プライバシー保護技術の動向
～主要技術の概要と組み合わせ事例の紹介～

LINEヤフー株式会社
Privacy & Trust Team, Manager

竹之内 隆夫

LINEヤフー

日本銀行金融研究所情報技術研究センター
第24回情報セキュリティ・シンポジウム
2024年２月15日

自己紹介

- 氏名：竹之内 隆夫 (たけのうち たかお)
- 所属：LINEヤフー株式会社
  Privacy & Trust Team, Manager
- 経歴：
  - 前前職・前職も含めプライバシーで10数年の研究開発
    - k-匿名化、秘密計算、差分プライバシー
  - 技術だけでなく普及促進・法制度議論も
    - 学会 企画イベント、特集記事企画、解説論文 等
    - 「秘密計算研究会」の設立
    - データ社会推進協議会「秘密計算活用WG」 設立
    - 「プライバシーテック協会」のアドバイザー

- その他： 博士(工学)、MBA、大学講師、書籍執筆、講演など

情報処理学会の会誌で
プライバシー保護技術の特集企画

本発表の趣旨

- プライバシー保護技術の重要性
- トレンドなプライバシー保護技術
  - 差分プライバシー、連合学習、MPC/TEE(”秘密計算”)
  - 組み合わせた事例が増加

- 今後の競争軸：「連携」
  - 組み合わせが重要 → 異なる技術分野の連携
  - 技術だけでない　→ 技術者と非技術者の連携

Contents

Contents

プライバシーとは
(非専門家・エンジニアの視点で)

- プライバシーとは、時代・地域・文化等の社会的背景や個人の感覚で変化する複雑な概念
  → 定義が不明確・変化するため、(ある程度は) 動向を追う必要がある（が大変）

■ プライバシー（の権利）の特徴
（大御所の堀部政男先生の文献引用※1）

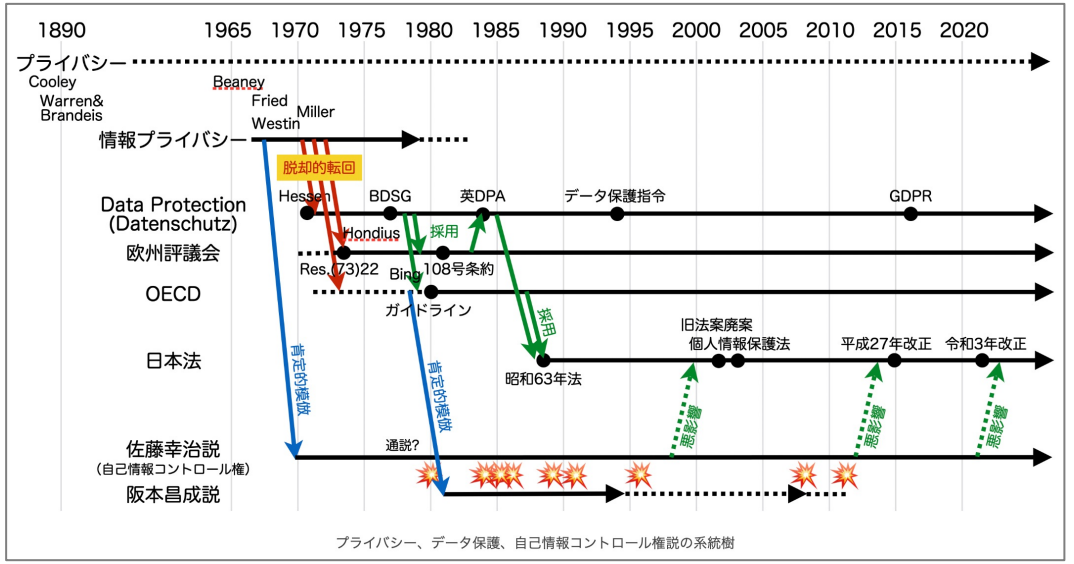
「プライバシーの権利」ないし「プライバシー権」の意味するところは、歴史的に異なる

■ プライバシー（の権利）の歴史的変化※1

- ひとりにしておかれる権利（right to be let alone）
  Samuel D. Warren and Louis D. Brandeis, “The Right to Privacy” (1890)
- 私生活をみだりに公開されないという法的保障ないし権利
  「宴のあと」 東京地裁 判例 (1960)
- 個人，グループ又は組織が、自己に関する情報をいつ、どのように、また、どの程度に他人に伝えるかを自ら決定できる権利
  Alan F. Westin, “Privacy and Freedom”(『プライバシーと自由』) (1967)

※1 出典・参考：堀部政男, "プライバシーを守ったITサービスの提供技術：1. プライバシー・個人情報保護論議の世界的展開と日本", 情報処理,54(11),1106-1114 (2013-10-15)

■ 「自己情報コントロール権」が主流な中、別の考えも



出典：高木浩光(語り手),小泉真由子(聞き手),宇壽山貴久子(撮影),"高木浩光さんに訊く、個人データ保護の真髄 ――いま解き明かされる半世紀の経緯と混乱", 情報法制研究所, https://cafe.jilis.org/2022/03/18/160/

プライバシー原則とプライバシー保護技術

- Privacyには様々な解釈があるが、プライバシー原則は一定の合意があると思われる
  - 例：OECDガイドライン※2、Privacy by Design※3等。各国法制度はEU GDPRを参考※1。
- 「data minimization」原則のためには、**プライバシー保護技術の継続的な適用**が必要※5

GDPRのプライバシー原則※1

原則	概要
Lawfulness, fairness and transparency	合法、公正、透明性ある方法で処理すること
Purpose limitation	特定された明示的で正当な目的で、収集・処理すること
Data minimization	目的達成のために関連※4する必要最小限のデータ収集・処理であること
Accuracy	正確なデータであること
Storage limitation	目的達成後は削除すること
Integrity and confidentiality	データの完全性、機密性を保つこと（セキュリティ技術）
Accountability	上記原則の遵守を説明・証明できること

