データ利活用を促進するプライバシー保護技術

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

未来社会を開拓する世界最先端のICT

NICT はICTを専門とする唯一の公的研究機関として研究開発を推進する役割と社会実装に取り組む役割を担っています。

センシング基盤分野

観る（み）

多様なセンサー等により高精度な観測等を行うための技術を研究開発します。

- ●ゲリラ豪雨を早期に捕捉する技術（フェーズドアレイ気象レーダ）
- ●衛星測位などに影響を与える宇宙環境を計測・予測する技術
- ●より正確な時刻を作る技術（超高精度原子時計）
- ●安全な電波利用を確保する技術（数値人体モデル用ソフトウェア）

データ利活用基盤分野

創る（つく）

人工知能やビッグデータ解析等により新しい知識・価値を創造するための技術を研究開発します。

- ●世界の「言葉の壁」をなくす実用レベルの多言語音声翻訳技術（VoiceTra）
- ●誰でも専門家のような高度知識を得られる人工知能技術（WISDOM X、DISAANA）
- ●脳活動を測ることで健康・福祉・生活の質を向上する技術（高次脳機能型情報処理技術）

サイバーセキュリティ分野

守る（まも）

多様化するサイバー攻撃から社会システム等を守るための技術を研究開発します。

- ●標的型攻撃等多様化するサイバー攻撃の対応技術
- ●防御技術の検証を実施するプラットフォーム構築活用技術
- ●IoT デバイスにも実装可能な軽量暗号技術

フロンティア研究分野

拓く（ひら）

ICTに新たなブレークスルーをもたらすための技術を研究開発します。

- ●盗聴を防止する量子情報通信技術
- ●テラヘルツ帯など未踏周波数領域を開拓する技術
- ●酸化物半導体や深紫外光を利用したデバイス技術（深紫外 LED）

統合ICT基盤分野

繋ぐ（つな）

通信量の爆発的増加等に対応するための技術を研究開発します。

- ●IoT 時代に求められる柔軟性の高いネットワークの実現
- ●異種ネットワークの統合に必要なワイヤレスネットワーク技術
- ●現在の千倍以上の通信量に対応する世界最高水準の光ファイバ技術（マルチコア・マルチモードファイバ）
- ●衛星通信を高速化・大容量化する技術

産学官連携による研究成果の創出支援と社会実装に導く取組

研究開発成果の最大化に向けて取組みを強化します。

- ●IoT テストベッド等を活用した「利用者・企業・大学・地域社会の出会いの場」の創出と、技術実証と社会実証の一体的推進
- ●オープンイノベーション創出に向けた産学官連携の強化と「オープンイノベーション推進本部」の設置
- ●耐災害 ICT の実現に向けた取組
- ●戦略的な標準化活動の推進
- ●研究開発成果の国際展開の強化

NICT第4期中長期計画 (2016-2020) における目標

【中長期計画】1-4. サイバーセキュリティ分野　(3) 暗号技術

機能性暗号技術

IoTの展開に伴って生じる新たな社会ニーズに対応するため、新しい機能を備えた機能性暗号や軽量暗号・認証技術の研究開発に取り組む

暗号技術の安全性評価

暗号技術の安全性評価を実施し、新たな暗号技術の普及・標準化に貢献するとともに、安心・安全なICTシステムの維持・構築に貢献

プライバシー保護技術

パーソナルデータの利活用に貢献するためのプライバシー保護技術の研究開発を行い、適切なプライバシー対策を技術面から支援

- 暗号技術の安全性評価における日本随一の研究拠点
  - ✓ 中立公平で信頼性の高い安全性評価情報の継続的発信
- セキュリティ＆プライバシー保護技術の研究連携拠点に
  - ✓ S&Pの新たな研究開発テーマでの国内外研究連携拠点を目指す
- 社会で活用される研究開発成果の創出と社会展開
  - ✓ パーソナルデータ利活用や改正個人情報保護法施行に資する技術開発

災害・危機管理に備えた　情報共有システム

» 災害が起きてから構築する/初めて使うものでは役に立たない
» 普段使っているシステムの延長にあるべき
» 災害時でもプライバシー漏洩には配慮すべき.
一度ネットに流出すると回復は困難.

