



데이터베이스

데이터베이스의 개념





학습목표

- ➔ 정보처리 시스템에 대하여 설명할 수 있다.
- ➔ 데이터베이스에 대하여 설명할 수 있다.



학습내용

- ➔ 정보처리 시스템
- ➔ 데이터베이스



정보처리 시스템



데이터와 정보

01 데이터와 정보의 차이점

데이터(Data)

관찰/측정을 통해 수집된
사실(Fact) 또는 **값(Value)**이
특정 기준에 의해 정리된 것

VS

정보(Information)

어떤 목적에 의해 데이터를
해석하거나 **가공**한 형태

오늘의 온도, 습도, 강수량

VS

오늘의 온도, 습도, 강수량

연간 강수일수
(일 강수량 1mm 이상)
109일



강수 확률
평균 30%

02 정보(Information)의 개념

- 특정 상황에서 적절한 의사결정을 할 수 있도록 지원하는 지식
- 컴퓨터 시스템과 같은 처리기를 통한 데이터의 유효한 해석이나 데이터 상호 간의 관계를 의미



데이터

해석 or 가공

(컴퓨터 시스템)



정보



정보처리 시스템



정보처리 방식

01 정보처리(Information Processing)

- 의사결정(Decision Making)을 위한 정보를 컴퓨터가 생성하기 위해 **데이터를 처리**하는 작업

02 의사결정(Decision Making)

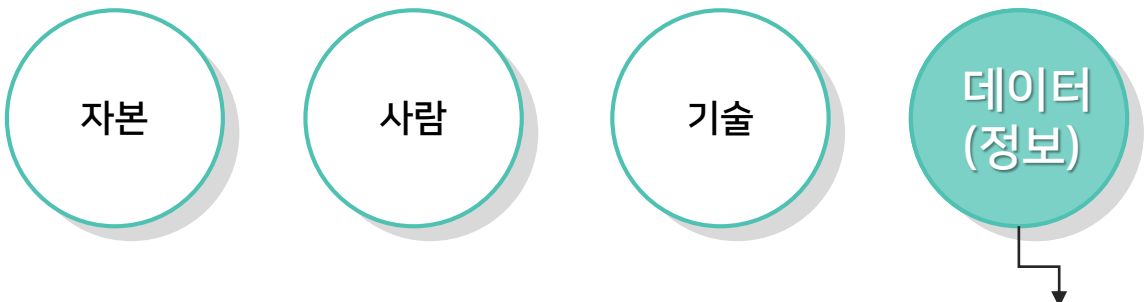
- 정보를 바탕으로 상황에 따라 주어진 문제를 해결하는 중요한 지능 활동
- 정보란 데이터의 유효한 해석이나 데이터 간의 상호 관계

03 정보의 특징



정보는 의사결정 과정에 반영되어
유용한 결과를 유발하므로 **정확성** 및 **현재성** 중요

[예: 기업의 4대 요소]



데이터는 사업 수행 방식을 묘사하고 계획에 대한 정보를 제공하여
기업 경영의 의사결정에 중요한 역할을 수행



정보처리 시스템



정보처리 방식



정보관리의 중요성

정보화 사회에서 적시에 정확한 정보를 얻을 수 있도록
데이터를 **효율적으로 관리**하는 것은 매우 중요



정보 시스템을 통한 관리 필요



정보 시스템



필요한 데이터를 수집, 조직, 저장해 두었다가 필요 시 처리하여
의사결정에 유효한 정보를 생성 분배하는 수단



배치 처리
(Batch Processing)

VS

온라인 처리
(Online Processing)

| 배치 처리(Batch Processing)

- 처리 방식
 - ▶ 데이터의 수집 및 분류
 - ▶ 유사한 일들을 모아서 **일정 시간 한꺼번에** 처리
- 장점: 컴퓨터 활용률(효율성) 매우 탁월
- 단점: 개별적인 처리 결과를 얻기 위해 대기 필요(컴퓨터 중심 처리)
- 응용 분야: 고지서, 공과금 계산, 이자 계산 등

| 온라인 처리(Online Processing)

