プライバシー強化技術の概説と動向

（Privacy Enhancing Technologies : PETs）

2021年11月22日

株式会社日本総合研究所
先端技術ラボ

＜本件に関するお問い合わせ＞
近藤浩史 エキスパート（kondo.hirofumi@jri.co.jp）**間瀬英之** シニア・リサーチャー（mase.hideyuki@jri.co.jp）**森毅**（mori.takeshi@jri.co.jp）

目次

1．1　データ利活用とプライバシー保護の重要性

- デジタル化によるデータ量の増加とAIの進化によって、データ利活用への期待が高まっている。
- 一方、プライバシー情報を利活用する側においては、プライバシー侵害の懸念等から、収集・蓄積したデータをどのように保護・活用していくかが課題に。

データ利活用の広がり

日常生活のデータ化

- スマートフォンの普及やIoTなどの技術的進展により、インターネット閲覧・購買履歴など、１日のほぼ全ての行動がデータ化されつつある。それに伴い、**利活用できるデータ量も増大。**
- 総務省「令和2年情報通信白書」によれば、全世界の月間IPトラフィック（*1）は、2022年までに396エクサバイト（*2）に達し、2017年からの5年間で3倍に増加見込み。今後、データ流通量はさらに伸びると予想される。

（*1）一定時間内に通信回線を経由してやりとりされるデータ量
（*2）1エクサバイトは2^{60}バイト

情報通信技術やクラウドの進展

- 技術進展により、多種多様なデータの蓄積が容易・安価となり、さらに、AI技術の進展により、蓄積したデータを分析し、**より広く深い消費者ニーズを発見・新サービス提供が可能**となった。
- このように、データ利活用は進展し、その重要性が年々、高まっている。

プライバシー保護の重要性

- **データ利活用が進むことで、個人にとっては、パーソナライズ化されたレコメンデーションなど、生活の利便性が高まる**といった利点がある。一方で、自らの情報が取得、蓄積、分析されることによるプライバシー侵害が危惧される。
- プライバシー情報を利活用する側においては、連携先や委託先からの漏洩によるプライバシー侵害の懸念、法規制遵守の難しさなどから、**収集・蓄積したデータのプライバシー情報をどのように保護し、活用していくか**が課題となっている。

プライバシー侵害に関するインシデントの例

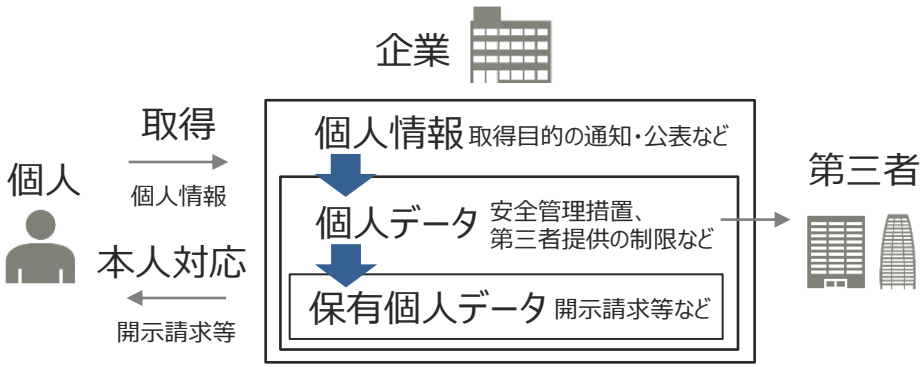
	年	概要
ベネッセコーポレーション	2014	ベネッセコーポレーションの委託先社員が顧客情報2,895万件を不正取得し転売
ケンブリッジ・アナリティカ	2016	Facebookの5,000万人分のユーザデータがデータ分析企業ケンブリッジ・アナリティカに不正利用された
リクルートキャリア（リクナビ）	2019	本人の十分な同意なしに、顧客企業から提供を受けた個人情報（Cookie情報など）とリクナビが保有する個人情報を突合して内定辞退率を予測

（参考）個人情報とプライバシー情報の関係

- 「個人情報」は、氏名や生年月日など個人が識別できるもの、および個人識別符号が含まれるもの。
- 個人情報は、プライバシー情報に包括される関係となる。企業として、個人情報保護について検討する際には、個人情報もプライバシー情報も同様に保護、取り扱い検討することが必要。

個人情報とは

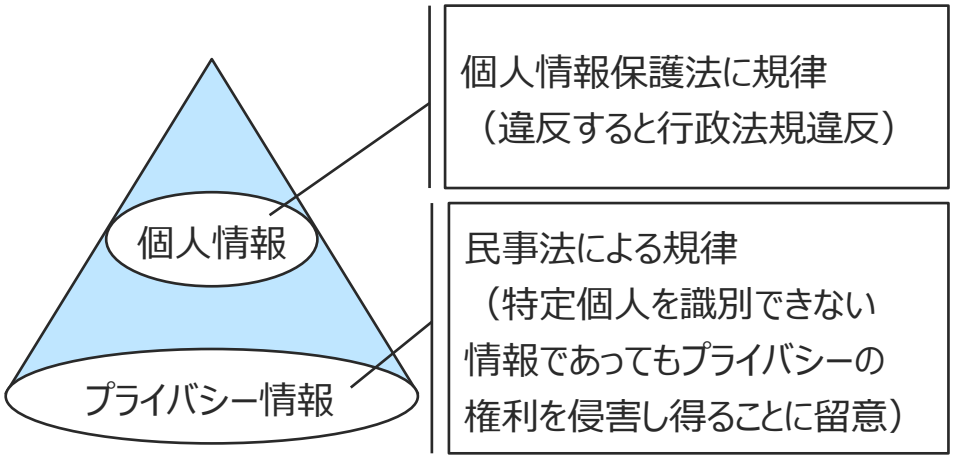
- 「個人情報」は、**氏名や生年月日など個人が識別できるもの、および個人識別符号が含まれる**もの。**①個人情報、②個人データ、③保有個人データに分類**される。
- 企業が①個人情報をもとに、個人情報データベースを作成した場合、そこに含まれる情報は、②個人データとなり、安全管理措置、第三者提供の制限などが対象となる。企業が6カ月を超えて保有している個人データは、原則として、③保有個人データとなる。保有個人データを取り扱っている企業は、本人からの請求に応じて、個人情報を開示、訂正、利用停止等しなければならない。なお、改正個人情報保護法において、6カ月以内に消去する短期保存データは保有個人データに含まれる。



個人情報とプライバシー情報の関係

- 個人情報保護の対象は、プライバシー保護の対象と重複しており、「**個人情報」は「プライバシー情報（プライバシー権に係る情報）」に包括される関係**となる。
- 利活用の観点からは、個人情報保護は、プライバシーの一部を保護するものにすぎない。しかしながら、**企業として、個人情報保護について検討する際には、個人情報もプライバシー情報も同様に保護、取り扱いを検討することが必要**である。

個人情報とプライバシー情報の関係図



参考：渡邊涼介「データ利活用とプライバシー個人情報保護」（青林書院、2020年）

1．2　プライバシー強化技術（Privacy Enhancing Technologies：PETs）の注目

- 世界中でプライバシー保護規制の動きが拡大。この動きは今後も止まらず全世界に広がることが予想される。
- プライバシー強化技術は、プライバシー保護規制の基となるプライバシー原則を実現・強化する技術として注目。

世界で広がるプライバシー保護規制

- デジタルデータは国境に関係なく流通するため、**世界中で法整備の動きが拡大。この動きは今後も止まらず全世界に広がることが予想される。**
- 各国・地域ごとにプライバシー保護の法規制の策定が進んでいるため、データ利活用の際には個別に対応する必要がある。

	法規制
米国	• カルフォニア州：CCPA（*1）（20年施行） CPRA（*2）（23年施行） • バージニア州：VCDPA（*3）（21年成立） • フロリダ州/ニューヨーク州/ワシントン州など：連邦法策定の動き
EU	• GDPR（18年適用開始）
ロシア	• 改正連邦法（21年施行予定）
日本	• 個人情報保護法（20年改正）
中国	• サイバーセキュリティ法（17年施行） • データセキュリティ法草案（21年9月施行） • 個人情報保護法草案（21年11月施行）
韓国	• 個人情報保護法等データ３法改正（20年施行） • 個人情報保護法改正案（21年意見募集）

（*1）CCPA：California Consumer Privacy Act（カリフォルニア州消費者プライバシー法）の略
（*2）CPRA：California Privacy Rights Act（カリフォルニア州プライバシー権利法）の略
（*3）VCDPA：Virginia Consumer Data Protection Act（バージニア州消費者データ保護法）の略

参考：個人情報保護委員会「個人情報保護法　令和２年改正及び令和3年改正案について」

プライバシー強化技術（PETs）

- プライバシー保護規制のベースには、OECD（*4）やISO（*5）のプライバシー原則があり（次頁参考）、ライフサイクル全体を通じてプライバシー原則を適用したデータ保護が必要になる。
- **プライバシー強化技術（PETs：Privacy Enhancing Technologies）は、プライバシー原則を実現・強化するための技術として注目**されている。

プライバシー原則	プライバシー強化技術で実現する機能
同意/個人参加	• 同意・ポリシーに基づく制御の実現 • 開示請求対応
正確性	• データの出元の保証と改ざん防止
透明性/利用等制限	• 履歴管理における改ざん防止と検証 • 廃棄の保証（忘れられる権利の対応）
最小化/収集制限	• 窃取されてもプライバシー漏洩を最小化
情報セキュリティ	• 機密性、完全性、可用性の確保 • 脆弱性対策

（*4）OECD：Organisation for Economic Co-operation and Development（経済協力開発機構）の略
（*5）ISO：International Organization for Standardization（国際標準化機構）の略

（参考）プライバシー保護規則について

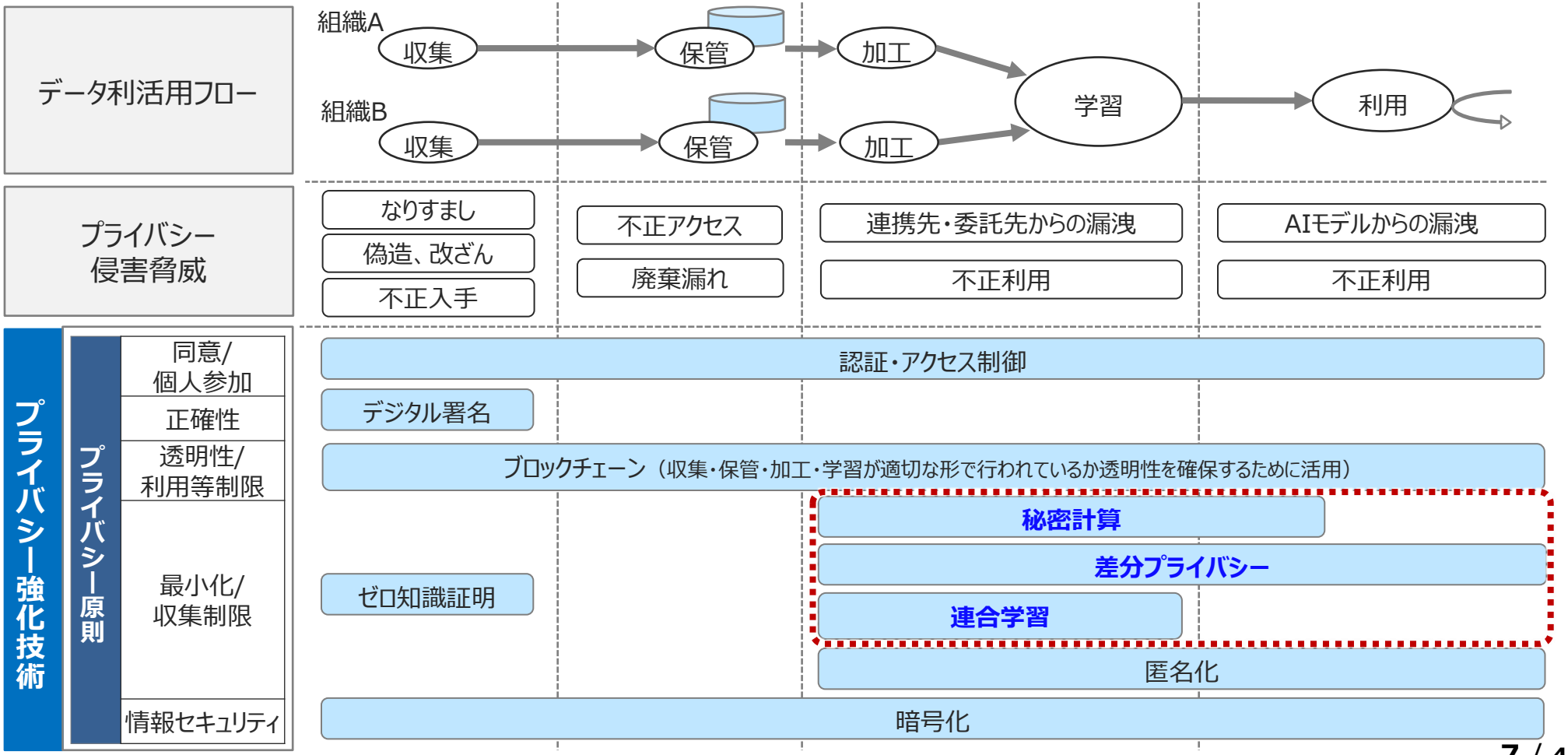
➢ GDPRや個人情報保護をはじめとしてプライバシー保護規制は、「OECD プライバシー8原則（1980年）」と「ISO/IEC 29100 プライバシーフレームワーク（2011年）」の考え方を採用している。

	ISO/IEC 29100	OECD	内容
1	同意及び選択	-	データ処理の同意をとる。分かりやすい選択の仕組みを提供する
2	目的の正当性及び明確化	目的明確化	収集目的を明確にし、データ利用は収集目的に合致させる
3	収集制限	収集制限	適法・公正な手段により、かつ情報主体に通知または同意を得て収集される
4	データ最小化	-	目的に沿ってパーソナルデータへのアクセス、収集するパーソナルデータ、個人の特定や属性推定などの処理を必要最小限にする
5	利用、保持及び開示の制限	利用制限	データ主体の同意がある場合、法律の規定による場合以外は目的以外に利用してはならない
6	正確性及び品質	データ内容	利用目的に沿ったもので、かつ、正確、完全、最新である
7	公開性、透明性及び通知	公開	データ収集の実施方針等を公開し、データの存在、利用目的、管理者等を明示する
8	個人参加及びアクセス	個人参加	自己に関するデータの所在および内容を確認させ、または意義申立を保証する
9	責任	責任	データ管理者は、上記の諸原則を実施するための措置に従う責任を有する
10	情報セキュリティ	安全保護	合理的安全保護措置により、紛失・破壊・使用・修正・開示等から保護する
11	プライバシーコンプライアンス	-	プライバシー保護に関する法令順守

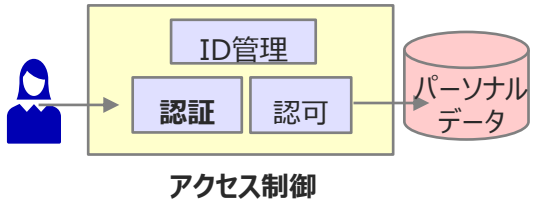
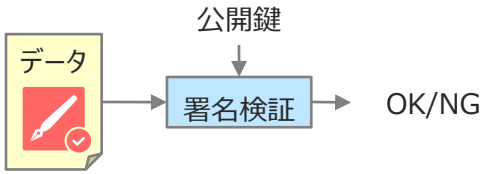
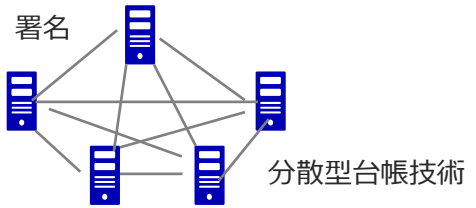
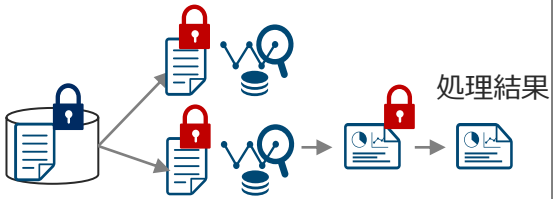
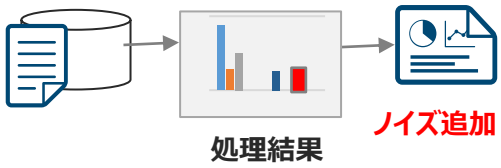
参考：個人情報保護委員会「個人情報の保護に関する法律等の一部を改正する法律案について」など

1．3　プライバシー強化技術の俯瞰図と注目技術

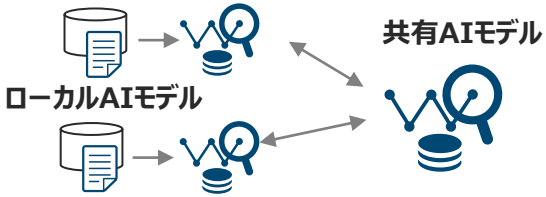
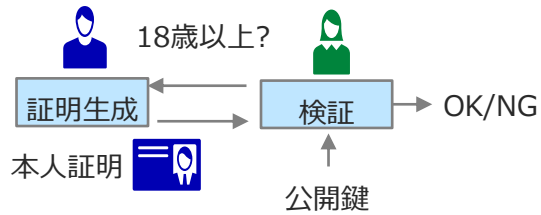
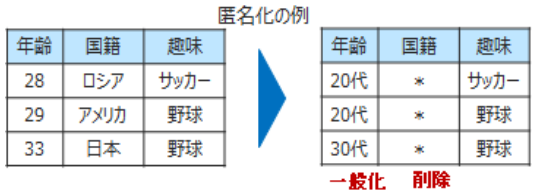
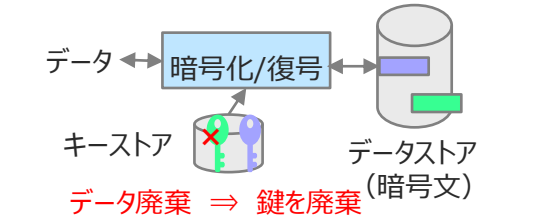
- プライバシー強化技術は、プライバシー保護の原則を実現・強化する技術の総称。個人情報の利用最小化、データセキュリティの最大化、個人に権限を与えることで、プライバシーを保護しながら、データ共有と分析を可能にする。
- 下図にて、データ利活用におけるプライバシー強化技術をプライバシー要件と合わせて俯瞰。学習・利用フェーズで、「最小化」の原則を実現する「秘密計算」「差分プライバシー」「連合学習」に注目（第２章で技術解説）。