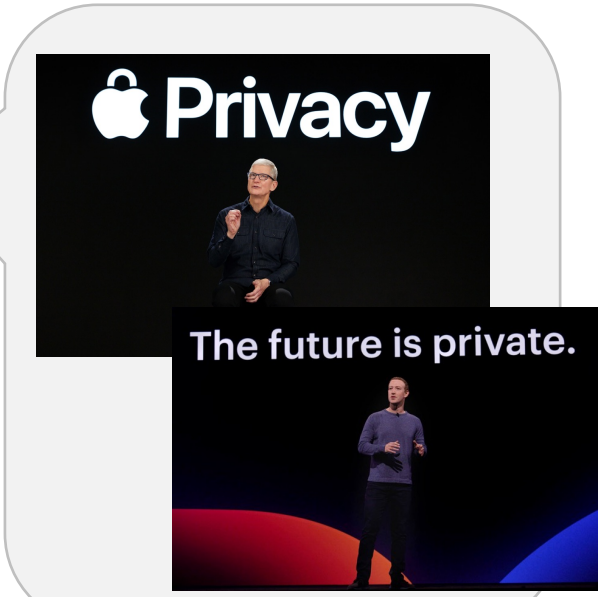
※1 EUのプライバシ関係の規則であるGDPR(General Data Protection Regulation)は、日本・米国・アジア圏の法制度に強く影響しているため、ここではGDPRのプライバシ原則（Privacy Principals)を抜粋。なお、原文ではminimisationであるが、本資料ではminimizationと表記している。
https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/
※2 “Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data”, OECD, 1980年制定,2013年更新
※3 "Privacy By Design", アン・カブキアン博士, https://www.soumu.go.jp/main_content/000196322.pdf
※4 OECDガイドラインの第２原則 “Personal data should be relevant to the purposes(略)”の意味
※5 例えば、EU GDPR第25条では「取扱いの方法を決定する時点及び取扱いそれ自体の時点の両時点において」（個人情報保護委員会の仮日本語訳から抜粋）、このような原則を満たすことを求めている。

プライバシーは経営戦略の一部に

- ユーザのプライバシー意識の高まりにより、プライバシーを経営戦略に位置付け
- 目的：ユーザ・企業からのデータ収集増 （データはBigTech各社の競争優位性の源泉）
- 目標：法令遵守は当然。それ以上のブランド構築
- 活動：先進的な技術導入と対外コミュニケーション (プライバシー技術は見え難いため)

プライバシー保護のレベル	行っている活動内容の例	
	技術	対外コミュニケーション
ブランド構築のレベル	➢ 先進的な技術開発と事業への導入	➢ 積極的なアピール ➢ 法制度・仕様等の検討リード
法令遵守のレベル	➢ 従来技術の導入	➢ 最低限の情報開示

BigTech企業の例

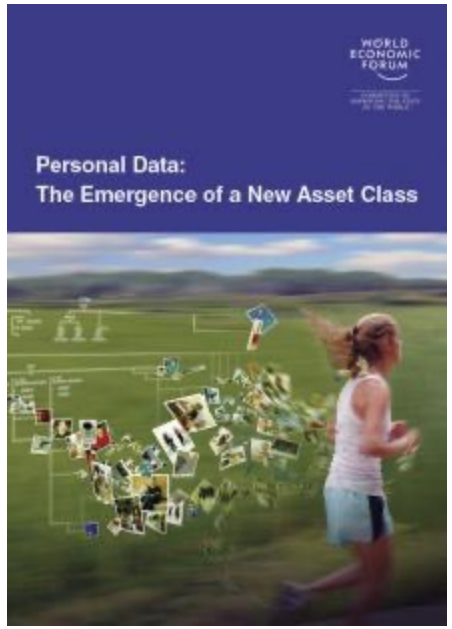


出典: https://time.com/6071901/apple-iphone-privacy-wwdc-2021-vpn/
https://about.fb.com/news/2019/04/f8-2019-day-1/

参考：“Personal data is the new oil” (約15年前)

- “Personal data is the new oil of the Internet and the new currency of the digital world.”

Meglena Kuneva, European Consumer Commissioner, March 2009



“Personal Data: The Emergence of a New Asset Class”,
An Initiative of the World Economic Forum, January 2011,
In Collaboration with Bain & Company, Inc.
https://jp.weforum.org/reports/personal-data-emergence-new-asset-class

参考：企業による戦略的なデータ活用と規模化

いわゆるBigTech企業は戦略的にデータを収集・活用し、そのサービス分野で”独占”の傾向※1

→ 今後は、単なる規模化ではなく**データを出さずに連携**することが重要となる可能性

約15年前と近年の時価総額ランキング（世界）の比較※2

2006	2007
Exxon Mobil	PetroChina
GE	Exxon Mobil
Microsoft	GE
Citi	China Mobile
Gazprom	ICBC
PetroChina	Microsoft
ICBC	Gazprom
トヨタ自動車	Royal Dutch Shell
Bank of America	AT&T
Royal Dutch Shell	Sinopec

2023	2024
Apple	Microsoft
Saudi Aramco	Apple
Microsoft	Saudi Aramco
Alphabet (Google)	Alphabet (Google)
Amazon	Amazon
Berkshire Hathaway	NVIDIA
Tesla	Meta (Facebook)
NVIDIA	Berkshire Hathaway
Exxon Mobil	Eli Lilly
Tencent	Tesla

日本企業との比較

出典：日本経済新聞, 2021年8月26日

GAFA、時価総額で日本株超え 安定収益が資金呼ぶ

GAFA（親会社のアルファベット含むグーグル、アップル、フェイスブック、アマゾン・ドット・コム）と日本株全体の時価総額が逆転した。(以下略)

参考※2： 2024年2月時点
東証プライム約1800企業 約900兆円
Apple 約400兆円(約3兆ドル)

※2 ランキングや企業価値情報は公開情報をもとに著者調べ

※1 参考：“ゼロ・トゥ・ワン 君はゼロから何を生み出せるか”, ピーター・ティール, ブレイク・マスターズ他, 3章や5章

Contents

トレンドなプライバシー保護技術の例

①差分プライバシー(Differential Privacy)、②連合学習(Federated Learning)、
③MPC/TEE(“秘密計算”) ※2が、BigTech企業等での採用も多い※3

主要なプライバシー技術の一覧※1

カテゴリ	技術名
プライバシー保護の「フレームワーク」	連合学習(Federated Learning)
	データ合成
	PIR(Private Information Retrieval)
プライバシー保護の「技術」（プライバシー保護の実現のためのBuilding Block）	差分プライバシー(Differential Privacy)
	MPC (Multi-Party Computation) / TEE (Trusted Execution Environment)
	k-Anonymization（K-匿名化）
	ゼロ知識証明