⇒ 平常時からプライバシー保護、セキュリティに配慮したシステムを構築しておくべき

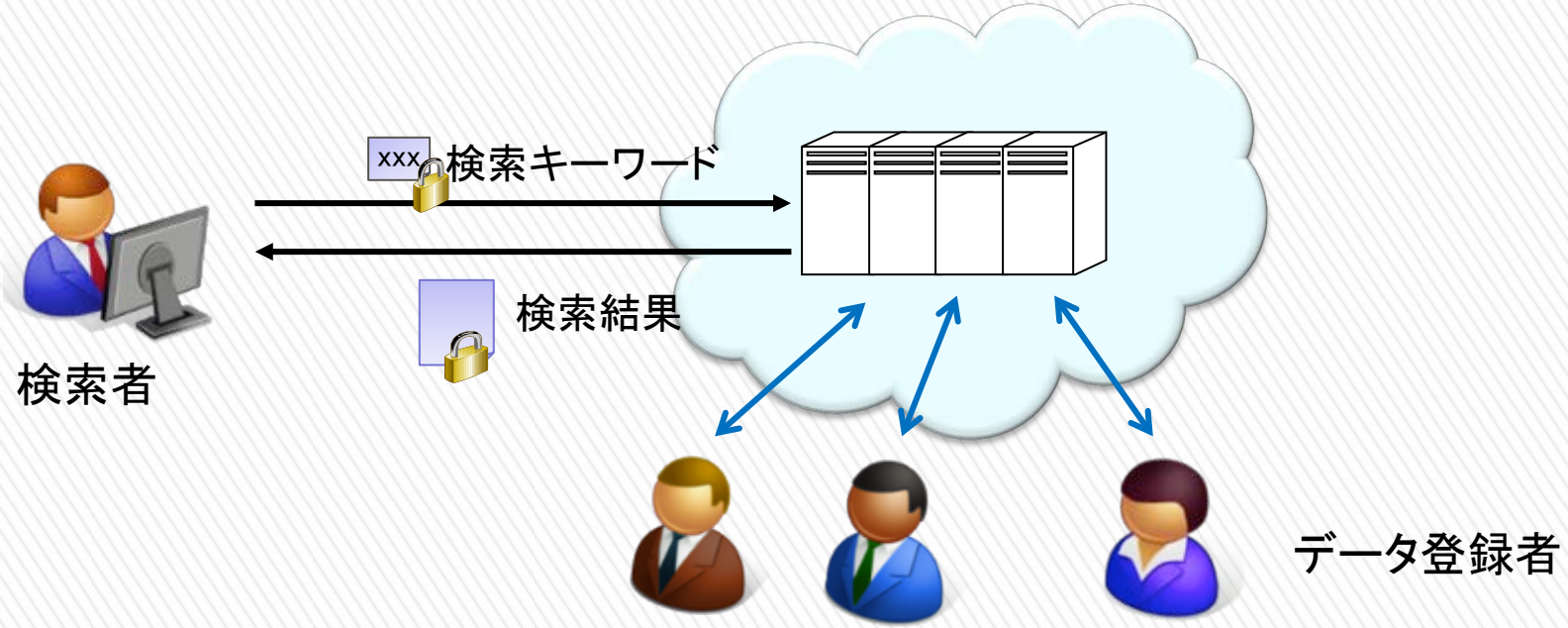
災害・危機管理に活用可能な暗号技術①
秘密分散

- データの分散バックアップ・BCP
- どこかの地域が壊滅的な被害を受けても、残りのデータで全てを復元可能
- 1つのサーバがサイバー攻撃を受けてデータが流出しても、情報は漏れない