（参考）プライバシー強化技術一覧（１／２）

技術名	概要	
認証・ アクセス制御	• ユーザの強固な認証とそれに基づくアクセス制御は、同意に基づく制御や開示要求対応など個人参加の安全性のベースになる。	
デジタル署名	• データ提供者の保有する秘密鍵によるデジタル署名によって、提供者からのデータであることが検証可能。	
ブロックチェーン	• 分散型台帳技術。 • 利活用の各フェーズにおけるデータ処理の履歴管理を改ざん不可、リアルタイムで検証・監視可能な形で実現。	
秘密計算	• 学習データ、クエリデータ、AIモデルなどを秘匿したまま処理できる。 • 連携先や委託先の内部犯行に対しても漏洩を強固に防止可能。	
差分プライバシー	• 利用フェーズにおけるAIモデルや推論結果から学習に用いたプライバシー情報の推測を防ぐために、データにノイズを追加して処理する。	

（参考）プライバシー強化技術一覧（２／２）

技術名	概要	
連合学習	• 組織がローカルに学習したAIモデルのパラメータや更新情報のみを共有し、統合したモデルを学習。 • 機密性の高い情報を組織間で共有することなく、高精度なモデルを作成。	
ゼロ知識証明	• 機微な情報そのものは明かさずに特定の事項を証明する暗号技術。 • 例えば年齢は明かさずに成人かどうかを証明するといった、プライバシーを保護した本人確認などに活用される。	
匿名化	• パーソナルデータを汎化やマスキングすることで、他のデータと紐づけても個人特定が不可能になるように加工できる。	
暗号化	• TLS（*1）などの通信暗号化とストレージやデータベースの暗号化で、秘匿と改ざん検出のデータ保護を実現。 • データは暗号化して保持し、廃棄時にはその秘密鍵を削除することで廃棄を保証する技術は、「crypto shredding」と呼ばれる。	

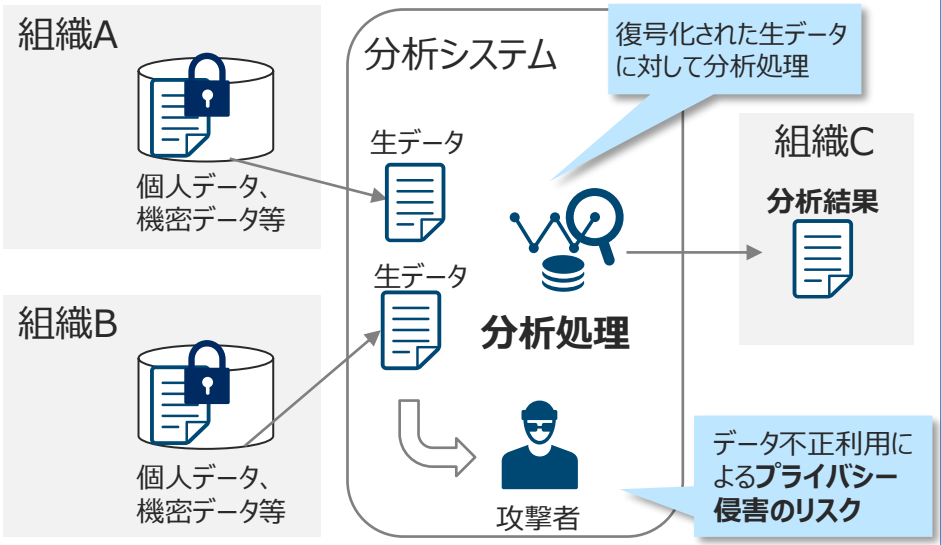
（*1）Transport Layer Securityの略。インターネット上での通信で通信内容を暗号化してやり取りするためのプロトコル

2．1．1　秘密計算の概要

- 従来のデータ分析は、生データを使って分析処理を行っていたため、分析システムからデータが漏洩するリスクやデータの不正利用によるプライバシー侵害の恐れがあった。
- 秘密計算は、複数人によりデータを秘匿化したまま処理する技術である。複数人の中に悪意を持って振舞う者（攻撃者）が存在しても安全性を保つことができ、処理中のデータ漏洩を防止する。

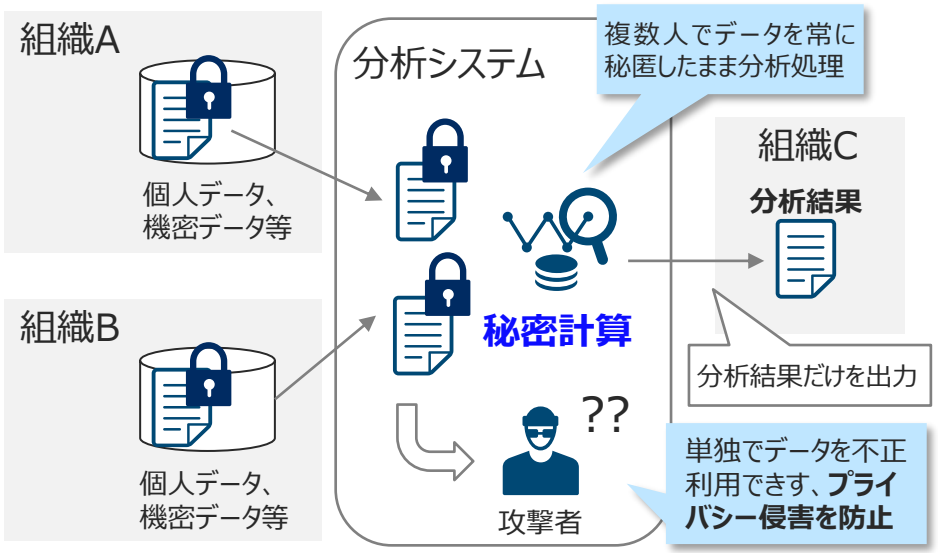
データ分析における従来の課題

- 従来より、個人情報や機密情報等のデータについては、サーバー・データベース内および通信中において暗号化されることが一般的である。一方で、分析処理を行う場合、生データに復号する必要があった。
- 複数組織のデータをもとに分析する場合、分析システムに生データを開示する必要がある。このとき、分析システムに攻撃者が存在すると、不適切な利用等のプライバシー侵害の恐れがあった。



秘密計算による安全なデータ活用

- 秘密計算とは、複数人で**データを秘匿化したまま処理する技術であり**、分析システムに攻撃者が存在したとしても、単独では生データを読み取ることができない。
- 秘密計算を用いることで、複数組織での安全なデータ活用が促進されることが期待される。



2．1．2　秘密計算の種類と特徴

- 秘密計算の実現方式は主に４つある。
- それぞれ特徴が異なるため、用途に合った方式を選ぶことが必要である。

実現方式	説明	応用範囲	ｽﾙｰﾌﾟｯﾄ	分散数	耐ｻｲﾄﾞﾁｬﾈﾙ攻撃	用途例
①秘密分散 (MPC：Multiparty Computation)	• 秘密をいくつかのデータ（シェア）に分散させ、シェアに対し計算を繰り返し実行し結果を得る方式 • 秘密を2カ所以上に分散して計算する。汎用性が高くデータ分析に適している	○	○	2以上	○	• データ分析 • 鍵管理
②Garbled Circuit	• 計算対象の論理回路の全てのゲートを、全ての入力パターンで暗号化することで秘密計算を実現する方式 • 2カ所以上に分散して計算する • 最初に技術が成熟し、鍵管理で多く商用化	×	△	2以上	○	• 鍵管理
③準同型暗号 (Homomorphic Encryption)	• 準同型暗号の暗号文のまま計算処理が可能な性質を利用する方式 • サーバークライアント間の様な二者での秘密計算を実現 • 生体特徴量の照合によく使われる	△	×	2以上	○	• 生体の特徴量照合
④ハードウェア実装	• Intel SGX等の隔離された計算環境「TEE（*1）」を用いて、秘密計算を実現する方式 • 最も処理性能が高いが、サイドチャネル攻撃（*2）に脆弱	◎	◎	無制限	×	• サイドチャネル攻撃が回避できる場合には用途に限定はない

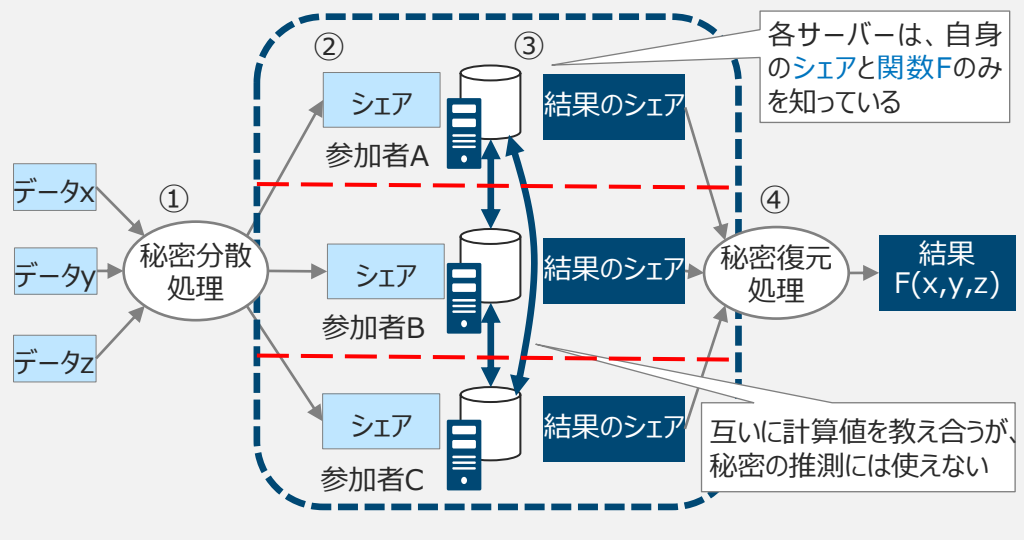
（*1）Trusted Execution Environmentの略　（*2）ハードウェアの動作状況を外から観測することで秘密を抜き出す攻撃

2．1．3　実現方式 - ①秘密分散

- 複数の参加者（マルチパーティ）が互いに協力して、分散計算する仕組み。
- 入力データを秘密分散により複数のシェアに分散し、複数のサーバーで計算を行う。各結果を集めて、計算結果を復元する。各サーバーはシェアしか持たないため、単一のシェアからは秘密は一切わからず、データ漏洩が防止される。

秘密分散を用いた秘密計算の流れ

①入力データx, y, zのそれぞれを複数個のシェアに分散し、参加者に送信（秘密分散処理）
②参加者(サーバー)は各関数F(行いたい処理)とシェアを受け取る
③他の参加者と互いに自身の計算値を教え合い、結果のシェアを得る
④それぞれのシェアから最終的な結果F(x,y,z)を得る(秘密復元処理)

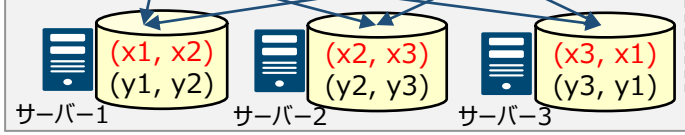


なお、各サーバーは入力・計算中の値・計算結果の全てに関してシェアしか持たず、シェアからは何も知ることができない。しかし、一定数以上のシェアを集めると、秘密を推測できるため、シェアへのアクセス権限を適切に管理することが重要である。

（参考）秘密分散の加算、乗算の実行例（*1）

事前準備：ランダム値の和で表現、2要素ずつ分散

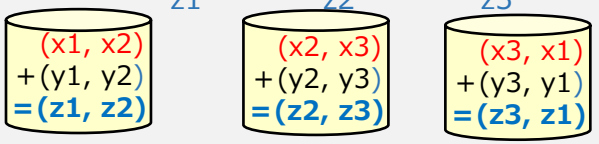
X = x1 + x2 + x3　　Y = y1 + y2 + y3



（例）
X=6
⇒1+2+3
Y=9
⇒2+3+4

加算：各サーバーが保持するシェアを加算

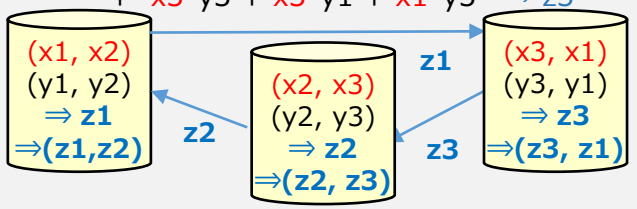
X + Y = (x1 + y1) + (x2 + y2) + (x3 + y3)
z1　z2　z3



（例）
z1=1+2=3
z2=2+3=5
z3=3+4=7
⇒3+5+7=15

乗算：以下のアルゴリズムにより計算。次の乗算へ備える為、データを共有

X·Y = (x1 + x2 + x3)·(y1 + y2 + y3)
= x1·y1 + x1·y2 + x2·y1 → z1
+ x2·y2 + x2·y3 + x3·y2 → z2
+ x3·y3 + x3·y1 + x1·y3 → z3



（例）
z1= 1·2+1·3
+2·2
z2= 2·3+2·4
+3·3
z3= 3·4+3·2
+1·4
⇒9+23+22=54

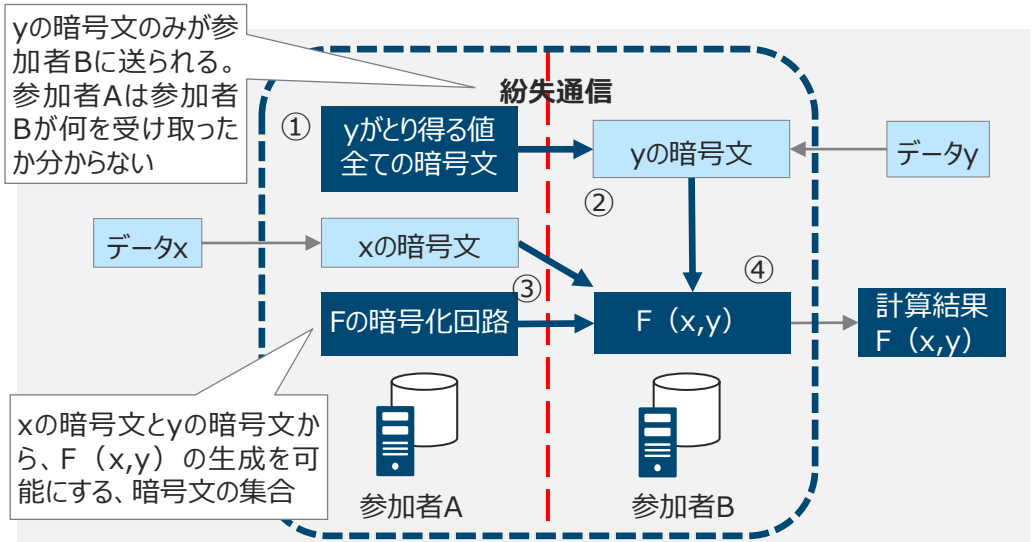
（*1）直感的な説明であり、実際には乱数の要素を入れるなどの仕組みがある

2．1．4　実現方式 - ②Garbled Circuit

- 計算したい関数を回路で表現し、各ゲートの出力の暗号文を暗号文の入力全ての場合（2x2通り）で暗号化して暗号化回路とする。入力の暗号文が与えられると、順に復号して回路の出力を得ることができる。
- 参加者間では暗号文しか交換されないため、関数の計算に全ての入力を知る参加者が必要なく、データ漏洩が防止される（2人の参加者によるものが代表的）。

Garbled Circuitによる秘密計算の流れ

① 参加者Aは、xとyが取り得る全ての値の暗号文を準備する
② 両参加者は、紛失通信（*1）と呼ばれる暗号プロトコルを実行し、参加者Bはyの暗号文を参加者Aの準備した暗号文の集合から選び取る
③ 参加者Aはxの暗号文と、関数Fを暗号化した回路を生成し、参加者Bに送る
④ 参加者Bはxの暗号文とyの暗号文を、Fの暗号化回路に入力しF（x,y）を得る



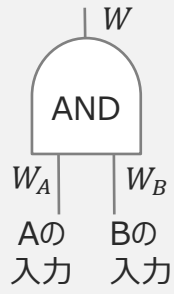
（*1）Oblivious Transferと呼ばれる暗号プロトコル。単純な形態では、送信者が2つのデータを送り、受信者がどちらか1つのデータのみを受け取る場合を考える。このとき、送信者は受信者がどのデータを受信した分からず、受信者も受け取らなかったデータについて何も知ることができないという性質を持つ。

（参考）計算例：ANDゲート

- 参加者Aと参加者Bが1bitのデータを保持し、お互いの生データを公開せずにANDを計算するという最も簡単な例
- 以下では参加者Aが0、参加者Bが1を保持しているとする

手順

① 参加者Aは、参加者Aと参加者Bの入力としてあり得る暗号文w_j^iを生成する（jは参加者AorB、iは入力値0or1を表す）
② 参加者Aはw_B^iを紛失通信によって送信する。参加者Bは保持するデータに従って必要なw_B^1のみを受信する。参加者Aは参加者Bが何を受け取ったか分からないことがポイントである
③ 参加者Aは暗号化回路（*2）を生成し、参加者Aの暗号文w_A^0と合わせて参加者Bに送信する
④ 参加者Bはw_A^0、w_B^1、暗号化回路から計算結果0を得る



W_A	W_B	W
0	0	0
0	1	0
1	0	0
1	1	1

ANDの真理値表

生成

暗号文		暗号化回路(*2)
W_A	W_B	
w_A^0	w_B^0	$Enc_{w_A^0}(Enc_{w_B^0}(0))$
w_A^0	w_B^1	$Enc_{w_A^0}(Enc_{w_B^1}(0))$
w_A^1	w_B^0	$Enc_{w_A^1}(Enc_{w_B^0}(0))$
w_A^1	w_B^1	$Enc_{w_A^1}(Enc_{w_B^1}(1))$

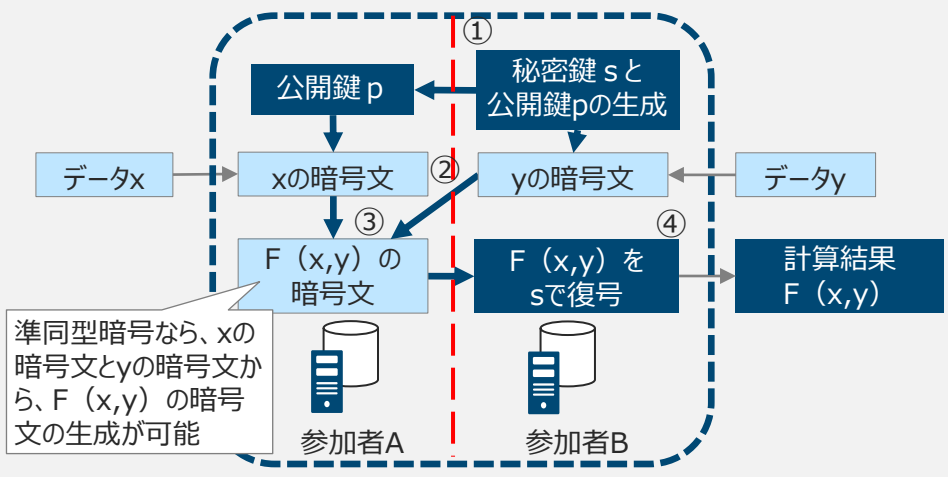
（*2）Enc_x（y）はyを鍵xで暗号化した暗号文を示す

2．1．5　実現方式 - ③準同型暗号/④ハードウェア実装

- 準同型暗号：二人の参加者によるものが代表的。両者の入力を公開鍵で暗号化して、暗号文のまま処理して結果の暗号文を得る。これを、秘密鍵を持つ参加者のみが復号して結果を得る。参加者間では暗号文しか交換されないため、データ漏洩が防止される。
- ハードウェア実装：TEE（Trusted Execution Environment）などの隔離された計算環境に、参加者から入力を集め、ここで処理して結果を参加者に返す。参加者はTEEの内部にアクセスできないため、データ漏洩が防止される。ただし、TEEはサイドチャネル攻撃に弱く、TEEからデータが漏れることもある。