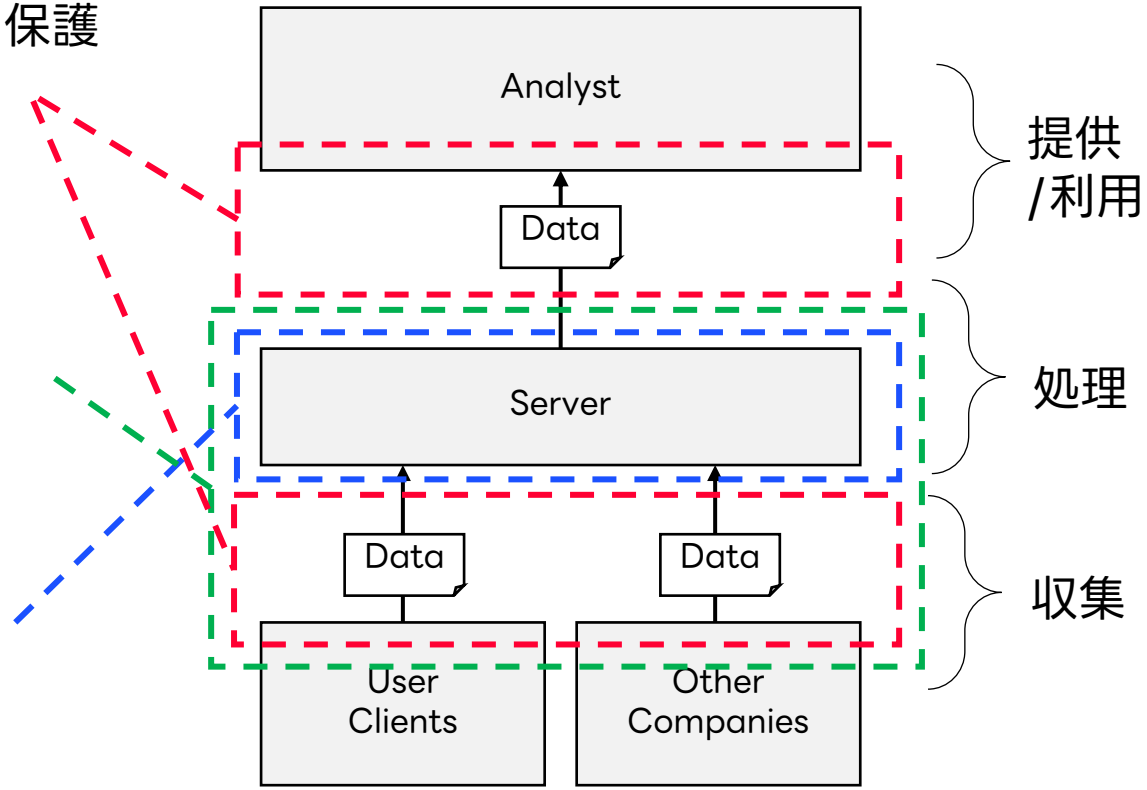
※1 参考文献： ENISA(The European Union Agency for Cybersecurity) Data Protection Engineering https://www.enisa.europa.eu/publications/data-protection-engineering
日本総研 プライバシー強化技術の概説と動向 https://www.jri.co.jp/page.jsp?id=101511
デロイトトーマツ, プライバシー強化技術の紹介動画「A day with PETs」, https://www2.deloitte.com/jp/ja/pages/deloitte-analytics/articles/a-day-with-pets.htm
※2 TEEはConfidential Computingとも呼ばれる。秘密計算は英語ではSecure Computationとも呼ばれ、日本では秘匿計算と呼ばれることもある。本資料では広くデータを秘匿したまま処理する技術という観点でMPCとTEEを同様な技術と捉えて記載。
※3 ※1の資料を参考に、発表者の視点で3つの技術が特にBigTech企業での採用が多いと判断。参考：竹之内, ”データ活用を促進するPrivacy Tech - データが競争優位となる時代の成長戦略 -”, 情報処理 2024年3月号, 情報処理学会

トレンドのプライバシー保護技術の概要

- トレンドの3つのプライバシー保護技術は、「収集」「処理」「提供」の各フェーズでデータを保護 → 組み合わせた活用が可能

技術名	技術概要
①Differential Privacy “ぼかす”	ビッグデータに適したデータ保護 • 数学的保証のある“ぼかし” • データ：大 → ぼかし：小 • プライバシーの定量化
②Federated Learning “減らす”	収集データの最小化 • クライアント端末で学習 • 更新情報だけを収集
③MPC/TEE “隔離する”	誰も関与できない秘密計算 • データを秘匿したまま処理