- 처리 방식
 - ▶ 사전 준비작업 없이 **곧바로 데이터를 처리하는** 방식
 - ▶ 생성되는 출처로부터 곧바로 컴퓨터에 전송되고 전송된 데이터는 컴퓨터가 즉시 처리해서 그 결과를 바로 전송
- 장점: 대기시간이 없으므로 사용자에게 매우 편리(사용자 중심 처리)
- 단점: 컴퓨터가 항상 가동 상태이어야 하므로 **테스트나 유지 보수**의 어려움
- 응용 분야: 핸드폰의 실시간 데이터 사용량 검색 등

과거의 배치 처리 방식이었던 것도 최근에는 온라인 처리 형태로 바뀜



중앙집중 처리 (Centralized Processing)

VS

분산 처리 (Distributed Processing)

| 중앙집중 처리(Centralized Processing)

- 처리 방식
 - ▶ 범 조직적 자료를 한군데 모아 놓고 처리하는 방식
- 장점: 전산요원의 효율적 이용, 업무 통제, 보안 등에 있어서 효율적

| 분산 처리(Distributed Processing)

- 처리 방식
 - ▶ Globalization에 따라서 데이터가 지리적, 시간적으로 분산되어 발생
 - ▶ 처리기를 데이터 발생 위치에 가까이 배치하여 효율성 증가
- 장점
 - ▶ 실세계의 데이터의 발생, 관리를 유연한 대처 가능

중앙집중 처리 (Centralized Processing)

분산 처리 (Distributed Processing)