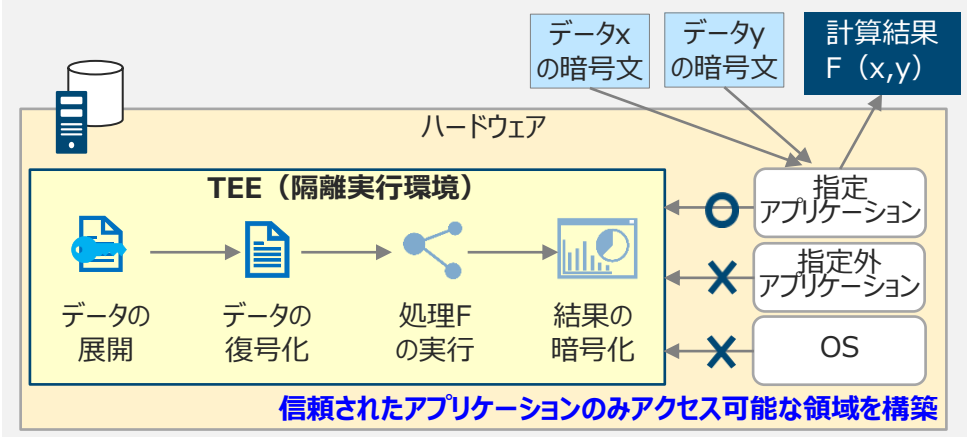
準同型暗号による秘密計算の流れ

① 参加者Bは、秘密鍵と公開鍵を生成して、公開鍵のみを参加者Aに送る
② 参加者Aは公開鍵pでxの暗号文を生成、参加者Bは公開鍵pでyの暗号文を生成して参加者Aに送る
③ 参加者Aはxの暗号文とyの暗号文から、F（x,y）の暗号文を生成して、参加者Bに送る
④ 参加者Bは、秘密鍵sを用いてF（x,y）を復号する



ハードウェア実装の概要

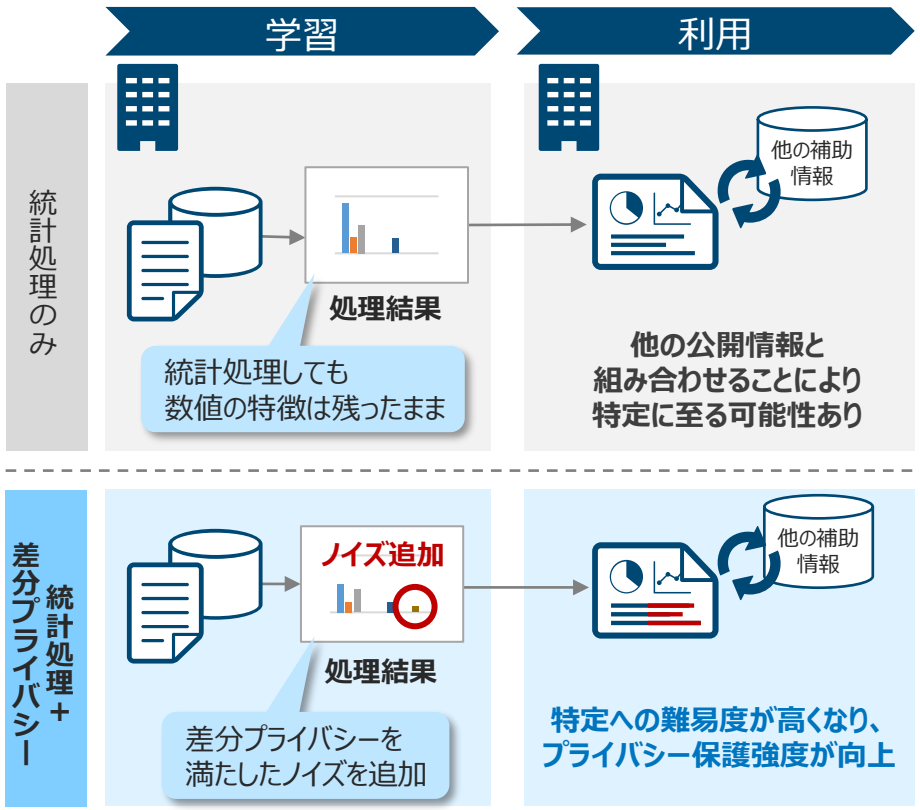
- ハードウェア上に隔離実行環境（TEE:Trusted Execution Environment、Enclave等と呼ばれる）を構築し、その上で計算を実行することでデータ保護を図る。
- TEEには限定された方法でしかアクセスできず、特権が侵害されたOSからも、内部の秘密や完全性を保つ。
- 計算参加者は各自データをTEEに入力し、計算結果を受け取る。参加者は許されたデータや計算結果のみにアクセスでき、決められた操作しかできない。



2．2．1　差分プライバシー（Differential Privacy）の概要

- 一見安全に見えるデータ（統計情報や匿名化データなどの処理結果）でも複数の情報を組み合わせることで個人の情報を特定できる可能性がある。
- 差分プライバシー（Differential Privacy、以降DPと略す）では、処理結果にあるノイズ（次頁詳細）を加えることで個人の情報の特定を防ぐことができる。

従来の統計処理と差分プライバシー



（参考）複数情報の組み合わせによる推測の例

- あるデータベースから平均年齢を計算して公開したとき、それに近いデータベースを持っている攻撃者がいた場合、元のデータベースの個人の年齢を特定されてしまう。

名前	年齢
Alice	10
Bob	25
Carry	**22**

平均 **19**

Carryの年齢は
22歳だと特定

攻撃者
外部情報

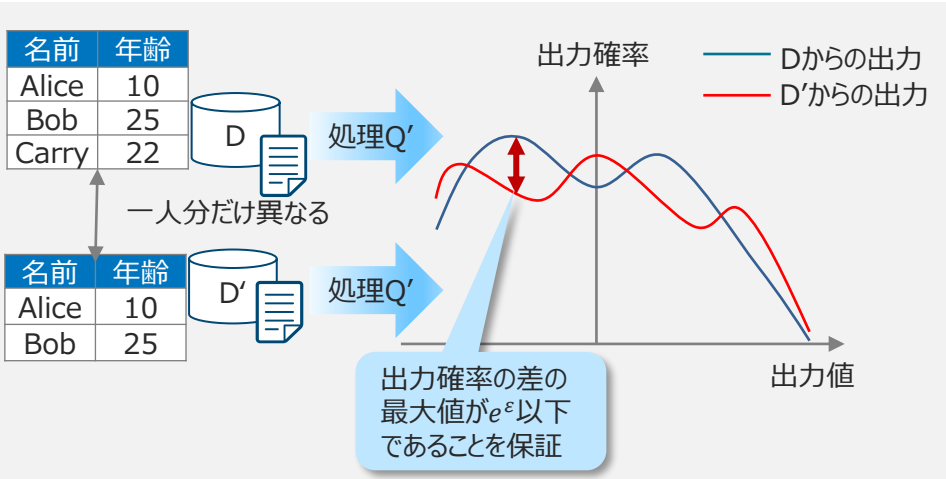
名前	年齢
Alice	10
Bob	25

2．2．2　差分プライバシーの定義とその解釈

- 差分プライバシーは、確率論の識別不能性（*1）に基づくプライバシー保護の安全性の尺度である。
- ある処理から得られた結果が、どのデータベースから導出されたかの特定を困難にすることによって、プライバシー保護を保証する。

差分プライバシーの定義

- 個人ごとのデータを記録し、各々に含まれるデータが一人分だけ異なる（隣接する）2つのデータベースをDとD′、データベースに対する処理をQ′とする。なお、処理Q′には確率的な要素が含まれるものとする。
- 2つのデータベースに対する処理Q′の結果がある値になる確率の対数差を考える。この対数差が、たかだかεに抑えられるとき、処理Q′はε-差分プライバシーを満たすという（*2）。



（*1）直感的には、2つの確率分布を見分けようとしたときに、その見分けがどのくらい困難かを定義するもの

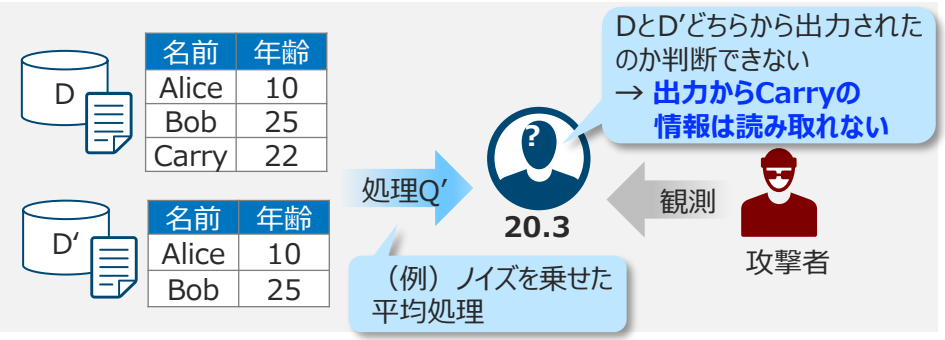
（*2）厳密には任意の隣接するD,D′の組に対して以下の式を満たす場合、Q′はε-DPを満たすという。Sは処理Q′の出力空間の任意の部分集合である。

$$\Pr[Q'(D) \in S] \leq e^{\varepsilon} \cdot \Pr[Q'(D') \in S]$$

差分プライバシーの定義の解釈および利点

差分プライバシーの定義の解釈

- ノイズを乗せた処理Q′による結果がDとD′どちらのデータベースから導出されたものかが区別できない場合、DとD′で異なる一人のデータは処理結果から推測できない。
- 任意のDとD′についてこのような性質が満たされれば、処理Q′の出力結果からは誰のデータも漏洩しないと言える。
- 差分プライバシーでは処理結果から2つのデータベースを区別できる最大の確率をパラメータεを用いて制限し、利用者が望む強度でのプライバシー保護を保証する。



差分プライバシーの利点

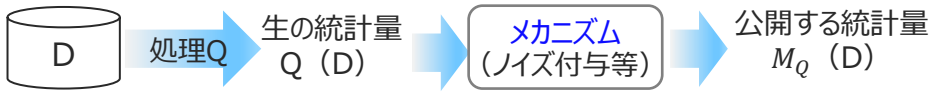
- プライバシー情報が推測されないことを、経験的でなく情報理論的に保証できるため、想定外の攻撃や外部情報が存在する場合でもプライバシー保護を実践できる。

2．2．3　差分プライバシーの実現方法と仕組み

- 多くのデータベース処理に対して差分プライバシーを満たすための技術（メカニズム）が提案されている。
- 単純な統計処理だけでなく、機械学習のような複雑な処理にも差分プライバシーが適用できる（次頁詳細）。

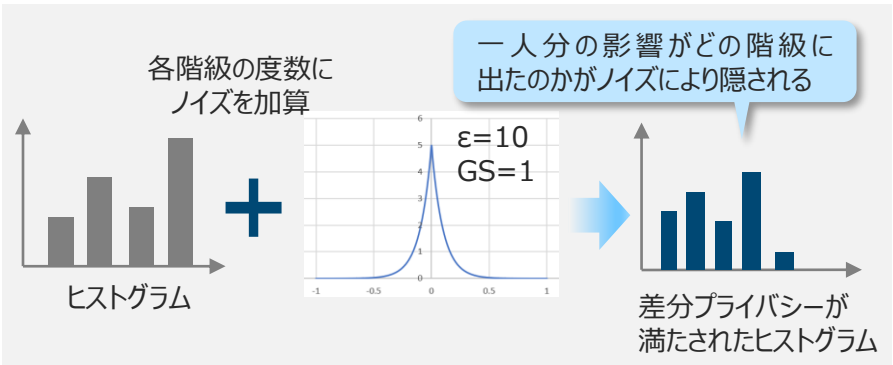
メカニズムとは

- メカニズムは差分プライバシーを満たすノイズやランダム性を処理結果に付与する方法のこと。
- 統計処理の結果を公開する際には、生の統計量ではなく、差分プライバシーを満たすメカニズムを通じて統計量を公開する。



代表的なメカニズム（ラプラスメカニズム）

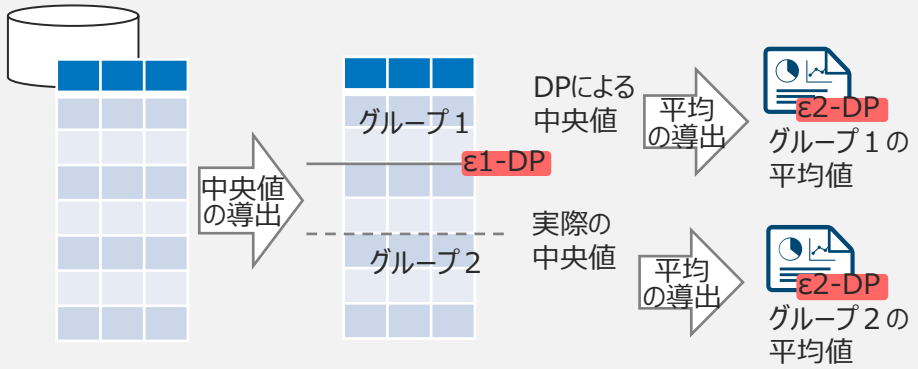
- 「データベース中の情報が１つ変わった際に処理の出力が最大でどれだけ変化するか」という指標である大域的感度GSを求め、平均0、スケールパラメータλ=GS/εとしたラプラス分布に従うノイズを出力に加算する方法。
- たとえば、ヒストグラムはデータベースの中の情報が1つ変わった際に、ある階級の度数は最大で1しか変わらない。そのため、大域的感度は１となる。



複雑な処理に対応したメカニズム

- 機械学習のような複雑な処理に対して大域的感度を求めることは困難である。
- そこで複雑な処理をパーツに分解し、個々のパーツに対して差分プライバシーを保証する。差分プライバシーの合成法則に基づき、全体の差分プライバシーが保証される。

中央値で分割したグループに対して平均を求める例



全体として，ε１+ε２-DPを満たす処理を構成

差分プライバシーの合成法則

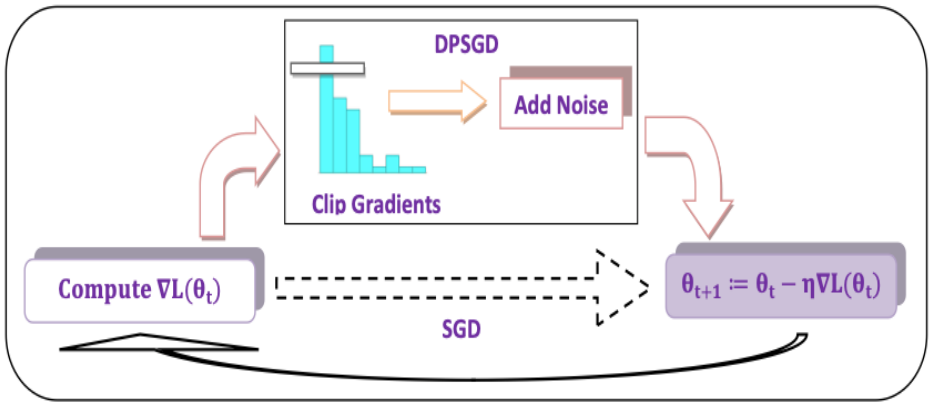
- ε1-差分プライバシーとε2-差分プライバシーをそれぞれ満たす２つのメカニズムがあった場合、組み合わせたメカニズムは（ε1+ε2）-差分プライバシーとなる。

2．2．4　差分プライバシーの機械学習への適用

- 機械学習モデルには学習データの情報が含まれており、モデルを通じたプライバシーの漏洩が懸念されている。差分プライバシーは、複雑な処理にも汎用的に利用できる性質から、機械学習への適用が注目されている。
- 一方で、差分プライバシーではノイズを付与する操作が加わるため、単純に適用すると予測精度が大きく低下してしまう。そのため、プライバシーと予測精度を両立するための研究開発も進んでいる。

DP-SGD

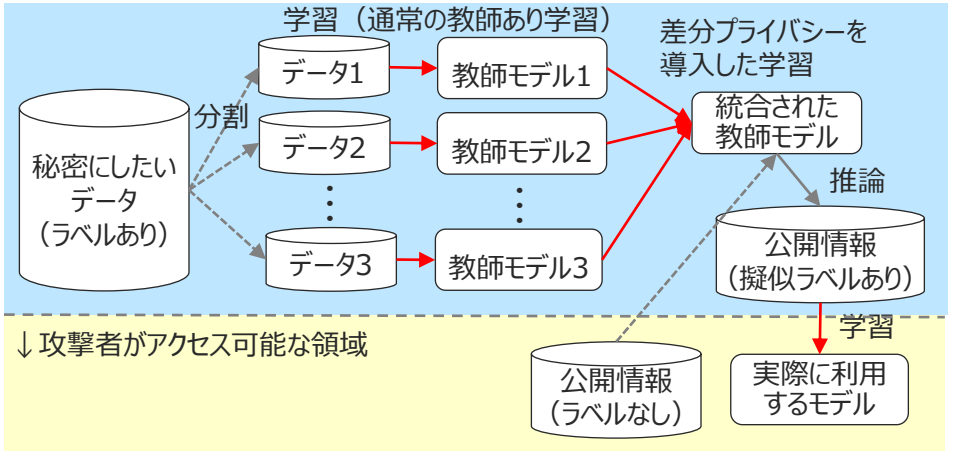
- ニューラルネットワークにおけるパラメータの更新方法であるSGD（Stochastic Gradient Descent:確率的勾配降下法）に対して差分プライバシーを保証。合成法則により学習全体の差分プライバシーを保証している。
- 工夫点として、SGDにおける更新量の上限に制約を設ける（クリップする）ことで大域的感度を抑え、差分プライバシーを保証するためのノイズを低減する。



