

# コンピュータ演習

AIリテラシー 03講 社会でどのようなデータが活用されているか

# 目次

- 第03講 社会でどのようなデータが活用されているか
  - 人の動線をめぐるデータ
  - 多くの機器のログとオープンデータ
  - 1次データ, 2次データ, メタデータ
    - 1次データ
    - 2次データ
    - HTML
    - メタデータの有用性
  - 非構造化データの増大

# 第03講 社会でどのようなデータが活用されているか

# 人の動線をめぐるデータ

## データ

何かが起これば必ずデータが発生する

- 風がふく
- 事故が起きる

データが今ほど記録されて、公開されている時代は初めて。

## 人のデータ

たくさんのデータが収集されて公開されると、それを応用したさまざまなサービスや製品が登場します。

特に注目されているのが人のデータです。

## カーナビの例

1. 現在地と目的地までのルートを調べるシステム
2. どの道が混んでいるか、目的地まで何分かかりそうかを導く  
総合情報システムへ
3. 災害時には、道が利用できるかを判断することができるように
4. 避難やロジスティクス(物流)に役立つ
  - 東日本大震災でのインターナビによる取り組み「通行実績情報マップ」が「2011年度グッドデザイン大賞」を受賞
  - 通れた道MAP

## 監視カメラの例

1. 監視カメラは安全管理・防犯目的で導入
2. お客様の回遊情報を分析可能に
3. 商品の販売促進や企業の競争戦略の教科にまで活用される

プライバシーには気をつけなくてはならないようですね。

- [GREENAGES Citywide Surveillance](#)





## 多くの機器のログとオープンデータ

## SNSのデータ分析

みなさんは個人としてSNSを利用する分には炎上だけ気をつけてくれれば良いですが、企業等にとっては、どの投稿が、いつ、どの位見て貰えたか、ということが利益につながる情報となります。

## Twitterの場合

- インプレッション (ユーザーがツイートを見た回数)
- エンゲージメント総数(ユーザーがツイートを見た上でいいねやリツイートなどの行動をした回数)
- リンクのクリック数
- いいね
- リツイート

Twitter分析ツール10選！ 無料ツールを使ってアカウント分析してみた結果

自分のアカウントでなくトレンドを見るにはTwittrendのツールもあります。

## オープンデータの活用

- 他の企業や個人、行政機関にもデータを使ってもらおうとする**オープンデータ**も取り組みが始まっている
- 使いやすいフォーマットでないといけないため、  
CSV(Comma Separate Values)や  
エクセル(セル結合なし)等が利用されている

## ログによるアクセスデータ

- 多くの機器はログ(動作記録)を取得・保存している
- Macの動作記録はコンソール.appで確認可能
- NASも接続情報などを保存している

## NTP

なぜここ...

Macでは、コントロールパネルの日付と時刻でNTP(Network Time Protocol)サーバを指定しています。

# 1次データ， 2次データ， メタデータ





# 1次データと2次データ

## 1次データ

事象や出来事の情報直接記録したもの

## 2次データ

1次データをまとめてみやすくしたり、何かの目的のために1次データを取捨選択して説明を行ったりしたもの

一般的には1次データを参照するべきと言われることが多いが、1次データの正確性についての判断力を求められる。鵜呑みにせず、学者によって書かれ、学術的な出版社によって出版された二次資料は信頼できると考えられるため、使い分けが必要。

## メタデータ

データを説明するデータのこと

セリフに対して感情などを記載したト書はメタデータとなる。

# HTML

文字情報に意味づけをしてホームページを作成するための言語

```
<!DOCTYPE html>
<html>
  <head>
    <title>タイトル</title>
  </head>
  <body>
    <h1>見出し</h1>
    <p>段落の文章...段落の文章...段落の文章...
    段落の文章...段落の文章...段落の文章...
    段落の文章...段落の文章...</p>
  </body>
</html>
```

## メタデータの有用性

メタデータが充実されていると、自動処理しやすくなる

50,40,60

だとよくわからないが

50(英語の点数),40(数学の点数),60(国語の点数)

と書いてあれば、後の処理はやりやすくなりますね。



# 非構造化データの増大

## データベースとデータの構造化

**データベース**というデータを扱う専門のアプリがありこれまでデータをデータベースに記録してきました。

事前定義された、関連があるデータ項目の集合体である**リレーショナルデータベース**に保存してきました。

Twitterであれば

- ユーザ/投稿/いいね

などを別々にデータベースとして保存し、それぞれ関連させることによってシステムができています。

データベース



## 非構造化データ

多くのデータを扱えるようになったときに、構造化されたデータのみならず、構造化されていないデータも活用する方向に動いています。



[https://solution.toppan.co.jp/digital/contents/cdp\\_contents07\\_0831.html](https://solution.toppan.co.jp/digital/contents/cdp_contents07_0831.html)

