

◇ 教科書 p.14～p.43

● データサイエンスと情報倫理

- ・ = 個人や家庭内の私事・私生活。個人の秘密。他人から興味本位で見られたり干渉されたりすることがなく安心して過ごすことができる権利。
- ・ = 生存する個人に関する情報で、特定の個人を識別することができるもの。
(例) 「氏名」・「住所」・「生年月日」・「性別」・「マイナンバー」・「人の容姿」など。

* 扱う上での注意点

- ① 取得・利用 : 目的を特定して通知・公表し、その範囲内で利用
- ② 保管 : 漏えい等が生じないように、安全管理
- ③ 提供 : 第三者への提供は、あらかじめ本人から同意
- ④ 開示請求への対応 : 本人からの請求に対応



情報には、オリジナルの作成者（著作権者）が存在する。

→ その著作物に関する著作権は尊重しなければならない。

* 個人情報や肖像（顔や容姿）は、本人の許可なく使ってはならない。
使用が認められたとしても、以下のことには気をつける必要がある。

- ① 当初の目的と異なる形での利用 ()
- ② 正当な理由のない状態での利用 ()

- ・ = 誰でも許可されたルールの範囲内で自由に複製・加工・頒布ができるデータ
(例) e-Stat (イースタット) … 政府統計の総合窓口。都道府県別人口など、扱うデータは多岐に渡る。

RESAS (リーサス) … 地域経済分析のためのウェブシステム。

- ・ = クリエイティブ・コモンズ・ライセンス
→ これらの情報は比較的自由に使うことができる。

ただし、情報に記されている原作者に関するクレジット（名前、作品名、出典など）の明記といった規約は守らなければならない。



◎ データの作成、収集、分析の際に注意すること

- [1]
本物らしく見せかけた偽物の情報（フェイクニュース）の蔓延は、近年の社会問題となっている。
公式情報を見に行くなど、情報源の確認を心がけたい。
- [2]
統計モデルの品質や結果の解釈を確かなものにするために、そのモデルのもとになるデータに関する十分な理解が求められる。データの中にバイアス（偏り）が潜んでいないか、注意をする必要がある。
- [3] ・ … データの値が欠けている状態
 ・ … ありえないデータの値
 ・ … 他の値から大きく離れたデータの値
 → データの不整合性に対する対処 =

【課題】 e-Stat を用いて、現在の日本の総人口を調べてみよう。 答： 約 1 億 () 万人