災害・危機管理に活用可能な暗号技術②
秘匿検索

» 暗号化データ中のキーワード検索

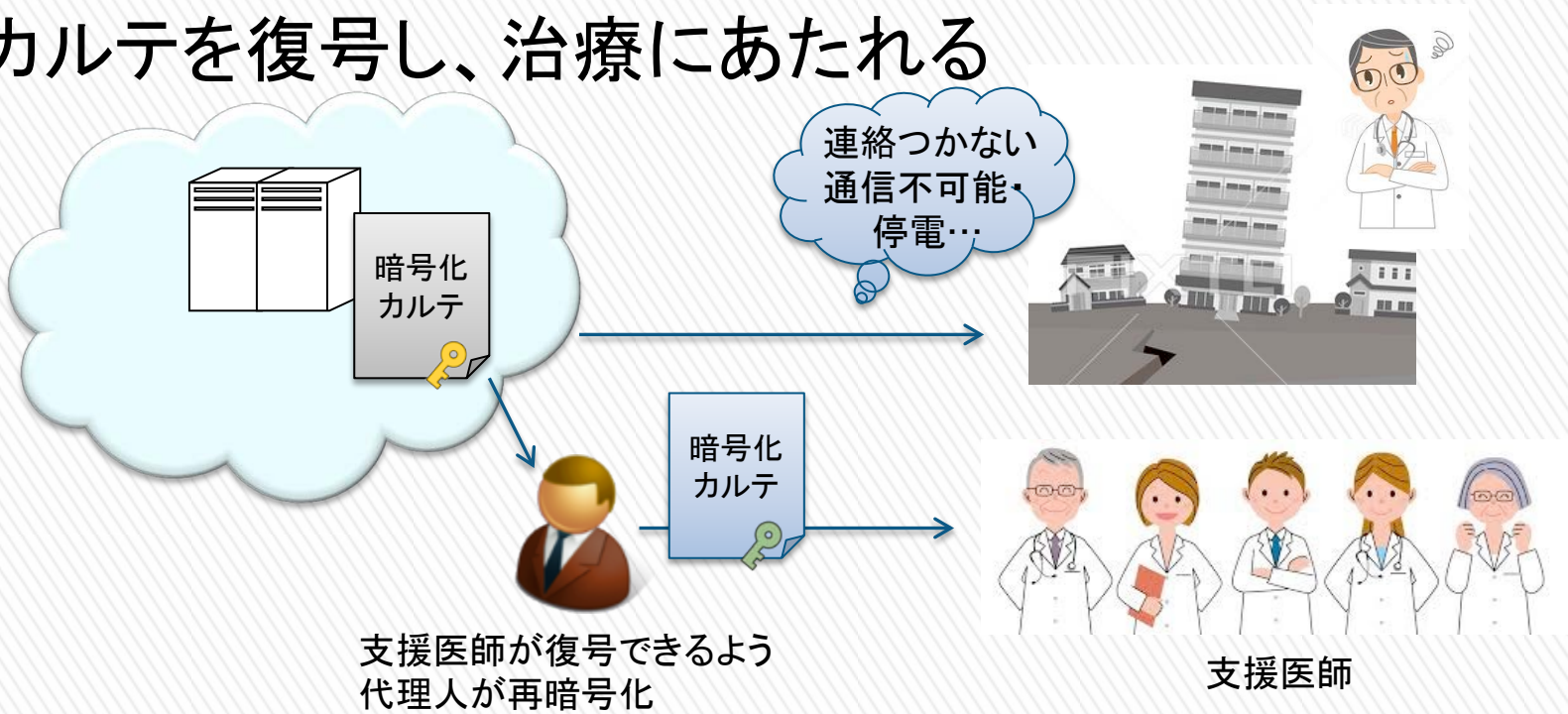
» 例）避難者/負傷者全員の名前を明かすことなく、特定の方が含まれているかを検索可能