出所：以下論文のFigure2より
M. A. Rahman他，Membership inference attack against differentially private deep learning model. Transactions on Data Privacy 11, 2018（http://www.tdp.cat/issues16/tdp.a289a17.pdf）

PATE（Private Aggregation of Teacher Ensembles）

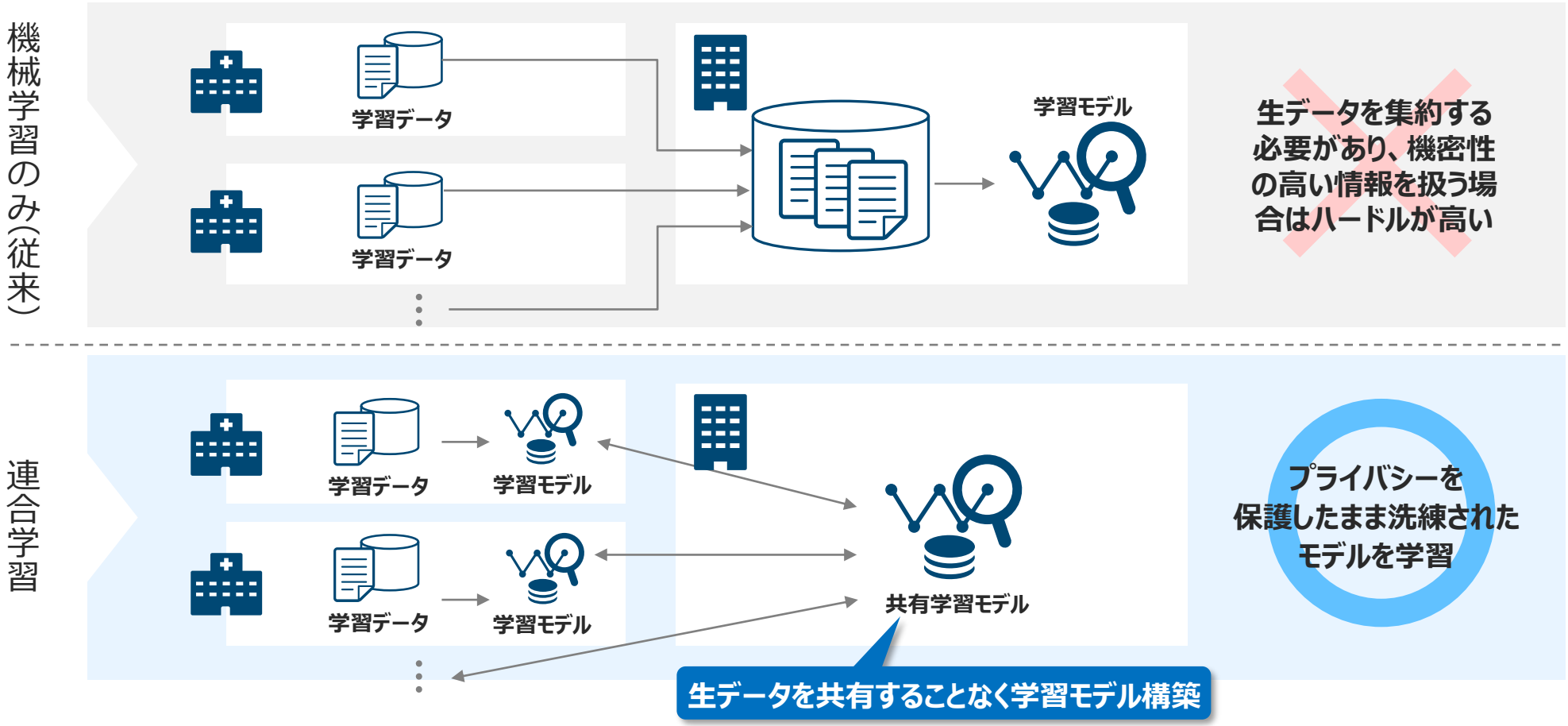
- 秘密にしたいデータで学習した教師モデル（差分プライバシーを導入）を用いて、ラベルなし公開情報の擬似ラベルを推論。この公開情報を用いて実際に利用するモデルを学習する手法。
- 公開情報の存在が前提となるが、モデルにノイズを加えるのではなく、実際に利用するモデルの教師データ（ラベル）にノイズを付加することで高精度な学習を実現。



参考：以下論文のFigure1
N. Papernot他，Semi-supervised knowledge transfer for deep learning from private training data, ICLR17, 2017, https://arxiv.org/abs/1610.05755

2．3．1　連合学習（Federated Learning）の概要

- 従来の機械学習は、すべてのデータを集約して学習を行う必要があり、組織が持つ個人情報などの機密性の高い情報を共有して高精度なモデルを学習することは困難であった。
- 連合学習は複数組織の学習モデルのパラメータや更新情報のみを共有し、統合したモデルを学習（連合学習）することで、機密性の高いデータを共有することなく高精度なモデルを学習できる。



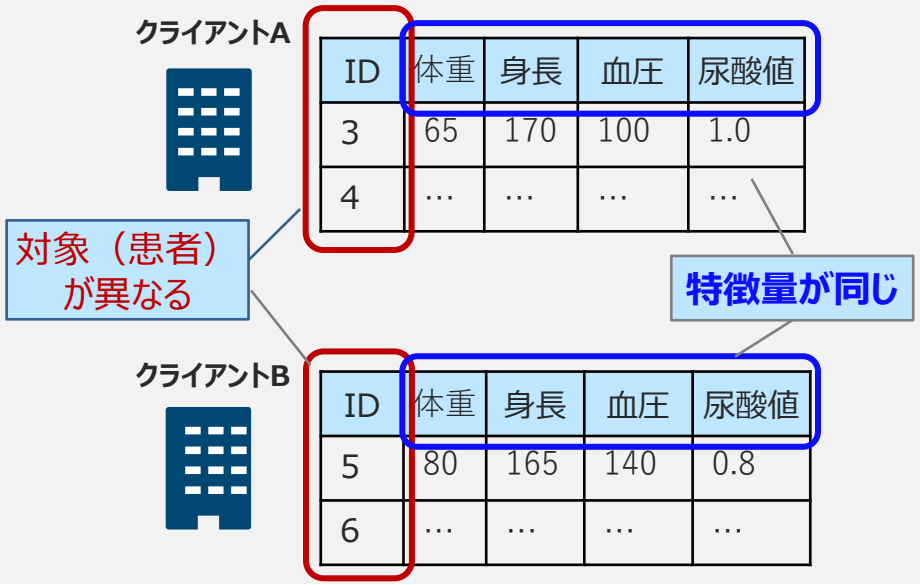
2．3．2　連合学習の種類

➢ 各企業が保有しているデータにどのような共通性があるかに応じて、適用できる連合学習の種類が異なる。

水平連合学習

- 全てのクライアントが**同じ特徴量**のデータを多く保有。
- 各クライアントが持つデータは異なる対象のもの。（下記の例では患者）
- 全体として学習データが増えることにより、モデルの性能向上が期待できる

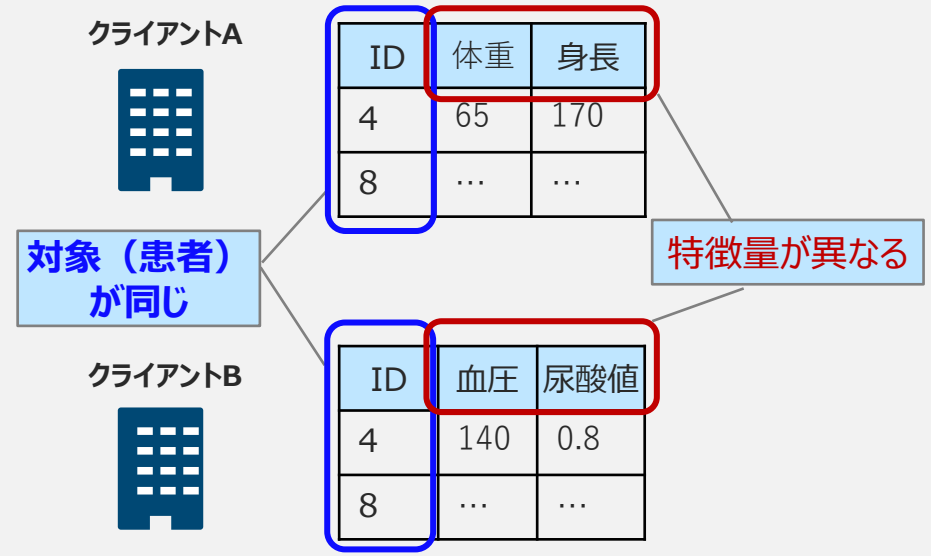
例：複数の病院で異なる患者のデータを用いたAIの作成



垂直連合学習

- 各クライアントが異なる特徴量のデータを多く保有。
- 全てのクライアントが持つデータは**同一の対象**のもの。（下記の例では同一IDの患者）
- 個人情報を使用する場合、現行の個人情報保護法では同意が必要となる手法。同意があったとしても生データを他組織に共有したくない場合に有効

例：ID連携しているサービス間でのAIの作成

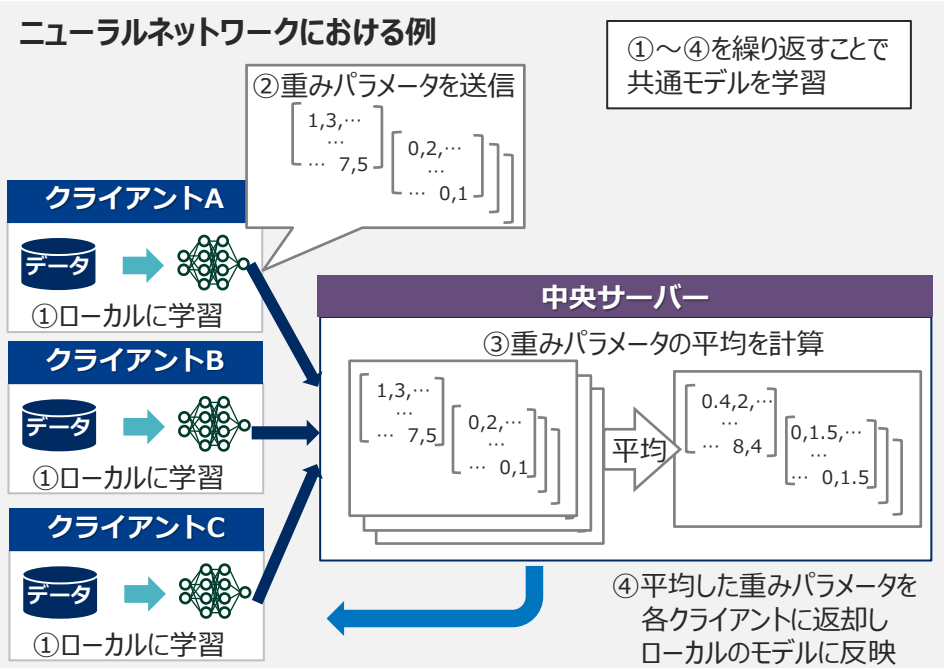


2．3．3　連合学習の仕組み

➢ 「水平連合学習」と「垂直連合学習」では学習の仕組みが異なる。「水平連合学習」はクライアント間で共通のモデルを学習し、「垂直連合学習」はクライアントとサーバー全体で一つのモデルを学習するように動作する。

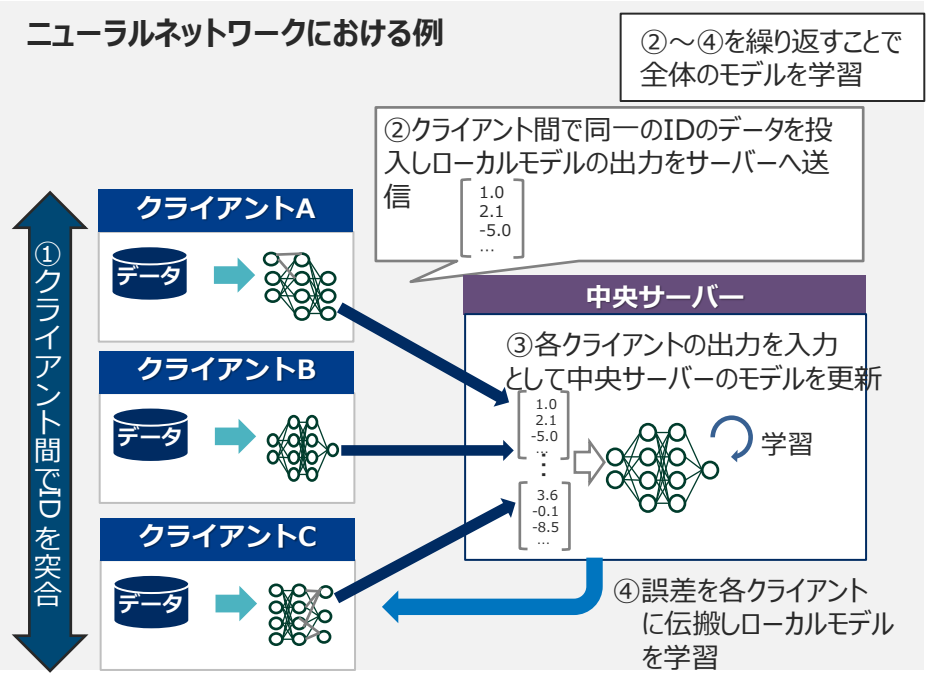
水平連合学習の仕組み（*1）

- 各クライアントが同一のモデル用いてパラメータをローカルに学習。
- 各層の重みパラメータのみをサーバーに送信して統合することを繰り返すことで、すべてのクライアントが所有するデータを用いて学習された共通モデルを生成する。



垂直連合学習の仕組み（*1）

- 各クライアントのデータに共通するIDを突合（*2）し、全体で一つのモデルを学習する。
- 各クライアントのモデルの出力を、中央サーバーが管理するモデルの特徴量として学習することで、全体のモデルを学習する。



（*1）　様々な手法が提案されているが、ここではニューラルネットワーク用いた典型的な仕組みを説明している
（*2）　垂直連合学習ではIDが共通するデータを対象に学習を行い、かつ、学習するデータの順番を各クライアントで揃える必要があるため、IDの突合が必要となる

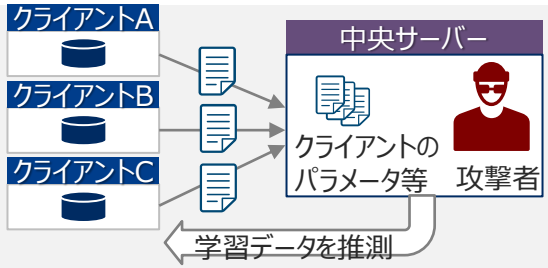
2．3．4　連合学習に対する脅威と対策

- 連合学習は中央サーバーやクライアントが攻撃者になることが想定されるため、学習の過程で得られる情報から、学習データが推測されるなどのプライバシー上の脅威を抱えている。
- 差分プライバシーや秘密計算を用いて、これらの脅威に対策を施した連合学習を「プライバシー強化連合学習」と呼び、様々な手法が提案されている。

連合学習における攻撃者

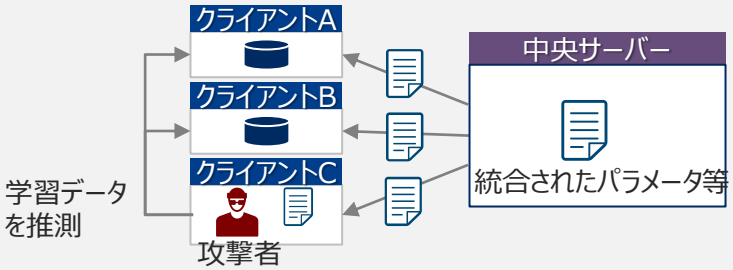
中央サーバー

- 中央サーバーがクライアントから得られる情報（重みパラメータやモデルからの出力など）を基に各クライアントの学習データに対する攻撃が可能である。



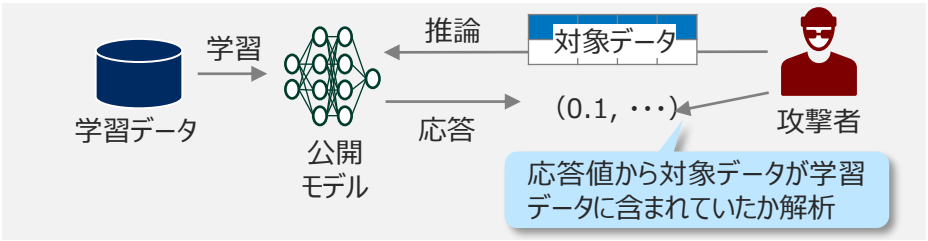
クライアント

- サーバーから各クライアントに提供される情報（統合された重みパラメータや伝搬される誤差など）には他のクライアントの学習データから導出された情報が含まれており，これに対する攻撃が可能である。



攻撃例：メンバー推測攻撃

- あるデータがAIの訓練に使われたデータか否かを判別する攻撃。
- 連合学習か否かによらず、推論時に攻撃が可能。攻撃者はモデルを直接入手しなくても、モデルの推論結果のみでも攻撃が可能。
- 連合学習では各クライアントから中央サーバーに送付されるモデルから、メンバー推測攻撃を仕掛けることが可能。攻撃者はモデルを直接入手できるためモデルの推論結果のみで攻撃するよりも、可能な攻撃の種類が多く、防御が困難となる。



対策：プライバシー強化連合学習

- サーバーおよびクライアントが得られる情報が、生のモデルパラメータやモデルの出力である場合に攻撃が成立しやすくなる。
- 生の情報を得られないようにするため、差分プライバシーや秘密計算と組み合わせて連合学習を行う。このような対策を施した連合学習を「プライバシー強化連合学習」と呼ぶ。

