

# データは誰のものか？



NIHU

佐藤 洋 一 郎

(人間文化研究機構・理事／  
総合情報発信センター長)

学問分野(※)と1次資料の多様性(暫定)※新しい学術の体系(日本学術会議)

|           | 認識科学    | プログラム科学  | 設計科学<br>人文学 |
|-----------|---------|----------|-------------|
| 学問の性格     | あるものの探求 |          | あるべきものの探求   |
| 探求するもの    | 法則の探求   | プログラムの探求 | 価値の探求       |
|           | 現象の解明   | 予測       | 意思決定        |
| 背景にある現象   | 自然現象    | 自然・社会現象  | 人間文化        |
| コンテンツの多様性 | やや一様    | 比較的多様    | 多様          |
| コンテンツの客観性 | 客観的(※※) |          | 主観的／価値含意    |
| データの所有者   | 公共      | 公共／個人    | 個人／公共       |

※※ 観察者からまったく独立な真に客観的データなどというものはない

# 人文系データ(1次資料)の特質

コンテンツの多様さ(絵画資料なども)

☞ コンテンツによっては異分野が関心を持つことも

作出時期の多様さ

☞ 4000年に達するものもある(日本では1000年)

誰もが知る古典籍・・・公開という点では意外と先行的

可視化にはときに「翻訳」が必要

☞ 多言語、ないし変形文字

権利関係が輻輳

☞ 固い所有者が存在. 権利は付随しているのではなくて、データの特性をなす

フォーマットがない

☞ 文字から絵画まで

# 1次資料を公開する際の問題点

---

- ☞ 「わが家の蔵には大量の文書があり研究に使ってもらってよいが、何が書いてあるのかわからない」
- ☞ 「所有の絵画を公表すれば税務上問題が生じる」
- ☞ 「某国でのインタビューに際し、発言者の氏名や出自などを公開しないと約束している」
- ☞ ゲノム情報はもっとセンシティブ

# オープンデータと1次資料の権利関係

## 1次資料の所有権とオープンデータ

「所有者がいたからこそ今に伝わった」データもある

## オープンデータは安全か？

データや1次資料をだれが管理するのか？

1次資料の質保証(「コモンズの悲劇(※)」に通底)

公的機関が資料保全にいつも前向きとは限らない

維持・管理の責任は誰が負う？

組織的隠ぺいがないと言い切れる？

※ Hardin, Garrett(1968) The Tragedy of the Commons, Science 162: 1243-1248

[doi:10.1126/science.162:1243](https://doi.org/10.1126/science.162:1243)

# データ公開と社会問題

---

自然科学的なデータでも問題が起き得る

 **ハザードマップ**

「公開されたら地価が下がる！」

 **個人のゲノム情報**

病歴データと対で公開されたら、さあ大変！

 **希少動植物種の所在**

不心得者に位置情報を与えてしまう➡絶滅を誘発！

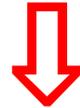
# データを生かすために

---

問題のデータ(※)を選別

問題のないデータはオープン化を推進  
(ただし「コモンズの悲劇」に注意)

問題のデータは適切に保護



国／機関の責任で実施  
実施体制の構築

※個人情報漏洩／諸権利の侵犯につながるようなデータ

# 結局、資料やデータは誰のものなのか？

オープンデータ化で研究はどのように進むか？  
オープン化は最終目的ではなく手段

研究はAIにとってかわられない？  
そもそも研究者の使命は？

Where is the life we have lost in living?

Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?

T.S.Eliot(1934)

Where is the information we have lost in data?

H. Inoue& J.R.Pierce