**데이터베이스**

주제 : OT

교제 추천 – 데이터베이스 시스템, 황규영 등 옮김

15시간 이상만 빠지지 말자

**제 1장. 데이터베이스와 데이터베이스 사용자 2025.09.04 ~ 2025.**

**데이터**: **의미**를 가지면서 **기록**될 수 있는 알려진 **사실**

* 의미가 없다면 데이터라고 볼 수 없다.
  + 의미는 주관적인 것이므로 누군가에게는 데이터일 수도 아닐 수도 있다.
* 기록될 수 없다면 의미가 있더라도 데이터라고 볼 수 없다.
* 사실이어야 데이터

데이터베이스: **관련 있는** 데이터의 모임

* 동일한 데이터베이스를 구축하더라도 구축한 사람에 따라 달라질 수 있다.
* 의미와 관련은 주관적인 영역이기 때문

데이터베이스 관리시스템: 데이터베이스의 생성과 관리를 담당하는 소프트웨어 패키지

데이터베이스 시스템: Database와 그를 관리하는 소프트웨어(DBMS, 응용 프로그램)를 칭하는 용어

작은 세계(mini-world): 데이터베이스 구축의 대상이 되는 실세계의 일부분

데이터베이스의 응용분야

전통 분야

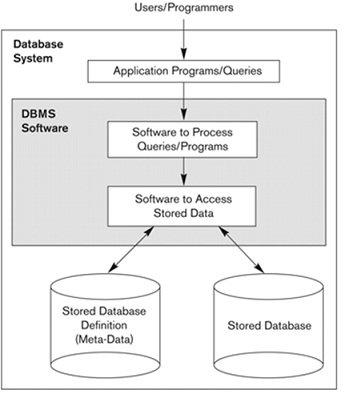
* 문서 DB / 예) 구글, 네이버 등

최신 분야

* Multimedia DB / 예) 네이버 렌즈 등
* GIS / 지도 정보를 통한 서비스
* Data Warehouse (데이터 창고)
  + 언젠가 사용할 데이터를 분류해서 모아놓는 곳
  + 지금 당장 의미가 없을지라도 일단 **KEEP**
  + 이 창고에서 관련 있는 데이터를 얻어내는 것을 데이터 마이닝이라고 한다.
* 실시간 DB
  + 시간적 제약이 중요한 분야가 존재한다
  + 시간 제약이 있는 경우, 대략적인 결과를 내보낸다.
* 빅데이터

//=================================================================//

**A Simplified Database System Environment**

**Application Programs/Queries**

데이터 베이스에 접근이 가능하도록 해주는

**응용 프로그램**

**DBMS Software/질의 처리기**

**Software to Process Queries/Programs**

사용자가 원하는 것을 **해석**하는 부분

**Software to Access Stored Data**

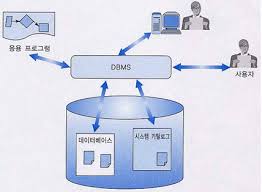
명령을 **실행**하는 부분

Stored Database: 데이터 베이스

Meta – Data: 다른 데이터를 설명하고 정의하여 찾기, 사용 및 관리를 더 쉽게 만드는 정보

//=================================================================//

**데이터베이스 시스템의 구성 요소**

**응용프로그램**을 통한 데이터베이스 액세스

컴퓨터를 통한 **직접** 데이터베이스 액세스

시스템 카탈로그 = Meta - Data

//=================================================================//

(데이터)파일이란?

* 물리적인 비트의 연속
* 순차적인 **레코드**들로 구성
  + **레코드**/Ex) [이름, 학번, 전화번호, 과목]
* 레코드는 연관된 **필드**들의 모임
  + **필드**/Ex) [이름], [학번], [전화번호], [과목]
  + 각 필드는 **동일**하다/[레코드1]->[필드1] = [레코드2]->[필드1]

파일 시스템이란?

* 파일을 이용하여 자료를 관리하는 방법

파일 시스템의 특징

* 데이터에 대한 프로그램의 **의존도가 높다**
  + 데이터를 정의하여 코딩하는 것
    - 데이터를 int로 정의한다면, 프로그램은 int를 통해서만 작동될 것이다.
* 데이터의 구조가 **수정(추가, 삭제)**될 때마다 프로그램을 수정해야 하는 문제가 생긴다.
* 이 문제는 데이터를 하드 코딩했기 때문이기에, 데이터의 구조를 미리 알 수 있다면 의존도를 낮출 수 있다.
* 이때 필요한 것이 데이터의 데이터, **Meta – Data**
* 하지만 Meta – Data를 저장하는데 필요한 메모리, 읽는데 필요한 시간이 증가한다.

**그렇다면 Meta – Data를 쓰는 게 이득일까?**

* 데이터에 대한 프로그램의 의존도를 낮추는 것은 성능 상의 이득이 더 크다.
* 밑에서 데이터의 구조가 변화하더라도 시스템은 **24시간 정상 작동**할 수 있게끔 한다.
  + 사용자에게 **완전한 자유를 보장**한다.

**그 밖의 문제점들**

데이터의 중복성

* 데이터 중복의 장단점
  + 장점
    - 백업 파일
  + 단점
    - 메모리 사용량이 많아진다.
    - 수정할 시 동일 데이터가 아니게 된다.
* 레코드는 다르더라도 안의 필드가 같을 경우, 각 필드에 따른 중복 파일이 존재할 수 있는데, 이 필드를 수정하게 되면 그 순간부터 동일한 데이터가 아니게 되므로 혼란이 생길 수 있다. 이때 어느 데이터가 정확한 데이터가 될까?
* **데이터의 부정확성**

각종 기능 부족

* 데이터 모델링 개념 **X**
  + 어떤 형태로 데이터를 만들어야 효율적인지 정해져 있지 않다.
  + 데이터 모델링: 데이터를 어떻게 관리해야 하는 가
* 질의어 **X**
  + 어떠한 질문을 해서 데이터를 끌어낼 수 없다.
  + 개발자가 결정해서 구현해야 한다.
* 동시성제어 **X**
  + 여러 사용자가 동시에 액세스 시 문제가 생긴다.
* 파손, 회복 **X**
  + 중복이 없으므로 백업이 존재하지 않는다.
* 보안 **X**
  + 권한에 따라 파일을 액세스 여부를 결정할 수 있다.
* 밑의 3개는 약간의 기능을 제공한다.

데이터베이스의 특징

* 데이터베이스 시스템의 자기 기술성
  + Meta – Data를 통해 데이터의 구조를 직접 알고 있지 않아도 데이터를 액세스 가능
* 프로그램과 데이터의 분리
  + 데이터베이스 내 데이터 저장 구조가 변경돼도 데이터베이스 응용 프로그램은 영향받지 않음
  + 대신 퍼포먼스가 떨어진다.
* 데이터 추상화
  + 데이터 모델을 사용함으로써 저장구조와는 별도로데이터의 의미를 표현하는 방법을 제공
* 데이터에 대한 다양한 뷰
  + 사용자는 전체 데이터베이스 보다는 관심이 있는 데이터베이스의 일부를 뷰로 정의할 수 있음
* 여러 사용자의 동시 사용