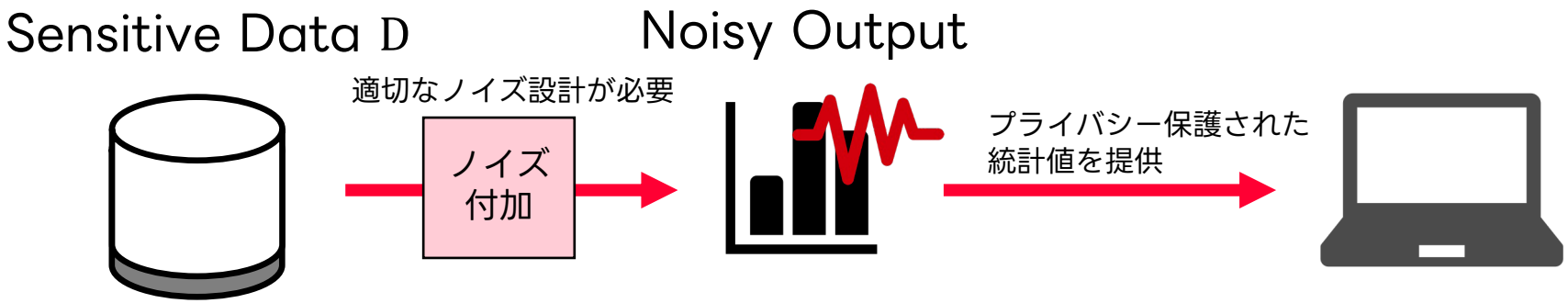
プライバシー保護の箇所



MPC : Multi-Party Computation
TEE : Trusted Execution Environment

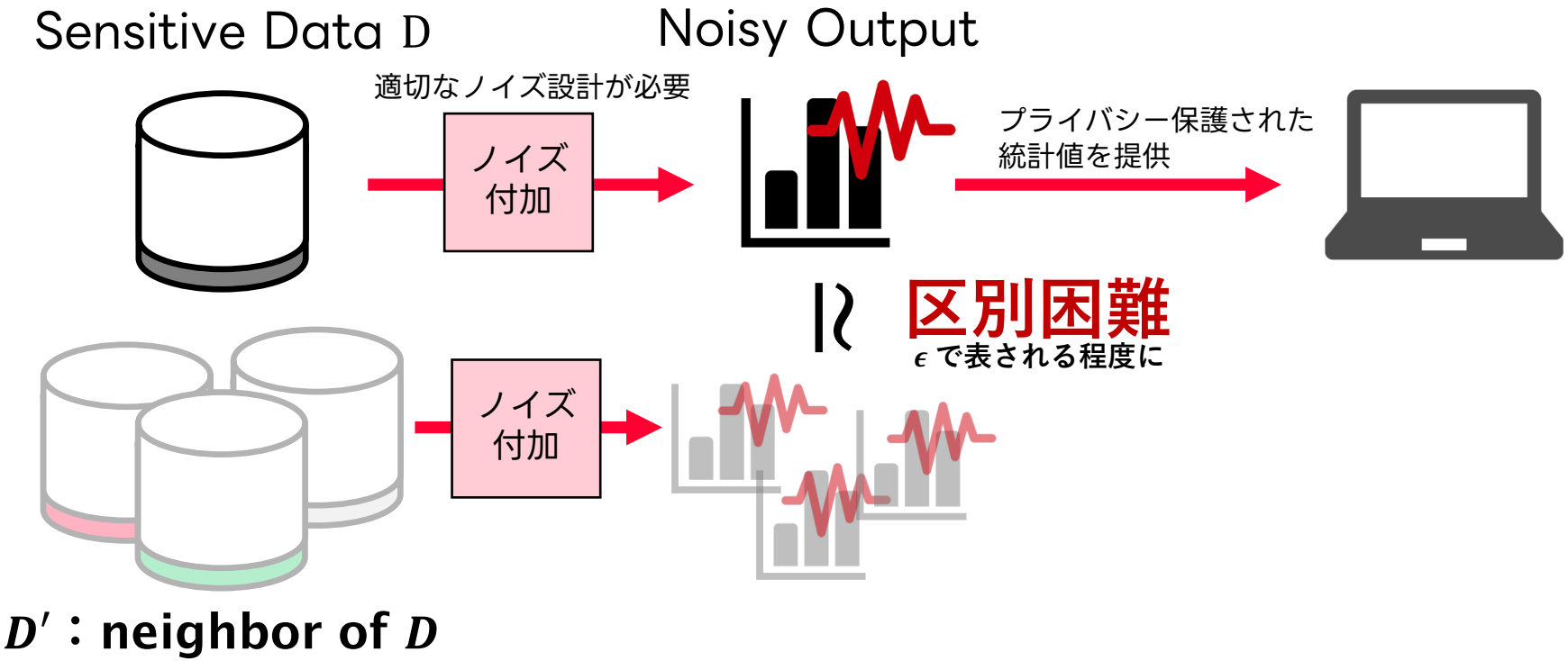
①Differential Privacy(差分プライバシー)とは

- データ収集・解析の結果に対してプライバシーの水準を統計的に表現した尺度
  - 統計的に「どれだけ他人と見分けがつかないか」をプライバシーパラメータ ε で表現
- （ノイズの加算により）いかなる知識との突合にも頑健なプライバシーを提供



①Differential Privacy(差分プライバシー)とは

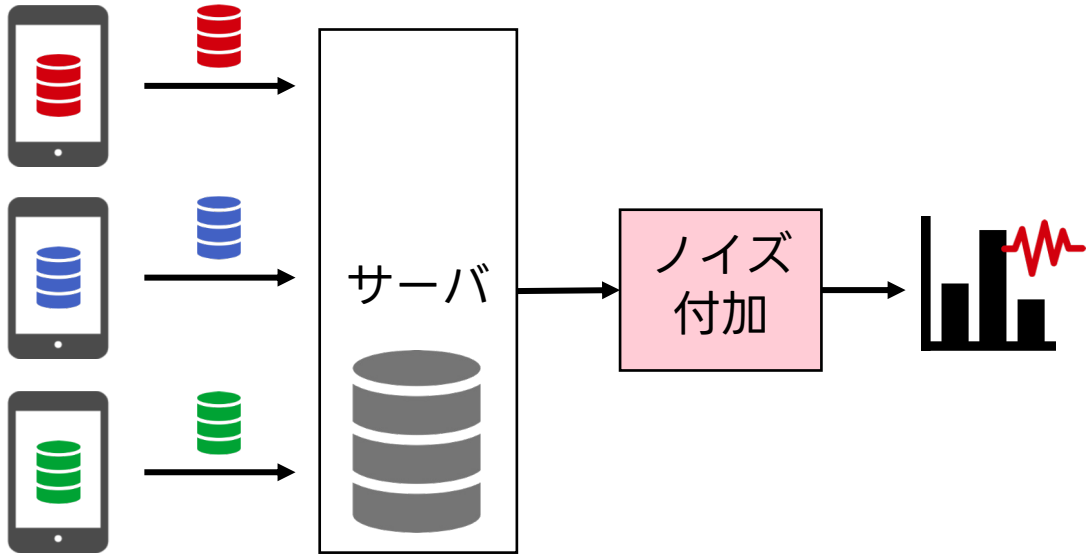
- データ収集・解析の結果に対してプライバシーの水準を統計的に表現した尺度
  - 統計的に「どれだけ他人と見分けがつかないか」をプライバシーパラメータ ε で表現
- （ノイズの加算により）いかなる知識との突合にも頑健なプライバシーを提供



差分プライバシーの2種類のモデル

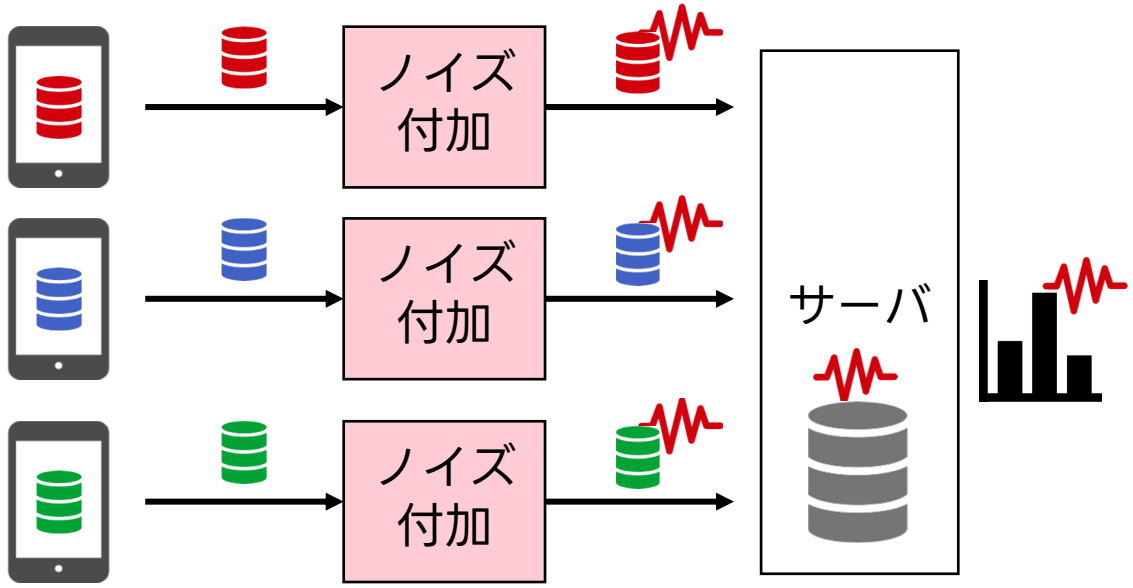
(A) Central Differential Privacy(CDP)