災害・危機管理に活用可能な暗号技術③
代理再暗号化

» 暗号化データを復号することなく、代理人が他の人が復号できるよう再暗号化できる技術

» 例）担当医と連絡が取れなくても、支援医師がカルテを復号し、治療にあたれる



データ統合利活用：
成長戦略・課題解決の鍵

データ漏洩対策は大丈夫？

交通

産業システム

脳

私のプライバシーは守られている？

この測定データは信頼できるのか？

移動履歴

購買履歴

興味・関心

検索キーワード

医療

農業

金融・経済

宇宙

環境

気象

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）

AI 技術等による
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

匿名加工情報

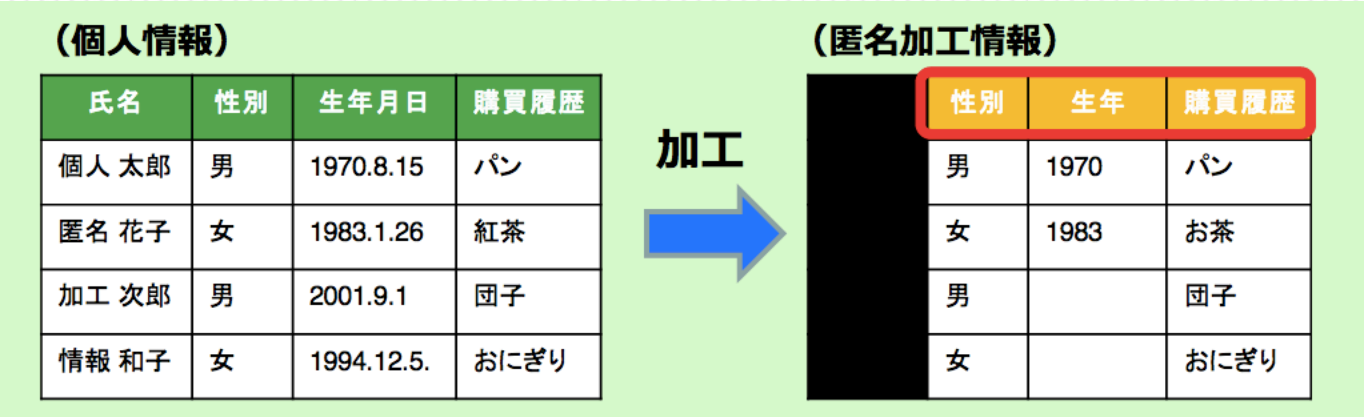
» 分野横断でのデータ利活用を推進するため
改正個人情報保護法で新設

【個人データの第三者への提供】
- 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」

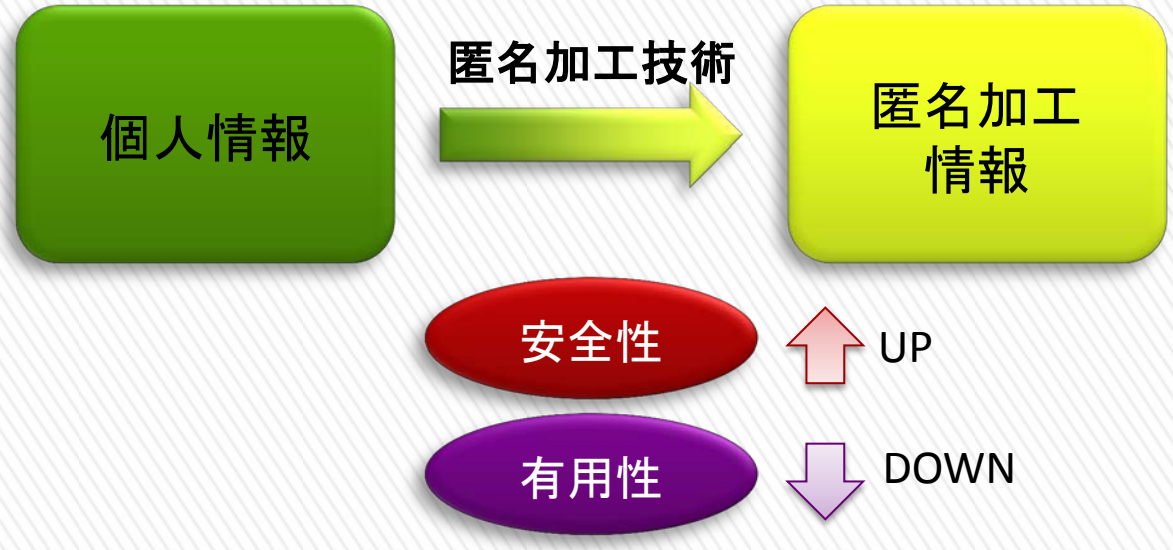
» 加工方法
  > 特定の個人を識別する項目の削除や、情報を”丸める”など
  > 「匿名加工情報「パーソナルデータの利活用促進と消費者の信頼性確保の両立に向けて」」(個人情報保護委員会, 2017.2)
  > 「匿名加工情報作成マニュアル」(経済産業省, 2016.8)