- ▶ 단점: 해킹 등 피해를 입었을 때
피해 규모가 큼
- ▶ 장점: 데이터가 한 곳에 집중되어
있어 보안 비용이
상대적으로 적게 듭

VS

- ▶ 단점: 데이터의 보안을 위해
많은 비용이 소요됨
- ▶ 장점: 해킹 등 피해를 입었을 때
특정 부분만 손실되어
피해 규모가 적음



정보처리 시스템



다양한 정보처리 시스템

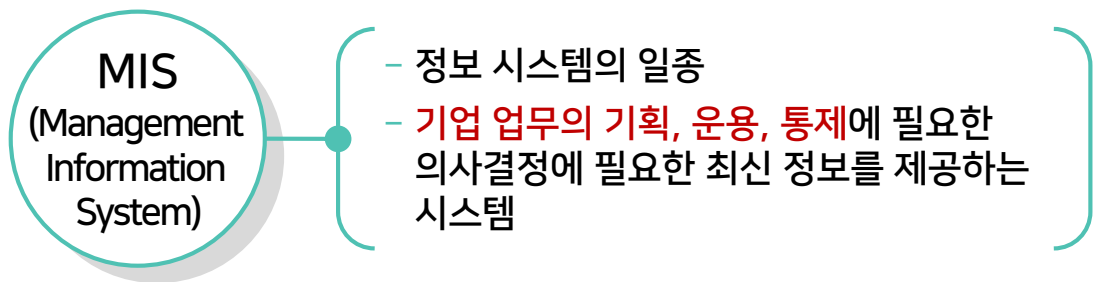
01

정보처리 시스템의 유형



| 정보 시스템

- 한 조직체의 활동에 필요한 데이터를 수집·조직·저장해 두었다가 필요 시 처리하여 의사결정에 유용한 정보를 생성 및 분배하는 시스템



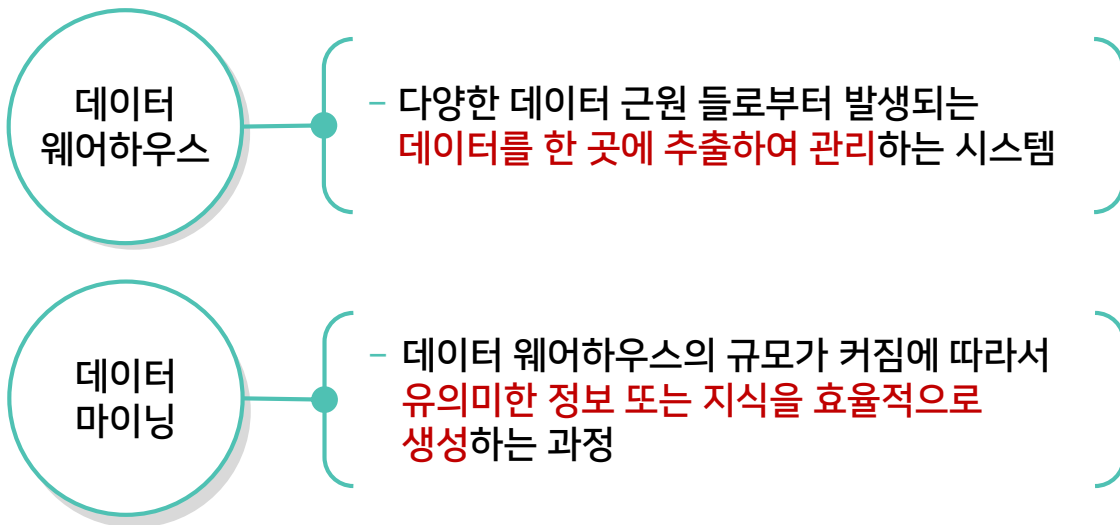


다양한 정보처리 시스템

01 정보처리 시스템의 유형


Ⅰ 의사결정 지원 시스템(DSS: Decision Support System)

- 여러 가지 복합적이고 광범위한 의사결정 목적으로 **통합적인 기능을 제공하는** 정보 시스템



데이터 마이닝

기저귀를 산 사람들이 왜 맥주도 함께 샀을까?

 + 

퇴근하는 남편들이 기저귀를 살 때 주말동안 마실 맥주도 함께 사는 경우가 많았음
→ **상품의 진열방식과 진열구역 선정에 반영하여 매출 증가**

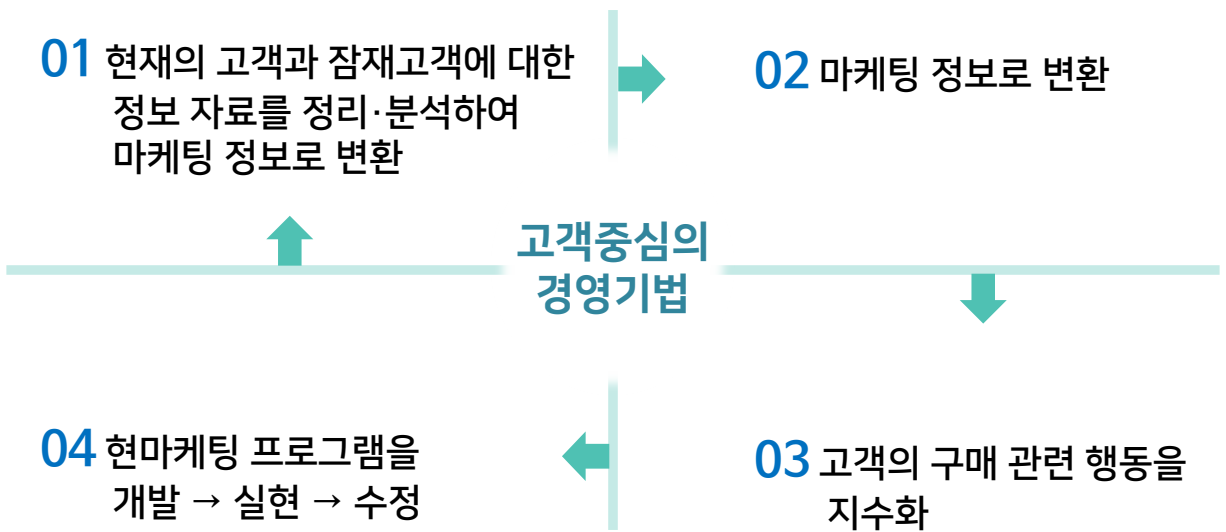


다양한 정보처리 시스템

01 정보처리 시스템의 유형

고객관리 시스템(CRM: Customer Relationship Management)

