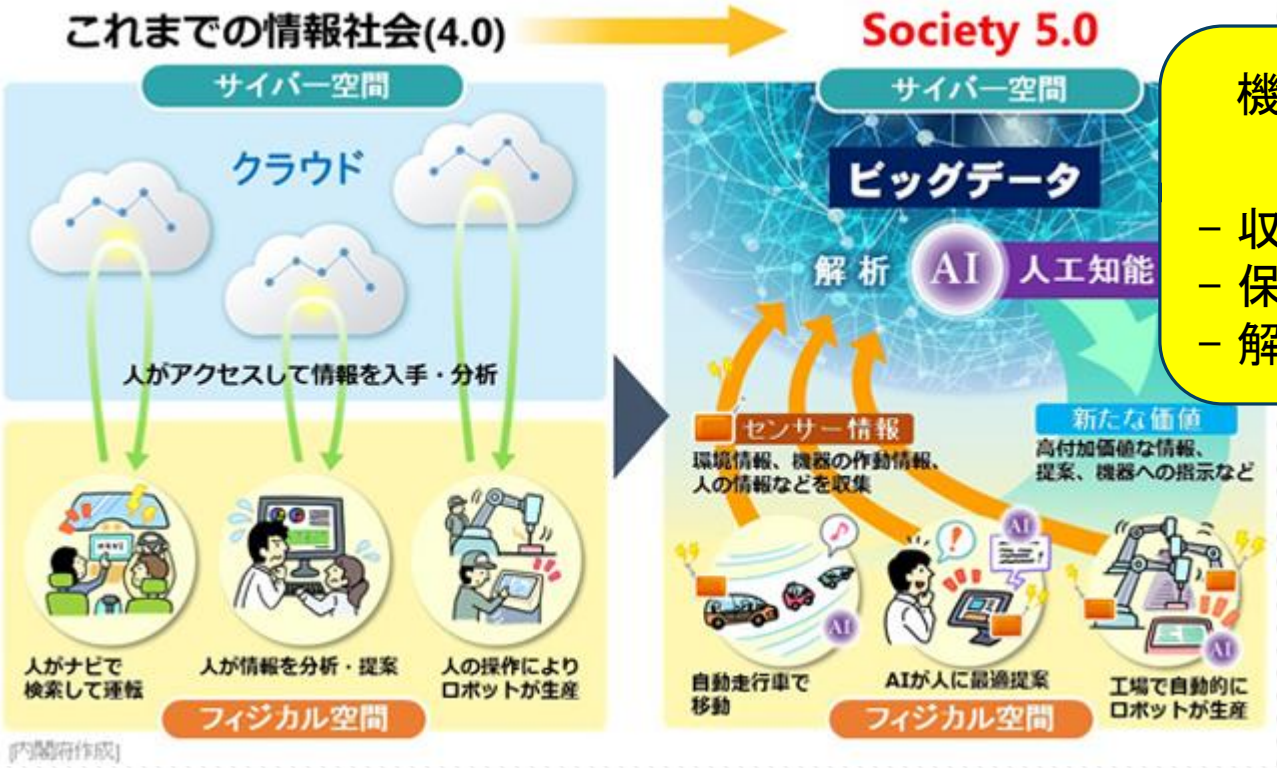
機微データの安全な利活用に向けたセキュリティ基盤研究室の取り組み

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
主任研究員 江村 恵太

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

NICTサイバーセキュリティシンポジウム2020

― Society 5.0に備えるセキュリティ技術 ―



機微データの取り扱い

- 収集
- 保管・管理
- 解析

» https://www8.cao.go.jp/cstp/society5_0/index.html

データの収集 (1)

» 提供者のプライバシー保護：身元を隠す

相反する要件

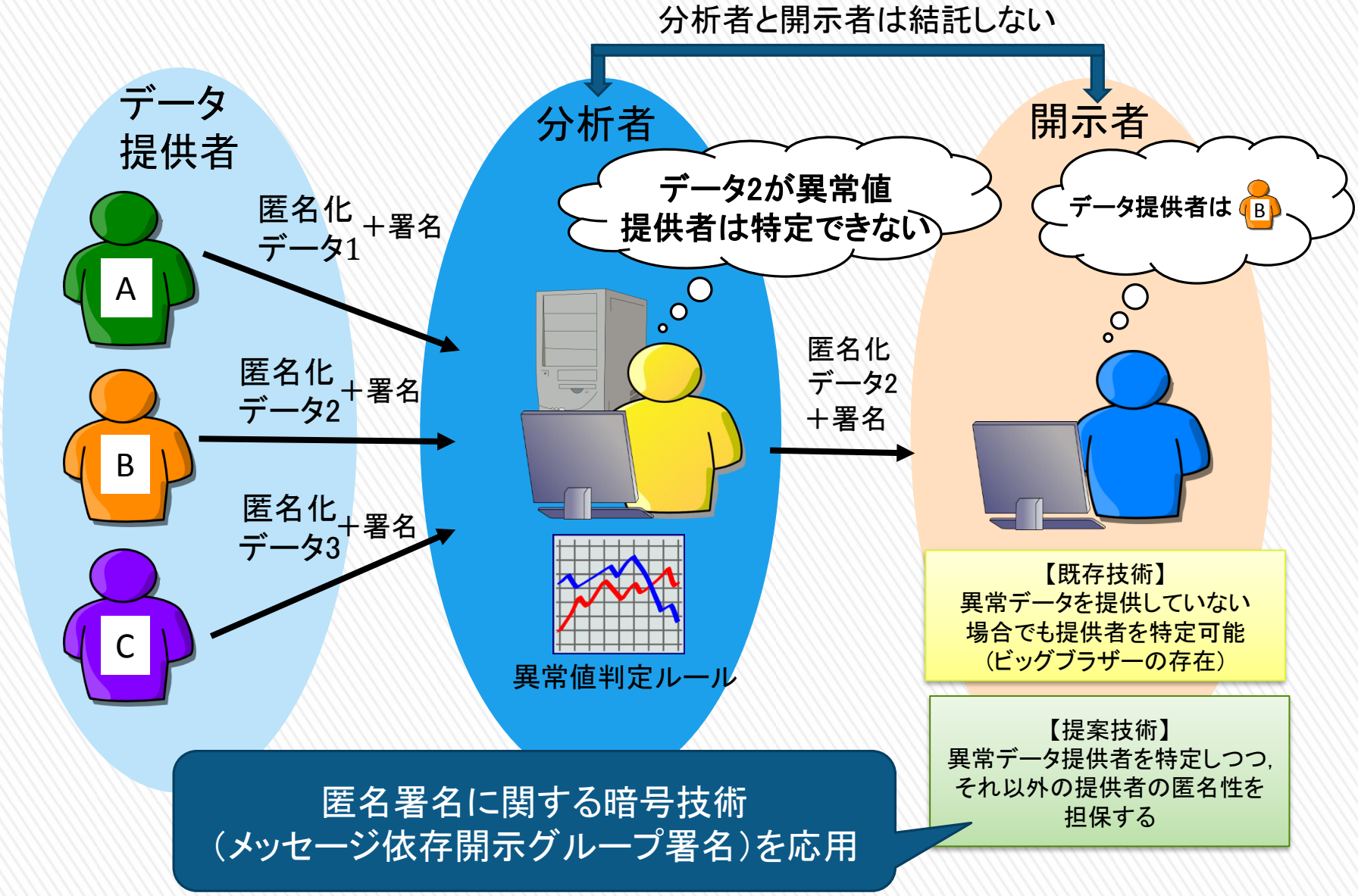
» 異常データの提供者は特定したい
  > 医療データを分析し，疾患がある場合に提供者に通知

