

자바(JAVA)활용 데이터 수집 및
분석을 위한 웹 시스템 구축 과정

데이터베이스

강 사 : 임 정 훈



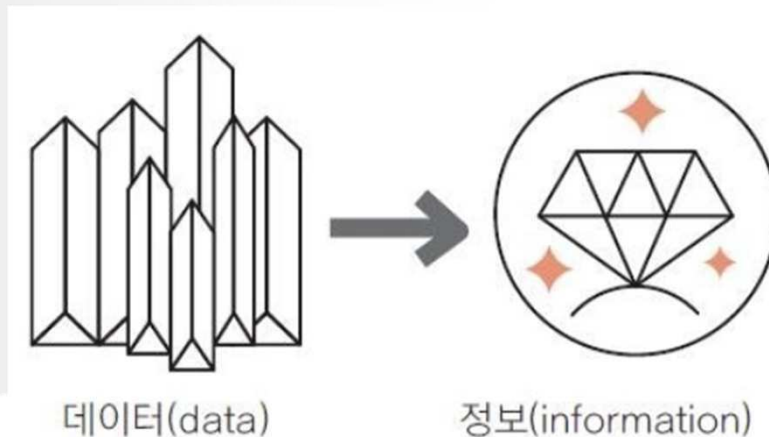
1. 데이터와 정보

● 데이터

- 어떤 필요에 의해 수집이 되었지만 아직 특정 목적을 위해 평가하거나 정제되지 않은 값이나 사실 또는 자료 자체를 말함.
- 데이터는 원석이라고 비유함.

● 정보

- 수집한 데이터를 어떠한 목적을 위해 분석하거나 가공하여 가치를 추가하거나 새로운 의미를 이끌어 낼 수 있는 결과를 말함.
- 정보는 보석이라고 비유를 함.

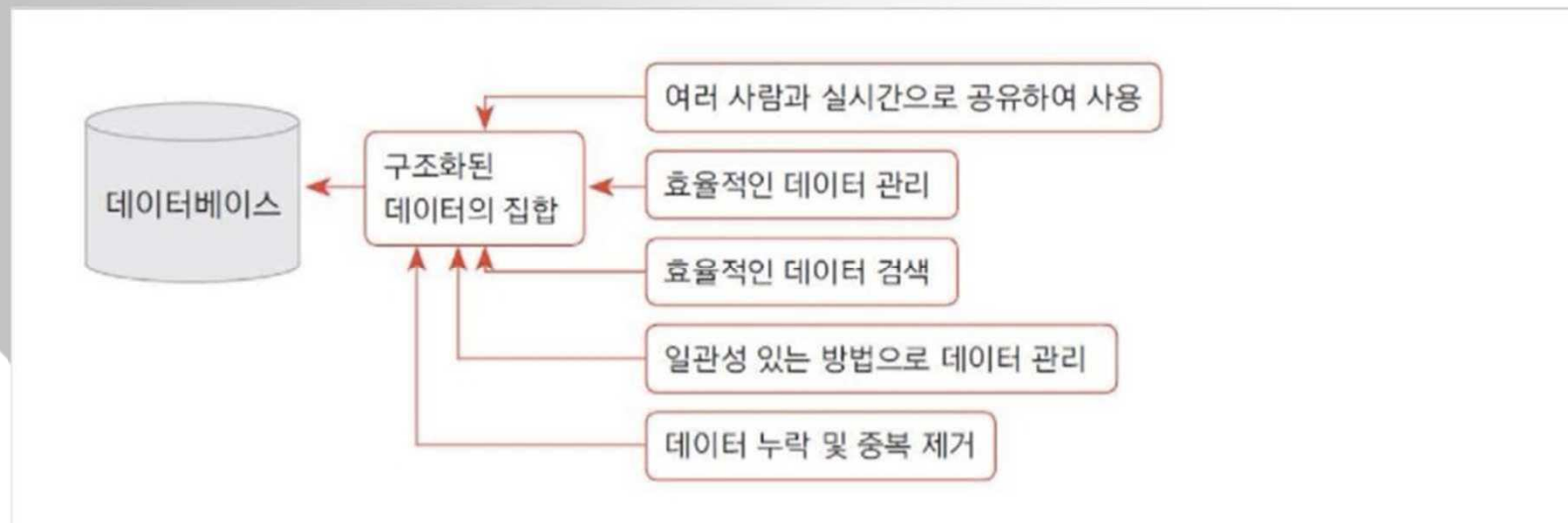




1. 데이터와 정보

● 효율적인 데이터 관리를 위한 조건

- 데이터를 통합하여 관리.
- 일관된 방법으로 관리.
- 데이터 누락 및 중복 제거.
- 여러 사용자(응용 프로그램 포함)가 공동으로 실시간 사용 가능.