- 기업이 고객과 관련된 내·외부 자료를 분석·통합해 고객 중심 자원을 극대화하고 이를 토대로 **고객특성에 맞게 마케팅 활동을 계획·지원·평가**하는 시스템



- 고객관리 시스템의 특징

01 고객 성향이나 취향을 먼저 파악한 뒤 이를 토대로 고객이 원하는 제품을 만들고 마케팅 전략을 개발

▶ 기존고객의 관리에 초점을 맞춤

02 CRM은 고객들의 행동패턴, 소비패턴 등을 통해 고객들이 원하는 것을 알아내야 하는 경우가 많아 고도의 정보분석 기술을 필요

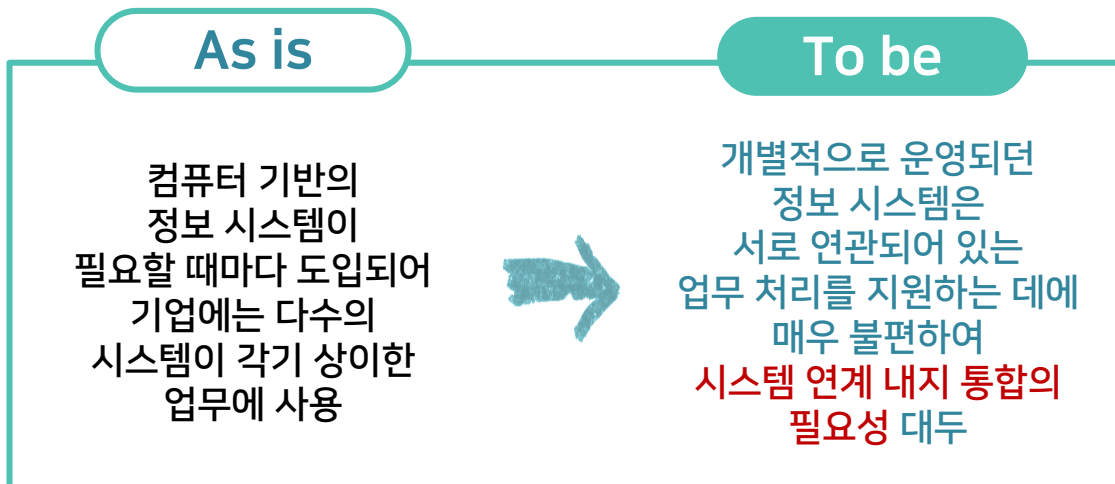


다양한 정보처리 시스템

01 정보처리 시스템의 유형

전사적 자원 관리 시스템(ERP: Enterprise Resource Planning)

- 전사적 자원 관리 시스템 구축의 배경



- 기업 내 생산, 물류, 재무, 회계, 영업과 구매, 재고 등 경영 활동 프로세스들을 **통합적으로 연계·관리**해 주며, 기업에서 발생하는 정보들을 서로 공유하고 **새로운 정보의 생성과 빠른 의사결정**을 도와주는 시스템