» 如何にして異常データの提供者“のみ”を特定し
それ以外のデータ提供者のプライバシーを保護するか．

» 異常検知方式ごとにプライバシー保護手法を提案するのはコスト高．汎用的なフレームワークの提案が望ましい．

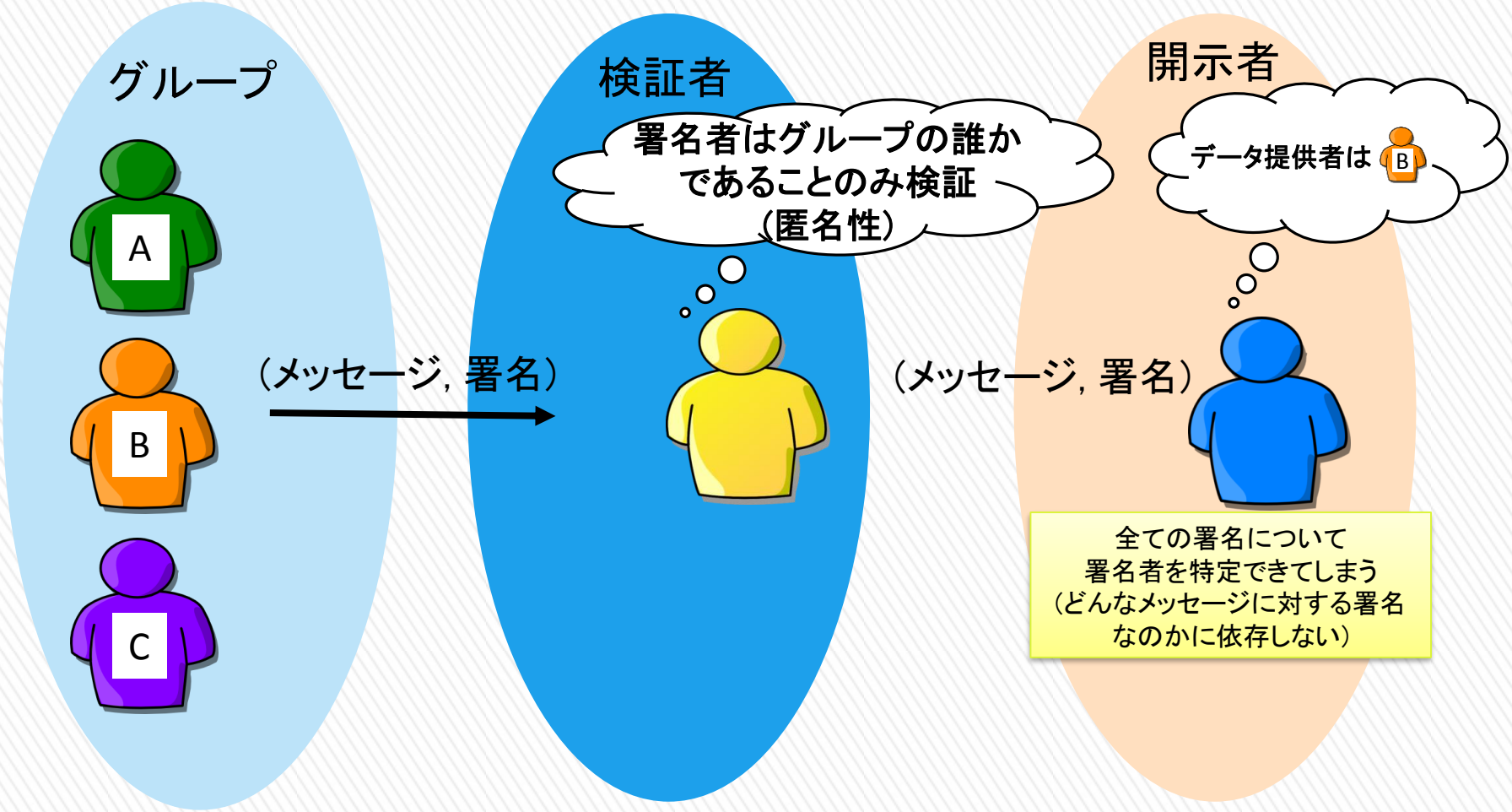
プライバシー保護フレームワーク

[Arai-Emura-Hayashi, WPES2017][PWS2017論文賞]



グループ署名

[Chaum-van Heyst, EUROCRYPT1991]