2．3．5　プライバシー強化連合学習の代表的な手法（1／2）

➢ プライバシー強化連合学習の代表的な手法としては、サーバー側による攻撃を秘密計算で、クライアント側による攻撃を差分プライバシーで保護する手法がある。

秘密計算と差分プライバシーを組み合わせた代表的な手法

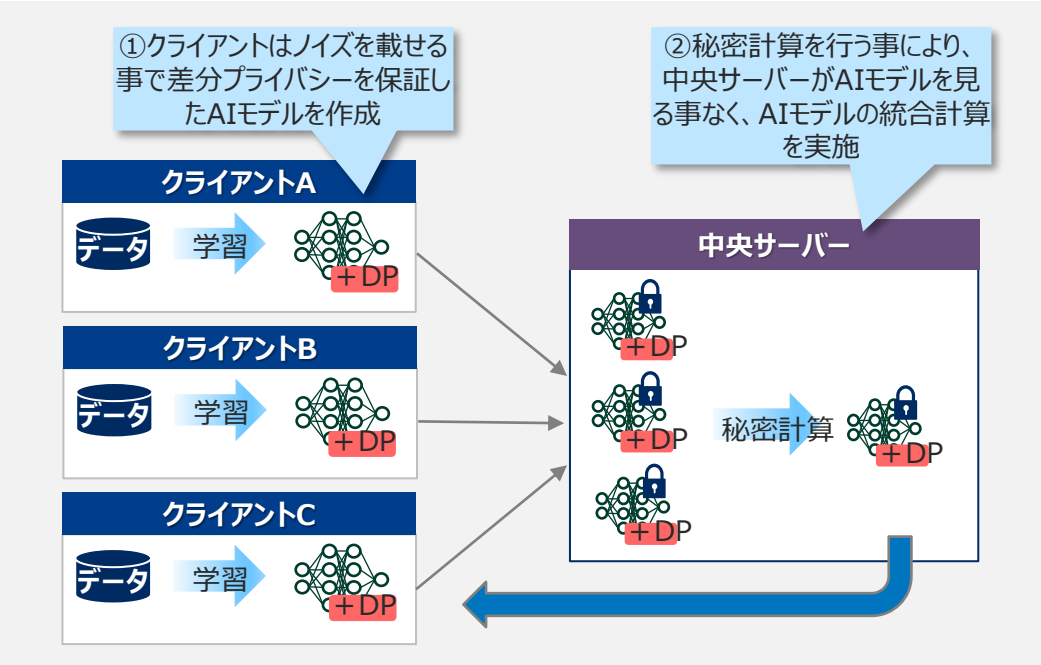
- 秘密計算と差分プライバシーを組み合わせることによりサーバー側・クライアント側双方での攻撃を防ぐ。
- 秘密計算を使わず差分プライバシーのみでも攻撃を防ぐことができるが、秘密計算を併用する事により差分プライバシーのノイズを小さくできるので、差分プライバシー単独で攻撃を防ぐ場合よりも精度劣化が小さくなるように工夫されている。

中央サーバーによる攻撃の防御

- 秘密計算により防御
- 中央サーバーはモデルを直接参照することができないので原理的に攻撃できない

クライアントによる攻撃の防御

- 差分プライバシーにより防御
- 例えばクライアントAのデータに関する情報は、Aのローカルモデルから統合モデルに反映され、最終的にクライアントBに伝わる
- クライアントBは受信した統合モデルを解析する事により、クライアントAが保持するデータを推測する攻撃が可能
- この攻撃をクライアントAが自身のローカルモデルに差分プライバシーをかける事で防御



参考：Stacey Truex他，A Hybrid Approach to Privacy-Preserving Federated Learning, arXiv preprint, 2019, https://arxiv.org/abs/1812.03224

2．3．5　プライバシー強化連合学習の代表的な手法（2／2）

➢ ラベルのない公開データを大量に持っているという状況において、各クライアントが学習するモデル自体にノイズを加えるのではなく、評価結果（ラベル）にノイズを乗せることで、精度劣化を抑える手法がある。

ラベルのない公開データを大量に持っている場合の手法

•サーバーおよび全てのクライアントが同一の（ラベルなしの）公開データを持っている事が前提となる手法。

学習の流れ

① クライアントは自身の秘密データでローカルモデルを学習する

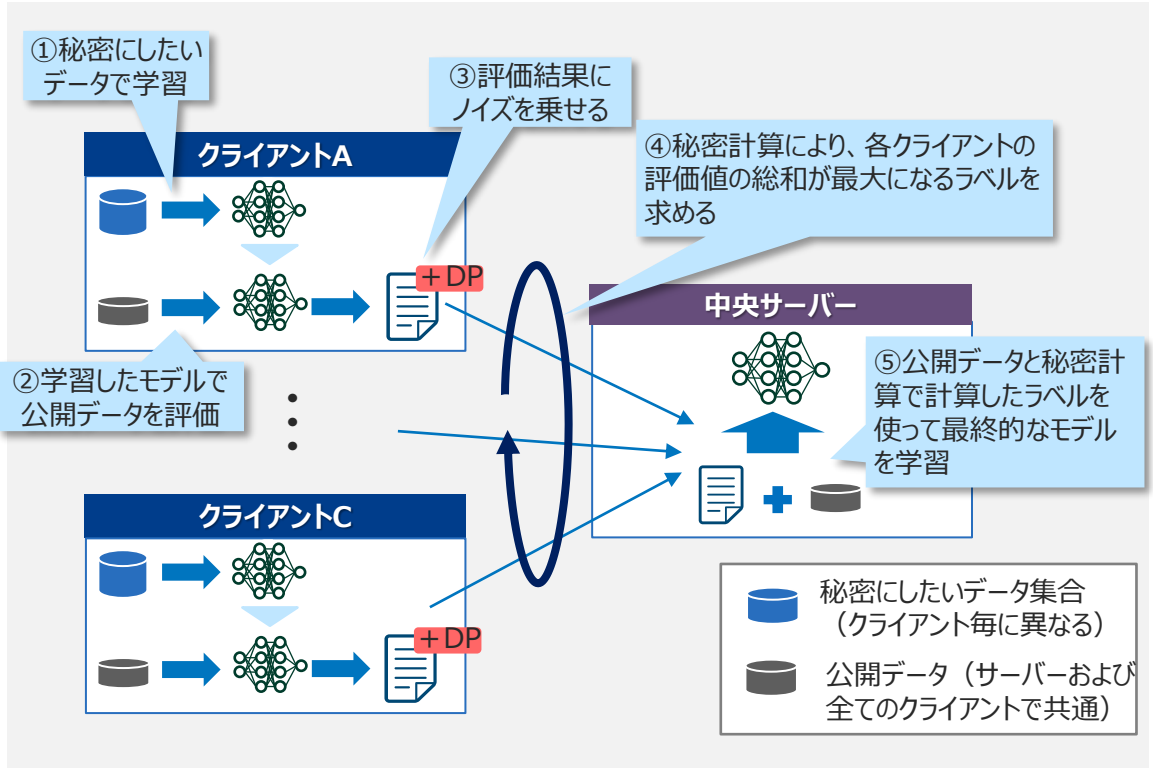
② ①で学習したローカルモデルに、ラベルなし公開データを入力し、ローカルモデルの評価結果を計算する

③ ②の評価結果に差分プライバシーを満たすノイズを加えた評価結果を生成する（通常の連合学習と異なり、ローカルモデルそのものを中央サーバーに送信しない）

④ 公開データの各レコードについて、各クライアントの評価結果の総和が最大になるラベルを秘密計算によって求める。

⑤ 中央サーバーは自身の持つラベルなし公開データと④で得られたラベルを用いて最終的なモデルをゼロから学習する。

⑥ 学習したAIモデルを各クライアントに配布する。



参考：Yuqing Zhu他，Voting-based Approaches For Differentially Private Federated Learning, arXiv preprint, 2021, https://arxiv.org/abs/1812.03224

３．１．１　国内動向（法規制）

- 個人情報保護法を中心に、「保護の強化」と「利活用促進」の両面から制度整備が進んでいる。
- 統計値算出など個人の選別を伴わない利活用では、組織間でも同意不要とするスキームの検討も進む。

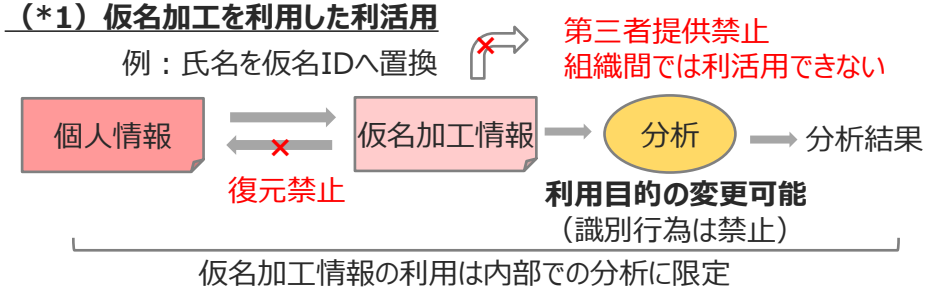
個人情報保護法

- 2015年の改正で、施行後3年毎の見直し規定が盛り込まれ、**保護強化とともにデータ利活用の促進**が図られている。

改正年	保護強化の例	利活用促進
2015年 (2017年 全面施行)	• 個人識別符号（生体情報、公的なIDなど）の個人情報を定義へ追加 • 保護対象であることを明確化	• 本人同意不要の匿名加工情報の条項を新設
2020年 (2022年 全面施行)	• 利用停止・消去等の請求権など個人の権利の拡充・漏洩等の発生で影響が大きい場合は、委員会への報告と本人への通知を義務化 • 提供先において個人情報となることが想定される情報の第三者提供の本人同意等義務	• **仮名加工情報の創設**（*1）内部分析に限定することを条件に利用目的の変更の制限、漏洩等の報告、開示・利用停止請求への対応等の義務を緩和

（*1）仮名加工を利用した利活用

例：氏名を仮名IDへ置換



一般財団法人 情報法制研究所（JILIS（*2））

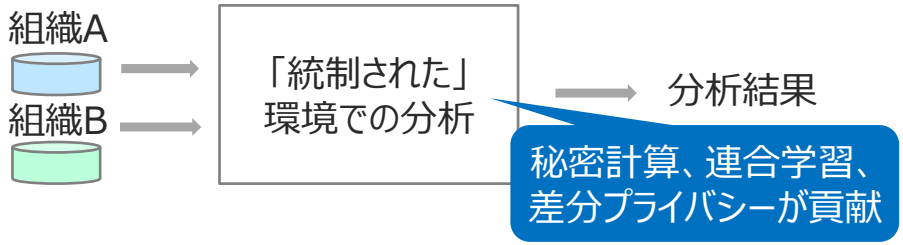
- 2016年に、情報法制に関する研究と政策提言を目的として設立。弁護士など法学系アカデミアを中心に構成。法人会員はKDDI総研、電通、NEC、LINEなど15社。
- プライバシー強化技術に関連する提言に向けた活動も実施

（１）秘密計算技術応用タスクフォース

秘密計算に関する法制度について議論。

（２）「統制された非選別利用」の提案

統計量への集計など「データによる個人の選別」を伴わない利用であれば、適切に「統制された」提供の下では組織間でも同意不要で利活用を可能にするスキームを提案。秘密計算、連合学習、差分プライバシーは「統制された」提供の実現に貢献すると期待される。



参考：JILISレポート Vol.3 No.10「情報法制学会第４回研究大会　講演録」

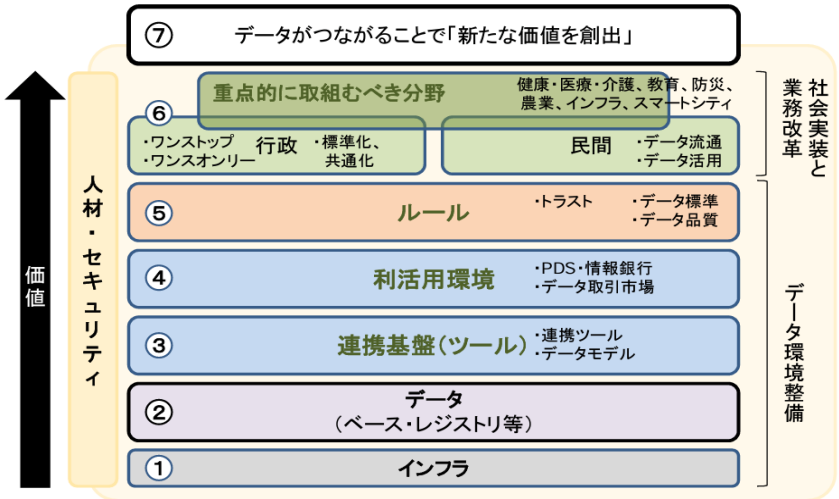
（*2）Japan Institute of Law and Information Systems の略

３．１．２　国内動向（国のデータ戦略、コンソーシアム）

- 情報銀行は国のデータ戦略で、パーソナルデータ利活用環境整備として、プライバシー強化技術も注目される。
- 秘密計算を中心に、複数のコンソーシアムが設立され、普及活動が進んでいる。

国のデータ戦略と情報銀行

- 政府は2021年6月に「包括的データ戦略」を発表し、GAFAに対抗する日本のデータ戦略の理念、**アーキテクチャーとその各層の指針**（下図）を示した。利活用環境の重点項目として、PDS（Personal Data Store）・情報銀行が挙がっており、パーソナルデータとの組み合わせによるデータ利活用の必要性が示されている（*1）。
- 今後の情報銀行の展開に向けた取り組みとしては、**新たなプライバシー関連技術(秘密計算など)への対応**が含まれる(*2)。



（*1）内閣府 データ戦略タスクフォース（第７回） 資料８－２「包括的データ戦略（案）」
（*2）総務省情報信託機能の認定スキームの在り方に関する検討会（第19回）会議資料「資料19-5 今後の情報銀行の展開に向けた取組み（総務省）」

コンソーシアム

秘密計算研究会

- 2021年2月に、NEC、デジタルガレージ、レピダムで設立。
- 秘密計算技術が広く社会実装され、クラウドサービスのデータ保護に対する不安の払拭や、組織や企業の枠を超えたデータ利活用により新たな価値が創出されることを目的として、**技術の安全性を客観的に評価するための基準作りや技術の理解促進のための情報発信**に取り組んでいる。

一般社団法人データ社会推進協議会（DSA）

- 2021年4月に、デジタル庁が推進予定の「データ戦略」や「内閣府・戦略的イノベー ション創造プログラム（SIP）」の後押しを受け、設立。
- 運営する連携基盤「DATA-EX」は、内閣府「包括的データ戦略」において分野間や海外との連携のハブになることが期待されているほか、利活用促進委員会には、**秘密計算活用WGが設けられ、具体的なユースケース検討**が行われている。

秘密計算コンソーシアム

- 2021年6月にAcompany、EAGLYSの共同運営で設立。
- ユースケースに焦点を当てて、勉強会などのイベント開催や情報発信を行っている。

３．２．１ 海外動向（法規制、世界経済フォーラム）

- 各国・地域毎にプライバシー保護の法規制の策定が進み、データ利活用の際には個別に対応する必要がある。
- 世界経済フォーラムは、DFFTやAPPAの概念を通じてプライバシーを尊重した利活用の推進を提唱している。

各国の法規制

欧州

- **一般データ保護規則（GDPR）**を2018年5月に施行。
- 個人の権利をベースにプライバシー保護の厳格化の流れを作った規制。
  - 保護対象となるデータの範囲が広い。
  - EU居住者のデータを扱う場合はEUに拠点がない事業者も対象。
  - アクセス権、訂正権、データポータビリティ、削除権などの権利の明示。
  - 域外移転のハードル：EU域内から域外の第三国等への移転のためには本人同意などが義務化。
  - **利活用に関して、最小化の観点から可能な限り、匿名化、仮名化の適用が求められている。**
  - 高額な制裁金：2020年1/28から1年間のGDPR違反の罰金の総額は1億5850万ユーロに上る。

米国

- **カリフォルニア州消費者プライバシー法（CCPA）**を2020年1月施行、**カリフォルニア州プライバシー権法（CRPA）**が2020年11月成立。
  - カリフォルニア州居住者のパーソナルデータを厳格に保護。
  - 一定以上の規模など事業者がある程度限定されている。
  - プライバシーポリシーの記載や削除・開示要求など細かい対応を義務化。
- **連邦政府においても、厳格な消費者オンラインプライバシー法（COPRA）の成立に向けた動き**がある。
  - 最短で2022/4月採択、10月施行。
  - 米国在住者の個人データは全対象、対象データはCCPAより広い。

中国

- 個人情報保護法を2021年8月に成立、11月に施行決定。
  - 2017年6月施行のサイバーセキュリティ法、2021年9月施行のデータセキュリティ法と併せて厳格な個人情報保護を規定。
  - 最重要インフラ事業者と大量に個人情報を処理する事業者は、**国内で収集した個人情報を国内で保存する義務**が定められている。

世界経済フォーラム（WEF：World Economic Forum）

- グローバルかつ地域的な経済問題に取り組むために、指導者層の交流促進を目的とした独立・非営利団体。データ利活用促進に向けた提言も活発に行われている。
- 2019年1月ダボス会議において、**プライバシーやセキュリティ・知財に関する信頼を確保しながら国際的に自由なデータ通信の促進を目指す Data Free Flow with Trust （DFFT）**の概念を日本が提唱。

- 2020年1月に、Authorized Public Purpose Access （APPA）のホワイトペーパーを公表し、個人の権利、データ保有者の利益、公益の三者のバランスの取れた、同意や匿名化不要のスキームを提唱している。**APPA実装のための手段の一つとして、秘密計算が挙げられている。**

出所：World Economic Forum White Paper「APPA – Authorized Public Purpose Access: Building Trust into Data Flows for Well-being and Innovation」

３．２．２　海外動向（標準化、コンソーシアムなど）

- 「ISO/IEC」の秘密分散（MPC）、「IEEE」の連合学習など、ガイドラインに向けた標準化が進む。
- 「MPC Alliance」などのコンソーシアム活動において、多くの企業、団体が参加。

標準化

ISO/IEC

秘密分散（MPC）

ISO/IEC 4922-1/2　Information security – Secure multiparty computation（2021年part 2 ドラフト作成中）

- 秘密分散（MPC）の主要な方式の標準化を進めている。

準同型暗号

ISO/IEC 18033　IT Security techniques – Encryption algorithms

- Part 6: Homomorphic Encryption（2019年発行）
  対象は加算のみ暗号化したまま可能な加法準同型の2方式。
- part 8: Fully Homomorphic Encryption　（2021年提案）
  完全準同型暗号の標準化プロセスも始まった。

IEEE

IEEE P3652.1　Federated Machine Learning Working Group

- 連合学習のアーキテクチャ、アプリケーションのガイドラインの発行を目指している。
- 「Draft Guide for Architectural Framework and Application of Federated Machine Learning」（2020年3月）8.2.1 Data privacy の考慮事項として、勾配情報から学習に使用した情報の推測を防ぐために秘密計算の適用を考慮すべきとの記述。

コンソーシアム・コンペティション

MPC Alliance

- 2019年設立、秘密分散（MPC）の開発・利活用企業のコンソーシアム。スタートアップを中心に、メンバー企業数は38。日本企業ではNTT、デジタルガレージが参画。
- MPCに関する豊富な技術資料をHPで公開している。
- 短期的にはコミュニティづくり、マーケティング、スタディなど、長期的には標準化への関与も想定されている。