ICT 관점

여러 부서에서 수행되는 업무 처리를 돕기 위해 일상 업무 작업을
자동화한 응용시스템(패키지 소프트웨어)들의 집합

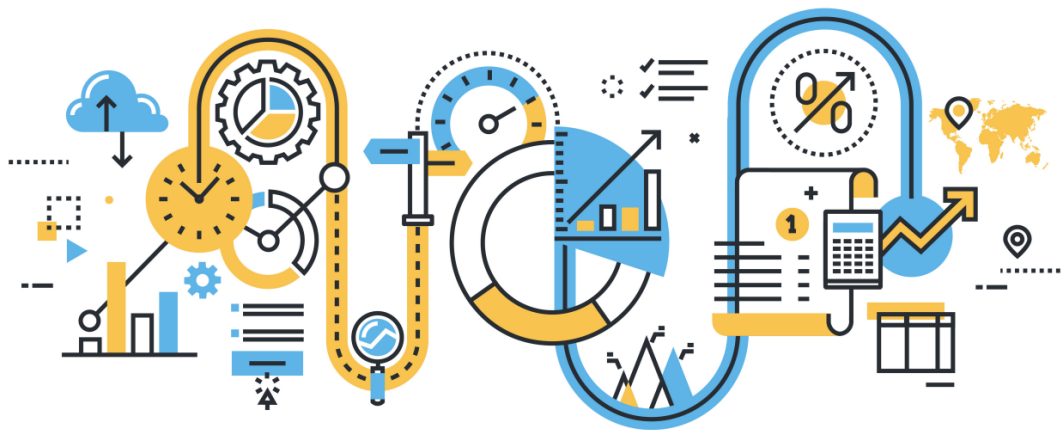


다양한 정보처리 시스템

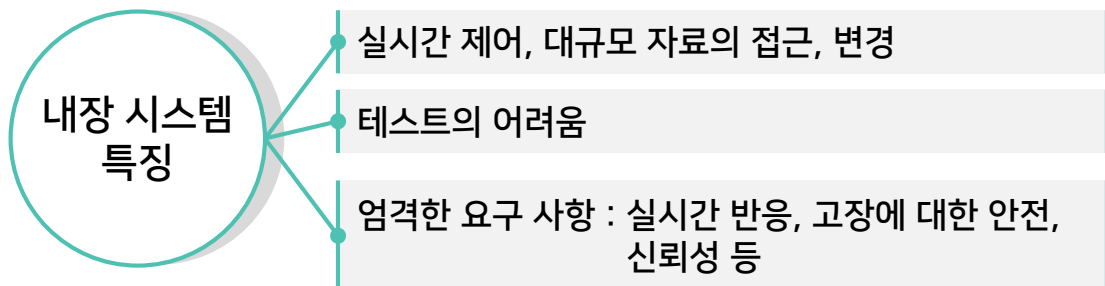
01 정보처리 시스템의 유형

기타 정보처리 시스템의 유형

- 제어 시스템
 - ▶ 사건을 감지하여 처리하고 자동적으로 보고
 - » 예 | 공정 제어, 수치 제어, 무기 제어, 항공 제어, 교통 제어



- 내장 시스템(Embedded System)
 - ▶ 데이터 처리가 주기능이 아닌 시스템의 구성 요소
 - » 예 | 비행기 유도, 레이타 추적 시스템, 스위칭 시스템 등





데이터베이스



데이터베이스의 정의



01 데이터베이스(Database: DB)의 정의

어느 한 조직의 다양한 응용 프로그램들이 공동으로 사용하는
데이터들을 통합하여 저장한 운영 데이터의 집합





데이터베이스의 정의

02 데이터베이스 정의에 함축된 개념

공용 데이터
(Shared Data)



- 한 조직의 여러 응용 프로그램을 공동으로 사용함
- 여러 사용자가 서로 다른 목적으로 공유

통합된 데이터
(Integrated Data)



- 여러 부서에서 사용하는 데이터를 한곳에 모아서 공동으로 관리함
- **최소한의 중복, 통제된 중복**

저장된 데이터
(Stored Data)



- 컴퓨터가 접근할 수 있는 디스크와 같은 저장 매체에 저장됨

운영 데이터
(Operational Data)



- 조직의 운영에 기본적으로 반드시 필요한 데이터를 저장함
- 조직의 고유한 기능을 수행하는데 필수적인 데이터를 저장함



데이터베이스



데이터베이스의 정의

02 데이터베이스 정의에 함축된 개념

| 통합된 데이터(Integrated Data)