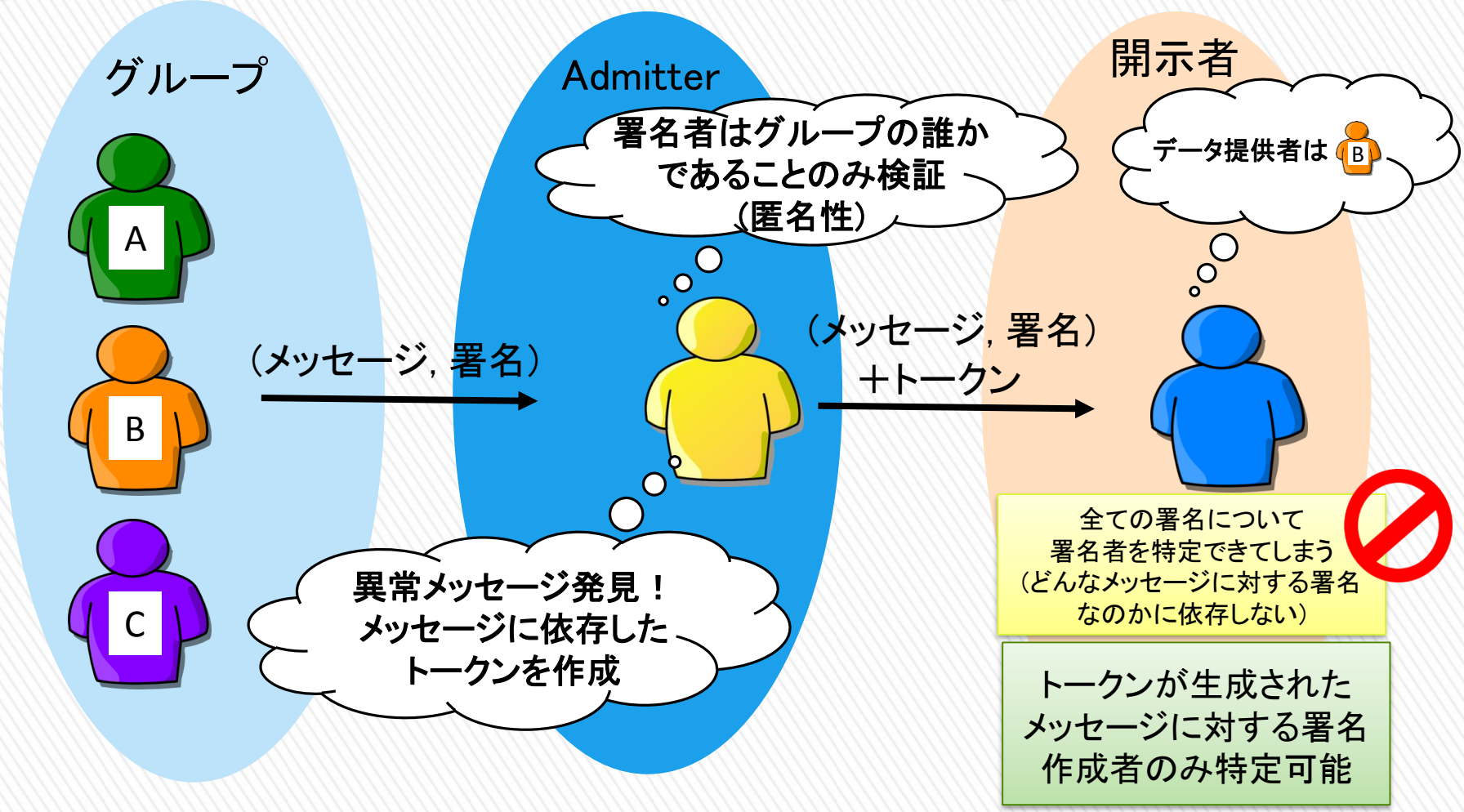


» 類似：リング署名（ブロックチェーン等で利用）

> 開示者が存在しない，署名鍵を署名者自身でセットアップ可能などの違い

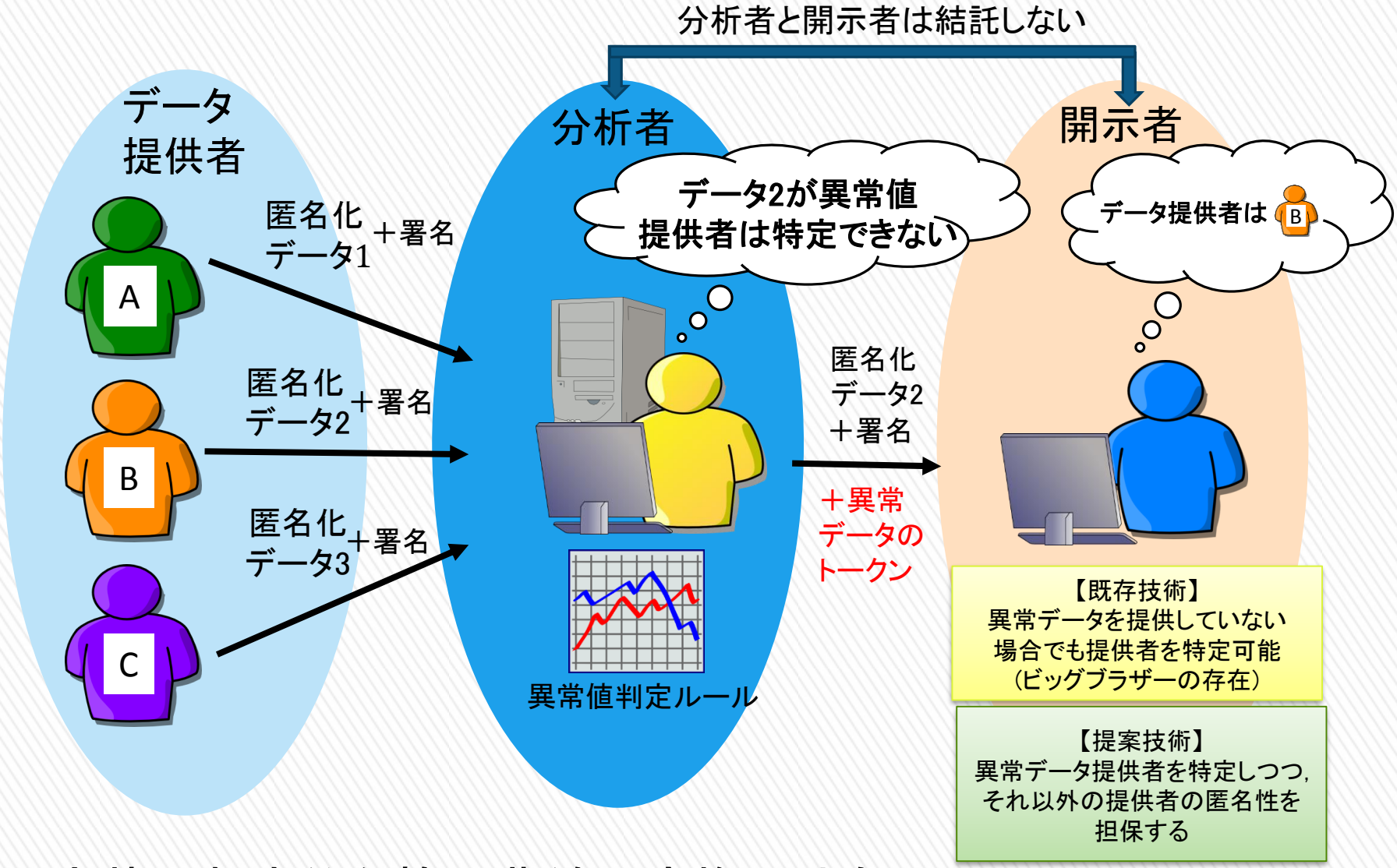
メッセージ依存開示グループ署名

[Sakai-Emura-Hanaoka-Kawai-Matsuda-Omote, Pairing 2012]
[Ohara-Sakai-Emura-Hanaoka, AsiaCCS 2013]



プライバシー保護フレームワーク

[Arai-Emura-Hayashi, WPES2017][PWS2017論文賞]



» 実装上効率的な楕円曲線の変換手法も利用

> Masayuki Abe, Fumitaka Hoshino, Miyako Ohkubo: Design in Type-I, Run in Type-III: Fast and Scalable Bilinear-Type Conversion Using Integer Programming. CRYPTO 2016

データの収集 (2)

» 路車間通信によるデータ収集
  > 渋滞情報, 経路情報, 道路の破損状況等
  > センシングデータとその署名を路側器に送信

相反する要件

» 余計な情報は提供したくない
  > 自宅から職場まで. 仕事帰りに寄り道等
  > 署名作成者は一意的に決まる (同じ車かどうか常にトレースされる (リンカブル))

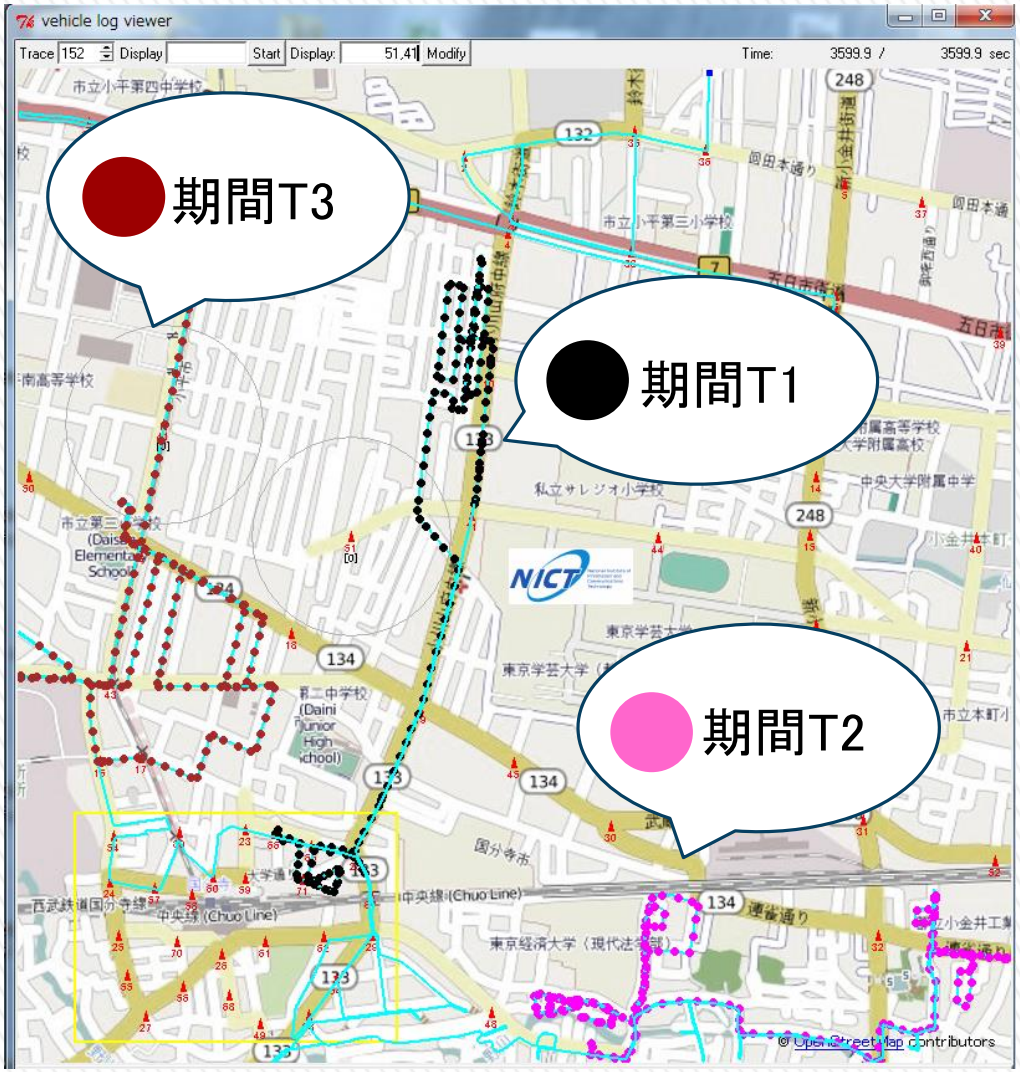
» 匿名署名 (グループ署名) を利用すると・・・
  > 同じ車が作成した署名か否かを暗号的に隠す
  > 経路情報 (どことどこが繋がっている) や速度情報 (ある車がどれだけ動いた), 渋滞情報 (ある車がどのくらいで渋滞を抜けるのか) 等が取得できない

» プライバシーを考慮しつつ, リンカブルな情報は取得できることが望ましい

期間に依存した匿名性を持つグループ署名

[Emura-Hayashi, IEEE Trans. Vehicular Technology 2018]

