

2022年度第4回一般社団法人世界メッシュ研究所「世界メッシュコード研究会」

Rによる

# 統計的データクリーニング

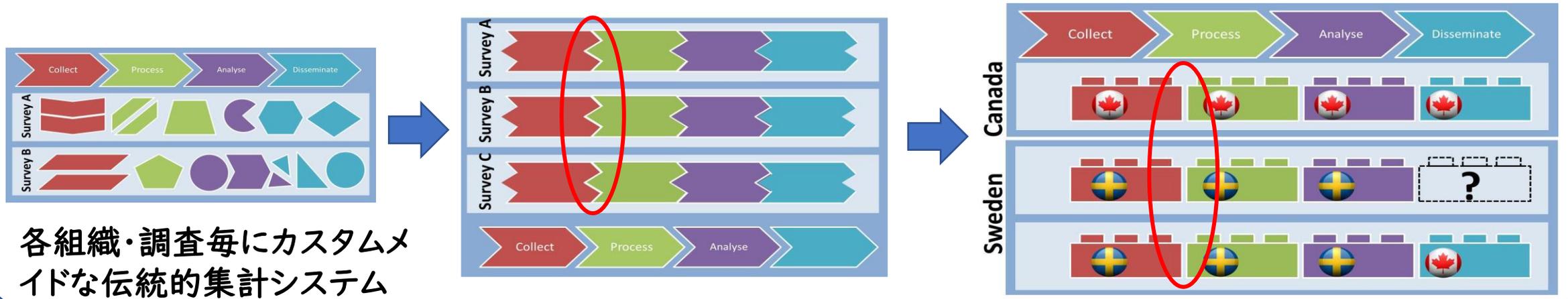
総務省統計研究研修所

和田 かず美

# 公的統計組織における近年の動向

GSBPM (一般統計ビジネスプロセスモデル) とそれに基づくCSPA (共通統計作成構造)

[図の出典: [https://ec.europa.eu/eurostat/cros/system/files/Enterprise%20architecture\\_UNECE.pdf](https://ec.europa.eu/eurostat/cros/system/files/Enterprise%20architecture_UNECE.pdf)]



各組織・調査毎にカスタムメイドな伝統的集計システム

国際的な業務の標準化  
集計モジュールの共有

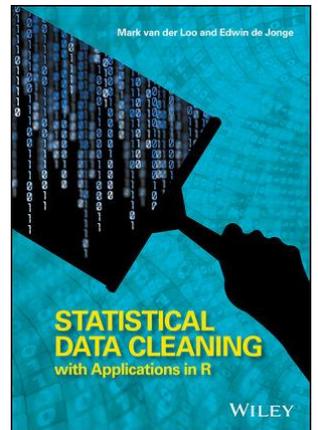
汎用集計システムの提案 ↓

**Awesome official statistics software**

<https://github.com/SNStatComp/awesome-official-statistics-software>



選択基準: オープンソースで無償 / 公的統計の作成に実用されている /  
インストール・利用が容易 / 安定バージョンがありメンテされている





## 当日の内容（予定）

この翻訳本で取り上げられている簡単な汎用集計システムのデモを予定しています。お手元で試していただけるよう、当日Rコードを配布します。

### 汎用集計システムの主な機能

- 検証ルールに基づくデータチェックと修正
- 検証ルールの管理
- 矛盾する観測データの修正箇所の特  
定（エラーの局所化）
- 欠測の可視化と補完