는 체제를 갖추어야 한다. 또한 이용자가 없는 데이터베이

스 역시 의미가 없으므로 이용자의 실적을 여러 각도에서 파악할 수 있는 틀을 데이터베이스 시스템에 포함시켜 이용 실적의 평가 분석을 위한 기초자료로서 활용할 수 있도록 하여야 한다.

(나) 정보 제공 체제

구축된 데이터베이스를 효율적으로 활용시키기 위해서는 데이터베이스 검색·원자료 제공에 대하여 그 절차와 방법을 결정하여 흐름도를 작성하여 체계화시켜야 하며, 검색 출력은 온라인·배치·병용방식 중에서, 원자료 제공은 열람·복사·대출 중에서 어떤 서비스 형태를 취할 것인지 고려해야 한다. 또한 검색·제공 대상에 대한 사용자그룹을 레벨화하여 사내정보에 대한 보안도 고려해야 하는데, 제공범위의 레벨에 대한 예는 <도표 II-8>과 같다.

(다) 데이터베이스 파일 관리 및 보안

데이터베이스 운영으로서 마스터파일 관리, 데이터베이스 보안대책 및 데이터 유지·보수에 대하여 검토해야 한다. 마스터파일 관리는 데이터보호 및 보안 등을 시스템 엔지니어와 함께 충분히 주의를 기울

여야 한다. 이는 사내컴퓨터를 활용할 경우와 타사 컴퓨터를 활용할 경우가 상당한 차이가 있으므로 별도의 대책이 필요하다. 데이터베이스 보안대책은 컴퓨터시스템 보안대책과 정보관리 보안대책으로 분류할 수 있는데, 컴퓨터시스템 보안은 시스템엔지니어가 시스템 구상시 대책을 강구하여야 하며, 정보관리 보안은 DBMS의 한 기능으로 보안기능이 준비되어 있다 하더라도 제작과정이나 정보제공후에 보안이 이루어지지 않으면 소용이 없으므로 패스워드관리나 정보제공처에 대한 모니터링이 필요하다.

패스워드관리에 있어서 상위 패스워드와 ID번호는 기록하여 두지 않으며, 정기적 혹은 비정기적으로 패스워드나 ID번호를 교체시킨다. 정보제공처에 대한 모니터링의 필요성은 가령 정보가 전담부서에 있을 때에는 보안대책에 주의를 기울인다 하더라도 정보제공처에서 정보취급을 느슨하게 하면 문제가 발생할 수 있기 때문이다. 따라서 정보배포선에 관한 리스트관리나 보고서 복사처에 보안지도를 하는 등 정보유통경로에 주의를 기울여야 한다.

데이터베이스 유지·관리는 두가지 측면

<도표 II-8> 사내정보제공 보안레벨의 예

보안 레벨	내 용
최고 기밀	최상급의 패스워드를 갖는 부류만이 접근가능
사내 한정	정보제공부서의 결정에 의해 사내 일부 계층만이 접근가능
사내 전체	사원전체가 자유롭게 정보입수 가능
관련 회사	정보내용에 따라 관련 회사에 제공가능
외부 발표	학회나 협회에 발표하여 이미 공개

에서 고려되어야 하는데, 데이터 추가·수정·삭제·백업 등 데이터 자체에 대한 유지·관리와 DBMS 버전업, 시스템 버전업 등 데이터베이스 주위 환경에 대한 유지·관리이다.

3. 추진 조직 및 자원

(1) 조직

(가) 개념설계 단계의 조직

데이터베이스 개념설계 단계에서는 각 사업부나 사업단위의 의견을 대표할 수 있는 인력의 협조가 필요하게 되는데 어떠한 조직체에서도 부서마다 정도의 차이는 있겠지만 나름대로의 규모와 방법으로 정보를 수집·관리할 것이므로 이러한 부서로부터 인력을 선발하여 프로젝트팀을 구성하는 것이 효율적이다.