- 同じ期間に作成された署名はリンク付け可能
  - 同じ車が作成した署名であることを検証可能
  - 車両の特定まではしない
- 期間をまたげば同じ車かどうかの判定が不可能に



データの保管・管理

» 保管データの暗号化による保護
  > 暗号文からデータに関する情報は漏れない
  > クラウドストレージ等に安全に保管

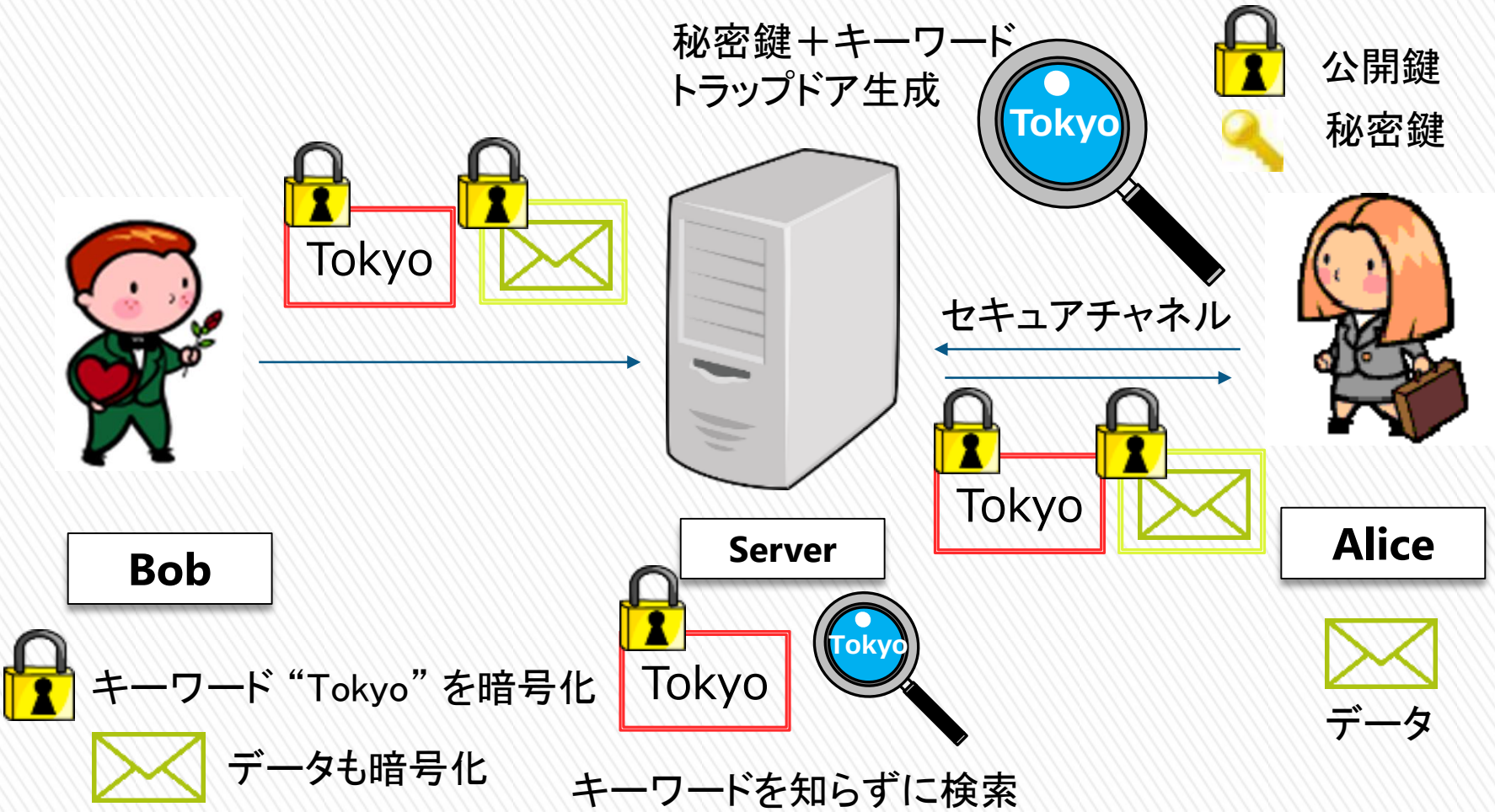
相反する要件

» 使用する際に検索を行いたい
  > 暗号文に対する検索が必要
  > インデックスをつければよい？（キーワード, データの暗号文）
  > この暗号化されたデータはこのキーワードに関連するという情報が漏れる

» (公開鍵) 検索可能暗号
  > 暗号化されたキーワードが検索可能に(キーワードの暗号文, データの暗号文)
  > トラップドア情報を用いて検索者がキーワードを知ることなく検索が可能

検索可能暗号

[Boneh-Crescenzo-Ostrovsky-Persiano: EUROCRYPT 2004]



» セキュアチャネルが必要（トラップドアがあると誰でも検索可能）

» データ秘匿に関して“CCA安全性”が保証されない

検索可能暗号

[Boneh-Crescenzo-Ostrovsky-Persiano: EUROCRYPT 2004]

電子政府における調達のために参照すべき暗号のリスト
（CRYPTREC暗号リスト）

平成25年3月1日
総務省
経済産業省

電子政府推奨暗号リスト

暗号技術検討会'及び関連委員会（以下、「CRYPTREC」という。）により安全性及び実装性能が確認された暗号技術[2]について、市場における利用実績が十分であるか今後の普及が見込まれると判断され、当該技術の利用を推奨するもののリスト。

技術分類		名称
公開鍵暗号	署名	DSA
		ECDSA
		RSA-PSS(注1)
		RSASSA-PKCS1-v1_5(注1)
	守秘	RSA-OAEP(注1)
	鍵共有	DH
		ECDH
共通鍵暗号	64ビットブロック暗号(注2)	3-key Triple DES(注3)
	128ビットブロック暗号	AES
		Camellia
	ストリーム暗号	KCipher-2

