


帝塚山学院大学
 TEZUKAYAMA GAKUIN UNIVERSITY

データベース概論

GoogleのDB,ビッグデータ

中野秀男
情報メディア学科

1 データベース概論 Google, Big Data 2018/1/29

今日の話

- ▶ 質問やコメントの回答
- ▶ ガーグルのDB
- ▶ ビッグデータ
 - ▶ ビッグデータ
 - ▶ ビッグデータの例
 - ▶ ビッグデータの利用方法

2 データベース概論 Google, Big Data 2018/1/29
 
帝塚山学院大学
 TEZUKAYAMA GAKUIN UNIVERSITY

Googleのデータベース(1)

- ▶ 世界中のウェブを自動巡回プログラム(ロボット)で収集
- ▶ 安価なコンピュータや補助記憶でDBを構成
- ▶ 列(フィールド)の単位で整理せずに、行の単位で整理
- ▶ 複数の行(文書)を複数のコンピュータに保存
- ▶ 検索要求(クエリ)はありそうな複数のコンピュータに要求
- ▶ 多くの同時にくる大量のクエリに平均的に対応
- ▶ シャーディング: データの断片化
- ▶ レプリケーション: 複製
- ▶ ページランク

3 データベース概論 Google, Big Data 2018/1/29
 
帝塚山学院大学
 TEZUKAYAMA GAKUIN UNIVERSITY

Googleのデータベース(2)

▶ Googleのデータベース説明の図が入ります。

▶ 4 データベース概論 Google.Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

Googleのデータベース(3)

▶ Googleのデータベース説明の図が入ります。

▶ 5 データベース概論 Google.Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

ビッグデータ (Wikipedia)

▶ 通常のデータベースでは扱えない量

▶ 構造化データや非構造化データ

▶ ペタ、エクサ、ゼタ、更に上

▶ K, M, G, T, P, E, Z

▶ 格納や活用(検索/共有/分析/可視化)

▶ いろいろな分野で先端技術を使う

▶ 6 データベース概論 Google.Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

ビッグデータの例

- ▶ 政府や自治体を持つ住民や組織情報
- ▶ 家電やシステムの操作ログ
- ▶ ライフログ
- ▶ センサーネット情報
- ▶ 顧客データ
 - ▶ 政府、自治体、企業、NPO
- ▶ ソーシャルメディアデータ
- ▶ マルチメディアデータ
- ▶ ウェブサイトデータ

7 データベース概論 Google.Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

ログ

- ▶ 操作ログ
 - ▶ 家電メーカーの思惑
 - ▶ B-CAS問題
 - ▶ フォレンジック(PC操作の記録)
- ▶ ライフログ
 - ▶ 歩数等の活動の記録
 - ▶ 読んだもの等の知識の記録
 - ▶ 写真等の記憶の記録
 - ▶ 体重や脳波等の身体の記録
 - ▶ レシート等の身体の記録
 - ▶ 位置情報等の環境の記録
 - ▶ 気分等の精神の記録

8 データベース概論 Google.Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

ビッグデータのステップ

- ▶ 多様な複数のデータを
- ▶ 総合的に収集・分析することで
- ▶ 可視化する/新たな知見を得る
- ▶ プラットフォームの構築(高速、大容量)
- ▶ サードパーティの参加(複数で多く)

9 データベース概論 Google.Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

時代の背景

- ▶ コンテンツ:
 - ▶ ソーシャルメディア: FBは300億/月
 - ▶ 動画配信: Youtube
- ▶ プラットフォーム
 - ▶ クラウド、Android OS
- ▶ ネットワーク
 - ▶ ワイヤレス通信、ブロードバンド
- ▶ デバイス
 - ▶ 各種センサーの普及、IoTの時代
 - ▶ ICT端末の多様化: 携帯, PC, TV, スマホ

10 データベース概論 Google, Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

仮説を検証

- ▶ 仮説を立て
- ▶ 複数の大量のデータを集め
 - ▶ 大容量/分散ストレージ
- ▶ データマイニング等の手法で推論(?)
 - ▶ 高速/並列計算、アルゴリズム
- ▶ 例
 - ▶ 紙オムツを買った人は缶ビールも買う
 - ▶ 冬にスキー用品を買った人は、夏にテニス用品を