- 한 조직 내에서 여러 부서가 **유사한 데이터 집합**을 사용

» 예 | 학사 관리 시스템



교무과

학생을 포함한 학교의
전반적 행정 사항 관리



학적과

학생의 주요 정보를 관리
(입학, 휴학, 복학, 자퇴 등)



학생과

학생의 장학금, 진학,
취업 등 지원



학생의 정보를 3개의 부서가 사용



데이터베이스의 정의

02 데이터베이스 정의에 함축된 개념

I 통합된 데이터(Integrated Data)

● 중복성 문제



각 부서가
독립적으로
학생 정보를 관리



학적과에서
휴학 중인 학생이
학생과에 의해
근로 장학생으로 선발



데이터의
일관성 문제 발생

Q

여러 부서가 사용하는 데이터를 통합하면?

- 중복성 제거 및 일관성 유지 가능

Q

각 부서가 따로 가지고 있으면?

- 검색 및 운영 시 편리

최소한의 중복과 통제된 중복이 필요함



데이터베이스



데이터베이스의 특징





데이터베이스



데이터베이스의 특징

01

실시간 접근성

- | 질의에 대한 **실시간 처리 및 응답** 가능함

* **실시간 처리** : 생성된 데이터를 컴퓨터에 보내어 그 처리 결과에 따라서 즉시 의사결정을 할 수 있는 것을 의미함

02

계속적인 변화

- | 한 시점에 데이터베이스에 저장되어 있는 내용을 데이터베이스 상태라고 한다면 시간의 흐름에 따라서 그 상태가 **지속적·동적으로 변동됨**

➡ 현실 세계를 표현하기 때문에 계속적으로 변화됨

03

동시 공유

- | 서로 다른 목적을 지닌 응용들의 **공용을 위하여 구축됨**

➡ 같은 내용을 여러 사용자가 동시에 공유할 수 있도록 지원

04

내용에 의한 참조

- | 데이터베이스에 저장된 내용은 데이터의 위치나 주소가 아닌 **데이터의 내용(즉, 값)에 따라서 검색됨**



데이터베이스의 특징

05 내용에 의한 참조와 위치에 의한 참조

프로그래밍 언어
: 위치에 의한 참조

데이터베이스 언어
: 내용에 의한 참조

- 검색방법 : 주소 검색
 - ▶ 특정 메모리 위치에 있는 값 제시
 - ▶ 메모리 주소
 - 0xFFFFFA18FF
 - 외우기에 어려움
 - ▶ 변수
 - 메모리 주소 대신 특정 이름(x, idx 등)으로 표시해 놓으면 기억하기 편리함

프로그래밍 언어
: 위치에 의한 참조

데이터베이스 언어
: 내용에 의한 참조

- 검색방법 : 내용 검색(SQL)
- » 예 | 학번이 100번인 학생의 이름을 검색하라.

```

- SELECT NAME
  FROM STUDENT
  WHERE STUDENT_NUM = 100