Homomorphic Encryption

- 2017年設立、完全準同型暗号の標準化コンソーシアム。
- 民間18社（MS、IBMなど）、政府6組織（NISTなど）、アカデミア22団体が参画。
- 毎年の標準化会合の他、国際学会でワークショップを開催。
- 標準文書の最新ドラフトは2019年7月公開。完全準同型暗号の仕様と安全性レベルに応じたパラメータ設定を含んでいる。

iDash

- カリフォルニア大学主催のゲノムや医療データの処理に関してのプライバシー強化技術のコンペティション。2014年から続いている。

（参考）　iDash2020の課題

課題1：準同型暗号を用いた腫瘍の分類
課題2：SGX上でのmRNAゲノムデータのクラスタリング
課題3：癌予測のための差分プライバシー連合学習

３．３　取り組み一覧

- 秘密計算：電子資産の鍵管理で実用化が進み、データ分析に適用する例もある。
- 差分プライバシー：ユーザデバイスの利用状況のデータを集める際に適用し、実用化されている。
- 連合学習：金融、医療、IT・通信など、幅広い分野で実証実験などの取り組みが行われている。

領域	活用技術	取り組み	取り組み企業・組織例
金融	秘密計算	暗号資産などのデジタル資産を保護（カストディ業務、オークション市場）	ATOMRIGS LAB
	秘密計算/連合学習	複数の金融機関データ・モデルを用いて、AML/CFT対策（金融不正シナリオのモデル構築・高度化）や与信審査を高度化	**NICT/神戸大/三菱UFJ銀行など,Consilient/Intel,WeBank,英国FCA**
	連合学習	金融機関とレンタカー会社のデータを用いて、自動車保険を高度化	**WeBank**
医療	秘密計算/連合学習	複数の医療機関のデータ・モデルを用いて、癌・糖尿病等のリスク予測や病院内の死亡率予測、ゲノム解析の高度化、創薬開発を実施	**MELLODDY,NEC,Intel,NVIDIA**,IBM,OWKIN,TNO,inpher,MountSinai,BioLizard
IT・通信	秘密計算	秘匿データにして外部リソース（クラウドなど）を活用	Galois
		データを暗号化した状態で、機械学習（生体認証、学習用データ作成など）	**NEC,Apple,**Microsoft,EAGLYS,inpher
	連合学習/差分プライバシー	音声データをローカルに残したまま、AIアシスタントや予測変換等を改善（ウェイクワード誤作動防止、健康アプリ分析、フォーム自動入力等）	**Google,Apple,**Snips,Beam Data
	連合学習	連合学習用のライブラリ・フレームワークを開発・提供	Facebook,Baidu,Alibaba,Sherpa.ai
		通信効率の向上（ネットワーク帯域幅の必要性を減らすなど）	Technica,Fraunhofer
		性能・障害管理データ等のデータを用いて、ハードウェア障害予測	ERICSSON
セキュリティ	秘密計算	鍵管理、電子資産、情報銀行（暗号鍵、署名鍵、ブロックチェーンの鍵の保全）	**Unbound,Sepior**,i4p,Fireblocks,Zengo,AMIS,IJS,METACO,QREDO,Secata,Partisia,Cybernetica
		場所や個人情報等に適用し、悪用を排除し、治安を守る（政府系）	Cape Privacy
	連合学習	メールのスパム分類のモデル精度の向上	Kaspersky
		スマートビルディングの異常検知（センサーデータを保護、処理効率化）	Binaize
製造	連合学習	製品の欠陥検知、自動運転車の開発（物体検知、目視検査等）	WeBank,Extreme Vision,VisonX,Nvidia
		複数の工場データを用いて溶接品質評価モデル向上、ロボット予知保全	MUSKETEER

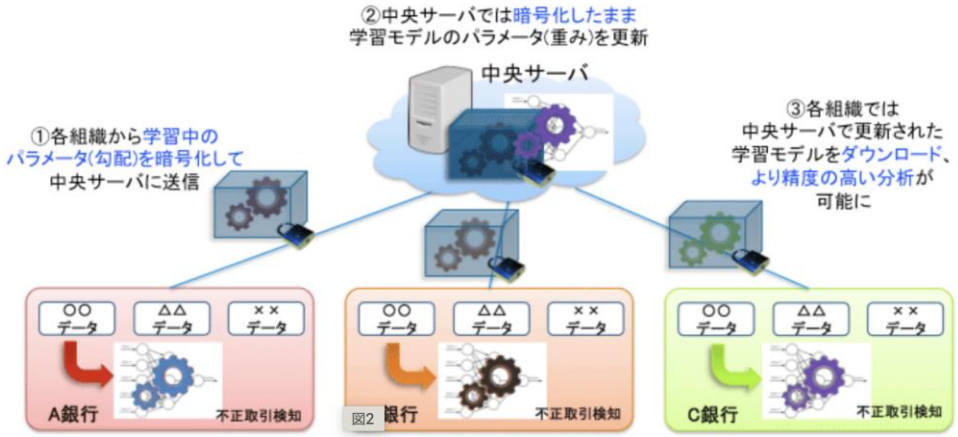
（参考）青字は次頁以降に詳細解説

３．４　金融分野での取り組み事例（１／２）

- 金融分野で秘密計算や連合学習が最も適用されている分野は、金融取引の不正検知である。
- 例えば、アンチマネーロンダリング(AML)やテロ資金供与の検知などがある。

不正送金の検知モデル（NICT/神戸大/三菱UFJ銀行など）

- NICT（情報通信研究機構）、神戸大学および三菱UFJ銀行をはじめとする金融機関5行が、不正送金の検知のためのモデル開発に共同で取り組んでいる。
- １つの金融機関ではモデルを学習させるために必要な不正な取引データが不十分であり、複数の金融機関が集まることで、学習に十分なデータを確保している。
- 不正送金の未然の検出率が現在は5割であるが、共同開発により8割まで高めることができ、22年度以降に実用化するとの報道もある（*1）。

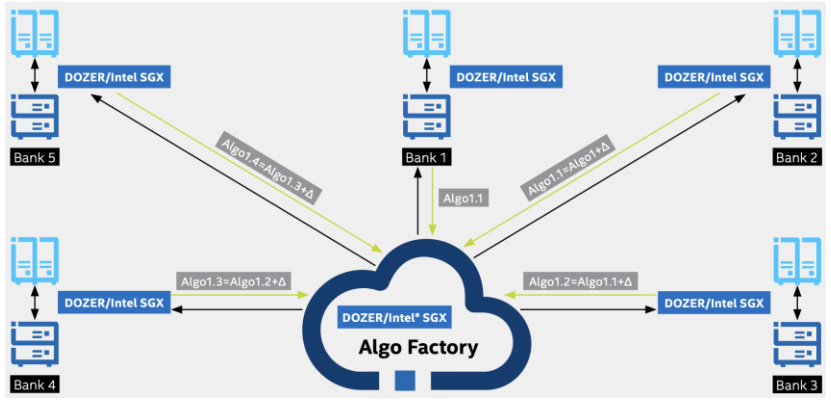


出所:
情報処理通信機構, プライバシー保護深層学習技術を活用した不正送金検知の実証実験において
金融機関5行との連携を開始, https://www.nict.go.jp/press/2020/05/19-1.html

AML/CFT対策（Consilient/Intel、WeBank）

Consilient/Intelの事例

- AML/CFT（*2）の対策（取引が正常か異常かを判定するモデル）に連合学習を使用。正常な取引を誤って異常と判定する率を12%削減。
- なお、連合学習だけでなく、秘密計算のTEE（Intel SGX）を用いることでよりセキュアな環境で実行している点が特徴。



画像出所: Consilient, Federated Learning through Revolutionary Technology,
https://consilient.com/wp-content/uploads/Q4_Intel_Consilient_Whitepaper_V04-1.pdf, 2021年10月18日参照

WeBankの事例

- 中国WeBankにおいてもアンチマネーロンダリングの対策に連合学習を使用。既存モデルと比較し、性能（AUC）が14%向上し、人手で検証する件数も大きく削減されたとのこと（*3）。

（*1）日本経済新聞, Suica履歴を「秘密計算」 データの中身、読まずに分析, https://www.nikkei.com/article/DGXZQOUC233NT0T20C21A6000000/
（*2）Anti Money Laundering/ Countering the Financing of Terrorism の略。マネー・ローンダリング及びテロ資金供与対策
（*3）FedAI, Utilization of FATE in Anti Money Laundering Through Multiple Banks, https://www.fedai.org/cases/utilization-of-fate-in-anti-money-laundering-through-multiple-banks/

３．４　金融分野での取り組み事例（２／２）

- 英国FCA（*1）においては、プライバシー強化技術が金融犯罪に関する情報の共有をいかに促進するかに焦点を当てたイベント（*2）を2019年に開催。大手外銀も参加し、ソリューションを提案している。
- 不正送金の検知など以外の事例では、保険を実行する際のリスク推定やクレジットリスクの推定などがある。なお、金融機関においては様々な取り組みが進められていると推測するが、現時点での公開情報は少ない。

金融犯罪対策（英国FCA）

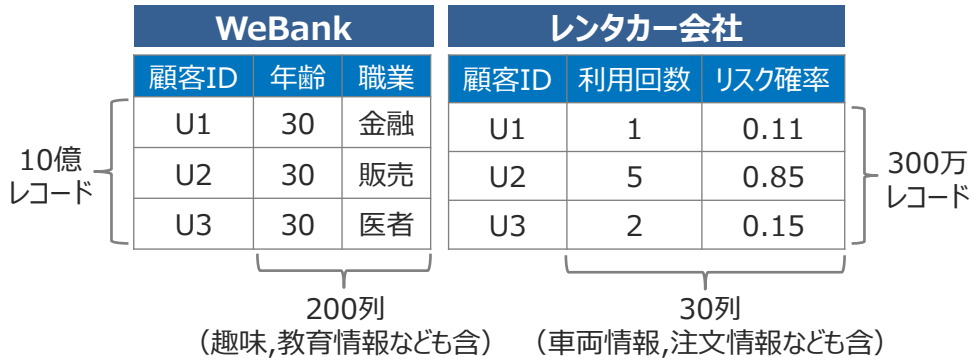
- 英国FCAは、プライバシー強化技術を用いて、金融機関や規制当局などの間で金融犯罪に関する情報の共有をどのように促進するかに焦点を当てた実証実験のイベントを開催。
- 本イベントではIBMなどのベンダーをはじめ、ベンチャー企業や大手外銀が参加しており、プライバシー強化技術を用いた全10件のソリューションが提案されている。

提案されたソリューション例	提案者
顧客のKYC業務において、秘密計算（準同型暗号）を使用して、リスクの高い顧客について銀行間でクエリを実行し、顧客のリスク情報を収集できるプラットフォームを提案	Enveil, EY, BAE Systems, Refinitiv, HSBC, Barclays, ING
秘密計算を使用し、送金を実際に実行する前に、送金・着金の金融機関が、アカウント名などの不一致を特定し、不正な送金をストップするソリューションを提案	Partisia, Sedicii, Goldman Sachs, Ex Ante Advisory, UBS, Deloitte

参考：FCA, 2019 Global AML and Financial Crime TechSprint, https://www.fca.org.uk/events/techsprints/2019-global-aml-and-financial-crime-techsprint

保険を実行する際のリスク推定（WeBank）

- 交通違反に関する保険に適用した事例。WeBankと、あるレンタカー会社のデータを使用し、連合学習によりモデルを構築。従来モデルと比較し、約1.5倍の収益向上に貢献。
- 金融機関とレンタカー会社という異業種での連携のため、連合学習のうち垂直連合学習を使用している。
- 個人情報保護法の観点から、日本において同様の取り組みを行うためには、顧客の同意が必要となる可能性がある点に注意が必要である。



参考：FedAI, A case of traffic violations insurance-using federated learning, https://www.fedai.org/cases/a-case-of-traffic-violations-insurance-using-federated-learning/

（*1）金融行動監視機構）　（（Financial Conduct Authority：金融行動監視機構）は、英国の金融を規制する機関
（*2）TechSprintsと呼ばれる、金融分野の内外から参加者を集め、業界の課題に対処するための技術ベースのアイデアを実装したり、実証実験を行うイベント

３．４　医療分野での取り組み事例（１／２）

➢ インテルは脳腫瘍を特定するためのAIモデル、NVIDIAはCOVID-19患者に酸素補充療法を必要とするかどうかを判断するAIモデルを、複数の医療・ヘルスケア関連機関でパートナーシップを組んで開発。

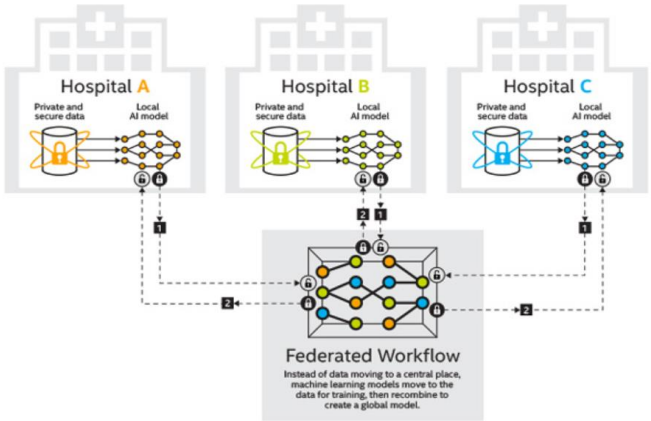
脳腫瘍を特定する予測モデル（Intel）

背景

- 脳腫瘍を治療可能な時期に早期発見するための予測モデルを構築するには大量の医療データを使用する必要があるが、これらはプライバシー上アクセスが困難であった。

手法・成果

- 2019年、ペンシルバニア大学とインテルは世界中の29のヘルスケアや医療の研究機関とパートナーシップを締結。
- 連合学習を用いてパートナシップに参加したヘルスケア関連機関がプライベートデータをシェアする事なく高精度なAIを構築。



出所： Intel, Intel Works with University of Pennsylvania in Using Privacy-Preserving AI to Identify Brain Tumors, https://newsroom.intel.com/news/intel-works-university-pennsylvania-using-privacy-preserving-ai-identify-brain-tumors/#gs.anm8yq,

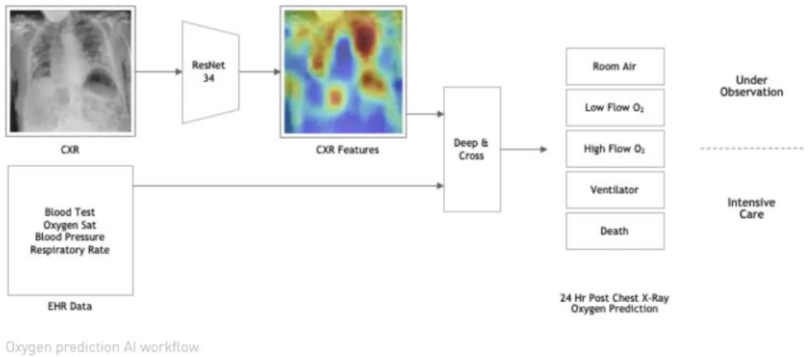
COVID-19患者対応AI（NVIDIA）

背景

- WHOはCOVID-19の患者に酸素補充療法を推奨しているが、どのような患者に、どのくらい投与するべきかの指針がなく、この問題に対処するAIを開発するには多くの病院のデータが必要であった。

手法・成果

- NVIDIAはMassachusetts General Brighamや世界中の20の病院と協力。
- 連合学習を使用し、個々の病院の胸部X線、患者のバイタル、臨床検査値を共有する事なくモデルが構築できた。



出所： Synced, NVIDIA & Mass General Brigham Hospital Federated Learning Project Predicts COVID-19 Patient Oxygen Need Using 20 Days of Data From 20 Hospitals, https://syncedreview.com/2020/10/07/nvidia-mass-general-brigham-hospital-federated-learning-project-predicts-covid-19-patient-oxygen-need-using-20-days-of-data-from-20-hospitals/

３．４　医療分野での取り組み事例（２／２）

- 創薬業界は、競合の観点からデータ共有が難しかったが、コンソーシアムを形成し、連合学習を用いることで共通モデルの構築を目指すプロジェクトが行われている。
- ゲノムなどの非常に機微な情報に対して秘密計算を用いて分析する事例なども存在する。

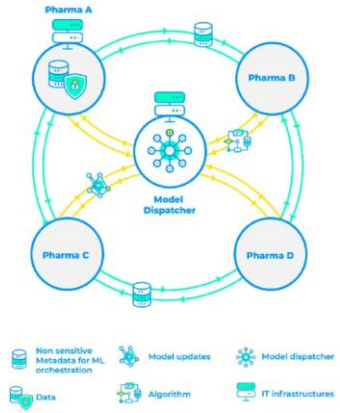
創薬AI（MELLODDY）

背景

- 新薬の開発には平均1.9億ユーロのコストがかかっており、製薬企業にとって機械学習による新薬開発のコスト削減は重要なチャレンジとなっている。
- 有用なAIモデルを学習するには大量のデータが必要だが、事業の競合の観点から製薬企業間でデータを共有することは難しかった。

手法・成果

- AMGENやAstraZenecaなどの10の製薬企業からなるコンソーシアムであるMELLODDY（The Machine Learning Ledger Orchestration for Drug Discovery）では、連合学習を用いて個々の企業のデータをシェアすることなく共通した創薬AIの学習を行った。



出所：Jaak Simm他, Splitting chemical structure data sets for federated privacy – preserving machine learning, https://www.researchgate.net/publication/353525271_Splitting_chemical_structure_data_sets_for_federated_privacy-preserving_machine_ learning

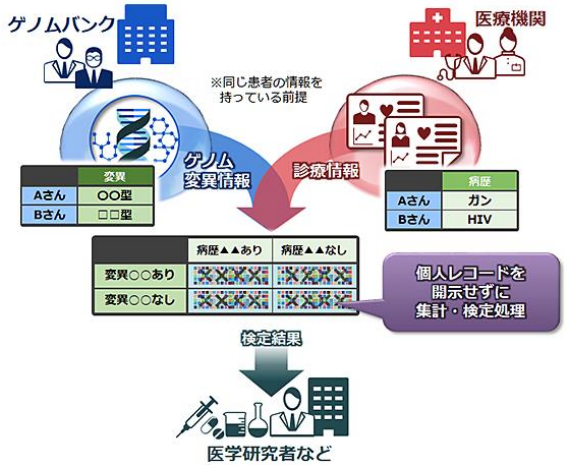
ゲノム分析（NEC）

背景

- ゲノム解析には、大量のゲノム情報が必要とされているが、ゲノム情報は非常に機微性の高いデータであり、罹患しやすい病気等も判明する可能性があり、漏洩によるリスクが高い。

手法・成果

- 秘密計算（秘密分散）を用いて、複数の医療機関がもつゲノム情報を統合解析するアプリケーションの実用性を検証。
- 約8,000人分のゲノム情報を約1秒で解析できるという処理速度や、開発のしやすさを実証した。



出所: NEC,情報を秘匿したままデータ解析ができる 秘密計算技術, https://jpn.nec.com/rd/technologies/201805/index.html,