CRYPTREC 電子政府推奨暗号リスト
http://www.cryptrec.go.jp/list/cryptrec-ls-0001-2016.pdf

データ守秘目的

RSA-OAEP

CCA安全な暗号

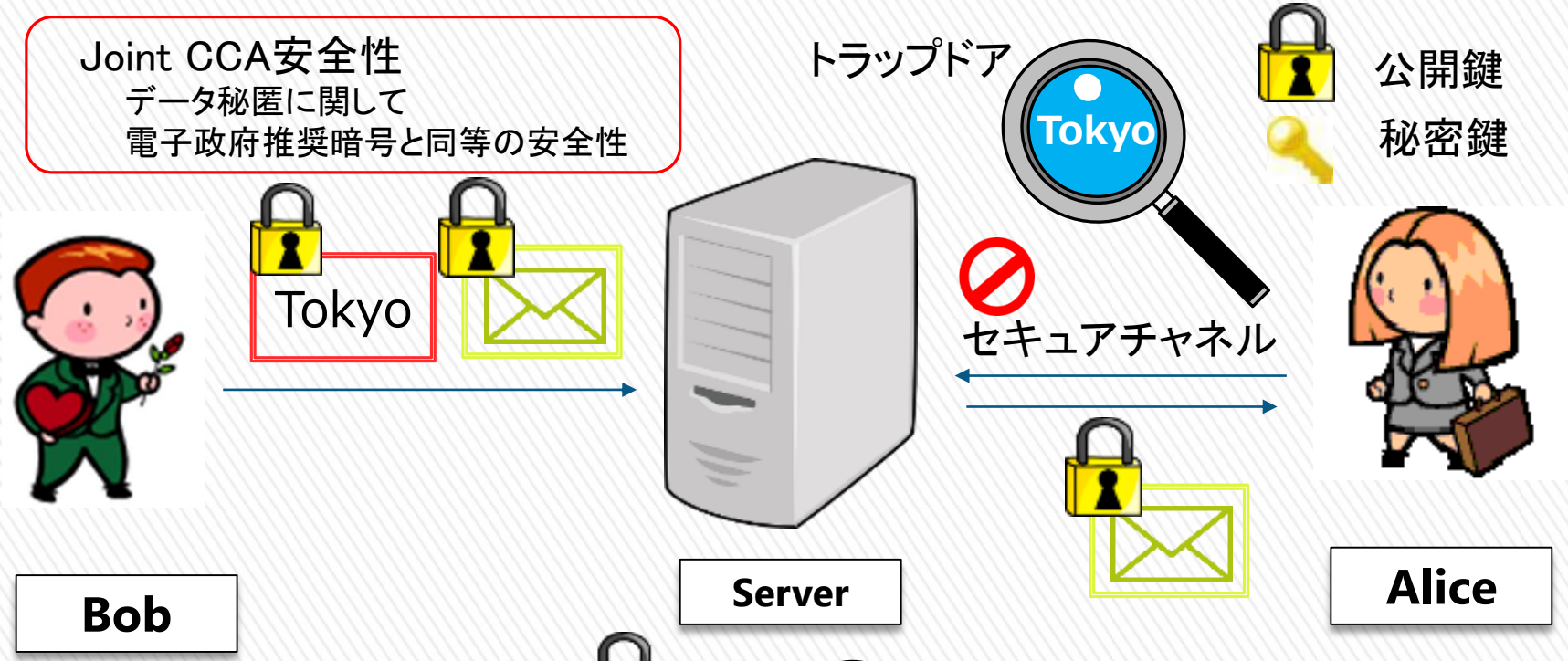
» セキュアチャネルが必要（トラップドアがあると誰でも検索可能）

» データ秘匿に関して“CCA安全性”が保証されない

検索可能暗号の安全性強化

[Suzuki-Emura-Ohigashi, Journal of Medical Systems 2019]