11 データベース概論 Google, Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

質問やコメント(1)

- ▶ 第一志望のゼミに入れてよかった
- ▶ 後期にするミニゼミ
- ▶ これから先、Accessを使って授業をすることはあるか
 - ▶ 3回生や4回生で使うか
- ▶ 1,2回休んだ程度だとレポートを書いたほうがいいのか
- ▶ レポートを書く人の指定が
- ▶ 2月9日のVODは10月23日の枠に書けばいいか
- ▶ 課題でクエリやSQLの設定は必要か
- ▶ 関数がAccessで使えないのは文字や数値で正確に情報を記録するためか
- ▶ AccessはExcelでやっているテーブルみたいで、Accessでやる意味はあるか
- ▶ Writerというアプリ

12 データベース概論 Google, Big Data 2018/1/29  帝塚山学院大学
TEZUKAYAMA GAKUIN UNIVERSITY

質問やコメント(2)

- ▶ リッチテキストとプレーンテキストの長所や短所
- ▶ マルチメディアが少しわかった
- ▶ マルチメディアに主キーは設定できるか
- ▶ マルチメディアはあらゆる情報をひとまとめにできるメディア
- ▶ 映画の情報が3GBとは実感がわからない
 - ▶ わかりやすい例は
- ▶ ほとんどのゲームでマルチメディアとDBが使われている
- ▶ 「身代わり」もデータベースの一つか
- ▶ メタデータはマルチメディアのデータの何を説明しているか

▶ 13

データベース概論 Google.Big Data 2018/1/29



質問やコメント(3)

- ▶ メタ:新しい言葉を覚えた
- ▶ メタ:新しい言葉を覚えた
- ▶ メタの意味がポケモンのメタモンのたとえでわかった
 - ▶ 実際にメタはどのように利用されているか
- ▶ メタモンの由来が意外
- ▶ 附带情報をメタデータという
- ▶ メタの意味がポケモンのメタモンのたとえでわかった
 - ▶ 実際にメタはどのように利用されているか
- ▶ メタモンの由来が意外
- ▶ 附带情報をメタデータという

▶ 14

データベース概論 Google.Big Data 2018/1/29



質問やコメント(4)

- ▶ テーブルの書き方
- ▶ DBをみてみてわかるようになるのは発達したことか
- ▶ テーブル設計マスターしたい
- ▶ ゲームのDBは面白いが実際は
- ▶ ポケモンは複雑なDBを使っている
- ▶ データの構造が難しい
- ▶ 標準とは
- ▶ どうしてLotus1-2-3という名前
- ▶ TwitterやFacebookに写真をあげると撮った場所が
- ▶ ゲームの課金制度
- ▶ ホームページは毎日更新されている

▶ 15

データベース概論 Google.Big Data 2018/1/29



質問やコメント(5)

- ▶ 今の時代、簡単にメアドが作れる
- ▶ Appleが7より前の性能を落としていると聞いたが
- ▶ iPhone Xの顔認証と使ってアニメキャラの動画配信ができるアプリができるみたい。MSのKinectみたい
- ▶ iPhone Xでホームボタンをつける付属品が
 - ▶ 便利なものを追求して全てなくすのは良くない
- ▶ 佐藤先生のThe Labの展示。人が多かった
- ▶ AmazonGOは当然、日本に来ない

質問やコメント(6)

- ▶ Googleが訴えられたらどうするか
- ▶ 税理士さんも伝票渡されて大変
- ▶ 全ての企業がDBを使っていない
- ▶ 小遣い帳の1200円は高いが、細かく記入すると日記に
- ▶ 中野のiXpensItは便利そう
- ▶ 中野はGoogle Homeを持っているか
- ▶ オラクルと任天堂のDB
- ▶ 任天堂がコロプラ社を起訴
- ▶ AI同士が話をする

質問やコメント(7)

- ▶ ポケ森はDSのように通信できるようにならないか
- ▶ どうぶつの森は毎日やっているのか
 - ▶ クラフトの材料が多くて羨ましい
- ▶ ポケ森、周りがやらなくなった
 - ▶ 移動範囲が狭い。DSの方と比べてしまう
 - ▶ アップデートで増えればいいのに
 - ▶ ポケモンGOよりは面白かった
- ▶ ポケ森、楽しいか