サーバーから第三者への統計値の提供時
（サーバー側でノイズ付加）



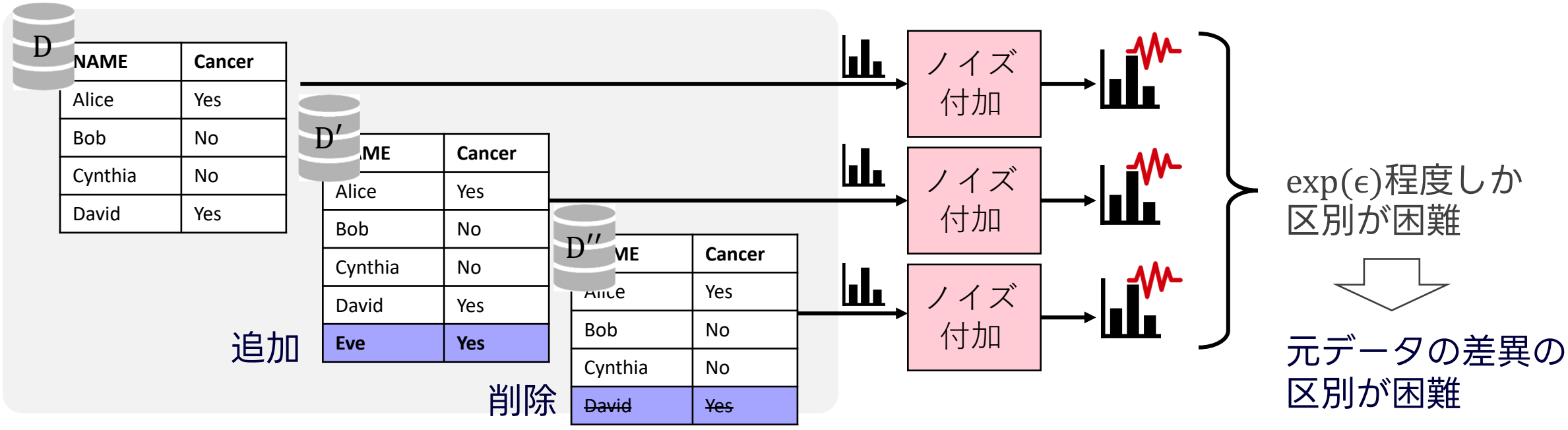
(B) Local Differential Privacy(LDP)