```



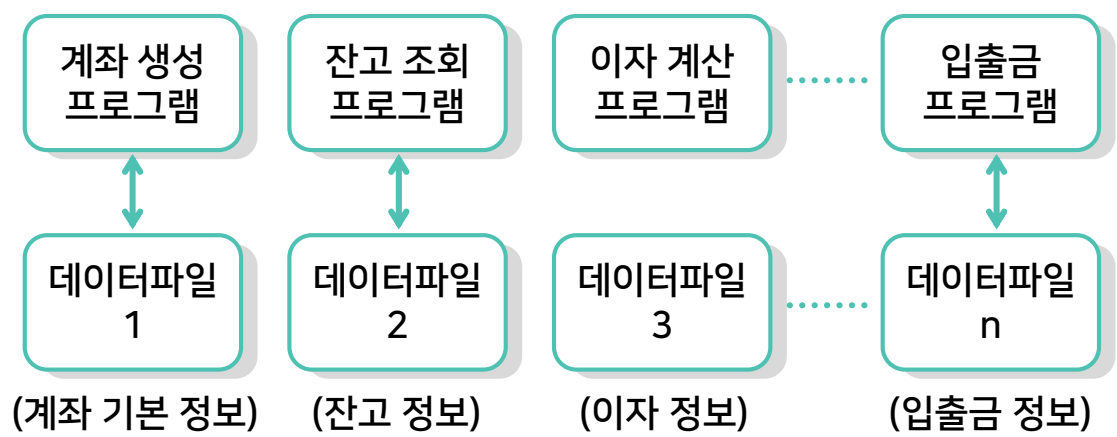
파일 시스템과 데이터베이스



파일 중심의 자료 처리의 특징

- | 각 응용 또는 기능마다 프로그램을 개발함
- | 각 프로그램은 자신이 사용할 데이터를 운영체제상의 파일로 관리함

» 예 | 은행 업무





파일 시스템과 데이터베이스

02 파일 중심의 자료 처리의 문제점

| 데이터의 중복과 비일관성(Inconsistency)

- 각 파일이 **다른 형식**을 지닐 가능성 존재
- 동일 정보가 여러 파일에 **중복 저장될 가능성 존재**
- **저장 공간의 낭비 및 비일관성 유발**

» 예 | 한 고객의 주소가 일반예금 정보와 적금예금 정보에서 상이하게 저장될 가능성

| 데이터 고립

- 데이터가 여러 파일에 흩어져 있고 파일 형식이 다르기 때문에 원하는 데이터를 검색할 수 있는 프로그램을 개발하기 어려움

| 무결성 문제

- **저장된 정보에 대한 제약 조건을 만족시키기 어려움**
- 새로운 제약 조건이 추가될 때마다, 모든 응용 프로그램에 이를 반영하는 새로운 코드를 추가하기 쉽지 않음

» 예 | 계좌 정보에서 3개월 동안 입출금 내역이 없으면 휴면계좌로 지정

| 원자성 문제

- **하나의 트랜잭션(논리적인 일의 단위)**은 이를 구성하는 내용이 전부 수행되거나 전부 수행되지 않거나 둘 중의 한 상태만을 만족해야 함

» 예 | 계좌 이체 - A통장에서 출금, B통장에 입금

- 기존 파일 시스템에서는 이를 보장하기 어려움



파일 시스템과 데이터베이스

02 파일 중심의 자료 처리의 문제점

| 동시 접근 문제

- 동일 시점에 동일한 데이터를 여러 사용자가 동시에 변경하는 경우 문제 발생
- 파일 공유 시스템에서 자주 발생됨

| 보안 문제

- 모든 사용자가 데이터베이스 상의 모든 내용을 접근할 필요 없음
- 보안 제약 사항을 각 응용 프로그램에만 추가해야 하므로 **보안성 유지가 어려움**