Joint CCA安全性
データ秘匿に関して
電子政府推奨暗号と同等の安全性



» セキュアチャネルの撤廃

» データ秘匿に関する“CCA安全性”の保証

データの解析

» 保管データの暗号化による保護
  > 暗号文からデータに関する情報は漏れない
  > クラウドストレージ等に安全に保管

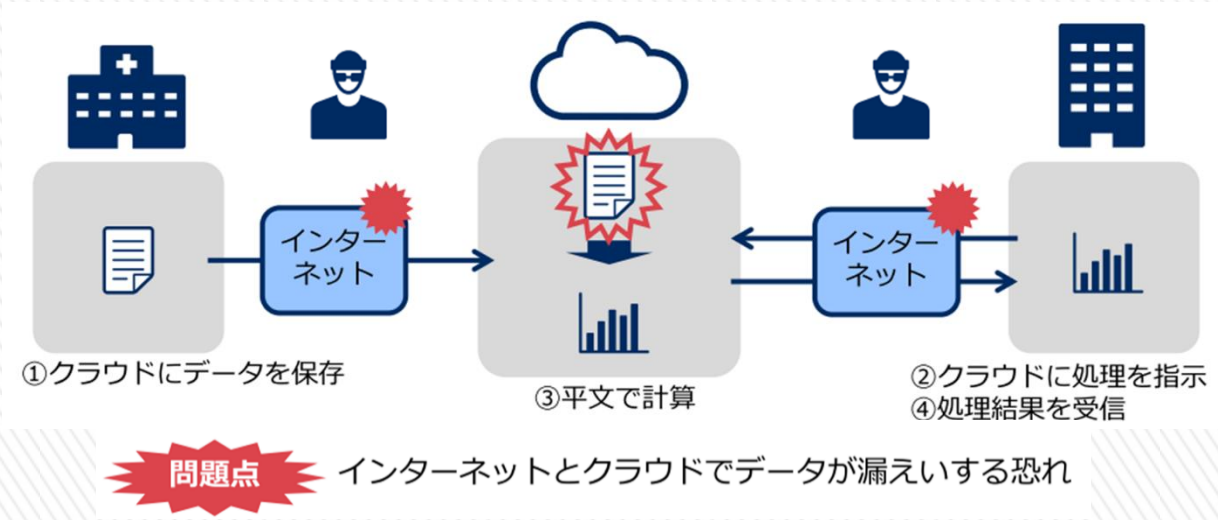
相反する要件

» データの解析のためには一度暗号文を復号する必要がある

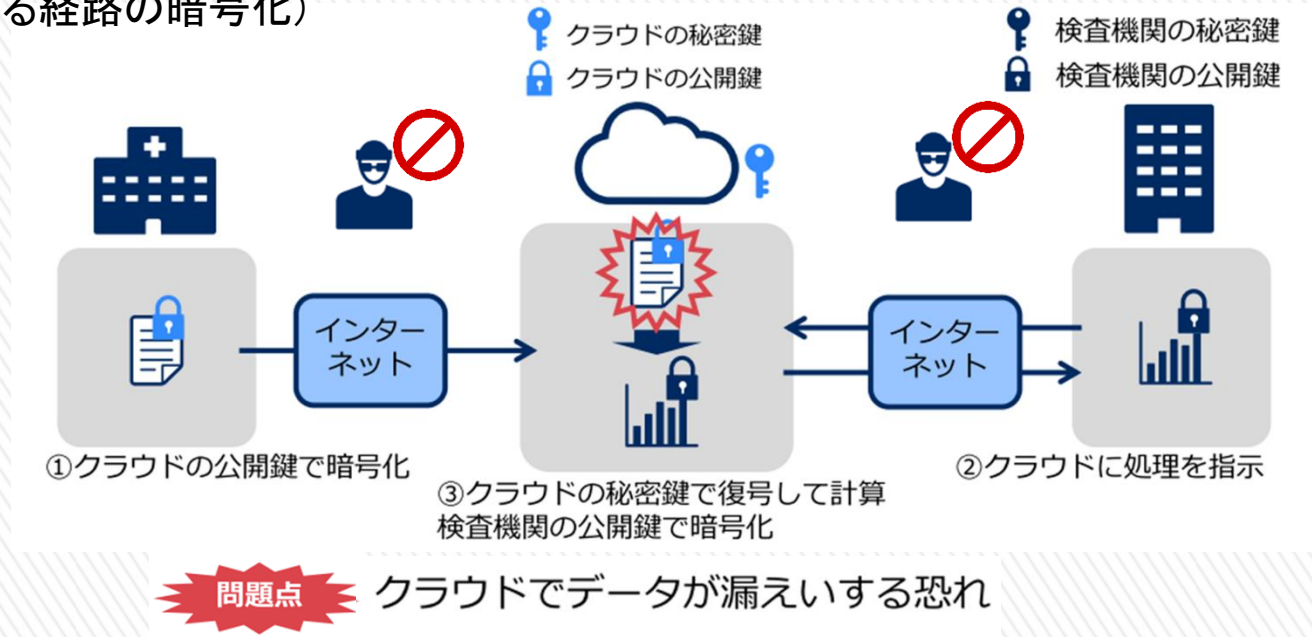
» 準同型暗号を用いて“暗号化したまま”データ解析

“暗号化したまま” データ解析 ~ 既存技術との違い~

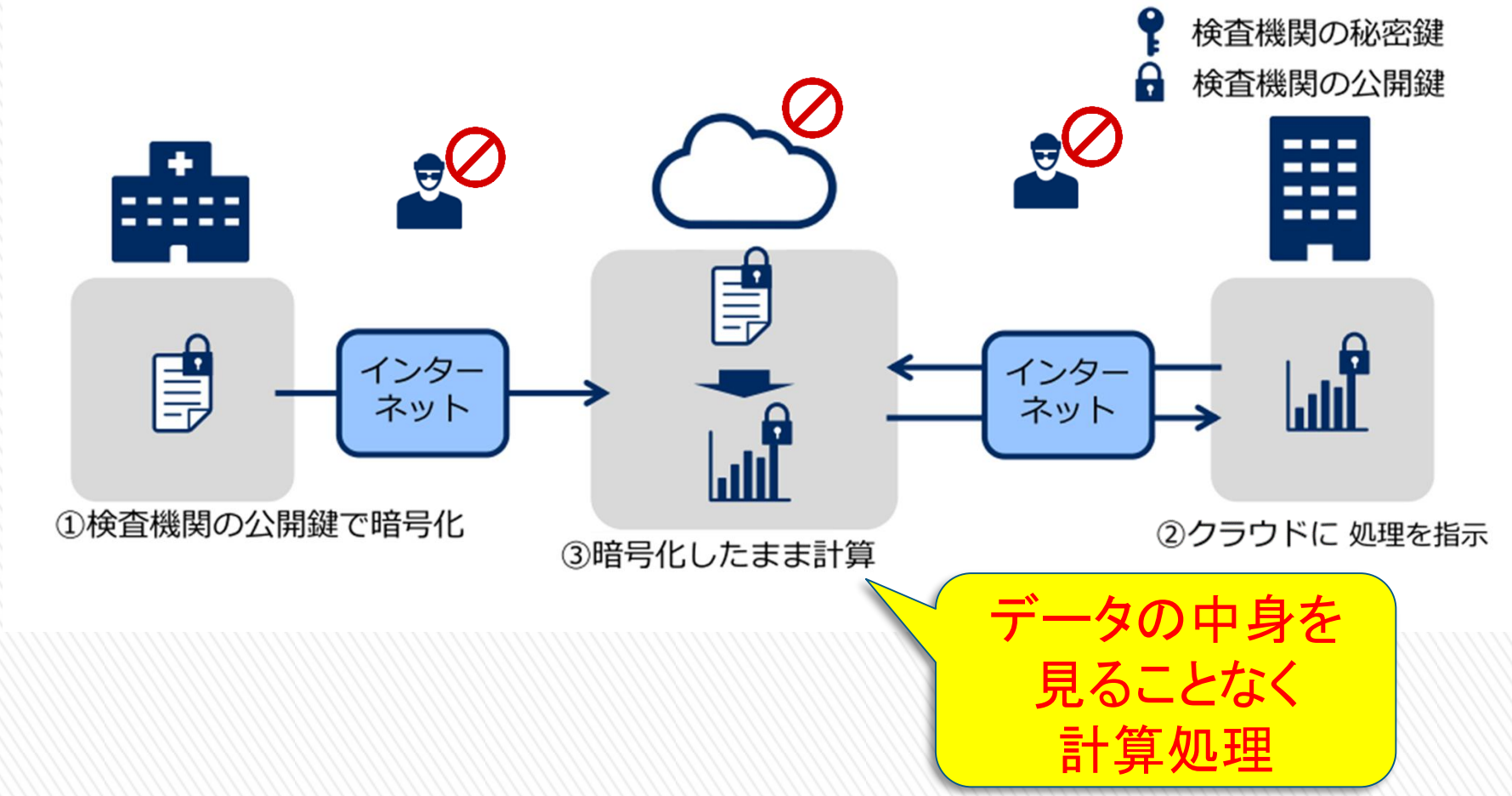
1. ノーガード



2. 既存技術
（SSL/TLSによる経路の暗号化）



“暗号化したまま” データ解析



例 投票: 1 もしくは0の暗号文. 投票結果を知ることなく集計が可能

“暗号化したまま” データマイニング

銀行（複数）

銀行ごとの
金融取引データ

データの暗号文

統計値の
暗号文

データの中身を
見ずに統計処理

JST CRESTプライバシー保護データ解析技術の社会実装
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

“暗号化したまま” データマイニング

– 各銀行が持つデータを統合
大きなデータベースを作成
（集合知）
– パーソナルデータ解析の外注
新たな産業創出の種!!!

準同型暗号の安全性強化 (1)

[Emura-Hanaoka-Nuida-Ohtake-Matsuda-Yamada, PKC2013, DCC 2018]
[SCIS2012イノベーション論文賞]

» 公開鍵暗号として望ましい安全性
  > CCA安全性
    + 暗号化された中身が改変できない
» 準同型性暗号の安全性
  > CCA安全性の実現不可能性
    + 暗号化したまま演算が可能という性質と矛盾

» 鍵付き準同型暗号 (Keyed-Homomorphic Encryption)
  > 公開鍵, 秘密鍵 (復号用), 秘密鍵 (準同型演算用)
  > 準同型演算用秘密鍵では復号できない
  > 準同型演算用秘密鍵を持たない攻撃者に対してCCA安全性を保証

電子政府における調達のために参照すべき暗号のリスト
(CRYPTREC暗号リスト)

平成25年3月1日
総務省
経済産業省

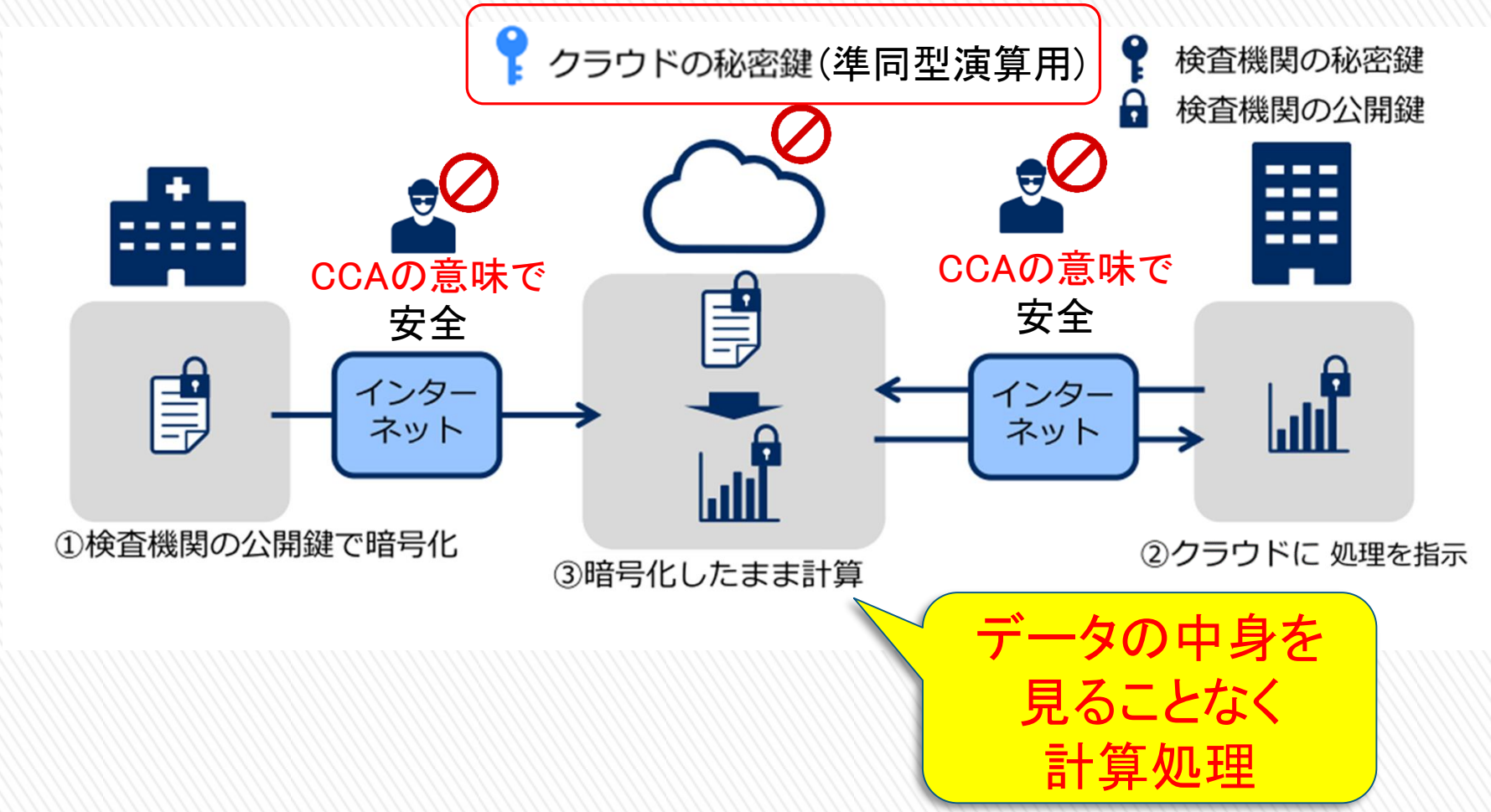
電子政府推奨暗号リスト

暗号技術検討会[1]及び関連委員会(以下、「CRYPTREC」という。)により安全性及び実装性能が確認された暗号技術[2]について、市場における利用実績が十分であるか今後の普及が見込まれると判断され、当該技術の利用を推奨するもののリスト。