クライアントからサーバーへのデータの収集時
（クライアント側でノイズ付加）



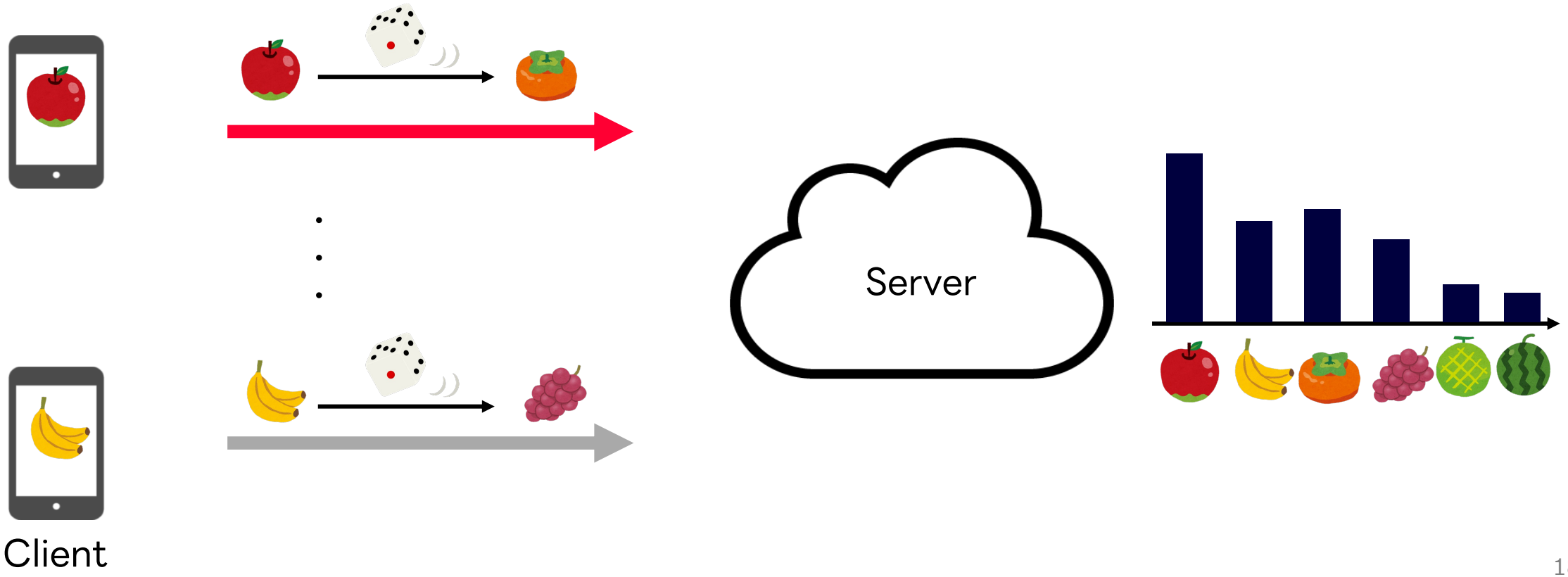
(A) Central Differential Privacy

- サーバーで収集したデータの集計結果にノイズを付加
- 1名のデータの存在/非存在が集計結果の傾向から区別できない



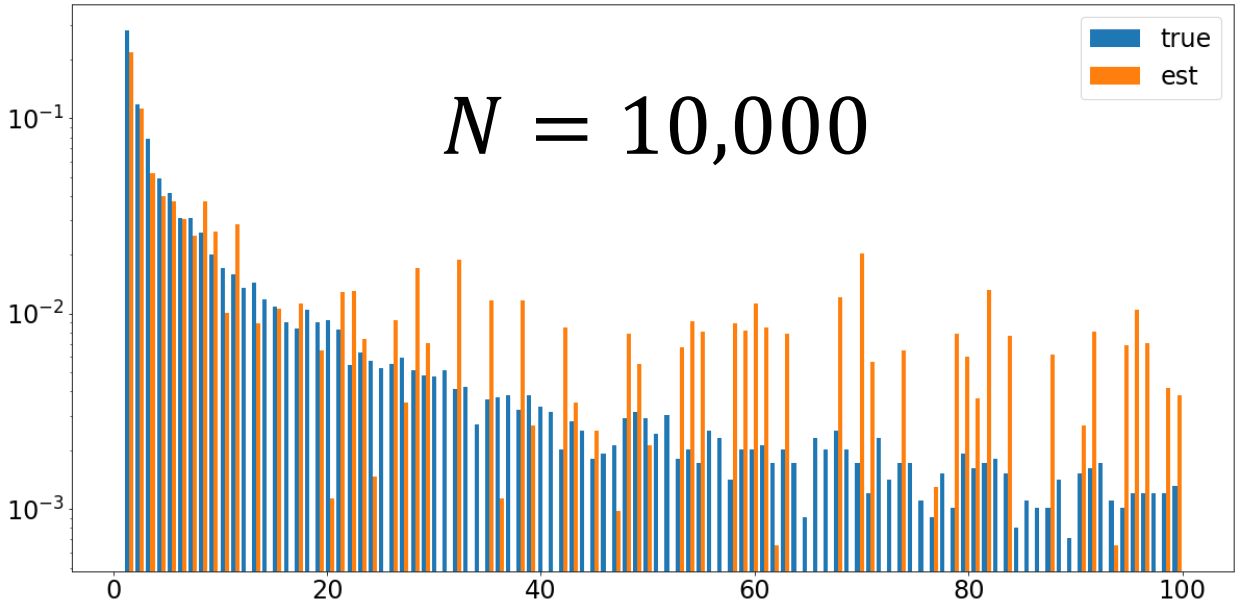
(B) Local Differential Privacy

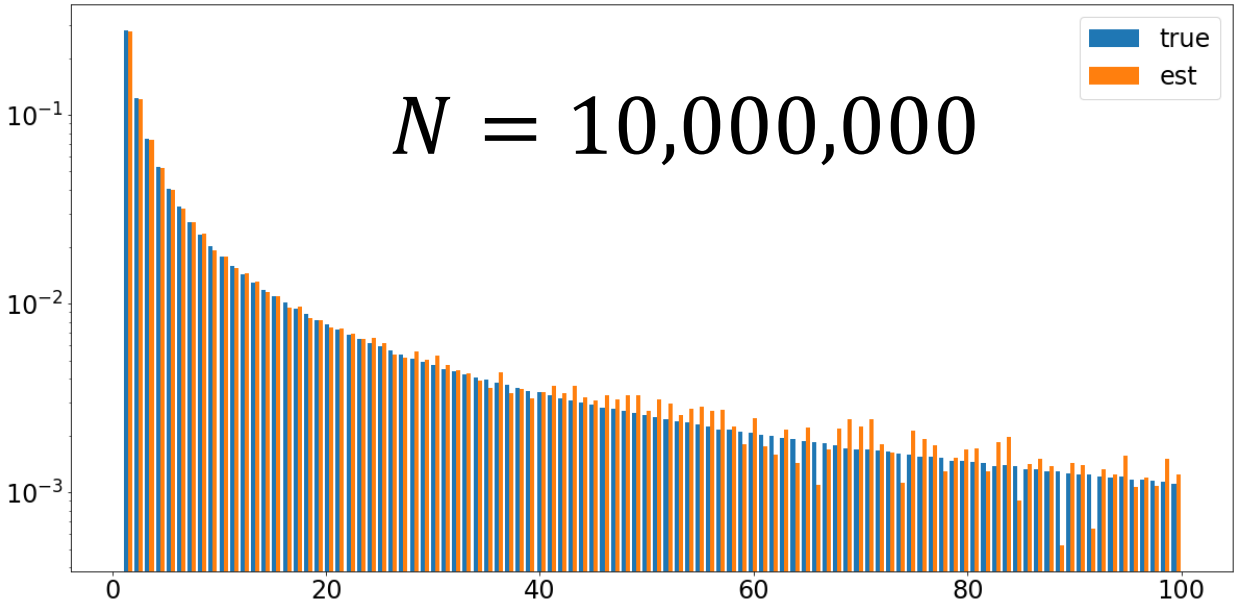
- クライアントは情報を隠しつつも, サーバは真の収集結果を推計する
- (Local Differential Privacyを満たした形で実現)