## ビッグデータとは

- 構造化データ
- 非構造化データ

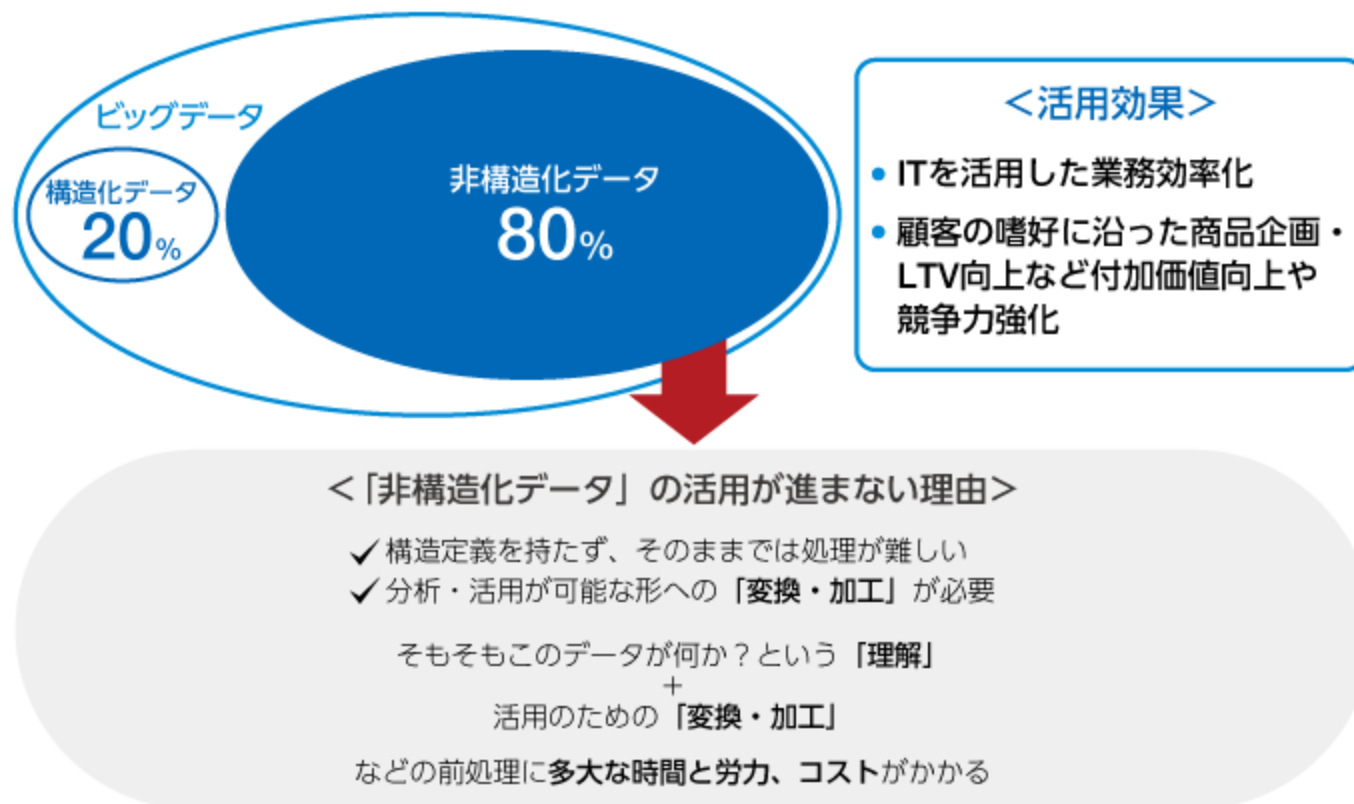
これらを総称してビッグデータと言えるでしょう。

## 非構造化データの活用

非構造化データは構造定義をもっていないため、そのままでは処理が非常に難しいです。

非構造化データを分析・活用するためには「変換・加工」が必要となります。

## 非構造化データの活用



[https://solution.toppan.co.jp/digital/contents/cdp\\_contents07\\_0831.html](https://solution.toppan.co.jp/digital/contents/cdp_contents07_0831.html)