「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
  - 匿名加工情報の再識別リスクに関する標準的な評価手法は未確立
- 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

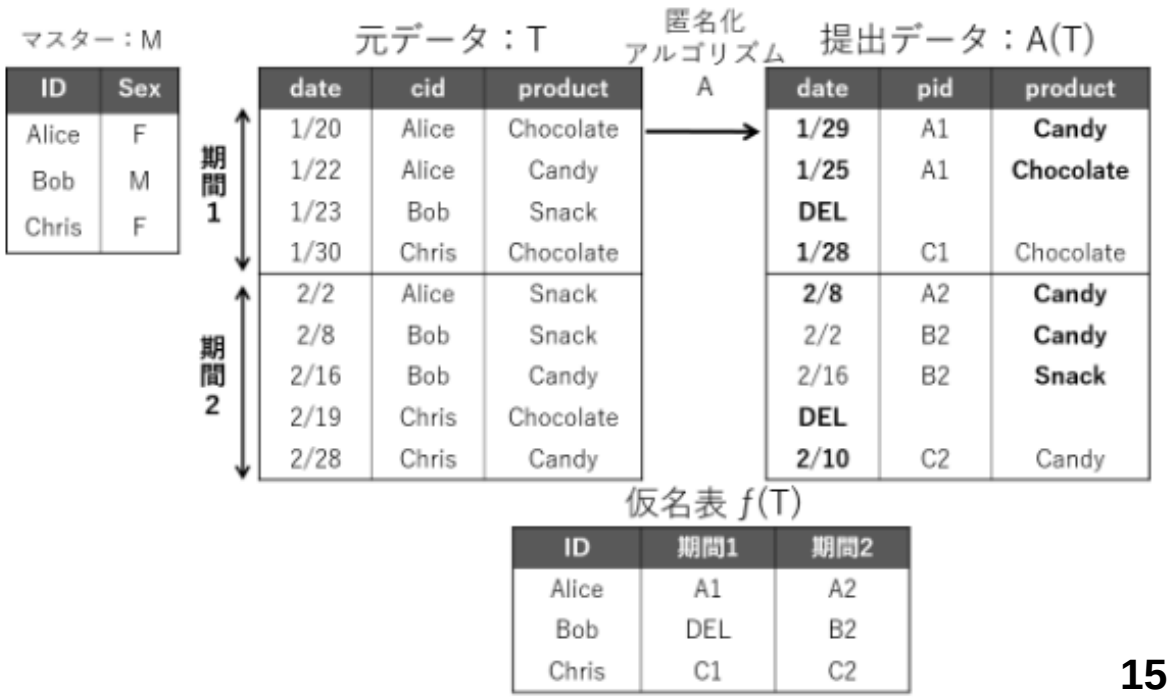
- 2015年から情報処理学会 コンピュータセキュリティシンポジウム(CCS)で開催
- 実行委員長：菊池浩明(明治大)
- 後援：個人情報保護委員会
- 目的：
  - 安全で有用性の高い匿名加工技術の開発を促進する

PWS CUP2017

» 「長期間の履歴データの再識別リスクを競う」
  > データに振られた仮名IDを長期間使用することで再識別される可能性が高まる
  > 復元することができる規則性を有しないように，定期的に変更するなどの措置が講じることが求められる

» 利用データ

UCI Machine Learning Repositoryで公開されている英国のオンラインショッピングサイトにおける2010年からの1年間の購買履歴



ID	Sex
Alice	F
Bob	M
Chris	F

date	cid	product
1/20	Alice	Chocolate
1/22	Alice	Candy
1/23	Bob	Snack
1/30	Chris	Chocolate
2/2	Alice	Snack
2/8	Bob	Snack
2/16	Bob	Candy
2/19	Chris	Chocolate
2/28	Chris	Candy

date	pid	product
1/29	A1	**Candy**
1/25	A1	**Chocolate**
DEL		
1/28	C1	Chocolate
2/8	A2	**Candy**
2/2	B2	**Candy**
2/16	B2	**Snack**
DEL		
2/10	C2	Candy