2. 파일시스템과 DBMS

● 파일시스템을 통한 데이터 관리

- 데이터베이스 개념이 등장하기 전에 사용하던 데이터 관리 방법.
- 서로 다른 여러 응용 프로그램이 제공하는 기능에 맞게 필요한 데이터를 저장, 관리.
- 각 파일에 저장된 데이터는 서로 연관성이 없음.
- 연관성이 없으므로 중복 또는 누락이 발생할 수 있음.



2. 파일시스템과 DBMS



이러한 현상은 학생의 재학 상태를 관리하는 데이터가 각 응용 프로그램 별로 흩어져 있기 때문에 발생함.

학사 프로그램

학번	이름	학과	학년	학기	상태
2018-00001	홍길동	철학과	1	2	군휴학
2018-00002	이수선	컴퓨터공학과	4	2	졸업
2018-00003	이지수	경영학과	2	1	재학
2018-00004	김연아	사회체육학과	3	1	휴학

장학금 신청 프로그램

신청 장학금 종류	신청 일자	학번	재학 상태	장학금 신청 가능 여부
국가	20180409	2018-00001	군 휴학	신청 불가
성적	20180310	2018-00002	재학	신청 가능
동문	20180223	2018-00003	재학	신청 가능
근로	20180213	2018-00004	휴학	신청 불가

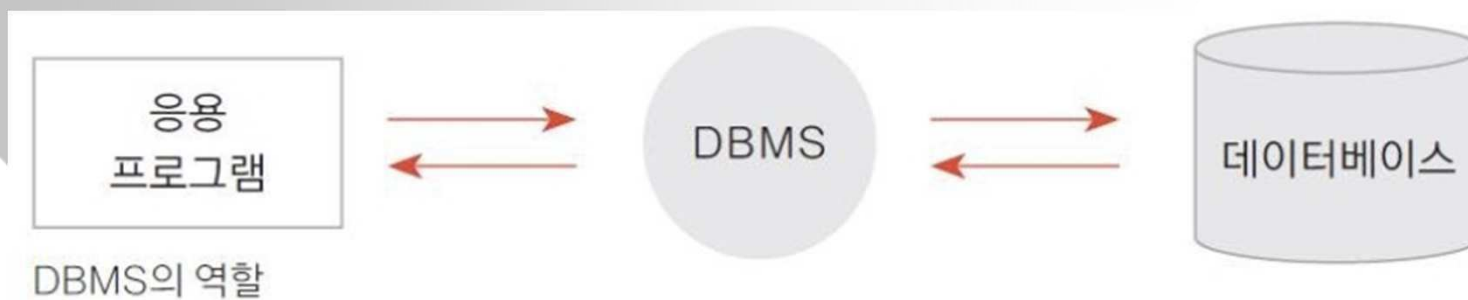
장학금 신청 프로그램에 졸업 정보가 누락된 경우



2. 파일시스템과 DBMS

● DBMS를 통한 데이터 관리 – 데이터베이스 관리 시스템

- 앞서서 보았던 파일 시스템의 단점을 보완하여 등장함.
- 효율적인 데이터 관리 조건을 만족하며, 서비스 제공의 효율성을 높이기 위해 등장함.
- 데이터베이스 관리 시스템은 데이터베이스의 데이터 조작과 관리를 극대화함.
- 여러 목적으로 사용할 데이터의 접근 · 관리 등의 업무를 전담함.
- 응용 프로그램이 필요한 데이터 작업을 요청하면, 자신이 관리하는 데이터베이스로 관련 작업을 수행하고 결과값을 제공함.





2. 파일시스템과 DBMS

● 파일시스템 방식과 데이터베이스 관리 시스템의 차이점

파일시스템 방식의 문제점	DBMS를 통한 데이터 관리
데이터의 중복.	하나의 소프트웨어가 데이터를 관리하므로 데이터의 중복을 피할 수 있음.
응용 프로그램이 개별 데이터를 직접 관리.	여러 응용 프로그램이 하나의 DBMS를 통해 데이터를 사용하므로 데이터를 동시에 공유할 수 있음.
응용 프로그램이 데이터를 쓰는 방식이 각각 다름.	하나의 DBMS를 통해 데이터를 관리하기 때문에 각각의 응용 프로그램이 데이터를 관리하는 방식이 통합됨.
데이터가 특정 응용 프로그램에 종속되어 있으므로 응용 프로그램을 변경하면 기존 데이터를 사용할 수 없음.	응용 프로그램과는 별도로 데이터가 DBMS에 의해 관리·보관되기 때문에 응용 프로그램의 업데이트 또는 변경과 관계없이 데이터를 사용할 수 있음.



