

### 3 データベースの種類

関係(関係型, リレーショナル型)データベース(RDB)

- 蓄積したデータを複数の表で整理して管理する。
- データをフィールドとよばれる複数の項目で表現。
- データの集合をテーブルとよばれる表で表す。
- 1行に1件分のデータを記録。
- 追加・削除などは行単位。
- ID番号や名前等のキーとなる項目を利用してデータの検索や抽出。

表(リレーション) ← テーブル (属性) ← フィールド (属性) ← タブル (レコード) ← 属性名

図書番号	書名	著者名	出版社	分類記号
100070408	パスワード忍びの里	松原秀行	A社	913
100020411	福島の少年	あさのあつこ	B社	913
100108216	そのときは彼によろしく	市川拓司	C社	913
100018233	晩夏のプレイボール	あさのあつこ	D社	913

図2 関係データベース  
右図の名称は覚えること。

### RDBのデータ処理(関係演算の例1)

関連: pp.120-121

**選択 (selection)** 1つの表において、指定した条件を満たす**レコード(行)**を取り出す処理。  
例 ジャンル ⇒ 文化

**射影 (projection)** 1つの表において、ある条件を満たす**フィールド(列)**を取り出す処理。  
例 記事タイトル, 日時

No	記事タイトル	記事本文	ジャンル	地方	日時
1	関東地方の経済	本文...	経済	関東	12月1日
2	関西地方の文化	本文...	文化	関西	12月1日
3	東北地方の文化	本文...	文化	東北	12月15日
4	東北の政治	本文...	政治	東北	1月5日

### RDBのデータ処理(関係演算の例2)

関連: pp.120-121

**結合 (join)** 複数の表において、それぞれの表の特定のフィールドに着目して表を結合させ、新たな表を作成する処理。  
※フィールドの**関連付け**を**リレーションシップ**と呼ぶ

地方ID	地方	ジャンルID	ジャンル	記事ID	記事タイトル	記事本文	ジャンルID	地方ID	日時
A	関東	10	経済	1	関東地方の経済	本文...	10	A	12月1日
B	関西	20	文化	2	関西地方の文化	本文...	20	B	12月1日
C	東北	30	政治	3	東北地方の文化	本文...	20	C	12月15日
				4	東北の政治	本文...	30	C	1月5日

No	記事タイトル	記事本文	ジャンル	地方	日時
1	関東地方の経済	本文...	経済	関東	12月1日
2	関西地方の文化	本文...	文化	関西	12月1日
3	東北地方の文化	本文...	文化	東北	12月15日
4	東北の政治	本文...	政治	東北	1月5日

※表内の赤字で示したフィールドは、主キーの役割を果たす。

### 正規化(第一正規形)

関連: pp.120-121

**正規化** 構造を見直し、データの重複を排除して作りなおすこと (データを効率よく管理できるように複数の表に分割する)

よく行われる作業: 第一正規形・第二正規形・第三正規形

**第一正規形: ひとつの行・列に、ひとつの値を入れる。**

生徒氏名	貸出日	情報とは	図書館
神戸太郎	12月22日	情報とは	ハッカー
御影花子	12月22日	肖像権	

生徒氏名	貸出日	図書館
神戸太郎	12月22日	情報とは
御影花子	12月22日	肖像権

### 正規化(第二・第三正規形)

関連: pp.120-121

**第二正規形: 異なる従属の関係を確認して、表を分割する。**

レシートID	商品番号	商品名	単価	個数
abc111	1234	神戸牛乳	120	5
abc111	1278	兵庫茶	100	2
abc112	1234	神戸牛乳	120	3

ID	レシートID	商品番号	個数
abc111-1	abc111	1234	5
abc111-2	abc111	1278	2
abc112-1	abc112	1234	3

商品番号	商品名	単価
1234	神戸牛乳	120
1278	兵庫茶	100

**第三正規形: 従属の推移を確認して、表を分割する。**

氏名	部活動	出身中学	出身中学
22110	神戸太郎	野球部	精道中
22305	御影花子	柔道部	住吉中

氏名	部活動	出身中学	出身中学
神戸太郎	野球部	精道中	芦屋市
御影花子	柔道部	住吉中	神戸市

中学ID	出身中学	出身中学
15	精道中	芦屋市
20	住吉中	神戸市

※赤字は主キー。青字は主キーにならず重複の可能性。

### データベース管理システム(蓄積から活用へ)

**データベース管理システム(DBMS)**

- データベースを効果的に運用管理するためのシステム

**データベース管理システムの機能**

- 問い合わせ(クエリ)処理  
ユーザーの要求に応じてデータを返す機能。
- データ表現と記憶装置の管理(データ資源管理機能)  
データを共有するための記録方法・操作の共通化。
- データの整合性(整合性制約機能)  
複数の処理要求に対応。
- アクセス制御(セキュリティ機能)  
データの機密保護と安全確保。
- 障害の防止や回復(トランザクション管理機能)  
障害発生時にDBを復旧。  
排他制御・ロールバック処理等。

