

7章 情報検索とデータベースの 指導法

情報の価値と信頼性

- 情報の**価値**
 - 必要な情報を、必要なときに、速やかに取り出すことができること
- 情報**検索**
 - 必要な情報を取り出してくる技術
- **データベース**
 - データを定式化し保管・管理する技術

情報の価値と信頼性

- 情報
 - 目に見えない
 - 固有の価値がある(無償ではない)
 - より正確であるほど価値が高い
- 例
 - 天気情報サービス
 - 気象庁、気象情報サービス会社から提供
 - 農作物のコントロール
 - 国勢調査
 - 数年に1回、全世帯対象
 - 正確な情報に基づいて行政を実施

情報の整理と検索

- 情報の**整理**
 - 必要な情報をすぐに見つけられる(検索できる)状態
- 分類
 - テーマごとに
 - 内容ごとに

↓

 - 順に調べる**範囲を限定**
 - **短時間**で検索できる
- 情報の整理には**労力やコスト**がかかる

情報の整理と検索

- 「超」整理法
 - 項目別に分類 → 維持が大変(労力に合わない)
 - 最近使ったものから順番に袋に入れて棚に並べる
- 検索したい情報
 - 比較的最近にアクセスした情報である確率が高い
 - テーマ別の分類 → 使用時点別の分類
- 分類基準の例
 - 複数の分類基準(書籍など)
 - テーマ別、著者別、出版年順

ディレクトリサービスと検索エンジン

- ディレクトリサービス

- 人がチェック
- 比較的精選されている
- ページ数はあまり多くない
- 作成者が申請しなかったページは登録されない

- 検索エンジン

- ロボットにより自動収集
- 高度な検索条件の指定
- 関係ないページも検索されてしまう

データベースとDBMS

- データベース(Database)
 - 収集するデータに対してある程度**決まった形や構造**を持っている
 - 計算機システムによる**管理・検索・内容更新**を効率よく行える
 - **範囲**の定まった、明確に管理されている
 - **汎用性**のあるデータの集合体である

データベースとDBMS

- データの独立性
- データの整合性
 - データベース管理システム(DBMS)
 - バックアップ
- スキーマ
 - データの種類や形式を決める情報
 - CODASYL3層アーキテクチャ
 - 概念スキーマ、内部スキーマ、外部スキーマ

関係モデルと関係データベース

- RDB(Relational DataBase)
 - 表 → 関係(Relation)
 - 表に属する個々のデータ → 組(Tuple)
(または Tuple)
 - 列に対応するデータ → 属性(Attribute)
- 関係データベースの例(図7.2)
 - 社員データ
 - 取引先
 - 担当

表 関係 (Relation)

社員データ

社員番号	氏名	年齢
u182	田中太郎	24
u241	鈴木由佳	33
u007	中原祐二	28

取引先

取引先番号	社名	年商額
t0142	A商事	300000000
t1988	C興産	120000000
t2401	B電業	90000000

表に属する個々のデータ 組(Tuple)

社員データ

社員番号	氏名	年齢
u182	田中太郎	24
u241	鈴木由佳	33
u007	中原祐二	28

取引先

取引先番号	社名	年商額
t0142	A商事	300000000
t1988	C興産	120000000
t2401	B電業	90000000

列に対するデータ 属性(Attribute)

社員データ

社員番号	氏名	年齢
u182	田中太郎	24
u241	鈴木由佳	33
u007	中原祐二	28

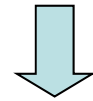
取引先

取引先番号	社名	年商額
t0142	A商事	300000000
t1988	C興産	120000000
t2401	B電業	90000000

第一正規形

- 項目には複数の値を入れることはできない

社員番号	氏名	社名	年齢
u182	田中太郎	B電業, C興産	24
u241	鈴木由佳	A商事	33



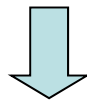
第一正規化

社員番号	氏名	社名	年齢
u182	田中太郎	B電業	24
u182	田中太郎	C興産	24
u241	鈴木由佳	A商事	30

第二正規形

- 主キーの一部だけで一意に指定可能な列を別のテーブルとして分離

社員番号	氏名	社名	年齢
u182	田中太郎	B電業	24
u182	田中太郎	C興産	24
u241	鈴木由佳	A商事	30



第二正規化

社員番号	氏名	年齢
u182	田中太郎	24
u241	鈴木由佳	30

第三正規形でない例

社員データ

社員番号	氏名	年齢
u182	田中太郎	24
u241	鈴木由佳	33
u007	中原祐二	28

← 退職

取引先担当

社員番号	取引先番号	社名	年商額
u182	t2401	B電業	90000000
u007	t1988	C興産	120000000
u241	t0142	A商事	300000000

→ 取引先のデータ
が無くなる

第三正規形

- データ項目同士の**従属関係の分離**

社員番号	取引先番号	社名	年商額
u182	t2401	B電業	90000000
u007	t1988	C興産	120000000
u241	t0142	A商事	300000000



第三正規化



社員番号	取引先番号
u182	t2401
u007	t1988
u241	t0142

取引先番号	社名	年商額
t0142	A商事	300000000
t1988	C興産	120000000
t2401	B電業	90000000

関係データベースの例

社員データ

社員番号	氏名	年齢
u182	田中太郎	24
u241	鈴木由佳	33
u007	中原祐二	28

取引先

取引先番号	社名	年商額
t0142	A商事	300000000
t1988	C興産	120000000
t2401	B電業	90000000

担当

社員番号	取引先番号
u182	t2401
u007	t1988
u241	t0142

問合せ例(検索)

- SQLによる記述

- 年齢が40以上の社員の社員番号と年齢

```
select 社員.社員番号, 社員.年齢 from 社員  
where 社員.年齢 >= 40
```

- 年商が一億以上の取引先を担当している社員の氏名

```
select 社員.氏名 = 社員,担当,取引先  
where [ 社員.社員番号= 担当.社員番号 and  
       担当.取引先番号 = 取引先.取引先番号 and  
       取引先.年商額 >= 100,000,000 ]
```