



데이터베이스 Database

01 데이터베이스와 DBMS

데이터베이스 개요

데이터베이스 관리 시스템

데이터베이스 발전과 종류

데이터베이스 개념

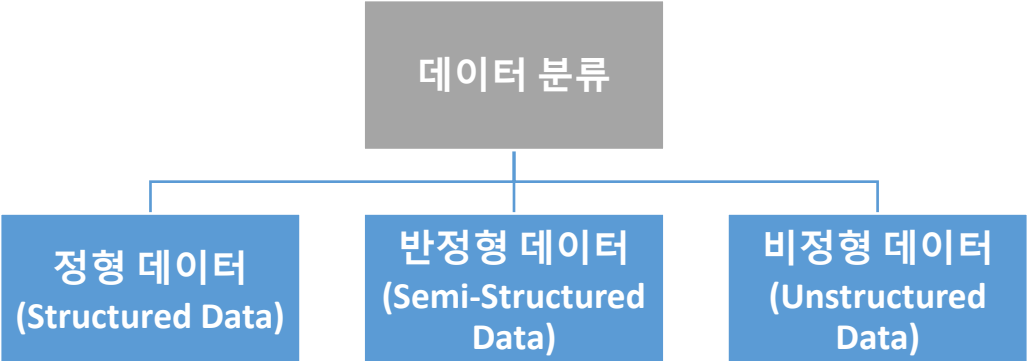
- 방대한 데이터를 효율적으로 관리하기 위해 컴퓨터에 통합·저장한 것
 - 특정 조직의 여러 사용자가 **공유**하여 사용할 수 있도록 **통합**해서 **저장**한 **운영** 데이터의 집합
 - 데이터베이스 관리 시스템(DBMS)이라는 프로그램을 이용하여 관리
-
- 데이터베이스에 저장된 데이터의 특징

공유 데이터 (shared data)	통합 데이터 (integrated data)	저장 데이터 (stored data)	운영 데이터 (operational data)
<ul style="list-style-type: none">• 특정 조직의 여러 사용자가 함께 소유하고 이용할 수 있는 공용 데이터	<ul style="list-style-type: none">• 최소의 중복과 통제 가능한 중복만 허용하는 데이터	<ul style="list-style-type: none">• 컴퓨터가 접근할 수 있는 매체에 저장된 데이터	<ul style="list-style-type: none">• 조직의 주요 기능을 수행하기 위해 지속적으로 필요한 데이터

- 데이터베이스의 특징

실시간 접근성 (Real-time accessibility)	계속 변화 (Continuous evolution)	동시 공유 (Concurrent Sharing)	내용 기반 참조 (Content reference)
<ul style="list-style-type: none">• 사용자의 데이터 요구에 실시간으로 응답	<ul style="list-style-type: none">• 데이터의 계속적인 삽입, 삭제, 수정을 통해 현재의 정확한 데이터 유지	<ul style="list-style-type: none">• 서로 다른 데이터의 동시 사용뿐만 아니라 같은 데이터의 동시사용지원	<ul style="list-style-type: none">• 데이터가 저장된 주소나 위치가 아닌 내용으로 참조• (예) 성적이 70점 이상인 학생

데이터 분류



- 정형 데이터(Structured Data)
 - 관계형 데이터베이스 시스템의 테이블과 같이 고정된 컬럼에 저장되는 데이터와 파일
 - 지정된 행과 열에 의해 데이터의 속성이 구별되는 스프레드시트 형태의 데이터
 - 관계형 데이터베이스 시스템의 정형 데이터를 비정형 데이터와 비교할 때 가장 큰 차이점은 데이터의 스키마를 지원하는 것
 - 데이터의 스키마 정보를 관리하는 DBMS와 데이터 내용이 저장되는 데이터 저장소로 구분

스키마에 의해 정의된 컬럼

column1	column2	column3	column4
data	data	data	data
data	data	data	data
data	data	data	data
data	data	data	data

