

# 1-01. 데이터 베이스의 필요성

## 데이터와 정보

데이터 : 현실 세계에서 단순히 관찰하거나 측정하여 수집한 사실이나 값

정보 : 의사 결정에 유용하게 사용할 수 있도록 데이터를 처리한 결과물



그림 1-1 데이터와 정보의 이해 : 원유와 가공 우유

## 정보시스템과 데이터베이스

정보시스템 : 조직운영에 필요한 데이터를 수집, 저장했다가

필요시 유용한 정보를 만들어주는 수단 (시스템 전체)

데이터베이스 : 정보시스템 안에서 데이터를 저장하고 있다가

필요시 제공하는 역할을 담당 (저장 담당)

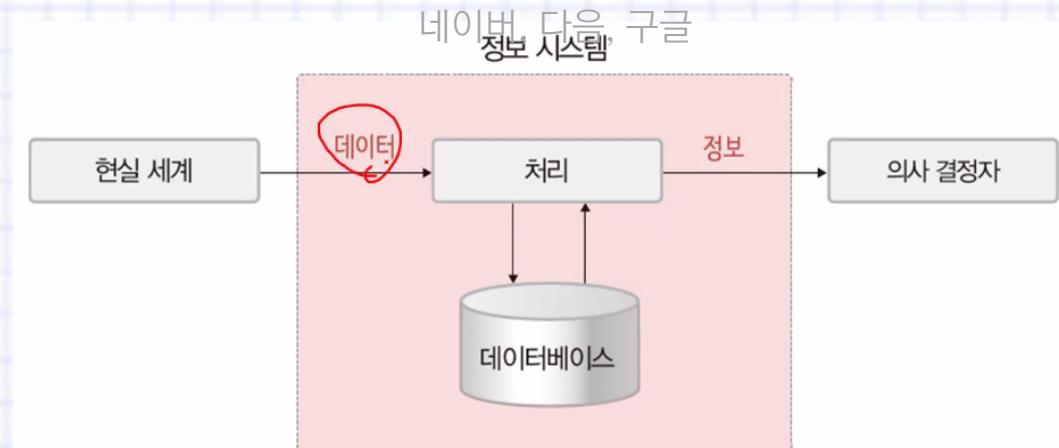


그림 1-3 정보 시스템의 역할과 구성

정보처리(information processing) : 데이터에서 정보를 추출하는 과정 또는 방법



그림 1-2 정보 처리의 예

## 1-02. 데이터베이스의 정의와 특징

### 데이터베이스(DB; DataBase)

특정 조직의 여러 사용자가 **공유**하여 사용할 수 있도록 **통합**해서 **저장**한 **운영** 데이터의 집합

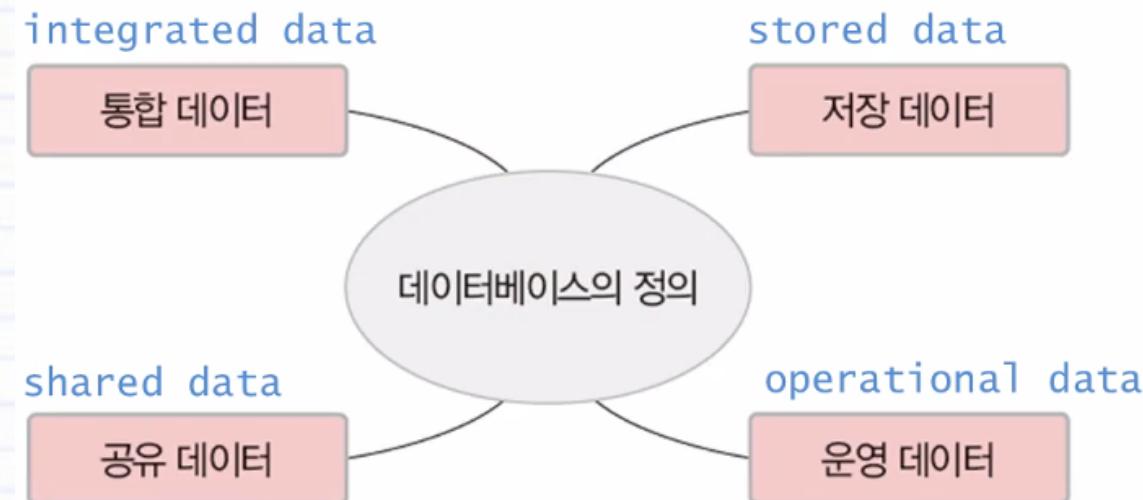


그림 1-4 데이터베이스의 정의

### 데이터베이스의 특징

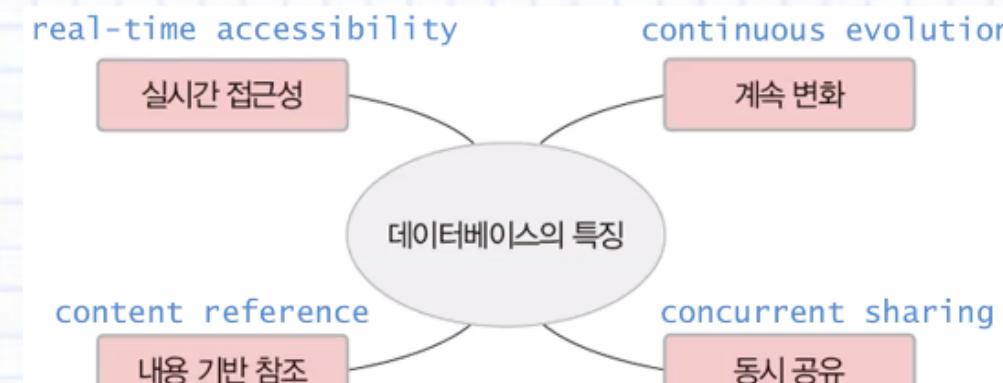


그림 1-5 데이터베이스의 특징