프로젝트팀의 구체적인 선발대상 기준은 사업부의 핵심단위, 연구소와 공장 등 업무성격이 판이한 단위, 물리적으로 거리가

떨어진 단위 등으로 생각할 수 있으며 직위면에서는 실제 정보의 현상을 파악할 수 있는 중간 레벨의 인력이 적절하다고 판단된다. 또한 프로젝트팀의 조직내의 위치를 명확하게 할 필요가 있는데 전체 조직을 통제·조정하는 본사의 기획부와 같은 부서가 주도적으로 수행하는 것이 바람직하며 향후 데이터베이스 구축 활용단계의 활동 중심부서와는 일치하지 않을 수도 있다.

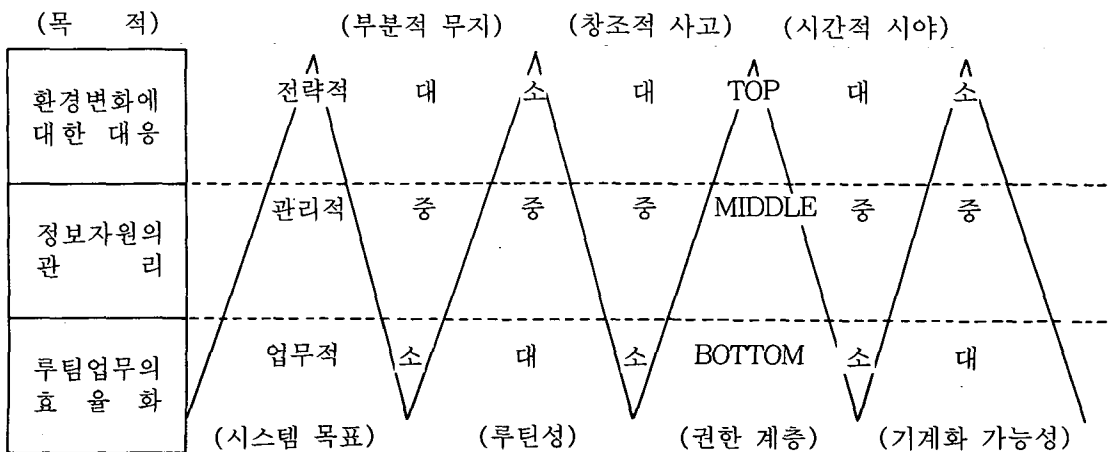
(나) 구축·활용단계의 조직

데이터베이스의 구축·활용의 필요성이 인지되어 동 활동을 정상적으로 수행하기 위해서는 정보 관리 조직이 필요하다. 정보 관리 조직의 기능이라면 전술하였던 데이터베이스 시스템의 구성 요소중 제작·유통지원 시스템의 기능인 수집, 입력, 정보가공(색인언어), 정보검색, 운영관리와 거의 일치한다고 볼 수 있다.

