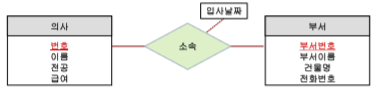
[빅데이터이론] 2019.03.11(월)

데이터베이스 설계

* 요구사항 수집 및 분석
  + 연산이 무엇이 필요한지 등등
* 개념적 설계
  + 데이터 타입, 관계, 제약조건들 기술
* 논리적 설계
  + 데이터베이스 시스템을 선택하여 그 시스템에 맞게 개념 스키마를 논리 스키마로 변환하는 과정

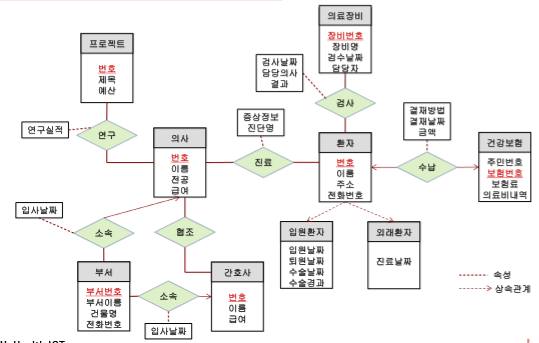
ER 모델 (ER(Entity-Relationship) 모델)

* 모델링 중에서 개념적 모델링 사용
* 객체들과 그들 간의 관련성을 그림으로 표현
* 데이터베이스 설계 시에 많이 사용됨



* + 입사날짜 🡪 속성
  + 번호, 부서번호 🡪 주키.
* 용어
  + 개체: 실 세계에서 다른 객체와 구별되어지는 객체
  + 속성: 집함에 의해 표현
  + 관계: 개체 간에 발생하는 행위

유헬스 데이터베이스 설계를 위한 ER 다이어그램



관계형 모델 🡪 대부분의 DB

* 데이터와 이들 사이의 관계(또는 릴레이션)를 테이블들의 모임으로 표현
* 릴레이션의 스키마에는 릴세이션 이름, 각 필드의 이름, 각 필드의 타입이 기술됨.
* 용어
  + 스키마: 데이터베이스의 논리적 구조를 기술하는 것
  + 인스턴스: 실제 내용, 변수(value)의 값
  + 키(key): 릴레이션에서 특정 레코드를 식별할 때 사용하는 속성
    - 주 키: 사용자가 선택한 것
    - 후보 키: 슈퍼키 중에서 길이가 짧은 것(1, 2, 3)
    - 슈퍼 키: 키가 될 수 있는 모든 것
      1. 부서번호
      2. 부서명
      3. 전화번호
      4. 부서번호 + 부서명
      5. 부서번호 + 건물명
      6. 건물명 + 부서명 + 건물명

관계형 모델의 예

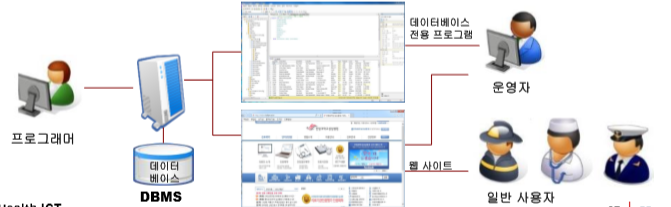


SQL 질의 언어

* SQL 질의 언어 🡪 비절차식 언어  
  (일반 프로그램 언어 🡪 절차식 언어)
* DB 전용 입출력 format이 자유로움
* SQL 질의 언어 분류
* DDL 데이터 정의어
  + create table, char 등
* DML 데이터 조작어
  + 데이터의 검색, 삽입, 삭제, 수정하는데 사용
  + SELECT, INSERT, DELETE, UPDATE 문