### 데이터베이스의 정의

공유데이터 : 특정 조직의 여러 사용자가 함께 소유하고 이용할 수 있는 공용데이터  
통합데이터 : 최소의 중복과 통제 가능한 중복만 허용하는 데이터  
저장데이터 : 컴퓨터가 접근할 수 있는 매체에 저장된 데이터  
운영데이터 : 조직의 주요 기능을 수행하기 위해 지속적으로 꼭 필요한 데이터

실시간 접근 : 사용자의 데이터 요구에 실시간으로 응답  
계속 변화 : 데이터의 계속적인 삽입, 삭제, 수정을 통해 현재의 정확한 데이터를 유지  
동시 공유 : 서로 다른 데이터의 동시적인 사용뿐 아니라 같은 데이터의 동시 사용도 지원  
내용 기반 참조 : 데이터가 저장된 주소나 위치가 아닌 내용으로 참조  
ex). 재고량이 1,000개 이상인 제품의 이름을 검색하시오.

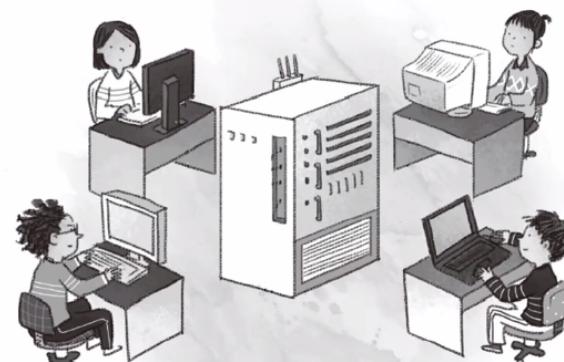


그림 1-6 데이터베이스의 이용

# 1-03. 데이터와 데이터베이스

## 데이터의 분류

- 정형데이터
- 반정형데이터
- 비정형데이터

## 정형 데이터(structured data)

구조화된 데이터, 즉 미리 정해진 구조에 따라 저장된 데이터

ex). 엑셀의 스프레드시트, 관계 데이터베이스의 테이블

	A	B	C	D
1	일자	배송업체	배송 건수	전일대비 상승률
2	2019-03-02	빠르다 택배	100	0%
3	2019-03-02	한빛 택배	200	10%
4	2019-03-02	안전 택배	50	3%
5	2019-03-02	당일 택배	30	-10%

