情報通信審議会　情報通信技術分科会　技術戦略委員会（第14回）
平成29年2月27日

資料１４－５

データ統合利活用におけるプライバシー保護とデータセキュリティ

国立研究開発法人情報通信研究機構
サイバーセキュリティ研究所
セキュリティ基盤研究室
盛合　志帆

国立研究開発法人
情報通信研究機構
National Institute of Information and Communications Technology

データ統合利活用：新たな成長戦略の鍵

このセンサーデータは信頼できるのか？

データ漏洩対策は大丈夫？

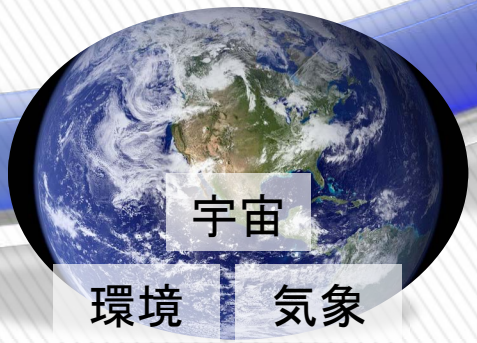
移動履歴

購買履歴

興味・関心

検索キーワード

私のプライバシーは守られている？



セキュリティ・プライバシー対策: データビリティの必須要件

Datability is all about the ability to use large volumes of data **sustainably** and **responsibly**.

[CeBIt 2014, held in Hannover, Germany]

データビリティとは、大規模なデータを持続可能かつ責任ある形で活用する能力のことです。

[CeBIt 2014（ドイツ、ハノーバー）にて提唱]

【sustainable】

- 継続的な対応を可能にするリソース整備
- データ・マネジメント＆ガバナンスの確立
- セキュリティ対策

【responsible】

- 社会問題／環境問題の解決（スマートシティ、ヘルスケア、エネルギー活用、経済活動等）
- プライバシー／個人情報の問題

情報通信審議会　情報通信技術分科会　技術戦略委員会（第13回）
資料13-6「データビリティサイエンス」（大阪大学 理事・副学長 八木教授）より

データ統合利活用における プライバシー/個人情報保護

パーソナルデータ利活用には
プライバシー保護が不可欠
cf. 改正個人情報保護法
全面施行（2017.5）

次世代AI 技術に
よる分析・解析

新たな知見・イノベーション
多様な経済分野でのビジネス創出

改正個人情報保護法のポイント

（一部）

①個人情報の定義の明確化
個人識別符号の概念

【個人情報】

生存する個人に関する情報であって、

1) 氏名、生年月日、住所等により特定の個人を識別することができるもの（他の情報と容易に照合でき、それにより特定の個人を識別することができるものを含む）

例：データベース化されていない書面・写真・音声等に記録されているもの

2) **個人識別符号**（①又は②）が含まれるもの

① 特定の個人の身体の一部の特徴を電子計算機のために変換した符号

例：顔認識データ、認証用指紋データ等

② 対象者ごとに異なるものとなるように役務の利用、商品の購入又は書類に付される符号

例：旅券番号、免許証番号等

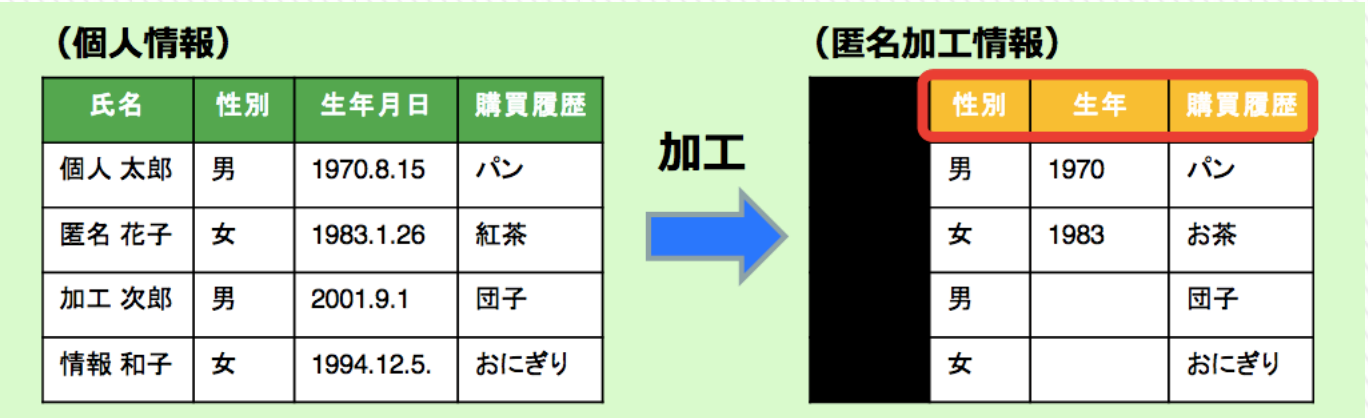
②匿名加工情報の新設

【第三者への提供】分野横断利活用のポイント

- 本人の同意を取れば提供可能
- 委託、事業承継、共同利用に伴って提供する場合には、「第三者」に提供するものとはされない
- **「匿名加工情報」**に加工すれば、本人の同意をとらなくても自由に利活用可能
  → 新事業や新サービスの創出、国民生活の利便性の向上を期待

