

広島大学高等教育研究開発センター 2024年度公開研究会

数理統計を用いた個人情報とプライバシーに関する問題と研究公正・研究倫理・高等教育研究への含意

趣旨説明

研究データ公開性の光と影：現状認識の必要性



野内玲（広島大学）

2024. 5.7 於 広島大学きてみんさいラボ（ハイブリッド開催）

趣旨（概要）

- ・ **問題提起**：オープンデータの推進が研究対象者の権利保護の課題になりうる。研究公正と研究倫理の噛み合わない現状がある。

オープンデータ・オープンサイエンスの推進

公的統計データ・研究データの公開や共有
→データ資源の共有=コスト削減
→**再現性の確保、捏造・改ざん防止（研究公正）**

例：高等教育

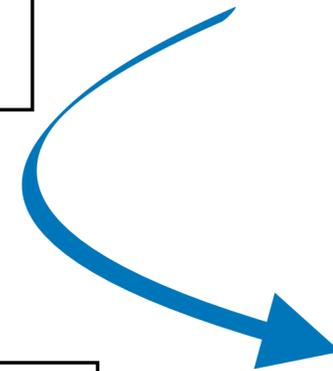
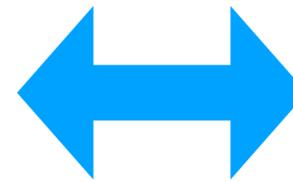
IRの必要性、学生に関する各種データの利活用
→**大学の戦略・経営・研究活動に展開**
→**研究倫理、個人情報保護というハードル**

データサイエンス

機械学習モデル・データフュージョン技術による分析
→**個人情報やプライバシー情報の暴露・漏洩の可能性**

個人情報・プライバシーの保護と分析技術

- ・ **差分プライバシー**：ノイズをのせ、個人情報をマスキしつつ予測の信頼性を確保
- ・ **連合学習**：データを組織間で収集せずにモデルのみを共有



データを組み合わせることそれ自体は問題ではない。

情報のやり取りの適切な手続きを踏んでいるか

?

参考：技術開発における データフュージョン

自動運転に必須なセンサーフュージョン
カメラや速度センサー、GPS情報



必要なこと！

位置情報や運転経路情報に関する
個人情報・プライバシー情報の管理
(ELSI的要素)



自動運転のための必要技術には、デジタル地図生成、自己位置推定、周辺環境認識、走行軌道生成・車両制御、ヒューマンマシンインターフェースなどが挙げられ、これらを統合して自動運転システムを形成しています。私たちの研究室では、これらの要素技術を幅広く研究し、自律型自動運転自動車のシステム開発を行っています。