(B) Local Differential Privacyの集計例

- 集計するランダム化レポートの数が多いほど正確な統計を推定できる





Contents

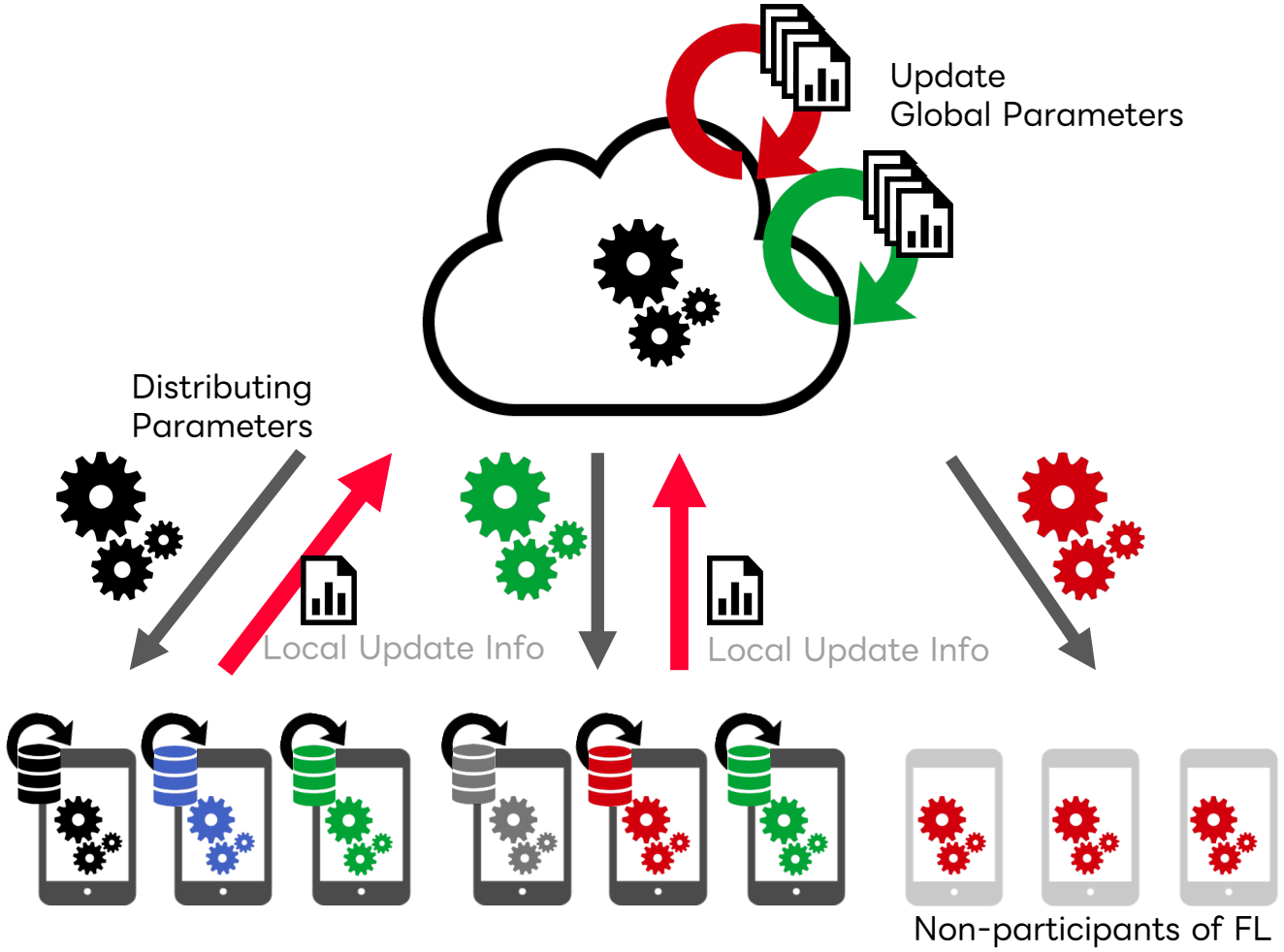
Federated Learning(FL, 連合学習)とは

Federated Learning (FL) とは

- クライアントで機械学習を実施して
  更新情報だけをサーバーが収集
  (データはクライアントから出ない)

解決する課題

- クライアントでしか扱いを許容されない
  機微データの活用を実現
- サーバーのデータ管理コストの削減



Federated Learningにおけるプライバシーリスク

学習モデルの更新情報（勾配）から
訓練データ (画像) を復元できる
➔プライバシー保護が必要

Figure 3: Single-Image Reconstruction from the parameter gradients of trained ResNet-152. Top row: Ground Truth. Bottom row: Reconstruction. We check every 1000th image of the ILSVRC2012 validation set. The amount of information leaked per image is highly dependent on image content - while some examples like the two tenches are highly compromised, the black swan leaks almost no usable information.

Figure 6: Information leakage for a batch of 100 images on CIFAR-100 for a ResNet32-10. Shown are the 5 *most* recognizable images from the whole batch. Although most images are unrecognizable, privacy is broken even in a large-batch setting. We refer to the supplementary material for all images.

(出典)
“Inverting Gradients - How easy is it to break privacy in federated learning?”
https://arxiv.org/abs/2003.14053

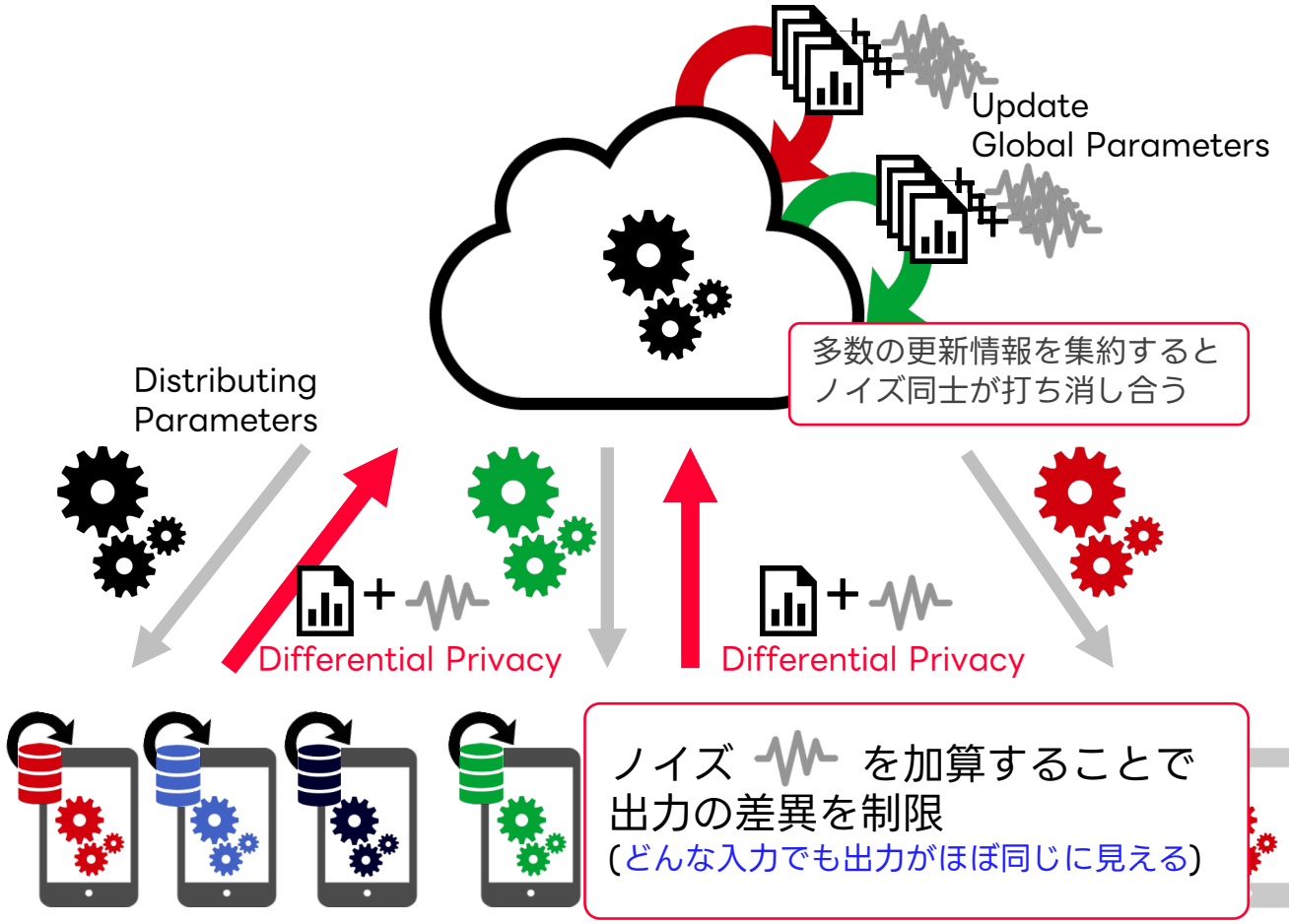
Federated LearningにDifferential Privacyを適用