匿名加工情報

- 個人の特定性を低減したデータ
  - 「個人情報を加工して、通常人の判断をもって、個人を特定することができず、かつ、加工する前の個人情報へと戻すことができない状態にした情報」
- 加工方法
  - 特定の個人を識別する項目の削除や、情報を”丸める”など
  - 「匿名加工情報作成マニュアル」(経済産業省, 2016.8)
  - 個人情報保護委員会でも匿名加工情報に関する事務局レポートが作成される



「匿名加工情報制度について」(個人情報保護委員事務局 参事官 小川久仁子) 資料より

匿名加工情報:
社会実装に向けた研究開発課題

» 匿名加工技術の評価技術　有用性指標と安全性指標
  > いかに再識別のリスクを低減し、データの有用性を保ったまま加工するか
  > NICTでも第4期中長期計画にて取り組み

» PWS CUP　匿名加工・再識別コンテスト
  > 2015年から情報処理学会 コンピュータセキュリティシンポジウムで開催
  > PWS組織委員会委員長：菊池浩明(明治大)
  > 後援：個人情報保護委員会
  > PWS CUP 2016
    + 匿名加工部門：顧客情報データと購買履歴データを有用性を残して安全に匿名加工する
    + 再識別部門：元の顧客データをヒントにして、匿名加工された購買履歴から顧客を識別する

データ統合利活用における データセキュリティ

暗号・認証技術により
データ機密性・データ信頼性を
確保することで
分野横断でのデータ利活用を促進

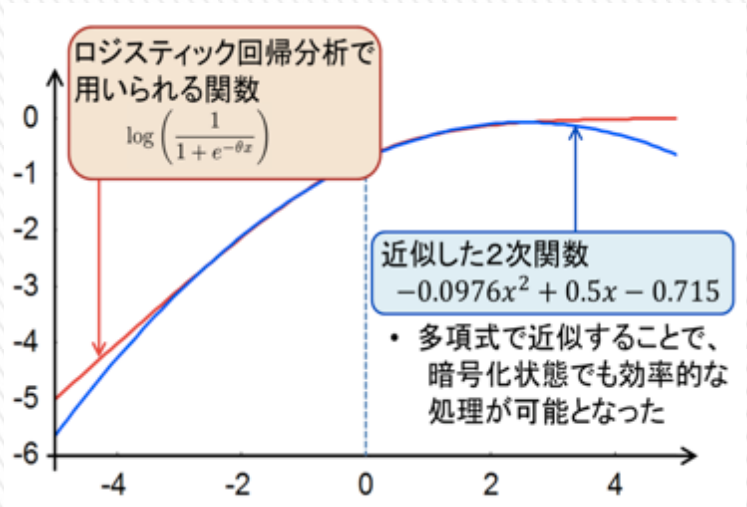
次世代AI 技術による分析・解析

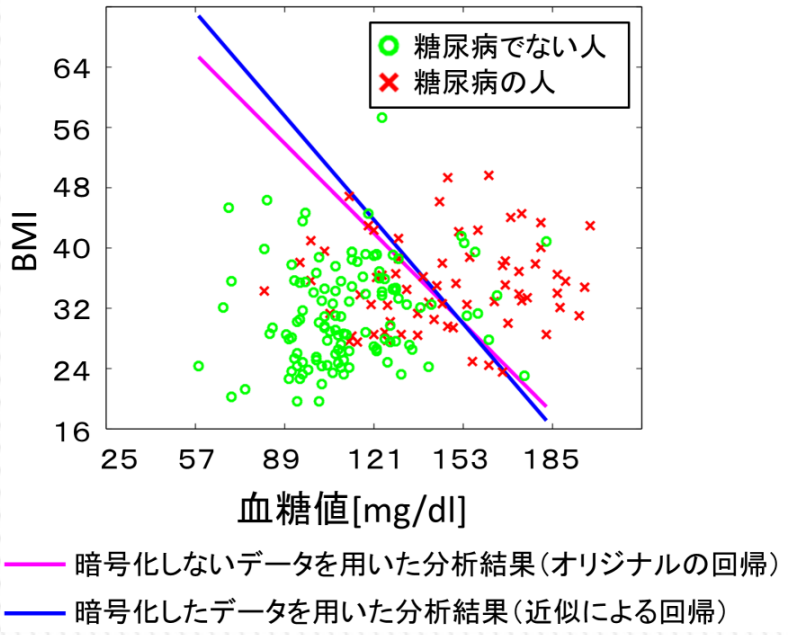
暗号化したまま分析・解析!?