データの蓄積 → 情報の活用

データベースソフトウェア

データベースサーバ

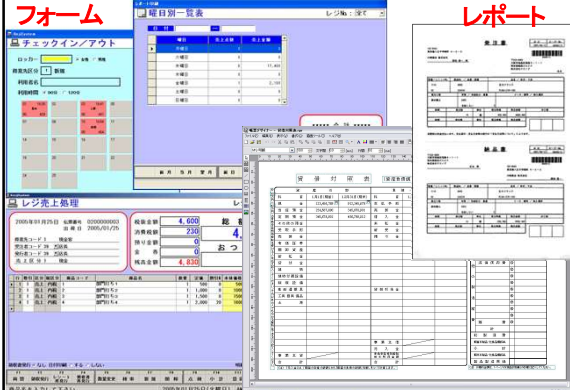
クライアント: ウェブブラウザ

サーバ: ウェブサーバ, データベース管理システム, データベース

### RDBと表計算ソフトとの違い、ファイル形式

- 表計算**:ワークシート(表)の中の**個々のデータ**に対して直接編集を加えて、必要な形にする。  
⇒ その時点で最もふさわしい形式に情報を整理しやすい。
- ※ エクセルにおける「フィルタ」や「並び替え」、VLOOKUP関数等は、シートの一部をRDBと見なした機能である。
- RDB**:**表全体にルールを適用**して格納するデータを統一する。  
⇒ 必要なときに**必要な形**で情報を取り出しやすい。
- ※ RDBでは、**複数の表**に格納されたデータから**必要な部分を取り出して、使いやすいフォーム**を設計して利用する。
- ファイル形式**:他のソフトウェアとのデータのやりとりには、**CSV形式**(データをカンマで区切る)や**XML**(タグでデータ構造を定義)が利用される場合が多い。

### データベースソフトを見る フォーム・レポート



### データの種類と尺度

関連: pp.122-123

- 入手したデータは、**質的データ**(性別、血液型、成績順位等)と**量的データ**(温度、身長、売上金額等)の、2種類に分類できる。
- 質的データは「**名義尺度**」と「**順序尺度**」に、量的データは「**間隔尺度**」と「**比率尺度**(比尺度、比例尺度)」に、分類できる。

分類	定義(基準)	例
質的データ	<b>名義尺度</b> 対象の属性の「分類」。記号や数値は区別のためのためなので、数字計算(加減乗除等)は無意味。	組、性別、血液型、商品分類等
	<b>順序尺度</b> 対象間の「順序」。データ比較は可能だが、順番以外に意味はなく計算不可(中央値は使用有)。	成績や好みの順位、満足度等
量的データ	<b>間隔尺度</b> 目盛が「等間隔」。尺度間の計算(和差)は可能。絶対的な原点がないので、比(乗除)は無意味。	時刻、日付、温度、知能指数等
	<b>比率尺度</b> 絶対的な「原点が存在」する尺度。間隔に加えて比率にも意味があり、四則演算全てが可能。	体重、時間、速度、売上金額等

### 情報収集の目的・手段

関連: pp.219,223

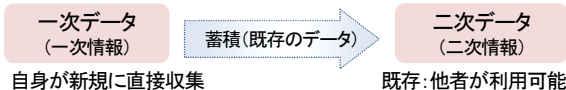
#### 情報収集の目的

- 問題を明確化して、問題の原因や解決方法を調べる
- 仮説を検証する

情報の見極めが大切

#### 情報収集の手段による分類

- **一次データ(一次情報)**:ある目的のために、新規に直接収集したデータ
- **二次データ(二次情報)**:既に蓄積された、既存のデータ(外部データ)
- **オープンデータ**:機械判読(機械可読)に適したデータ形式で公開された、二次利用を可能とするデータ。



### 調査活動(一次データの収集例)

関連: pp.122,163

#### インタビュー調査: 少ない人数の対象者と直接会話する方法

- 直接対面するため、表情などの情報が得られる。
- 臨機応変に質問内容を変更できる。
- 他の出席者の話が刺激となり発言しやすくなる。

#### アンケート調査: 質問をまとめて提示し、回答を得る方法

- 調べたい事柄について、数値としてデータを得られる。

#### アンケート調査の種類

- **全数調査**: 対象となるすべての人を調査する。
- **標本調査**: 一部の対象者を無作為に決めて調査する。

### 収集した情報の検証(評価・整理・分類・加工・分析)

#### 情報源 check

- 必ず情報源を確認(記録)する
- 書き手(発信側)の**立場**を意識する
- 情報が発信された**時期**を確認する

#### 内容の吟味

- 発信側の意見か憶測か、事実か?
- 大げさな表現等になっていないか?
- あいまいなところ、欠けているところは?
- 論理的におかしくないか?

#### 内容の分析

- 比較: **クロスチェック**(複数の情報源から確認)
- 分析方法: **定量的分析**(量的データ)  
**定性的分析**(質的データ)
- 表現: 表・図解・統計・グラフ・シミュレーション等