유헬스 데이터베이스 시스템

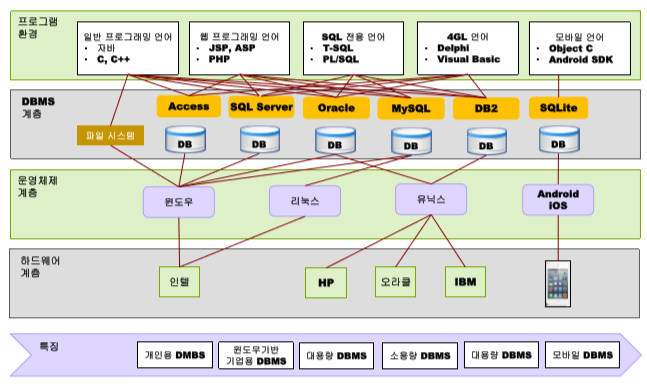
* 유헬스 시스템 운영 환경



* + 일반 사용자와 접근 방식이 틀림
  + 운영자, 개발자들은 두 가지 방법으로 접근 가능

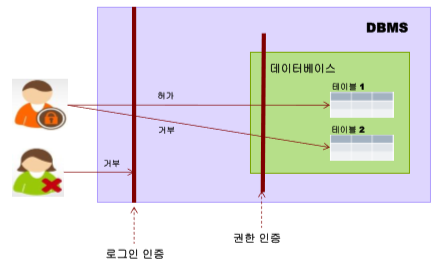
데이터베이스 프로그래밍

* 데이터베이스 프로그래밍
  + 호스트 언어, DBMS 종류, 운영체제, 컴퓨터 기종에 따라 다양
  + 최근 모바일 환경에서는 SQLite와 같은 경량 DBMS가 사용됨
* 데이터베이스 프로그래밍 방법
  + SQL 전용 언어를 사용
    - SQL 자체의 기능을 확장하여 변수, 제어, 입출력 등의 기능을 추가한 새로운 언어를 사용하는 방법
    - SQL server는 T-SQL 언어 사용, Oracle은 PL.SQL이라는 언어 사용
    - 데이터베이스를 다루는 능력이 뛰어나고, 다루는 방법이 쉽다
    - 그러나, 사용자 인터페이스 구축 기능이 없어 독립적으로 사용 안 함.
* 일반 프로그래밍 언어에 SQL 삽입하여 사용 🡪 java, C, C++ 등 사용
* 웹프로그래밍 언어에 SQL을 삽입하여 사용 🡪 JSP, ASP, PHP 등 사용
* 4GL(4th Generation Language)을 사용 🡪 GUI 기반 소프트웨어 개발 도구 사용

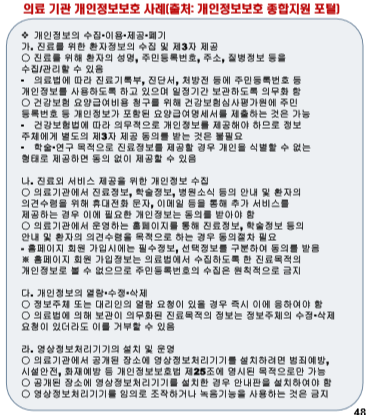


데이터베이스 권한과 보안

* 데이터베이스 보안
  + 데이터베이스는 기업의 중요한 무형의 자산
  + 데이터베이스는 구축/관리 비용이 매우 큼
  + 개인 정보 유출 및 오남용에 따른 사건/사고가 빈번히 발생
  + 어떤 유형의 위험으로부터 반드시 보호해야 함.
* 데이터베이스 보안 유형
  + 물리적 보호: 자연 재해, 디스크 손상
  + 운영 보호: 운영자의 실수
  + 권한 보호: **가장 많이 씀.** 특정 사용자만, 권한 인증



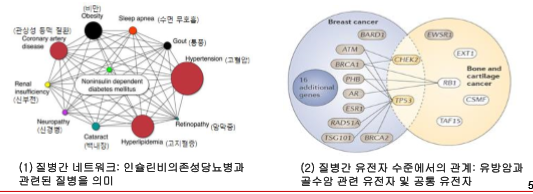
* 데이터베이스에 권한
  + 권한 허가
    - GRANT SELECT ON Doctor TO user1;
  + 권한 취소
    - REVOKE SELECT ON Doctor FROM user1;
* 개인정보 보호 규제
* 개인 보호법 🡪 기본적으로 진료와 관련된 것, 처방전 등, 연구 목적으로 사용 시



* 유헬스 시스템에서의 개인정보 보호
  + 유헬스 데이터(진료내역, 처방내역, 질병이력, 건강보험내역 등)은 매우 민감한 정보로 정보 유출을 최대한 방지 및 예방해야 함.
    - 개인 식별/기밀 정보 암호화:  
      주민 번호, 진료 내역, 건강보험내역 등 고유 식별 정보의 수집 제한 및 암호화
    - 개인정보 인증:   
      실명 확인, 성인 인증, 휴대폰 인증 등 다양한 개인 정보 인증 기법 개발
    - 모바일 보안 강화:   
      스마트폰 보안 솔루션 개발, 무선 통신망 보안 장비들에 대한 연구 필요
    - 통합 관제 모니터링:   
      개인 정보 접근 로그에 대한 모니터링 상시 관제 체계 구축
    - 통합 콜센터:   
      개인 정보가 유출된 경우 정보 주체에 관련 사실을 통지하여 피해 최소화

유헬스 데이터 분석

* 최근 연구 동향
  + 질병에 대한 연구



연구개발 사례

* 유전체 분석
  + 유전자는 아데닌(A),구아닌(G),시토신(C),티민(T)의 염기로 구성됨
  + 인간 유전자의 경우 이들 네 종류의 염기 30억 개가 일정한 순서로 배열되어 있음
  + 염기 서열에 따라 키와 피부색 등 생물학적 특성이 결정됨
  + 게놈 프로젝트로 인간의 염기서열이 밝혀짐에 따라, 다양한 종에 대해 염기서열을 완성하기 위한 연구가 진행되고 있음