ID	期間1	期間2
Alice	A1	A2
Bob	DEL	B2
Chris	C1	C2

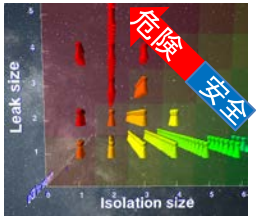
ユーザリスク評価システム　URANUS
(User Risk Assessment and NUllification System)

研究背景

- プライバシーデータの収集・利用が増加
- 収集したデータは単純に名前や住所などを削除した形式で保存，分析
- プライバシーリスクの評価への取り組みが不十分

URANUS

匿名加工処理 → CSV → ユーザリスク評価 → URANUS → 安全 or 危険

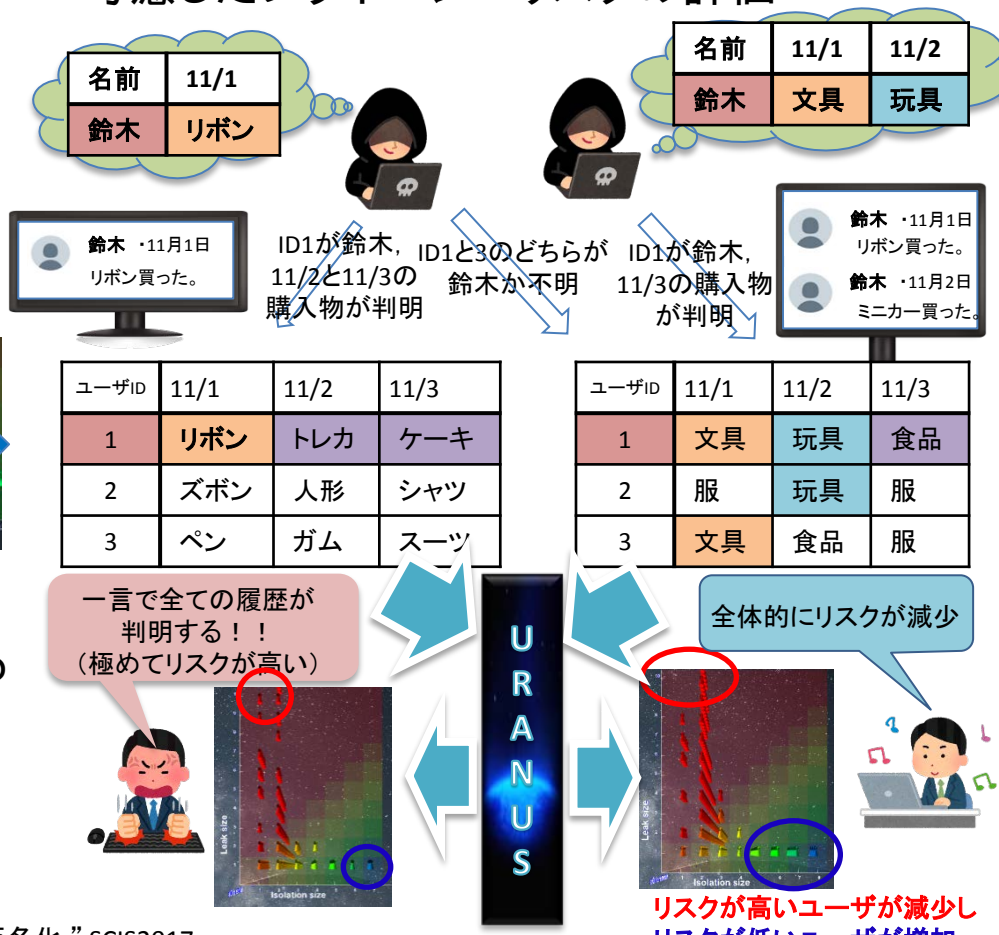


リスク評価軸

- **Isolation size/識別耐性**：個人を特定するのに必要な知識量（大きいほど安全）
- **Leak size/漏洩量**：個人特定時の被害の大きさ（大きいほど危険）