3. 데이터 모델

● 계층형 데이터 모델

- 데이터 모델 : 컴퓨터에 데이터를 저장하는 방식을 정의해 놓은 개념 모형.
- 나뭇가지 형태의 트리(tree) 구조를 활용함.
- 데이터의 관련성을 계층별로 나누어 부모와 자식 같은 관계를 정의하고 데이터를 관리.
- 데이터베이스 관리 시스템은 데이터베이스의 데이터 조작과 관리를 극대화함.



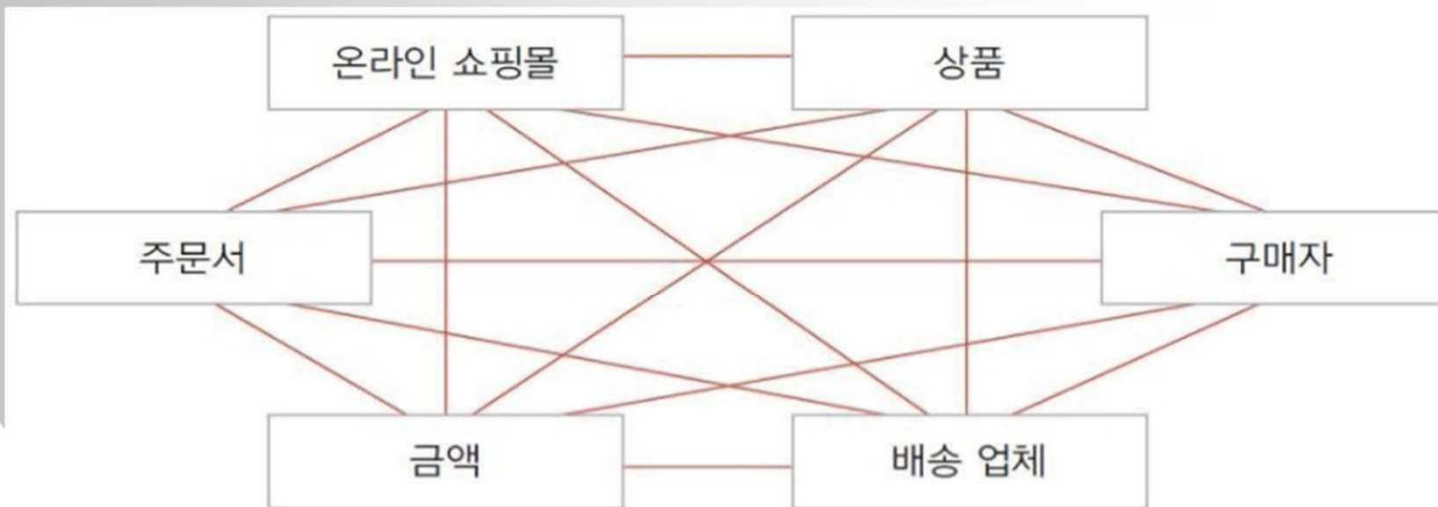
계층형 데이터 모델



3. 데이터 모델

● 네트워크형 데이터 모델

- 망형 데이터 모델이라고도 함.
- 그래프(graph) 구조를 기반으로 함.
- 객체 간 관계를 그래프 구조로 연결하므로 자식 개체가 여러 부모 개체를 가질 수 있다는 점에서 계층형 데이터 모델과 차이가 있음.