３．４　IT・通信分野での取り組み事例（１／２）

- NECは、秘密計算（準同型暗号）を用いて、ユーザー端末とサーバーの二者で互いに秘密を守りながら生体特徴量を照合する方式を開発。
- Appleは、秘密計算技術を用いて写真アプリ中の児童性的虐待コンテンツ（CSAM）を検出する仕組みを開発。

生体特徴量の認証（NEC）

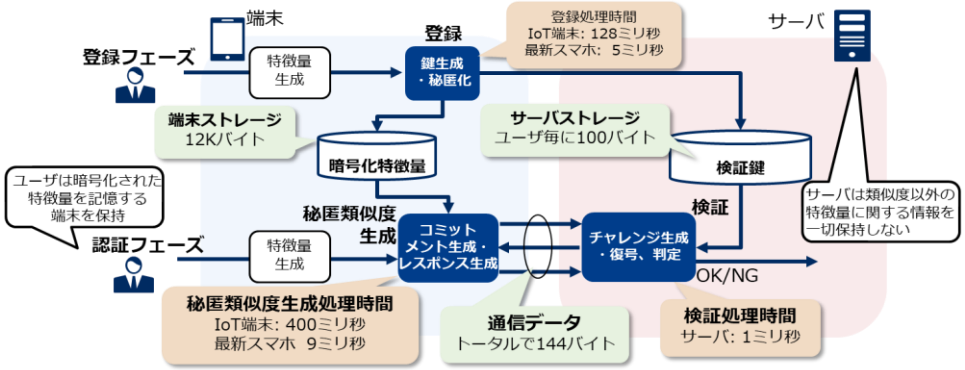
背景

- 端末認証において、認証サーバーに生体特徴量を出さず、かつ端末からも漏洩しないことが望まれる。

手法・成果

- FIDO（*1）のセキュリティ強化方式として、サーバーは類似度以外の生体情報に関する情報を一切保持しない、端末-サーバー間のセキュア生体認証方式を開発。

（*1）**F**ast **ID**entity Onlineの略で、従来のユーザ名/パスワードによる認証に代わる生体認証などの新たな認証技術のこと



出所：NECから提供

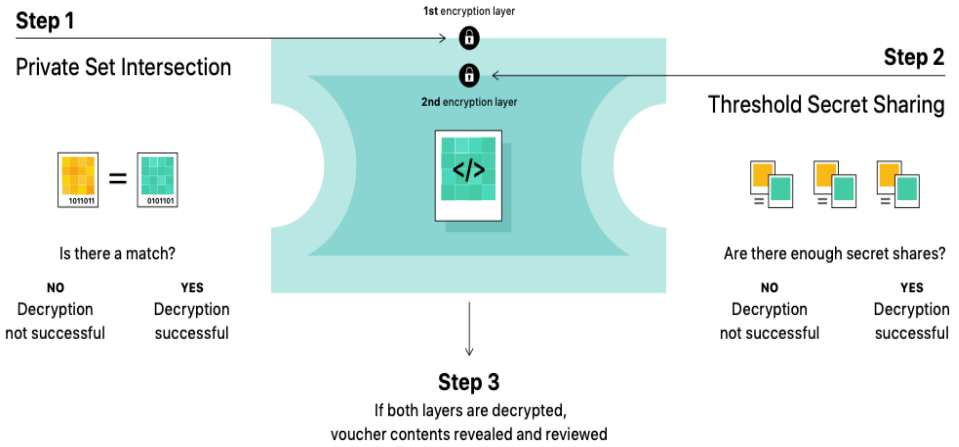
プライベート集合積によるCSAM検知（Apple）

背景

- 子供を守るために、利用者のプライバシーを守りつつも、CSAM（Child Sexual Abuse Material）の拡散を防ぎたい。

手法・成果

- 二者秘密計算の一種であるPSI（Private Set Intersection）およびThreshold Secret Sharingを用いて端末から一定数以上のCSAM画像がクラウドにアップロードされるのを検出(2021年10月時点で、実端末への本機能の適用は延期)。



出所：Apple, Expanded Protections for Children, https://www.apple.com/child-safety/pdf/Expanded_Protections_for_Children_Technology_Summary.pdf

３．４　IT・通信分野での取り組み事例（２／２）

➢ スマートフォンデバイスにおける予測変換や音声認識の精度向上への活用が進んでいる。

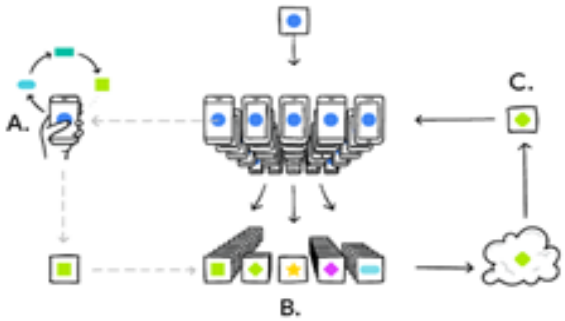
高精度な予測変換（Google）

背景

- スマートフォンのソフトウェアキーボードは予測変換機能を備えるが、快適な入力には予測モデルを常に改善する必要がある。
- 予測モデルの改善には利用者の変換履歴を使う必要があるが変換履歴はプライバシー情報であり収集が難しかった。

手法・成果

- スマートフォン内の変換履歴をもとに連合学習することで、生の変換履歴を共有することなく、精度が向上した予測変換モデルを端末間で共有。
- 連合学習はユーザがスマートフォンを使用しておらず、かつ通信料のかからない状態でのみ実行されるため、スマートフォンのパフォーマンスに影響無いようスケジュールされている。



出所： Google Blog, Federated Learning: Collaborative Machine Learning without Centralized Training Data, https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

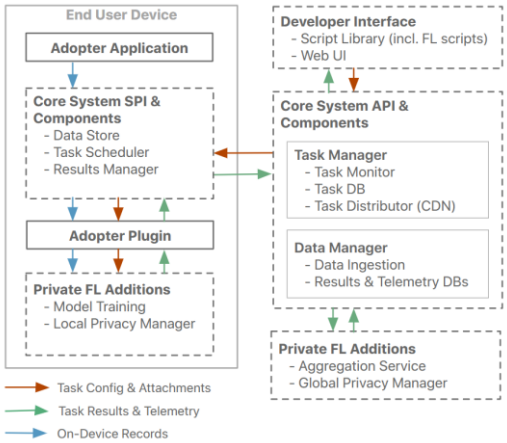
高精度な音声認識（Apple）

背景

- SiriはiPhoneの持ち主以外が起動フレーズを発したときでも起動してしまっていたが、持ち主のみが起動できるように改善を目指していた。
- そのためには大量の音声データから学習を行う必要があった。

手法・成果

- 連合学習を用いて、音声データを端末から出すことなく音声認識のモデルを学習した。



- 連合学習に加えて差分プライバシーにより学習データにノイズを乗せることで、悪意のある参加者が学習モデルから元の音声データをリバースエンジニアリングすることを防いでいる。

出所：Matthias Paulik 他, Federated Evaluation And Tuning For On-Device Personalization: System Design & Applications, https://arxiv.org/pdf/2102.08503.pdf

３．４　セキュリティ分野での取り組み事例

- Unboundは、最先端のGarbled Circuitを使う秘密計算による鍵管理を最初に売り出した会社。Sepiorは三者秘密計算を使った鍵管理も販売。
- ビットコイン等の電子資産や暗号鍵の保護は、秘密計算の事業として最も発展している。

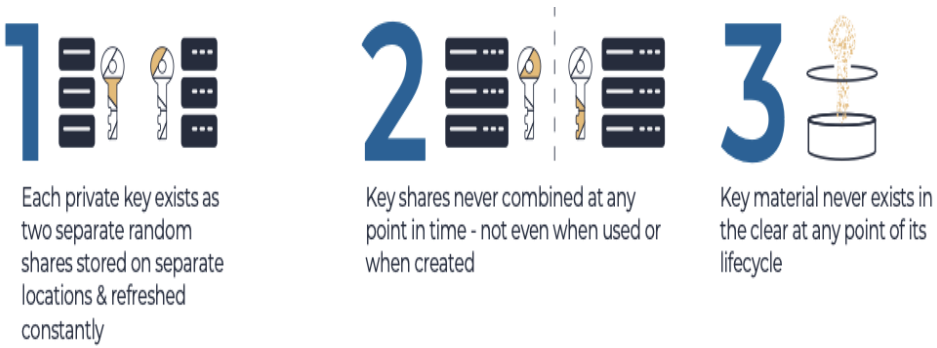
セキュア鍵管理（Unbound）

背景

- 暗号鍵の管理は秘密管理の要であり、単一障害点であった。

手法・成果

- 暗号鍵を秘密分散で二分割して生成して保存する。
- Garbled Circuitを使った秘密計算などで分割された暗号鍵を分割したまま利用するため、暗号鍵がそのライフサイクル上一度も存在しない。分割した鍵の両方を同時期に盗まれない限り、暗号鍵は一切漏れない。



出所：Unbound, Solution Brief, https://www.unboundsecurity.com/wp-content/uploads/2020/12/Cloud-Agnostic-SB-v7.pdf

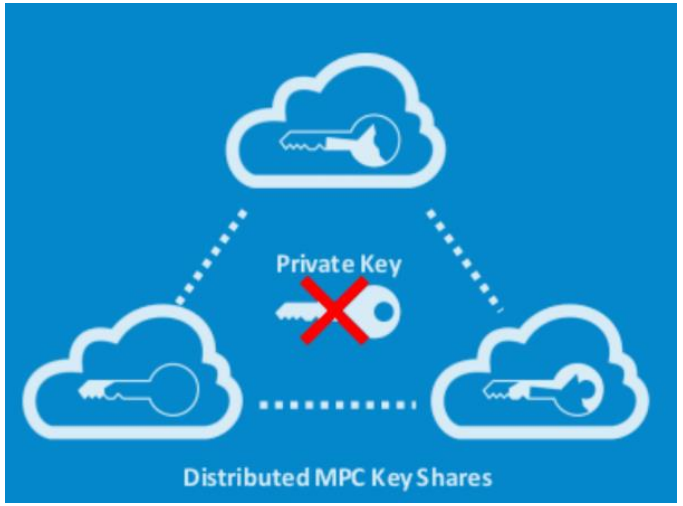
セキュア鍵管理（Sepior）

背景

- 暗号鍵の管理は秘密管理の要であり、単一障害点であった。

手法・成果

- 暗号鍵を秘密分散で三分割して生成して保存する。秘密計算により暗号鍵を利用する。三分割した暗号鍵のうち、二つを盗まれない限り、元の暗号鍵は一切漏洩しない。



出所：Sepior web-site, Threshold Key Management （Threshold KM）,https://sepior.com/products/threshold-km

４．１　現状の整理

- 秘密計算、差分プライバシーは一部実用化されているケースもあるが、広く実用化されるまでには至っていない。連合学習は実証実験のフェーズであり、実用化されているケースは限定的。
- また、プライバシー強化技術の活用に向けては、法律、ITシステム、人材の観点から検討・対応も必要。

各技術の進展度

- 技術的な課題を抱えているものの、それらの課題を前提として実用化に向けた取り組みが進んでいる。**秘密計算、差分プライバシーについては一部実用化されるフェーズ**に至っている。
- 一方、**連合学習（プライバシー強化連合学習）は、実証実験のフェーズ**であり、実用化に至っているものは少ない。

主な技術的課題

秘密計算	• 計算処理の高速化
差分プライバシー	• 特になし（*1）（秘密計算・連合学習との組み合わせた手法の開発が進む）
連合学習	• クライアント間でデータ分布が偏る場合に最適なモデルが学習できない • プライバシーとモデル性能のトレードオフ　など

活用に向けた検討事項

法律

- **データ保護に関する法律は発展途上**で、背景にある狙いは国や地域によって違い、年々変化する。そのため、**地域の法律や考え方の変化を丁寧に追う必要**がある。
- 第三者への開示は秘密計算や連合学習を用いても、当面は本人同意が必要である。今後法改正により不要になる可能性はあるが、提供先での利用方法に依存する点は変わらない見込み。

システム開発

- システム上の新たな攻撃界面が生まれる。ここに対する通常のセキュリティ施策も必要。よりプライベートなデータを扱うため、新たな運用上の注意点が生まれる。

人材・体制

- データ活用人材には、**AIやセキュリティに加え、プライバシー保護関連の法制度の知見が必要**。
- しかしながら、先進的なプライバシー強化技術を用いた開発を担える専門人材は限られているため、専門人材以外でも使える開発支援ツールの整備が期待される。
- また、異なる分野の担当者が横断的に連携して追跡するか、社外有識者の支援を受けることを推奨する。

（*1）技術課題ではないが、実用化に向けてはεの概念や設定方法が一般に分かりにくいことなどが課題としてあり得る

４．２　技術的課題と展望/考察

- 秘密計算：直接扱える演算種類の拡大など、秘密計算の処理性能を上げる（高速化）方法の研究が進む。
- 連合学習：クライアント間のデータ分布が異なる場合の学習手法の研究に加え、精度の劣化を抑えたプライバシー保証の方法などの研究が進む。

技術	課題	展望／考察
秘密計算	秘密計算処理の高速化 • 秘密計算中の通信や暗号化のため、通常の計算に比べて低速 • 秘密計算は単純な論理演算、整数演算、ソートなどの限られた演算のみを直接扱える。その他はこれらの演算の組み合わせより実現するため低速	• **浮動小数点演算、分岐処理などを直接扱うための研究開発が進んでいる。**これにより、データ分析において10-100倍の高速化も期待できる • 参加者数が４や５の場合の秘密計算や、効率的なゼロ知識証明を活用し、秘密計算の処理性能を上げる方法が研究される（ただし、画期的に処理性能は上がらないと考えられる）
連合学習	• 各クライアントが保有するデータ分布が異なる場合、連合学習で作成したモデルが、各クライアントにおいて性能が向上する最適なモデルとならない	• クライアント毎に異なるAIモデルを作るなど、**クライアント毎の精度を向上させる手法の研究開発が活発に進んでいる**
	• 大量のクライアントが存在する状況や、クライアントの通信・計算資源に制約がある状況では効率的な連合学習が困難である	• クライアント毎に**ソフトウェア・ハードウェアの性能が不均一な状況で効率的に連合学習を実行する方法の研究開発が進んでいる**
	実システムに適用する際に、例えば以下の運用上の課題がある • データが各クライアントに分散しているため、特徴量の設計など、連合学習の前処理をどのようにするか • クライアント間の利害関係を調整するため、クライアントの貢献度に応じたインセンティブ設計をどのようにするか	• **データを共有せずに特徴量設計を行う方法等の研究**が期待される • 技術的解決のみならず運用上の解決方法も模索すべきであり、**事例の共有などが進むことが望ましい**
	• プライバシーとモデル性能はトレードオフの関係にあり、プライバシーの保証強度を高めると精度劣化が大きくなる	• 単純な差分プライバシーの適用は精度劣化が大きい。**連合学習に特化した差分プライバシーや、それ以外のプライバシー強化技術の考案**が期待される

４．３．１ 活用に向けた検討事項 - 法律（１／２）

- プライバシー強化技術の利活用促進のためには、ガイドラインの整備を伴った法律面の対応が必要となる。
- プライバシー保護規制は発展途上であり、パーソナルデータ活用に関する国や地域の法律や考え方の変化を丁寧に追従する必要がある。

法律面におけるプライバシー強化技術

- 法律面において、プライバシー強化技術はまだ考慮されていない。例えば、個人情報保護法では、政府推奨暗号などの「高度な暗号化」を適用していれば漏洩時でも報告義務はない、と規定されているが、秘密計算が「高度な暗号化」に該当するのかは技術の進展や社会実装の動向も踏まえつつ、引き続き検討という見解が示されている（*1）。パーソナルデータを突合し、その結果を集計して統計値を算出するまでを一括して秘密計算で行えば、実質的にパーソナルデータは誰にも提供されないが、法律ではこれらはまだ考慮されていない。
- そのため、プライバシー強化技術活用に際しては、**安全性基準等のガイドラインの整備を伴った法律面の対応などが必要**。

  ①プライバシー強化技術の適切な選択・運用の判断は難しく、法律面の対応のためにも、**ガイドラインの作成と技術の進展に伴っての更新が不可欠。**

  ②**知見を求める学習とその結果の個人選別への利用を分けて考えることが重要**。前者はプライバシー強化技術に対する法律面の対応が望まれるが、後者は技術に関わらず慎重な判断が今後も必要と考えられる。

（*1）e-Govパブリック・コメント「個人情報の保護に関する法律についてのガイドライン（通則編、外国にある第三者への提供編、第三者提供時の確認・記録義務編及び匿名加工情報編）の一部を改正する告示」等に関する意見募集の結果について－【別紙２－１】意見募集結果（通則編）（2021/8/2）

地域、時間で変わるプライバシー保護規制への対応

- 各国のプライバシー保護規制は、プライバシー原則のベースは共通であるが、具体的な規制は異なり（下図）、また、時代に合わせて改正・変化していく。
- プライバシー強化技術の適用に関しても、**プライバシー保護に関する国や地域の考え方を深く理解し、変化に追従していく**必要がある。

項目	概要
保護対象範囲	• 日本の個人情報保護法と比較して、**GDPRやCCPA等はIPアドレスも含まれることが明記されるなど、保護の対象範囲は広い。** • フランスの自動体温測定に関する判決では、データを記録すらしていなくてもGDPR違反の判断がされた。**日本では、総務省と経産省が共同で本ケースの「配慮事項」（*2）を発表しており、一定の条件を満たせば実施できる**としている。
法令違反時の罰則金の上限	• GDPR(EU)：2千万ユーロもしくは前年度売上高の4% • CCPA(米)：1件ごとに2,500ドル。漏洩の場合、一人当たり100-750ドル • 中国：5千万元もしくは前年度売上高の5% • 日本：上限1億円

（*2）民間事業者によるカメラ画像を利活用した公共目的の取組における配慮事項

４．３．１ 活用に向けた検討事項 - 法律（２／２）

- 公平性などのプライバシーに係るAI規制にも考慮することが必要となる。
- 個人情報以外にもプライバシー保護に関わる法令は多数あり、案件に応じての対応が必要となる。

AI規制について

- AIを利用したプロファイリングによる望まない属性の推定や不正確な推定によるプライバシー侵害のリスクなどから、**プライバシーの観点を含めたAIの倫理に関する議論が活発**となり、規制の動きが始まっている。
- 品質への影響の評価に加えて、**AIの用途として公平性などプライバシーに関する倫理的な判断を適切に実行し、AI規制に対応**していくことが必要となる。

	概要
AI社会原則（内閣府）	• AIの社会的・倫理的・法的な課題を考慮して七つのAI社会原則を示した。以下、関連原則(3)プライバシー確保の原則(4)セキュリティ確保の原則(6)公平性と説明責任および透明性の原則
機械学習品質ﾏﾈｼﾞﾒﾝﾄｶﾞｲﾄﾞﾗｲﾝ(産総研)	• プライバシーに関連の深い、公平性に関しても品質評価軸を示している
包括的なAI規制案（EU）	• AIの中から規制対象になるものを列記して(1)受容できないリスク(2)ハイリスク(3)限定的リスク(4)最小限のリスク、4段階に分類 • (1)は禁止(2)では適切なデータの利用や消費者への説明など第三者機関による事前審査が想定