URANUSの適用事例

ツイッターなどのSNSによる情報発信を考慮したプライバシーリスクの評価



[関連発表]早稲田篤志，野島良，盛合志帆，菊池浩明，“良い仮名化　悪い仮名化，” SCIS2017

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

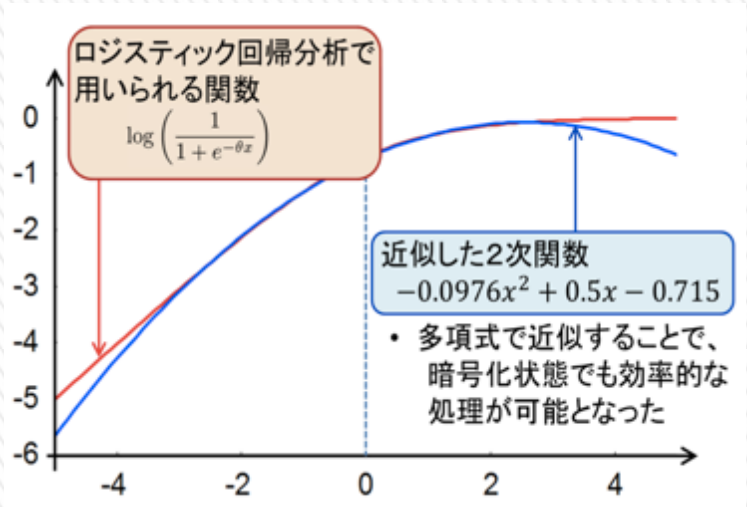
AI 技術等による
分析・解析

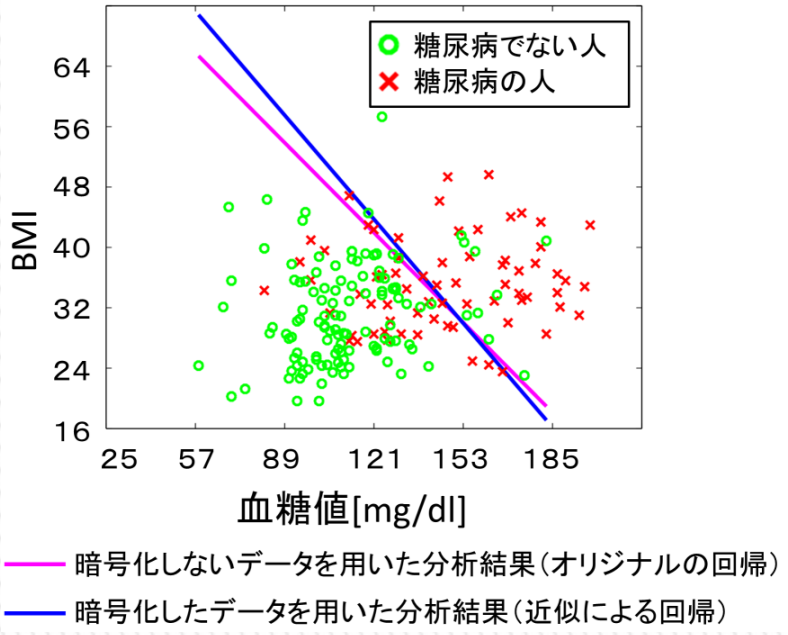
暗号化したまま
分析・解析!?

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)