参考： 次世代医療基盤法

各サイトで持っている情報を
うまく統合して活用する

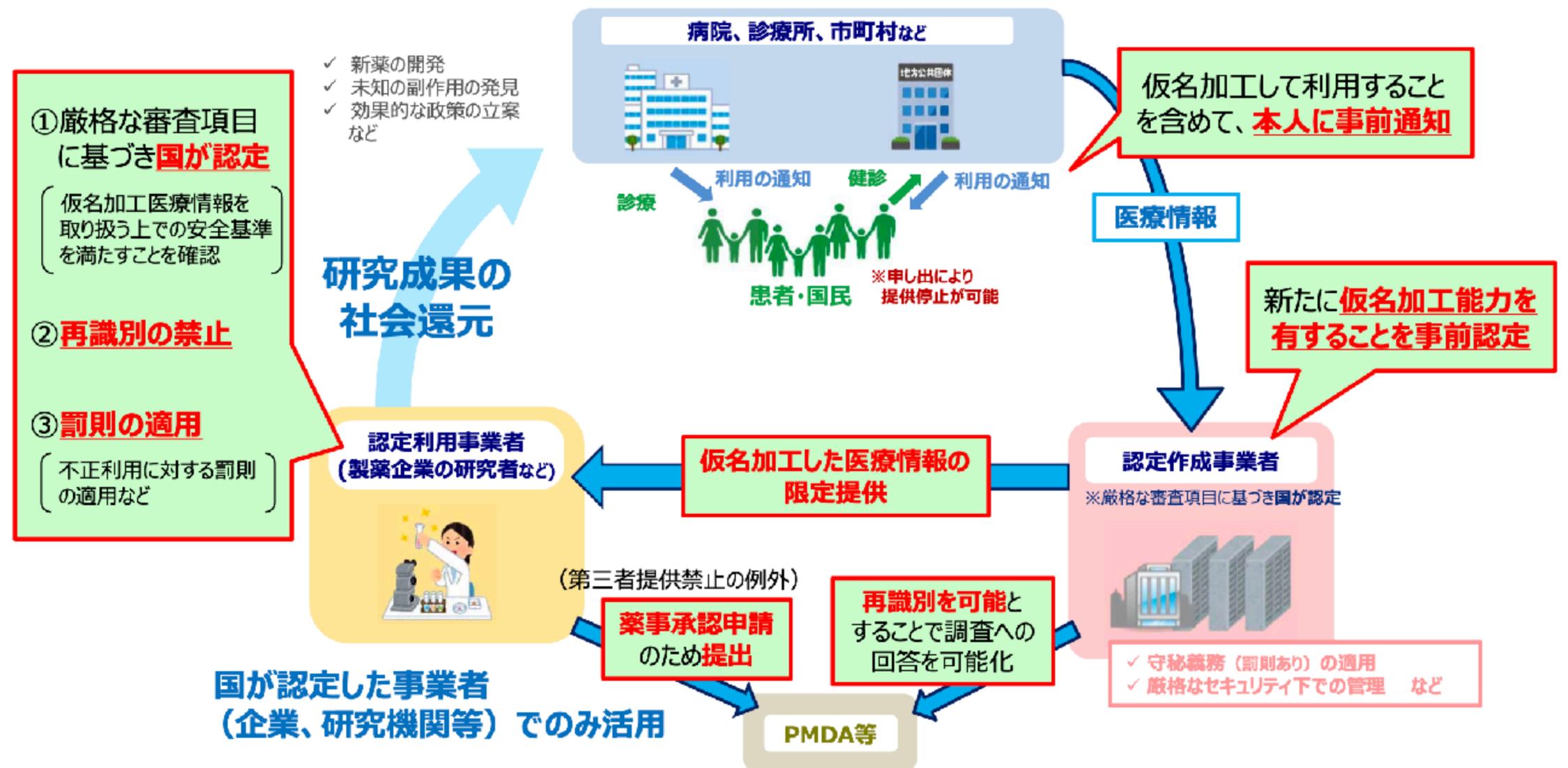
新薬の開発
副作用の発見
政策立案

必要なこと！
個人情報の保護

仮名加工医療情報の利活用に係る仕組みの創設

■ 仮名加工医療情報の利活用に係る仕組みの創設

- 医療情報の研究ニーズ、社会的便益の観点から、**新たに「仮名加工医療情報」の作成・提供を可能とする。**
- その際、**個人情報の保護の観点から、仮名加工医療情報の提供は国が認定した利活用者に限定。**



参考（続き）：tasokarena（タソカレナ）

NTTテクノクロスwebサイトより抜粋要約

<https://www.ntt-tx.co.jp/products/anontool/>

個人情報を法令にしたがって適切に匿名加工情報に加工することを支援するソフトウェア。医療・ヘルスケア分野に限らず、金融、コールセンター、自治体などの分野でも利用。使用するデータの特性や利用目的に応じて最適な加工方法の選択と加工結果を評価する環境を提供。

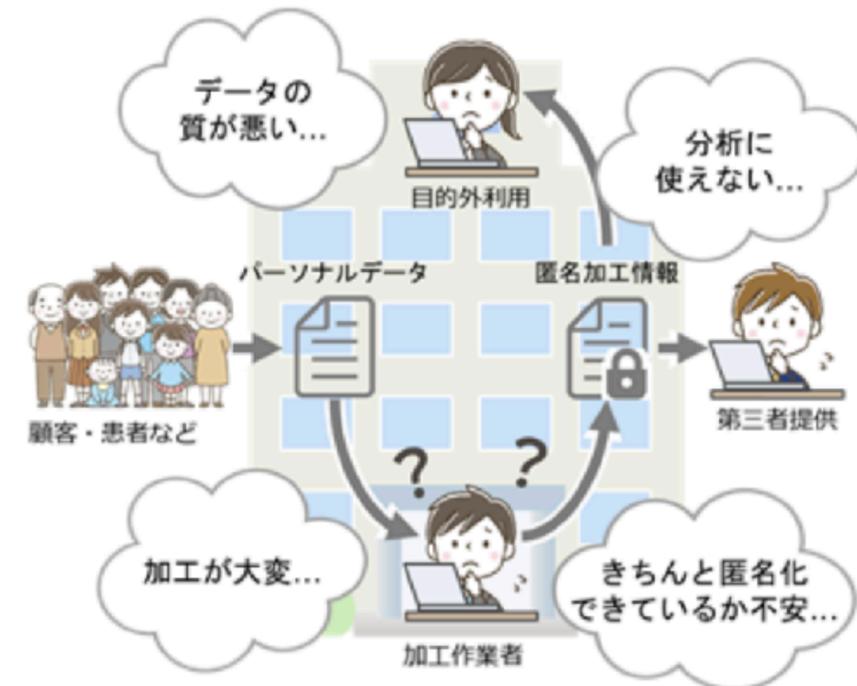
次世代医療基盤法に基づいた認定医療情報等取扱受託事業者や、仮名加工情報の作成においても使用可能

新技術

…匿名加工情報の作成では、元のパーソナルデータの数が少ない場合にはデータの匿名性が有用性のどちらかが低下するという課題があったり、緻密な分析を行う際にはデータの有用性の観点から匿名加工情報よりも実際のパーソナルデータが活用されたりするのが現状。

より高度なデータ利活用を実現するため、パーソナルデータの匿名性を維持したまま分析の精度や範囲を高めるためv4で搭載を開始した機能「合成データ生成」をさらに強化し、ベイジアンネットワークをもとに合成データを生成する機械学習的手法（差分プライバシー対応）を追加し、発売

BEFORE: 手作業で匿名化



AFTER: tasokarena™で匿名化



改めて問題提起

- 問題提起：確かに利点はあるものの、どの程度までのリスクが認識されていて、こういった対応がありうるかを知らずに「オープンデータの推進」を闇雲に推奨するのは危険
- 「同意を得た情報」同士を組み合わせ「同意を得ていない or 同意を得られない」情報を取り出すことができる。→**インフォームド・コンセントの反故** 【匿名化しても個人が特定できるのは伝統的な問題】
 - ✓ 回答への「同意」は研究者（データ処理者）による「その後の操作への同意」なのか？ICのあり方の再考を要する。研究倫理審査委員会における検討事項としての周知や、アカデミアでのルール形成（学問の自由、公的な利益との兼ね合い）といった形での対応が必要かもしれない。
- ➡ **【倫理面での配慮と信頼性のトレードオフ】** 秘匿処理や差分プライバシーといった対応策は漏えいリスクをどこまで軽減し、倫理面の懸念を払拭するか。一方で、ノイズを乗せるなど情報漏洩リスク対策を講じた場合、分析の精度はどの程度保たれるのか。
 - ✓ 利用する側の倫理意識だけでなく、公開する側も適切な対応（処理、手続き）が必要