컬럼에 의해 정의된 데이터

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa

데이터 분류

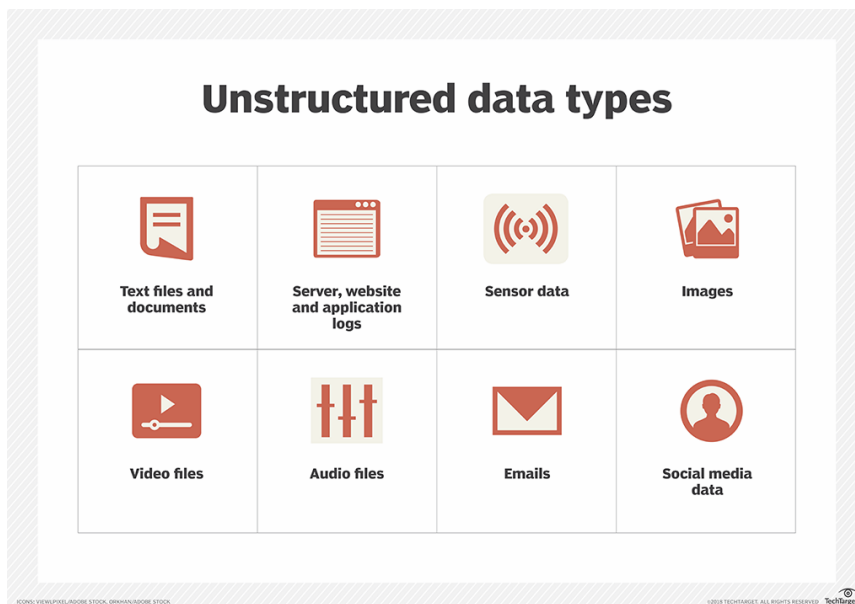
- 반정형 데이터(Semi-Structured Data)

- 데이터 내부에 정형데이터의 스키마에 해당되는 메타데이터를 갖고 있음
- 일반적으로 파일 형태로 저장
- 데이터 내부에 데이터 구조에 대한 메타정보를 갖고 있기 때문에 어떤 형태를 가진 데이터인지를 파악하는 것이 필요
- 데이터 내부에 있는 규칙성을 파악해 데이터를 파싱할 수 있는 파싱 규칙을 적용
- 반정형 데이터 예
 - URL 형태로 존재 - HTML
 - 오픈 API 형태로 제공 - XML, JSON
 - 로그형태 - 웹로그, IoT에서 제공하는 센서 데이터

```
[{ "Sepal.Length" : 6.8,  
  "Sepal.Width" : 3.2,  
  "Petal.Length" : 5.9,  
  "Petal.Width" : 2.3,  
  "Species" : "virginica" },  
{ "Sepal.Length" : 6.7,  
  "Sepal.Width" : 3.3,  
  "Petal.Length" : 5.7,  
  "Petal.Width" : 2.5,  
  "Species" : "virginica" }]
```

- 비정형 데이터(Unstructured Data)

- 데이터 세트가 아닌 하나의 데이터가 수집 데이터로 객체화
- 언어 분석이 가능한 텍스트 데이터나 이미지, 동영상 같은 멀티미디어 데이터가 대표적인 비정형 데이터
- 웹에 존재하는 데이터의 경우 html 형태로 존재하여 반정형 데이터로 구분할 수도 있지만, 특정한 경우 텍스트 마이닝을 통해 데이터를 수집하는 경우도 존재하므로 명확한 구분은 어려움
- 비정형 데이터 예
 - 이진 파일 형태: 동영상, 이미지
 - 스크립트 파일 형태: 소셜 데이터의 텍스트



파일 처리 시스템

- 데이터를 파일로 관리하기 위해 파일을 생성·삭제·수정·검색하는 기능을 제공하는 소프트웨어
- 응용 프로그램마다 필요한 데이터를 별도의 파일로 관리함
- 파일 처리 시스템의 문제점

같은 내용의 데이터가 여러 파일에 중복 저장 (데이터 중복성)

- 저장 공간의 낭비는 물론 데이터 일관성과 데이터 무결성을 유지하기 어려움

응용 프로그램이 데이터 파일에 종속적 (데이터 종속성)

- 사용하는 파일의 구조를 변경하면 응용 프로그램도 함께 변경해야 함

데이터 파일에 대한 동시 공유, 보안, 회복 기능 부족

- 하나의 파일을 동시에 공유해서 사용하기 어렵고, 누가 접근해서 사용하는지 보안 관리가 없으며, 시스템 문제가 발생하여 파일의 내용이 사라져도 회복이 어려움