» 예 | 은행 직원이 고객의 계좌 정보를 마음대로 열람 불가



파일 시스템과 데이터베이스

03 파일 중심 처리의 가장 근원적인 문제

각 응용 파일마다 **자신이 사용할 데이터**를 파일로 관리

- 데이터 중복성 문제 발생 ➡ 데이터의 **비일관성**

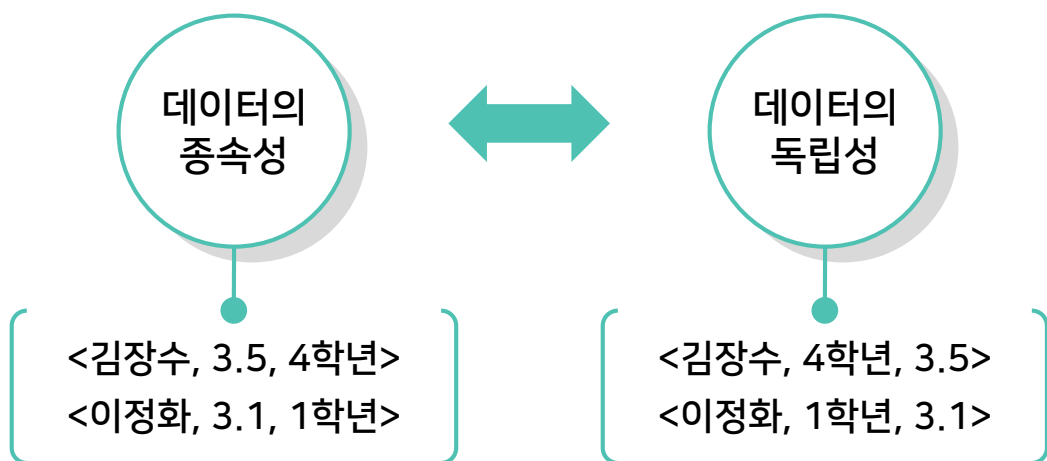
각 파일마다 **다른 형식**으로 내용을 구성

- **데이터 종속성(Data Dependency)** 문제 발생
- 파일 형태에 따라 다른 응용 프로그램 작성
- 파일 형태 변경 시 응용 프로그램도 변경됨

데이터 종속성(Data Dependency)

- 파일에 기록된 정보의 구조 변경 시 응용 프로그램 재작성 필요

➡ **위치에 의한 참조** 때문





데이터베이스



파일 시스템과 데이터베이스

04 데이터베이스의 장점

01 데이터 중복의 최소화

02 일관성 유지

03 데이터의 공유

04 무결성 유지

05 보안 보장

06 표준화 용이

07 전체 데이터 요구의 파악 조정



이와 같은 장점은 데이터베이스를 사용한다고 자동으로 얻어지는 것은 아님



파일 시스템과 데이터베이스

04 데이터베이스의 단점

01 운영비의 오버헤드

- ▶ 문제: DBMS는 **고가의 소프트웨어**이며 많은 컴퓨터 자원을 사용함
- ▶ 결과: DBMS 운영을 위해 **고사양의 컴퓨터** 필요

02 자료처리 방법의 복잡성

- ▶ 문제 : 다양한 유형의 데이터가 서로 연관되어 기록되어 있으므로 응용 프로그램은 여러 가지 제약사항 하에서 개발 및 수행됨
- ▶ 결과: **응용 프로그램의 구조나 로직이 복잡**해질 가능성 존재

03 백업과 회복의 어려움

- ▶ 문제: 다양한 사용자가 동시 공유로 데이터를 사용하고 있으므로 **장애 발생 시 정확한 사유나 상태 파악의 어려움**
- ▶ 결과: **정교한 백업 및 회복 기법**을 사용하게 됨

04 시스템의 취약성

- ▶ 문제: 통합된 시스템으로 **일부 장애가 전체 시스템을 정지**시킬 가능성 존재
- ▶ 결과: 시스템의 신뢰성과 가용성이 DBMS의 신뢰성과 가용성에 의존적이 됨

1 정보처리 시스템

- ✓ 데이터란 현실 세계에서 단순한 관찰이나 측정을 통해서 수집된 사실(Fact) 또는 값(Value)이 어떤 기준에 의해 정리되어 있는 것
- ✓ 정보란 어떤 상황에서 적절한 의사결정을 할 수 있도록 지원하는 지식으로, 컴퓨터 시스템과 같은 처리기를 통한 데이터의 유효한 해석이나 데이터 상호 간의 관계를 의미
- ✓ 정보처리란 사용자가 의사결정(Decision Making)을 함에 있어서 필요한 정보를 컴퓨터가 생성하기 위하여 데이터를 처리하는 작업
- ✓ 의사결정이란 정보를 바탕으로 상황에 따라 주어진 문제를 해결하는 중요한 지능 활동

2 데이터베이스

- ✓ 어느 한 조직의 다양한 응용 프로그램들이 공동으로 사용하는 데이터들을 통합하여 저장한 운영 데이터의 집합
- ✓ 데이터베이스의 특징
 - 실시간 접근성(Real-time accessibilities)
 - 계속적인 변화(Continuous evolution)
 - 동시 공유(Concurrent sharing)
 - 내용에 의한 참조(Content reference)