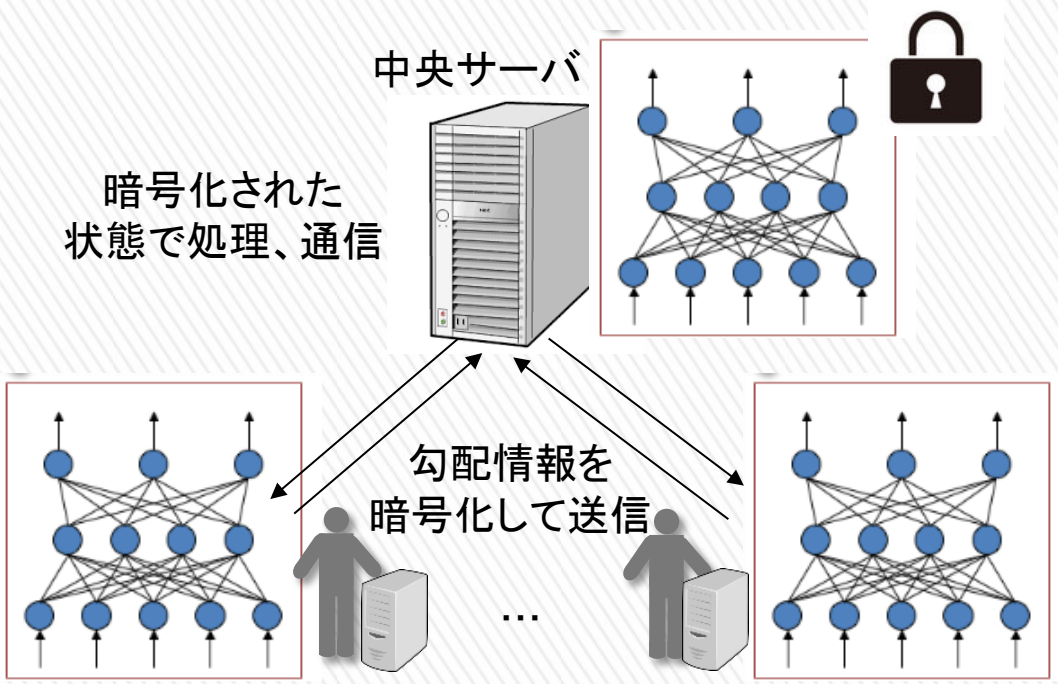


暗号化したまま深層学習*

* 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
- MNIST(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(阪大), 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

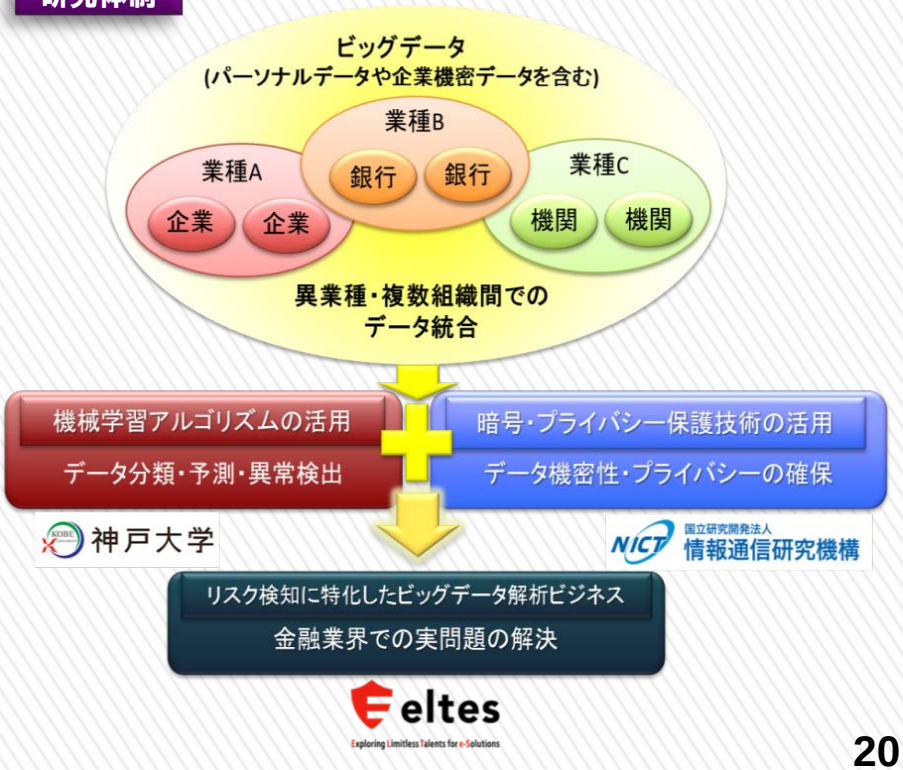
研究課題

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



まとめ

プライバシー保護

個人情報保護

安全管理措置

データセキュリティ

データの
機密性確保

データの
信頼性確保

匿名化技術

暗号技術