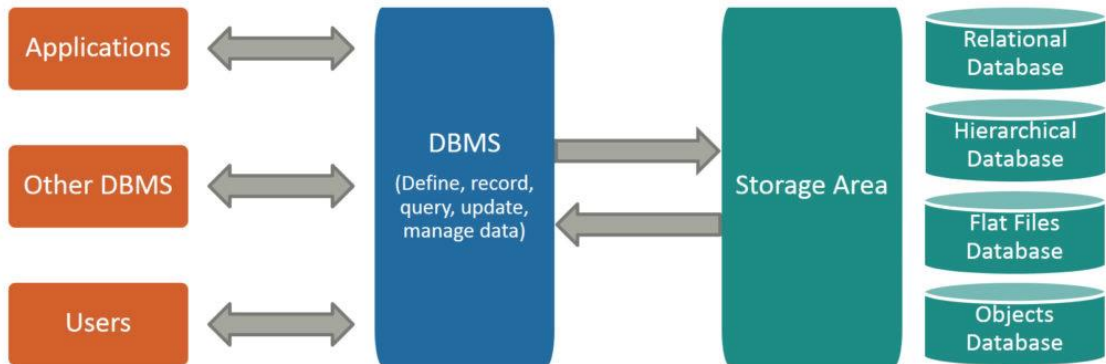
응용 프로그램 개발이 쉽지 않음

- 파일에 대한 처리 및 관리를 응용 프로그램에서 일부 해주어야 하므로 개발 자체가 쉽지 않음

데이터베이스 관리 시스템

- 데이터베이스 관리 시스템(DBMS: DataBase Management System)
- 파일 시스템의 문제를 해결하기 위해 제시된 소프트웨어
- 조직에 필요한 데이터를 데이터베이스에 통합하여 저장하고 관리함
- 사용자와 응용 프로그램에 편리하고 효율적인 데이터베이스 사용 환경을 제공하는 소프트웨어

Database Management System



- 데이터베이스 관리 시스템의 종류

ORACLE®



Informix®
SOFTWARE

SYBASE®



Microsoft
SQL Server

Microsoft
Access

MySQL®



MariaDB
Foundation

데이터베이스 관리 시스템

• 데이터베이스 관리 시스템의 주요 기능

정의 기능	조작 기능	제어 기능
• 데이터베이스 구조를 정의하거나 수정 가능	• 데이터를 삽입·삭제·수정·검색하는 연산 가능	• 데이터를 항상 정확하고 안전하게 유지

• 데이터베이스 관리 시스템 도입에 따른 장점

데이터의 중복과 불일치 감소	데이터 독립성 확보	데이터의 공유와 동시 접근이 가능함	데이터 보안 향상
데이터 무결성 향상	표준화 용이	시스템의 융통성 향상	응용 프로그램 개발 및 유지 비용 감소
사용자에게 더 나은 서비스 제공	요구 사항에 따른 조정 가능	시스템의 고장으로 부터 데이터베이스 복구 가능	데이터 중심의 중앙 집중 관리

데이터베이스 관리 시스템

• 데이터베이스 주요 기능

데이터 정의 언어(Data Definition Language, DDL)

- 데이터베이스 구조 정의 및 수정에 사용
- 데이터 저장 구조, 데이터 접근 방법, 데이터 형식 등의 정의 가능

데이터 조작 언어(Data Manipulation Language, DML)

- 데이터베이스에 저장된 데이터를 검색, 수정, 삽입, 삭제할 때 사용

데이터 제어 언어(Data Control Language, DCL)

- 데이터를 정확하고 안전하게 보호 관리
- 데이터베이스의 무결성 유지, 보안 및 접근 제어, 시스템 장애로부터의 복구, 병행 수행 제어 기능 등을 수행

• 데이터베이스 사용자

응용 프로그래머

- 프로그래밍 언어(C, COBOL 등)를 사용하여 특정 응용 프로그램이나 인터페이스를 구현하는 사람

최종 사용자

- 데이터의 검색, 삽입, 삭제, 갱신 등을 위해 DBMS를 사용하는 사람

데이터베이스 관리자

- 데이터 정의 언어(DDL)와 데이터 제어 언어(DCL)를 사용하여 데이터베이스 스키마를 생성하고 관리하는 사람



이수안 컴퓨터 연구소

suan computer laboratory