技術分類		名称
公開鍵暗号	署名	DSA
		ECDSA
		RSA-PSS(注1)
		RSASSA-PKCS1-v1.5(注1)
	守秘	RSA-OAEP(注1)
	鍵共有	DH
		ECDH
共通鍵暗号	64ビットブロック暗号(注2)	3-key Triple DES(注3)
	128ビットブロック暗号	AES
		Camellia
	ストリーム暗号	KCipher-2

CRYPTREC 電子政府推奨暗号リスト
http://www.cryptrec.go.jp/list/cryptrec-ls-0001-2016.pdf

“暗号化したまま” データ解析 ~鍵付き準同型暗号編~



仮に準同型用鍵がクラウドから漏洩しても，通常の準同型暗号と同等の安全性を保証

準同型暗号の安全性強化 (2)

[Emura-Hayashi-Kunihiro-Sakuma, AsiaCCS 2017]

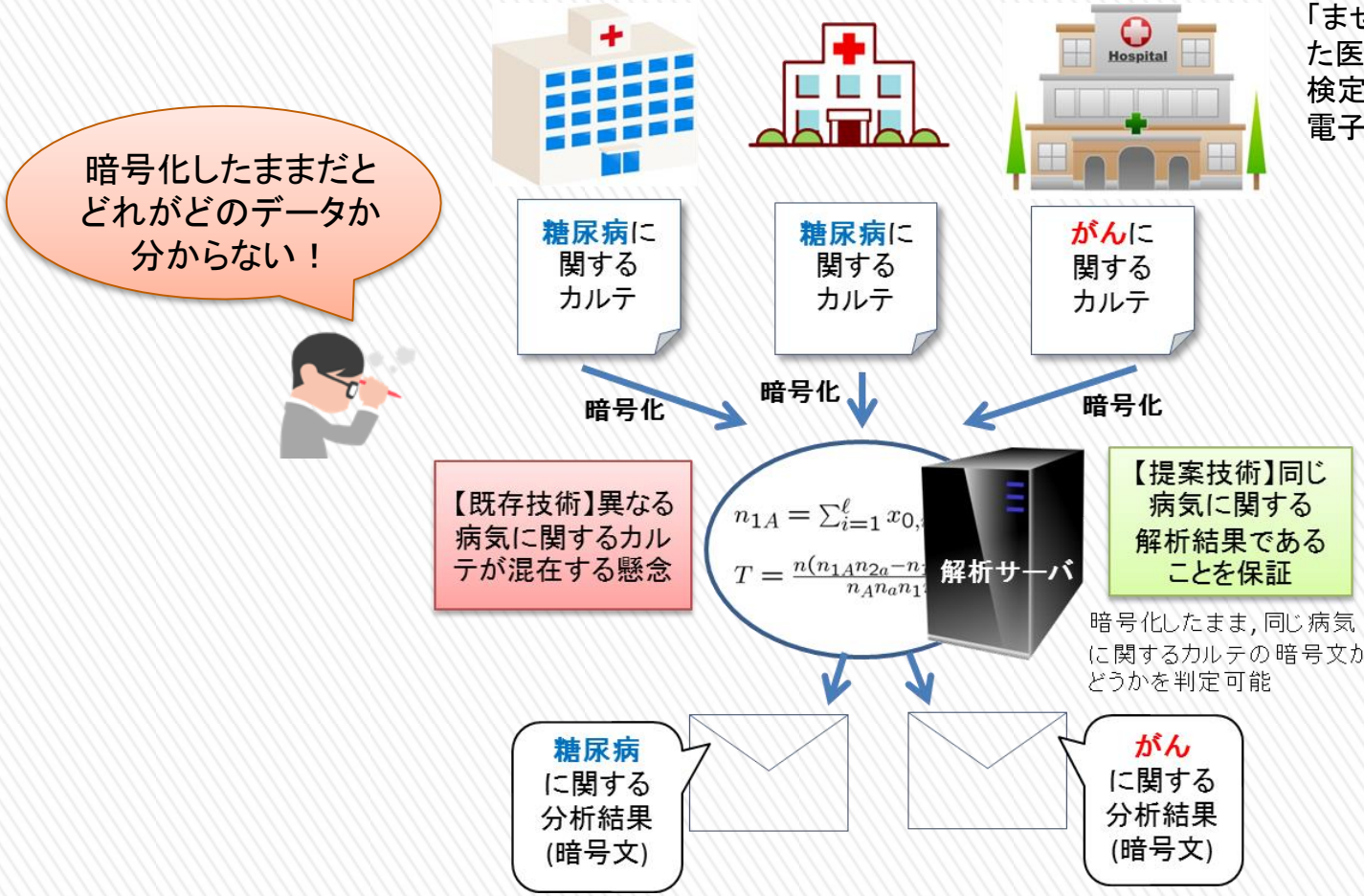
[CSS2016最優秀論文賞][2017年度山下記念研究賞]

» まぜるな危険準同型暗号

> 鍵付き準同型暗号＋検索可能暗号

2018.7.18 プレスリリース

江村, 林, 陸, 盛合, 佐久間, 山田, 「まぜるな危険準同型暗号を用いた医療データに対する χ^2 独立性検定」, 情報セキュリティ研究会, 電子情報通信学会



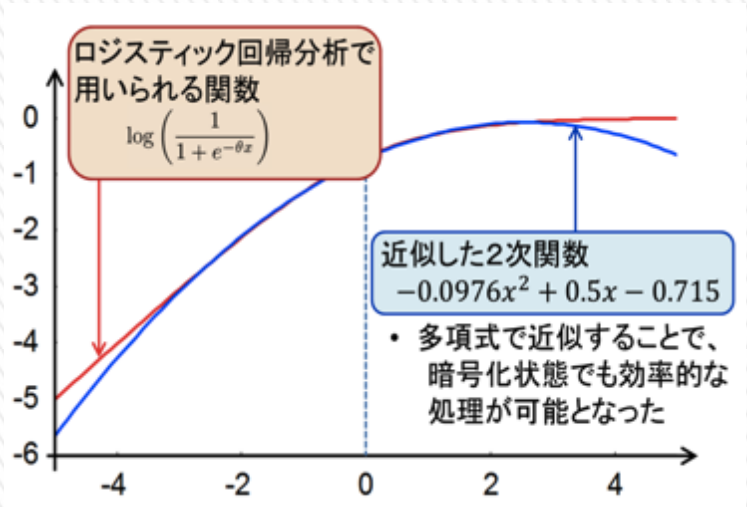
暗号化したままビッグデータ分類

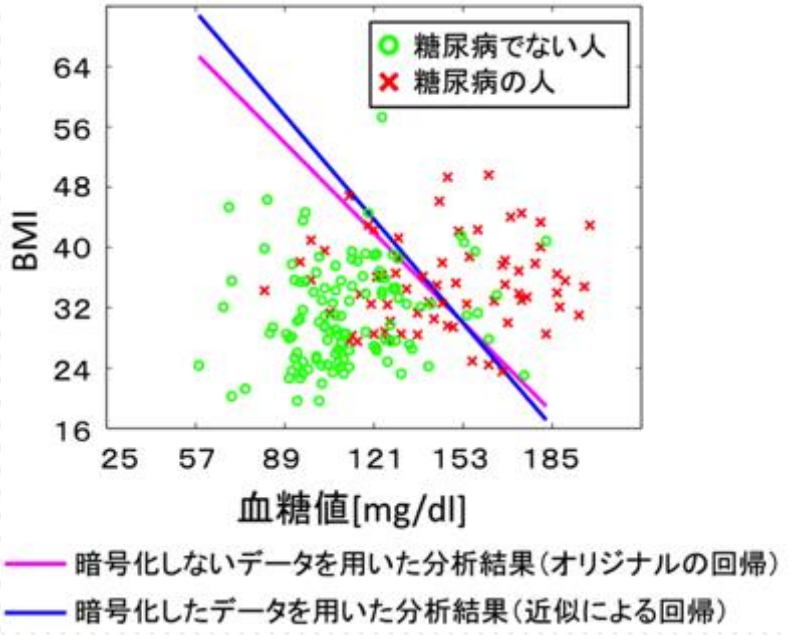
[Aono-Hayashi-Phong-Wang, IEICE Trans. 2016]

» 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)





プライバシー保護ディープラーニング

[Phong-Aono-Hayashi-Wang-Moriai, IEEE Trans. Information Forensics and Security 2018]