プライバシー保護に関わる法令の多さ

- プライバシー保護は個人情報保護法などだけではなく、案件に応じて様々な法令が関係することに注意する必要がある。
- 日本において関連する個人情報保護法以外の法律の例は以下の通り。個人情報保護法はこれらの法律に劣後するように作られている（例えば、法令に根拠があれば、本人の意思に関わらず第三者提供可能）。

個人情報保護法以外の法律例
刑法：不正指令電磁的記録に関する罪
電気通信事業法：通信の秘密
電波法
不正競争防止法：営業秘密の保護
消費者契約法
労働契約法
労働基準法
民法：契約関係、不法行為

参考：崎村夏彦「デジタルアイデンティティ」（日経BP、2021年）

４．３．２ 活用に向けた検討事項 - システム開発（秘密計算）

➢ 秘密計算を利用したシステムの実装には、従来のシステム開発における検討事項に加え、秘密計算を用いることによって発生する検討事項も存在する。専門家を交えての検討や、必要であれば実証実験を行う必要がある。

ライフサイクル	項目	課題	対応方針
計画	脅威分析	• 秘密計算は、データへのアクセス権限を分散させることで脅威（攻撃者）からデータを保護する技術であるため、想定する脅威が秘密計算で対抗できるか否かを確認しなければならない。	• **脅威を分析し、脅威に対抗可能な計算方式を採用** - 組織ぐるみを含む内部犯行か - システム内部への侵入であるか - 攻撃者はどこまで権限を獲得するか
設計	安全性設計	• 想定する脅威（攻撃者）に対して、十分な秘密計算の強度となるよう設計、運用をしないと秘密計算の十分な効果が得られない。	• **秘密計算の強度の決定** - 秘密計算を構成するサーバー台数の決定 - 攻撃者になり得るサーバー管理者の想定
導入	方式の選択とリソース設計	• 対象アプリケーションが秘密計算に適さない場合や適さない秘密計算の方式の場合、秘密計算の処理時間が長くなりすぎる可能性がある。 • アプリケーションへの秘密計算の実装には、未だに専門家の助けが必要である。	• **専門家のアドバイスによる適切な方式決定と実装** - 対象アプリケーションへの秘密計算の効果有無判断 - 対象アプリケーションに適する秘密計算の方式決定 • **専門家のアドバイスによる適切なリソース設計** - 必要なサーバー台数、ネットワーク容量の見積もり - 人的リソースの見積もり
運用	システム管理体制	• セミオネストな攻撃者のみに対抗する方式の場合、各サーバーに管理者が二人必要。全体の責任を担う監督者もいることが望ましい。	• 適切なサーバー管理者数の決定、監督者の設定 • 運用組織としての規則や契約の設定
	ｱﾌﾟﾘｹｰｼｮﾝ管理	• 秘密計算は、アプリケーションからの利用権限も複数参加者にて分散して管理できる（特権乱用の防止） • ただし、アプリケーションの特性（データの機密レベル、複数組織のデータ利用、他）によって、その権限設定が複雑になる可能性がある。	• 権限設定が複雑にならない工夫が必要
	監視、記録、廃棄	• セミオネストな攻撃者にのみ対抗する秘密計算の場合、クエリや入力値の真正性確認の仕組みが必要。異常を検知した場合、原因となる参加者の特定が可能な仕組みを作らなければならない。	• **真正性を保証する運用設計** - バックアップ、ログ取得 - データ廃棄のルール決定

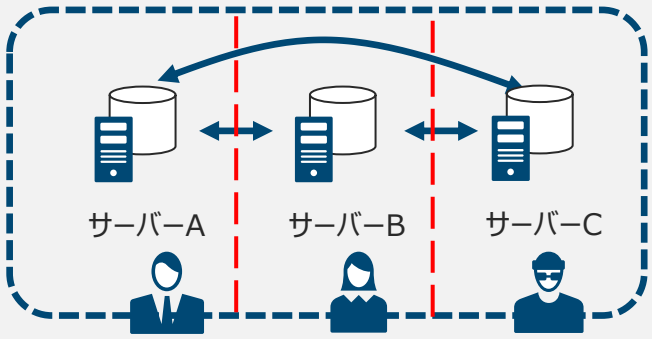
（参考）秘密計算が対抗する脅威

➢ 秘密計算が対抗する脅威は一部の参加者(サーバー)が攻撃者となること。攻撃者は大きく２種類に分けられる。
（１）悪意ある攻撃者：正直に自身の計算を実行せず、恣意的に作成したデータを他の参加者に送る。
（２）セミオネスト攻撃者：正直に自身の計算を実行するが、自身の扱うデータを盗み見て、秘密を復元する。

悪意ある（malicious）攻撃者

- 悪意ある攻撃者は、自身の持つデータを書き換えたり、自身の計算結果を偽る攻撃を行う。
- 強い権限を持った内部犯行者や、権限昇格に成功した外部からの侵入者も想定できるため、この様な攻撃者に対抗する**秘密計算は、非常に強い安全性を備える。**
- 例えば、システムの運用において管理業務が必要な場合は、各サーバーの管理者が如何なる不正を働いても、他のサーバーがこれを検知することが可能であるため、**各サーバーの管理者を一人**として、これを直接監視せずに済ませることが可能になる。各サーバー管理者の直接管理が不要になるので、**クラウドで**の利用などネットワーク越しの管理の監視が容易になる。

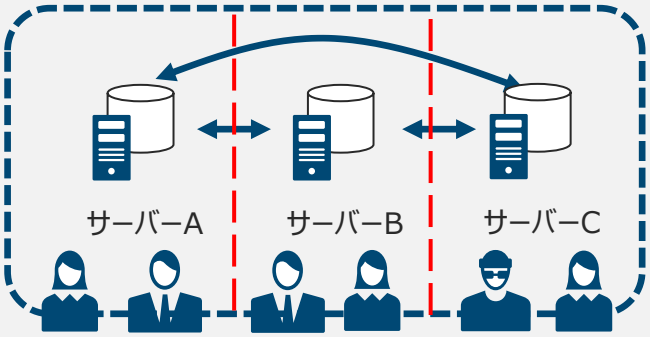
秘密計算（参加者が三人の方法での管理業務の例）



セミオネスト（semi-honest）攻撃者

- 参加者は攻撃者となっても、自分の持つデータを書き換えたり、秘密計算中の自身の計算において不正な動作をしない。
- 秘密計算以外の方法で、参加者の攻撃の抑制や監視がある程度行われていることを想定する。セミオネスト攻撃者に対抗する秘密計算は、悪意ある攻撃者に対抗する秘密計算と比較して、**安全性は弱い。**
- 例えば、システムの運用において管理業務が必要な場合、各サーバーの管理者の不正な振舞いを抑制する必要があるため、各サーバーの管理者を二人として互いに監視させるなどの対策が必要になる。また、悪意ある攻撃者に対抗できる秘密計算と異なり、他のサーバーが誤っても気付くことはできないため、サーバー間の責任の分担が難しい。

秘密計算（参加者が三人の方法での管理業務の例）



４．３．２ 活用に向けた検討事項 - システム開発（連合学習）

- 連合学習のうち、特にプライバシー強化連合学習を導入することで、更なるデータの利活用が可能となる。
- 一方、組織を跨いだデータ利活用の場合、通常のセキュリティ対策だけでなく、従来よりも機密レベルの高いデータを取り扱うことによる新たな運用やセキュリティ対策が必要となる。

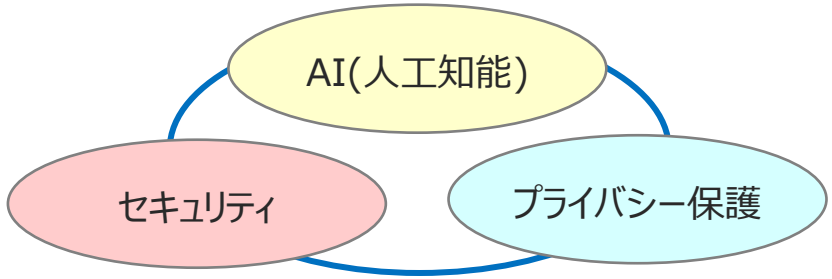
データ分析のライフサイクル	課題	対応方針
計画	• 適切なプライバシー強度に合意しておかないと、許可された以上のプライバシー情報が漏洩してしまう可能性がある。	• プライバシー強度の適切な評価 - 保護範囲の合意 - 利用の同意の管理　等
データ収集・保存	• 従来の分析より機密レベルの高いデータを学習で扱うことになるため、より適切な管理が必要になる。	• 高い秘密に適合したデータ管理 • データプライバシー管理システムの導入
データ分析	• 適切なモデルを学習するために、予め様々な側面からの分析が必要である。 • この分析フェーズでのプライバシー保護のため、プライバシー強化連合学習を含め別途、保護方式を検討する必要がある。	• 特別に許諾を得た限定的なデータで分析する 等
モデル学習	• クライアント側でのローカルな学習において、不正な学習アルゴリズムが使用された（悪意を持って入れ替えられた、等）場合、クライアント側の秘密情報が抜き取られる可能性がある。	• クライアントでの学習アルゴリズムの正しさの監視 • 適切なサーバー管理体制
	• 一部のクライアントでの学習データが正確でない場合、連合学習によって得られた全体の学習モデルも不正確になる。	• 全てのクライアントが必ず適切に学習するための仕組みを作る（クライアント間での事前合意、ガバナンス、法規制、等）
	• 学習するモデルによって、計算コストが大きく、学習に長時間かかる場合がある。	• 十分な計算資源を評価して準備
モデル利用	• 学習したモデルを利用する時、その学習モデルを通じてクライアントが保持する秘密情報にアクセスできてしまう。不正な学習モデルであった場合、秘密情報が盗まれる可能性がある。	• 学習モデルの正しさの監視 • 学習モデル利用の正しさの監視
破棄	• プライバシーデータが含まれた学習モデルや学習用データが適切に破棄されない場合、それらが漏洩してしまう可能性がある。	• 確実に完全な破棄を実行

４．３．３ 活用に向けた検討事項 - 人材・体制

- AIやセキュリティに加えて、プライバシー保護関連の法制度の知見が必要。
- 一方、プライバシーは取り扱う情報や技術、取り巻く環境によって変化しており、プライバシー保護の観点で考慮すべき範囲が拡大している。プライバシー問題全体を考えられる体制の整備が必要。

必要な人材

- プライバシー強化技術の活用には、AIやセキュリティに加えて、プライバシー保護関連の法制度の知見が必要。



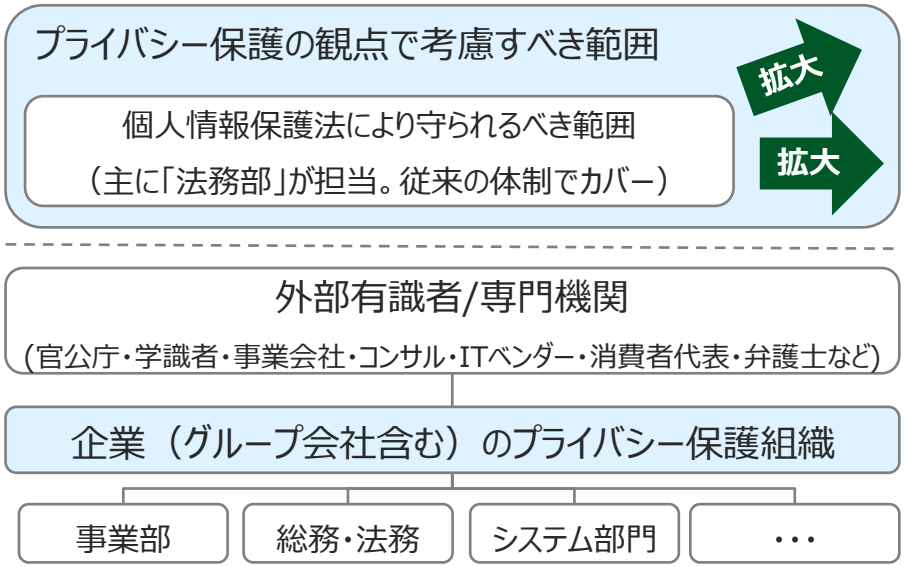
- 現在、プライバシー保護に関する認定資格が開始されている。今後、プライバシー強化技術の活用が進むことによって、それら関連資格も登場し始めることが考えられる。

資格名	発行団体
EXIN Privacy and Data Protection	EXIN（オランダの省庁が設立した機関）
CIPP/CIPM/CIPT	IAPP（米国の非営利団体）
個人情報保護士	一般財団法人全日本情報学習振興協会
CPA/CPP/CPO	一般社団法人日本プライバシー認証機構

プライバシーガバナンスの構築

- プライバシーは取り扱う情報や技術、取り巻く環境によって変化しており、プライバシー保護の観点で考慮すべき範囲が拡大している。そのため、プライバシー問題全体を考えられる体制（企業内で異なる分野の担当者が横断的に連携、産官学で社会的に監視・監督する体制の整備など）が必要。

プライバシー保護の観点で考慮すべき範囲と体制



参考：経済産業省HP プライバシーガバナンス（https://www.meti.go.jp/policy/it_policy/privacy/privacy.html）

４．４　今後の展望

- プライバシー強化技術（PETs）に関する法整備やプライバシーガバナンスの構築に関する検討と並行して、当面は、個人情報保護法の規制を受けない企業内データ（機密情報など）に対してPETsを活用し、組織間の共通課題解決／社会課題解決する取り組みが進展（次頁詳細）。
- 社会実装と法整備の進展により、将来的には、組織間でパーソナルデータに対するPETsの活用事例も登場し、新たな付加価値・新サービス創出が期待される。

現在 → 将来

PETs活用

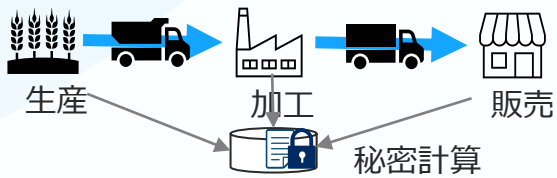
パーソナルデータ（含む、個人情報）活用

- 消費者の個人情報や行動履歴、購買履歴など**パーソナルデータに対してPETsを活用**

法律・規制を前提とした範囲内でのデータ活用（次頁詳細）

- [データ利活用]**機密保護のため、眠ったままになっている企業内データを開放。**PETsを用いて、**組織間の共通課題解決/社会課題解決**
- [安全強化]機密性の高いデータの保管にPETs活用。高セキュリティを確保

例）サプライチェーンの最適化



社会環境

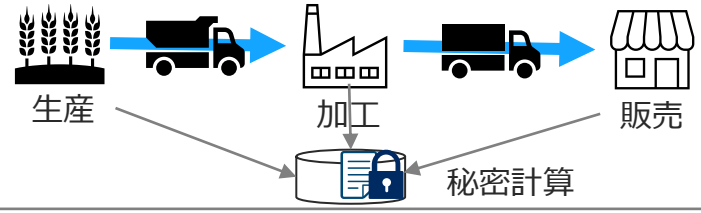
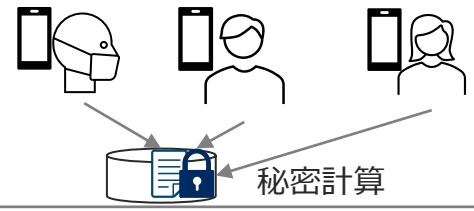
法整備・プライバシーガバナンスの構築

- プライバシー強化技術の**法律・規制面での位置づけ・活用範囲の検討、明確化**
- プライバシー保護のために**第三者機関（産官学）による監視・監督の体制構築**

日本総研 The Japan Research Institute, Limited

４．５　活用が進む想定ユースケース

- 足元、プライバシー強化技術（PETs）の活用が進むユースケースを整理すると、主に３つのパターンが考えられる。
- データ利活用の観点では、現時点の法律・規制を前提とした範囲において、プライバシー強化技術の活用が進展。例えば、個人情報保護法の規制を受けないプライバシー情報以外のデータ（企業の機密情報等）に適用。
- 安全管理の強化の観点では、情報銀行や鍵管理などのセキュリティ強化として活用が進む。

観点	パターン	要点	ユースケース例
データ利活用	① 組織間の 共通課題解決 /社会課題解決	法律上の制約を受けないデータ(企業の一般的な機密情報)を組織間で共有することで、共通課題を解決するためPETsが活用される	• サプライチェーンの最適化 生産/運送/販売計画等の機密情報をPETsで共有し最適化 
		組織間でパーソナルデータを利活用する場合、個人の同意取得のハードルが低くなると想定される社会課題の解決にPETsが活用される	• 伝染病の拡散防止 デバイスの接触履歴/ 位置情報から伝染病 罹患者との接触を特定 
	②プライバシー強化	パーソナルデータを収集・活用する際に、プライバシーの保護を保証する	• 個人の健康状態の収集・分析 • 個人の高精度な位置情報の収集・分析
安全強化	③セキュリティ強化	機密性の高いデータの保管にPETsを使用し、従来より高セキュリティを確保	• 情報銀行 • 電子資産の保護（鍵管理）

４．６　おわりに

- プライバシー保護の原則を実現・強化するうえでプライバシー強化技術（PETs）の普及・活用が重要であるが、技術開発の進展・成熟度のほかに、法律やシステム開発、人材・体制などの面からも検討が必要。今後は、パーソナルデータの利活用に伴い享受できるメリットとプライバシー侵害のリスクのバランスを取るべく、社会的な議論が進展すると予想され、その中で、プライバシー強化技術の活用に関しても、法律・規制面での位置付け・活用範囲が検討・明確化されるであろう。

- そのような中、当面は、個人情報保護法の規制を受けない企業内データ（機密情報など）に対してPETsを活用し、組織間の共通課題解決/社会課題解決する取り組みが進展することが予想される。これら社会実装と法整備の進展により、将来的には、組織間でパーソナルデータに対するPETsの活用事例も登場し、新たな付加価値・新サービス創出が期待される。

- プライバシー強化技術は、プライバシー侵害のリスクを低減させる有効な技術である一方、プライバシー強化技術を使えば全てが安心というわけではなく、正しくシステムを設計・実装・運用するためには、従来通り慎重な対応が求められる。実サービスへの適用にあたっては、現時点から検討・実証実験等を繰り返し、どのような事例に適用できるか見極めが必要である。その際、技術自体の理解だけではなく、セキュリティの知識、法律などの複合的な知識が必要なため、自社に十分な知見がない場合は、外部の専門家の知見を活用し、助言・指導などを求めることが肝要である。