新たな知見・イノベーション
多様な経済分野でのビジネス創出

暗号化したままビッグデータ分類

- ビッグデータ解析で多用されているロジスティック回帰分析をデータを暗号化したまま計算可能に
- 暗号化された1億件のデータを30分以内で複数グループに分類できることをシミュレーションで確認
  - NICTプレスリリース「**暗号化したままデータを分類できるビッグデータ向け解析技術を開発**」(2016.1.14)



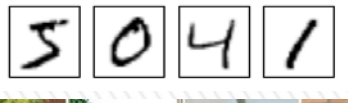


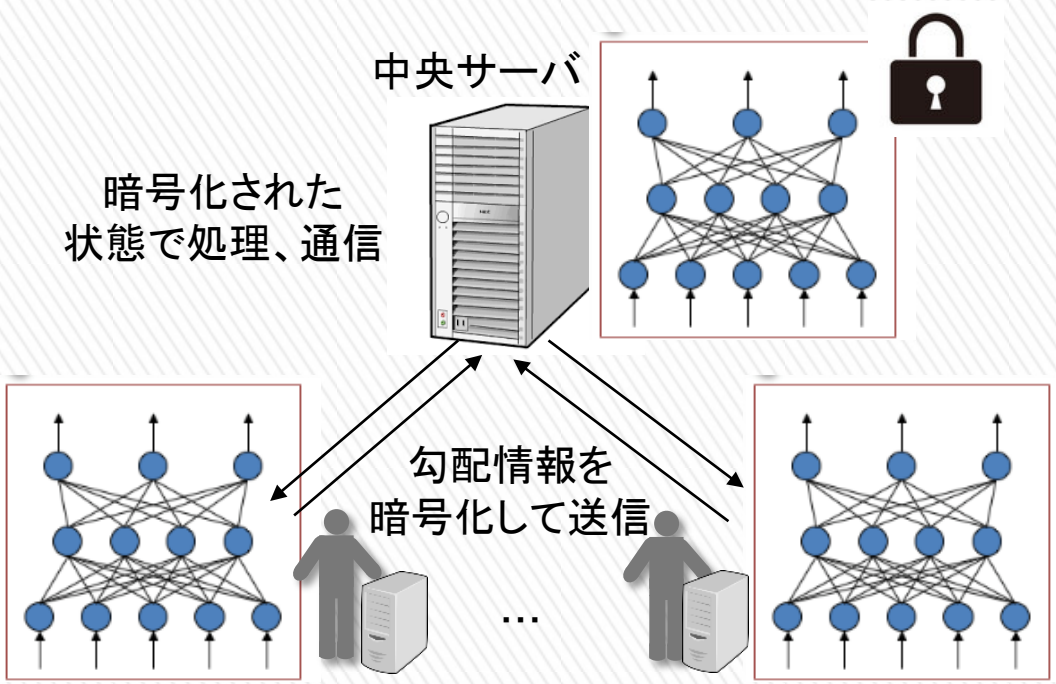
暗号化したまま深層学習*

* 深層学習(deep learning)
多層構造のニューラルネットワークを用いた機械学習

» 多数の参加者が持つデータセットを互いに秘匿したまま深層学習を行うプライバシー保護深層学習システムを提案

下記の機械学習用データベースで性能確認
- MINIST(手書き数字認識)
- SVHN (Googleストリートビュー写真から連続した数字を認識)
- Speechデータセット





N人の参加者と中央サーバ1台による深層学習
(分散協調学習)

「イノベーション創発に資する人工知能基盤技術の創出と統合化」
研究総括: 栄藤 稔(NTTドコモ), 文科省 AIPプロジェクトの一環として運営

JST CREST「人工知能」採択課題

» 「複数組織データ利活用を促進するプライバシー保護データマイニング」

> 研究代表者:盛合 志帆(NICT). 神戸大 小澤教授, (株)エルテスと連携

課題

複数の異なる業種・組織が有する実社会の
膨大なデータを統合して利活用する際、
プライバシー保護・データ機密性の確保が課題

研究課題

暗号技術や人工知能技術を活用し、
プライバシーを保護した状態で
高速にデータ分析や異常検知を行う技術を
研究開発

解決する社会問題

金融分野における社会問題の解決に活用.
金融機関以外がもつデータを利活用した
①インターネットバンキング不正送金の検知
②個人向け融資における 適正利率の導出
⇒ フィンテックにおけるイノベーション創出をめざす.

研究体制