» 組織が持つデータを外部に開示することなく
深層学習を行うプライバシー保護深層学習システム

オープンデータセットを用いた実用性検証

◆欧州のクレジットカード取引データ

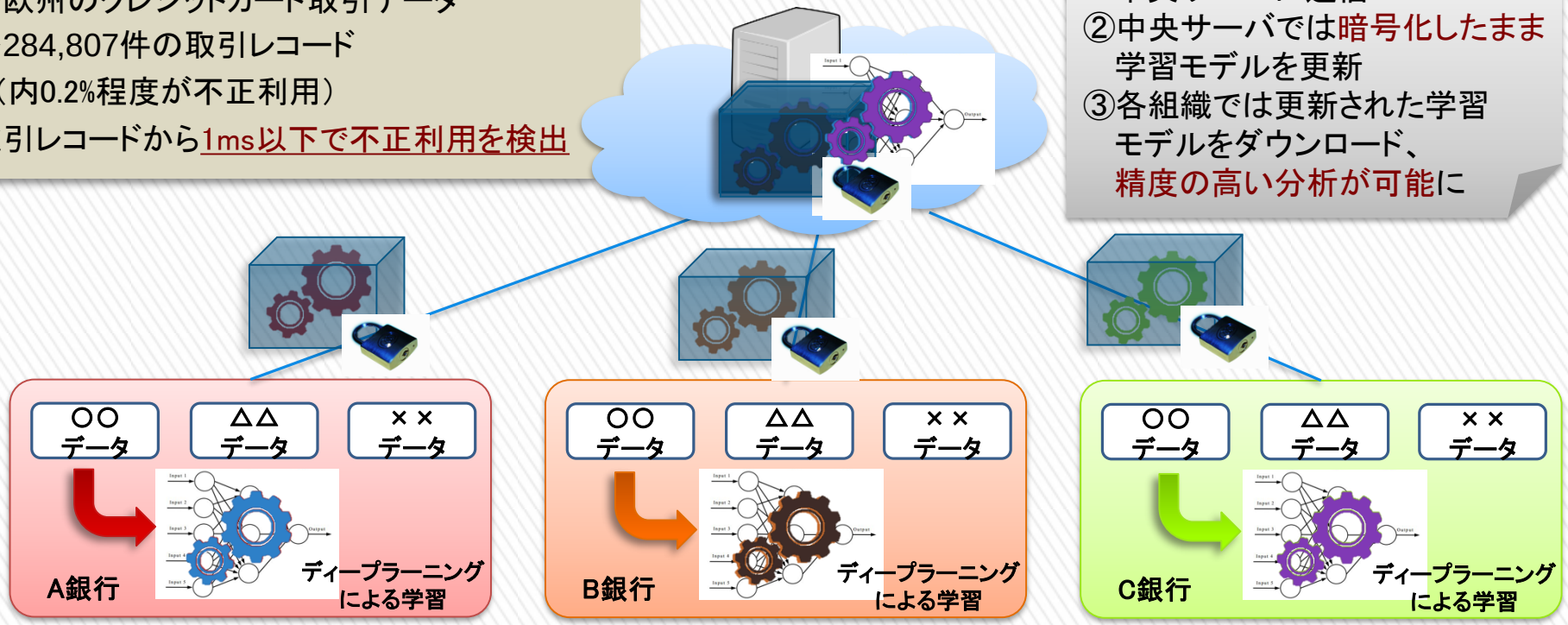
◆284,807件の取引レコード
（内0.2%程度が不正利用）

取引レコードから1ms以下で不正利用を検出

①各組織から学習済モデルの
パラメータを暗号化して
中央サーバに送信

②中央サーバでは暗号化したまま
学習モデルを更新

③各組織では更新された学習
モデルをダウンロード、
精度の高い分析が可能に



複数組織で連携した分散協調型の深層学習

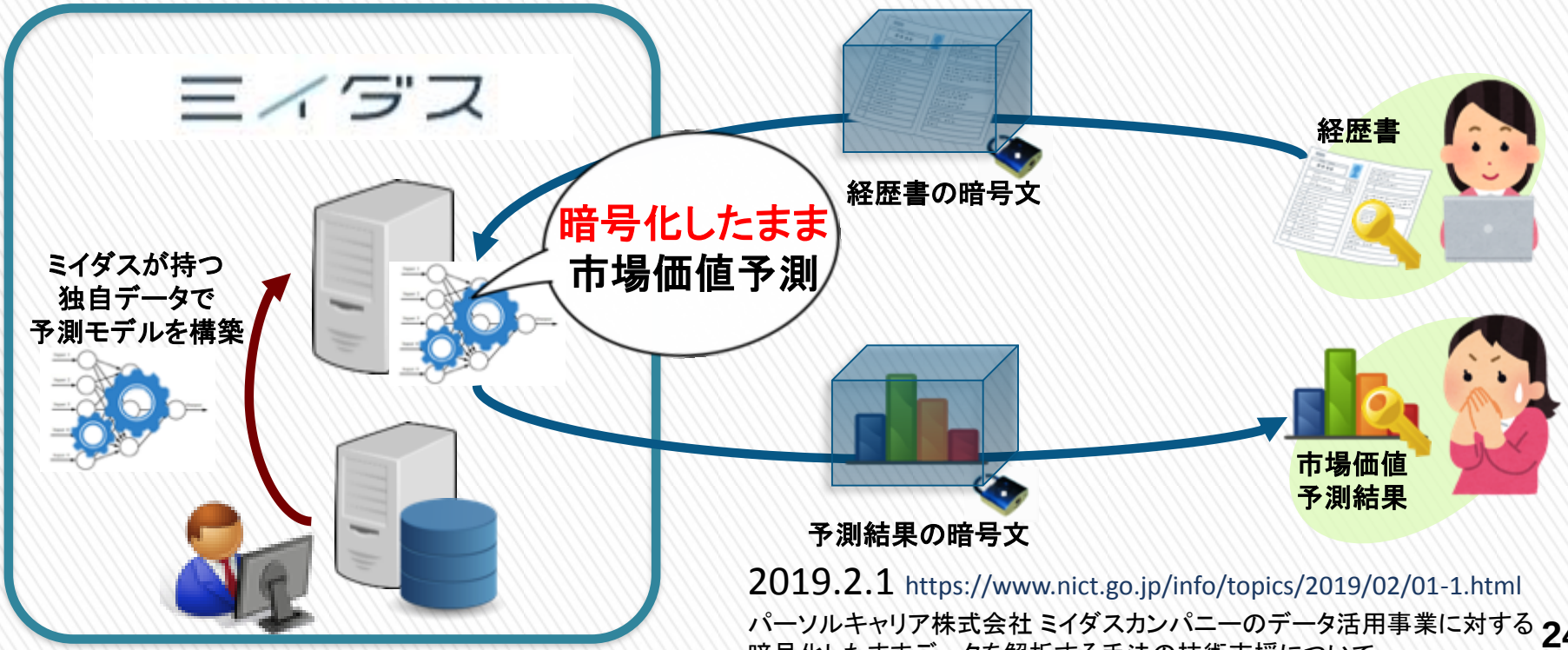
データ活用事業に対する技術支援

» ミイダス株式会社

➢ ICTを駆使した先進的な転職サービスを提供

例) プライバシー保護市場価値予測サービス

➢ ミイダスがユーザデータを一切見ることなく、ミイダスが持つ予測モデルを適用



2019.2.1 https://www.nict.go.jp/info/topics/2019/02/01-1.html

パーソルキャリア株式会社 ミイダスカンパニーのデータ活用事業に対する暗号化したままデータを解析する手法の技術支援について

秘密計算ハッカソン (2019.11.7-8)

» データ利活用の促進に向けて、KPMG Ignition Tokyo 主催の準同型暗号を用いた秘密計算ハッカソンに協賛
  > 当日は秘密計算に関するチュートリアルを実施

まとめ

- 機微データの安全な利活用に向けたセキュリティ基盤研究室の取り組み
  - 収集, 保管・管理, 解析
  - 個々の目的用にはある程度の成果
- 一般ユーザにも使いやすい暗号技術へ
  - 一気通貫な方式の提案：収集したデータを保管・管理し, 解析まで行う
- 量子コンピュータ時代でも安全な暗号技術へ
  - 楕円曲線/ペアリングベースから格子ベース（耐量子）へ