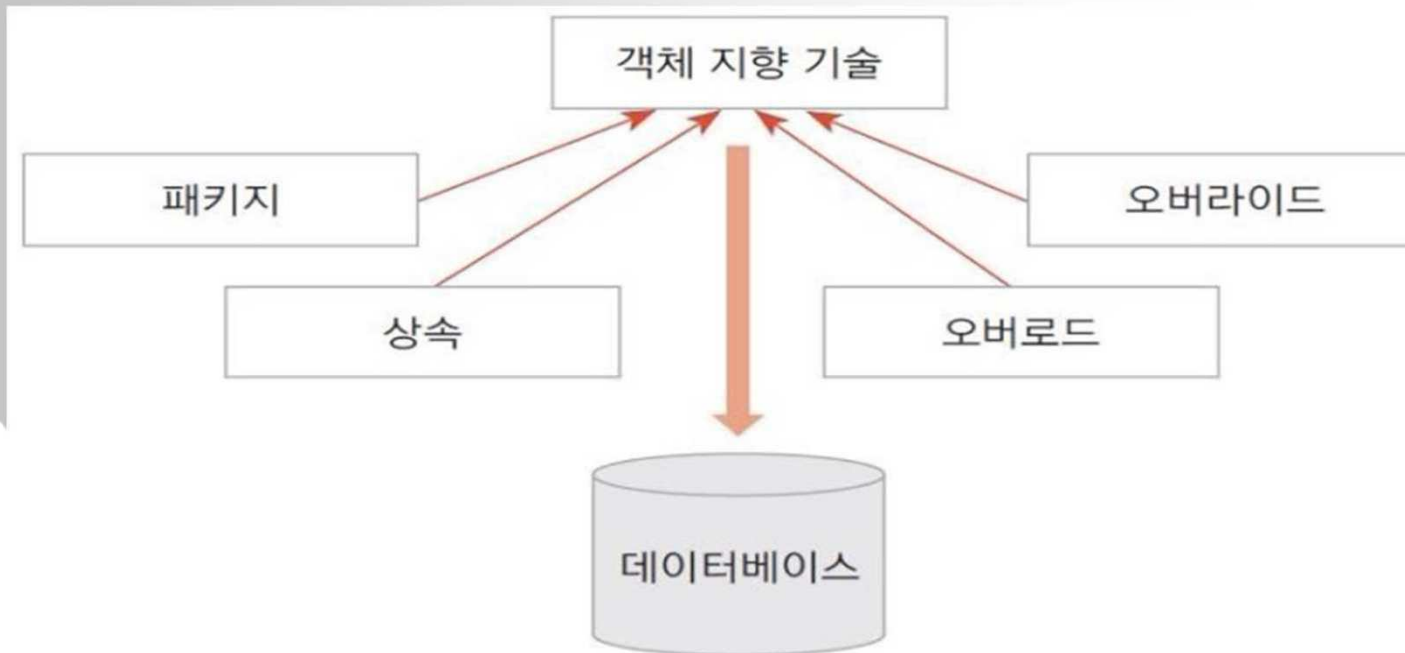
네트워크형 데이터 모델



3. 데이터 모델

● 객체지향형 데이터 모델

- 1980년대 등장한 모델 객체로 객체 개념을 기반으로 한 데이터 모델.
- 객체지향 프로그램처럼 독립된 객체로 구성하고 관리하며 상속, 오버라이드 등 객체 지향 프로그램에서 사용되는 강력한 기능을 활용할 수 있음.
- 객체 지향 모델 개념을 데이터베이스에 적용하는 것이 쉽지 않음. - 상용화가 힘들.



객체 지향형 데이터 모델



3. 데이터 모델

● 관계형 데이터 모델

- 현대에 가장 많이 사용하는 관계형 데이터베이스의 바탕이 되는 모델.
- 관계형 데이터 모델은 다른 모델과 다르게 데이터 간의 관계(relationship)에 초점을 맞춘 데이터 모델.

사원 정보	사원 번호	사원 이름	사원 직급	부서 이름	위치
사원 번호	0001	홍길동	과장	회계	서울
사원 이름	0002	성춘향	대리	연구	인천
사원 직급	0003	박문수	사원	운영	분당
부서 이름					
위치	0004	심청이	사원	회계	서울

데이터 중복 발생

사원 정보와 부서 정보를 하나의 데이터로 관리할 경우



3. 데이터 모델

사원 정보	사원 번호	사원 이름	사원 직급	부서 코드
사원 번호 사원 이름 사원 직급 부서 코드	0001	홍길동	과장	10
	0002	성춘향	대리	20
	0003	박문수	사원	30
	0004	심청이	사원	10
부서 정보	부서 코드	부서 이름	위치	
부서 코드 부서 이름 위치	10	회계	서울	
	20	연구	인천	
	30	운영	분당	



3. 데이터 모델

- 관계형 데이터 모델

이름	설명
개체 (entity)	데이터베이스에서 데이터화하려는 사물, 개념의 정보 단위임. 관계형 데이터베이스의 테이블(table) 개념과 대응되며 테이블은 릴레이션(relation)으로 표기하기도 함.
속성 (attribute)	개체를 구성하는 데이터의 가장 작은 논리적 단위. 데이터의 종류 · 특성 · 상태 등을 정의함. 관계형 데이터베이스의 열(column) 개념과 대응됨.
관계 (relationship)	개체와 개체 또는 속성 간의 연관성을 나타내기 위해 사용. 관계형 데이터베이스에서는 테이블 간의 관계를 외래키(foreign key) 등으로 구현하여 사용.



4. SQL

● SQL(Structured Query Language)

- 데이터를 다루고 관리하는데 사용되는 데이터베이스 질의 언어.

종류	설명
DQL(Data Query Language)	RDBMS에 저장한 데이터를 원하는 방식으로 조회하는 명령어.
DML(Data Manipulation Language)	RDBMS 내 테이블의 데이터를 저장 · 수정 · 삭제하는 명령어.
DDL(Data Definition Language)	RDBMS 내 데이터의 관리를 위해 테이블을 포함한 여러 객체를 생성 · 수정 · 삭제하는 명령어.
TCL(Transaction Control Language)	트랜잭션 데이터의 영구 저장 · 취소 등과 관련된 명령어.
DCL(Data Control Language)	데이터 사용 권한과 관련된 명령어.