그러나 이러한 정보 관리의 기능을 조직으로서 실현하는데에는 경영층의 정보활동

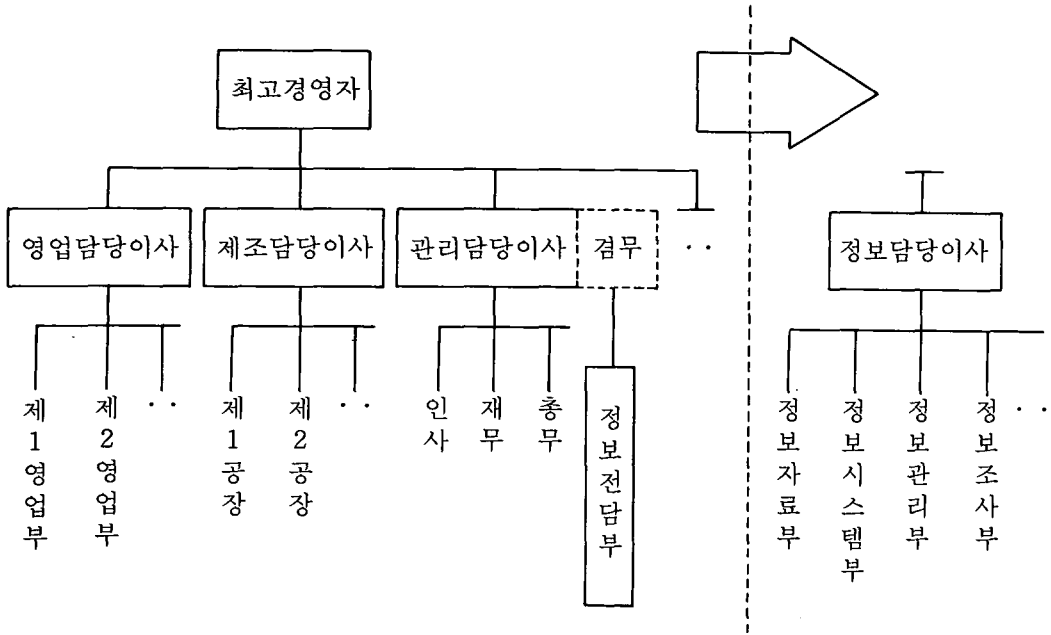
<도표 II-9>

정보 관리 조직의 목적과 특성



<도표 II-10>

정보 조직의 변화



에 대한 관심의 초점에 따라 여러형태로 나타나게 된다. 관심의 초점에 따른 정보전담조직의 목적은 루틴한 업무의 효율화, 정보자원의 관리, 환경변화에 대한 대응 등으로 구분되며 그 특성은 <도표 II-9>와 같다.

이러한 정보 관리 목적이 과거에는 '루틴 업무의 효율화나 정보자원의 관리중에서도 자료 수집 기능만이 강조되어 전산 시스템 부서 혹은 정보자료실 및 도서관 등에서 수행하는 경우가 많았으며 그 조직도 자연히 기획관리부문에 속하게 되었다. 그러나 현대 기업의 정보 관리 목적이 점차 환경변화에 대한 대응으로 변화해 감에 따라 <도표 II-10>와 같이 과거 3대 경영자원인 돈·물건·사람 중심의 조직체제로 부터 정보를 전담으로 하는 조직이 기존의

조직체제와 동등한 레벨로서 추가되고 있다.

따라서 조직이란 경영층의 고유 권한이므로 경영층의 정보 관리의 마인드 형성의 정도에 따라 전산 시스템 부문, 정보 자료 부문, 정보 관리 부문, 정보 전략 부문 등의 형태로 구성될 수 있으므로 현대 기업에 적합한 정보 전담부서를 조직하기 위해서는 평소 경영층이 정보에 대한 중요성을 인식할 수 있도록 노력해야 한다.

(2) 인 력

데이터베이스 제작에 있어서 활용할 수 있는 인력은 크게 사내인력과 사외인력 즉, 컨설턴트로 분류할 수 있다. 사내인력이 지원해야 할 부분은 우선 정보내용에 대한 지원으로 등록정보의 색인어 부여,



분류, 요약 및 초록을 총괄적으로 관리할 담당인력을 확보해야 한다. 이러한 업무는 사내에서 발생하는 모든 종류의 정보에 대해 망라적으로 취급할 수 있는 전담부서를 설정하거나 필요시 정보센터를 설립하여 그 역할을 수행할 수 있는 인력을 배치하여야 한다.

또한 시스템 지원인력 즉 정보시스템기술자(SE, System Engineer)를 확보해야 하는데, 범용컴퓨터의 관리는 퍼스널컴퓨터 관리와 달리 상시 시스템을 지원하는 SE를 데이터베이스 관리멤버에 소속시켜야만 한다. 데이터베이스 구축준비 단계부터 시스템 개발 및 구축후에도 상당기간 동안 소프트웨어 수정, 입출력 방식의 개선사항 등이 발생되므로 그 기간동안 동일 SE로

부터 지원받을 수 있는 체제가 효율적이다. 가령 구축 준비단계에서 시스템부로부터 파견된 SE가 정보전담부서의 계획대로 활동하지 않을 경우, 빠른 시간내에 조건에 합당한 SE로 교체시켜야 한다. 이는 SE의 역할이 전장에서 설명한 바와 같이 데이터베이스시스템 구축에 있어서 매우 중요한 비중을 차지하고 있기 때문이다.