그림 1-7 정형 데이터의 예

## 반정형 데이터(semi-structured data)

구조에 따라 저장된 데이터이지만, 데이터 내용 안에 구조에 대한 설명이 함께 존재

구조를 파악하는 파싱(parsing)과정이 필요 (정형데이터를 그대로 가져올 수 없으니)

보통 파일 형태로 저장

ex). 웹에서 데이터를 교환하기 위해 작성하는 HTML, XML, JSON 문서나 웹로그, 센서 데이터 등

```
{  
  "이름" : "오형준",  
  "나이" : 23,  
  "성별" : "남"  
}
```

〈친구정보〉  
 〈이름〉 오형준 〈/이름〉  
 〈나이〉 23 〈/나이〉  
 〈성별〉 남 〈/성별〉  
〈/친구정보〉

(a) JSON

(b) XML

그림 1-8 반정형 데이터의 예

## 비정형 데이터(unstructured data)

정해진 구조가 없이 저장된 데이터

ex). 소셜데이터의 텍스트, 영상, 이미지, 워드나 pdf문서와 같은 멀티미디어 데이터



그림 1-9 반정형 데이터의 예  
(Designed by S.salvador / Freepik)

활용신청 상세기능정보				
NO	상세기능	설명	일일 트래픽	미리보기
1	코로나19감염현황 조회 서비스	코로나19감염증으로 인한 일별 확진자, 완치자, 치료중인환자, 사망자등에 대한 현황자료	1000	<a href="#">확인</a>
요청변수(Request Parameter)				
항목명	샘	<body> <items> <item>		
ServiceKey	-	<accDefRate>1.4215047097</accDefRate> <accExamCnt>9319094</accExamCnt> <accExamCompCnt>9219878</accExamCompCnt> <careCnt>7883</careCnt> <clearCnt>121282</clearCnt> <createDt>2021-05-15 09:38:53.579</createDt> <deathCnt>1896</deathCnt> <decideCnt>131061</decideCnt> <examCnt>99216</examCnt> <resutINegCnt>9088817</resutINegCnt> <seq>512</seq> <stateDt>20210515</stateDt> <stateTime>00:00</stateTime> <updateDt>null</updateDt>		
pageNo	1	</item> <accDefRate>1.4193209445</accDefRate> <accExamCnt>9281291</accExamCnt> <accExamCompCnt>9219878</accExamCompCnt>		
numOfRows	10	<item> <accDefRate>1.4193209445</accDefRate> <accExamCnt>9281291</accExamCnt> <accExamCompCnt>9219878</accExamCompCnt>		
startCreateDt	20210515	<item> <accDefRate>1.4193209445</accDefRate> <accExamCnt>9281291</accExamCnt> <accExamCompCnt>9219878</accExamCompCnt>		
endCreateDt	20210515	<item> <accDefRate>1.4193209445</accDefRate> <accExamCnt>9281291</accExamCnt> <accExamCompCnt>9219878</accExamCompCnt>		

## 2-01. 데이터베이스 관리 시스템의 등장 배경

### 파일 시스템(기존 사용방식)

데이터를 파일로 관리하기 위해 파일을 생성, 삭제, 수정, 검색하는 기능을 제공하는 소프트웨어 (ex. 윈도우 OS)

응용프로그램마다 필요한 데이터를 별도의 파일로 관리



### 파일 시스템의 문제점

같은 내용의 데이터가 여러 파일에 중복저장됨 → **데이터 중복성**

- 저장공간의 낭비는 물론 데이터 일관성과 데이터 무결성을 유지하기 어려움

응용 프로그램이 데이터 파일에 종속적임 → **데이터 종속성**

- 사용하는 파일의 구조를 변경하면 응용 프로그램도 함께 변경해야 함



데이터 파일에 대한 동시 공유, 보안, 회복 기능이 부족

응용 프로그램 개발이 쉽지 않음