사내인력중 마지막으로 각사업부, 사업장 단위의 정보관리 책임자와 담당자를 검토해야 한다. 책임자는 그 부서에서 발생한 정보내용과 정보에 수반되는 사항에 전적으로 책임질 수 있도록 직제를 고려하여 검토하는 것이 바람직하다. 정보관리담당자는 정보의 종합분석능력, 비밀정보 취급 위치, 현장에서의 정보수집능력 등 개인적

성을 고려하여 배정한다.

이상 사내인력 확보에 관한 사항을 언급하였는데, 데이터베이스시스템의 규모와 내용에 따라 외부 컨설턴트 인력이 필요할 수 있다. 이와 같이 사내데이터베이스 구축에 컨설턴트를 활용하면 다음과 같은 장점이 있다. 즉,

- 외부의 객관적인 입장에서 현 시스템의 미비한 점을 발견할 수 있으며,
- 사내 구성원보다 사내조직에 더 큰 영향을 미칠 수 있고,
- 컨설턴트의 인맥이 활용 가능하며,
- 전문지식을 흡수할 수 있다는 것 등이다.

외부 컨설턴트를 활용하기 위해서는 우선 협력분야를 설정하여, 해당분야에 대한 일련의 방침을 세운후 정보관리 논문이나 관련 학회를 통하여 전문가 리스트를 확보하여 교섭한다.

(3) 장비(하드웨어)

데이터베이스 구축상 신규로 필요한 장비와 현재 보유하고 있는 장비에 대하여 체크해야 되는데 이하와 같은 장비가 필요할 것으로 판단된다.

- 범용컴퓨터
자사컴퓨터일 경우는 이용가능한 보조 기억장치의 용량을 조사하며,
외부 컴퓨터를 활용할 경우는 타임셰어링(Time Sharing)의 제반조건을 조사
- 전용단말기, 퍼스널컴퓨터
- 프린터
- 광과일링기기
- 모뎀
- 팩시밀리

- 복사기
- 마이크로필름기기 등

(4) 비 용

데이터베이스 구축 설계 단계에서 재산성에 대한 문제가 제기될 수 있다. 그러나 자사 고유의 정보는 전술한 보안문제 등으로 자체에서 관리하지 않을 수 없는 문제이며 정보관리의 중요성을 사내 전반적으로 인식된 상태에서 논의되어야 할 문제라고 생각된다. 그렇다고 지나치게 현실과 동떨어진 비용이 드는 시스템은 현실 가능



성이 없으므로 여러가지 경우에 적합하게 된다.

컴퓨터 하드웨어의 비용은 자사컴퓨터를 사용할 경우 컴퓨터 본체의 리스료, 마그네틱 운영 보관료, 인건비 및 일상 운용비 등의 고정비를 고려해야 하며 외부컴퓨터를 사용할 경우는 대개 타임셰어링 방식을 택하므로 CPU 사용 접속료, 디스크 및 TSS 사용료 등을 고려해야 한다. 소프트웨어 비용으로는 범용 DBMS를 사용할 경우 응용 소프트웨어 개발비를 산정하는 것이

중요하며, 그외 입력 및 유지, 관리비용을 산정하여 놓아야 한다.

4. 데이터베이스 구축 개념도

데이터베이스 시스템에 대한 개념설계가 완료되면 데이터베이스의 제작과 활용 과정을 쉽게 파악할 수 있는 개념도를 작성할 필요가 있다. 데이터베이스 구축 개념도의 일례로서 <도표 II-11>를 참고하기 바란다. ㉞

<도표 II-11> 데이터베이스 구축 개념도의 일례