FLにDifferential Privacyを適用

- 更新情報を他人と見分けがつかない形に
- モデルからの訓練データの推定を困難に
- 有効な学習には膨大なクライアントが必要

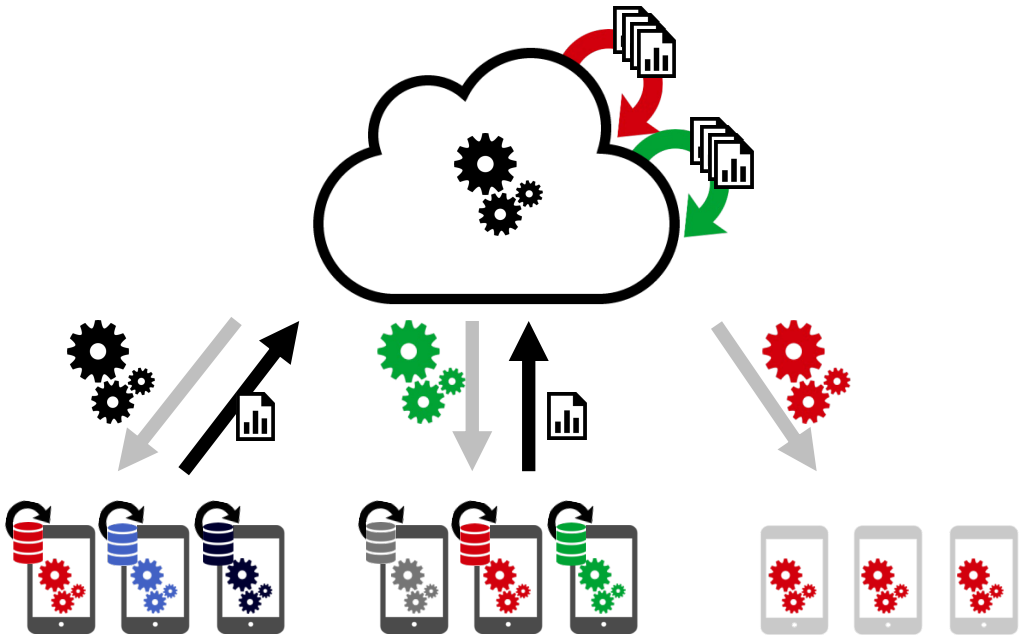
解決する課題

- 厳密なプライバシーの保証と管理



Federated Learningのバリエーション

Cross Device型



クライアント：多数
データサイズ：小
通信回線：従量課金 / wifiなど

Cross Silo型

クライアント：少数
データサイズ：大
通信回線：専用線など

Contents

TEE : Trusted Execution Environment
MPC : Multi-Party Computation

MPC/TEE(“秘密計算”)

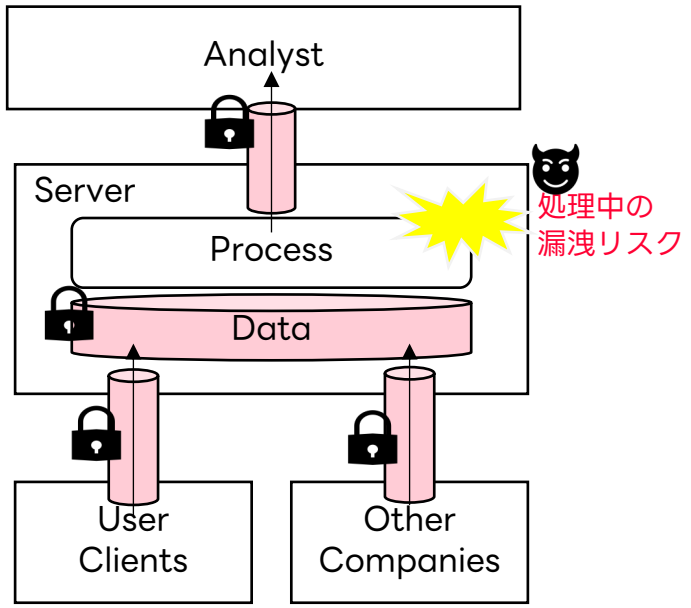
MPC/TEE とは

- データの「処理中」も暗号化できる暗号技術 (従来の暗号化は「通信中」と「保存中」のみ暗号化)
  - TEE: ハードウエアのチップを利用した方式
  - MPC: ソフトウエア（暗号理論）を利用した方式

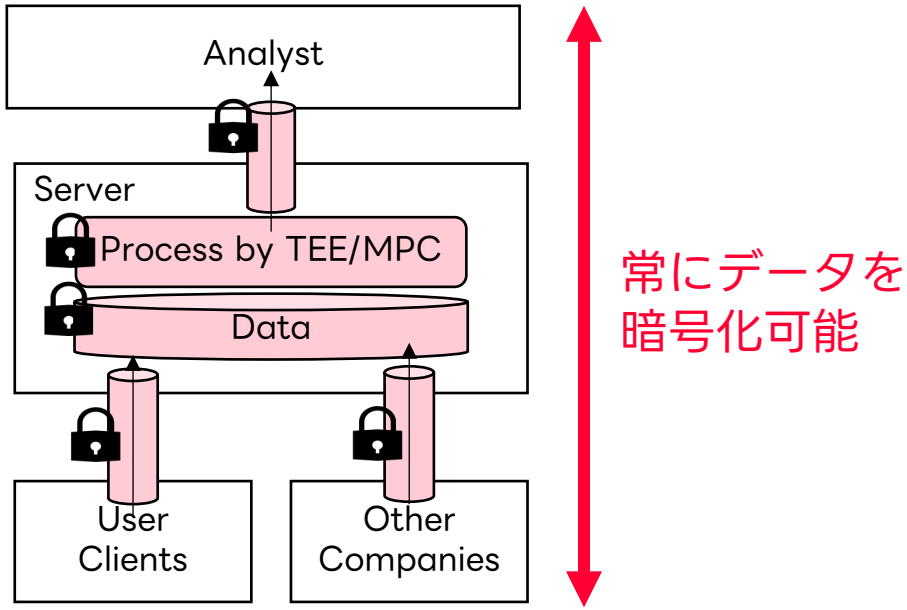
解決する課題

- 常にデータの暗号化を実現するため、管理者や不正者からの不正を防止

従来の暗号技術を使ったシステム



TEEやMPCを適用したシステム



機密性(Confidentiality)と完全性(Integrity)

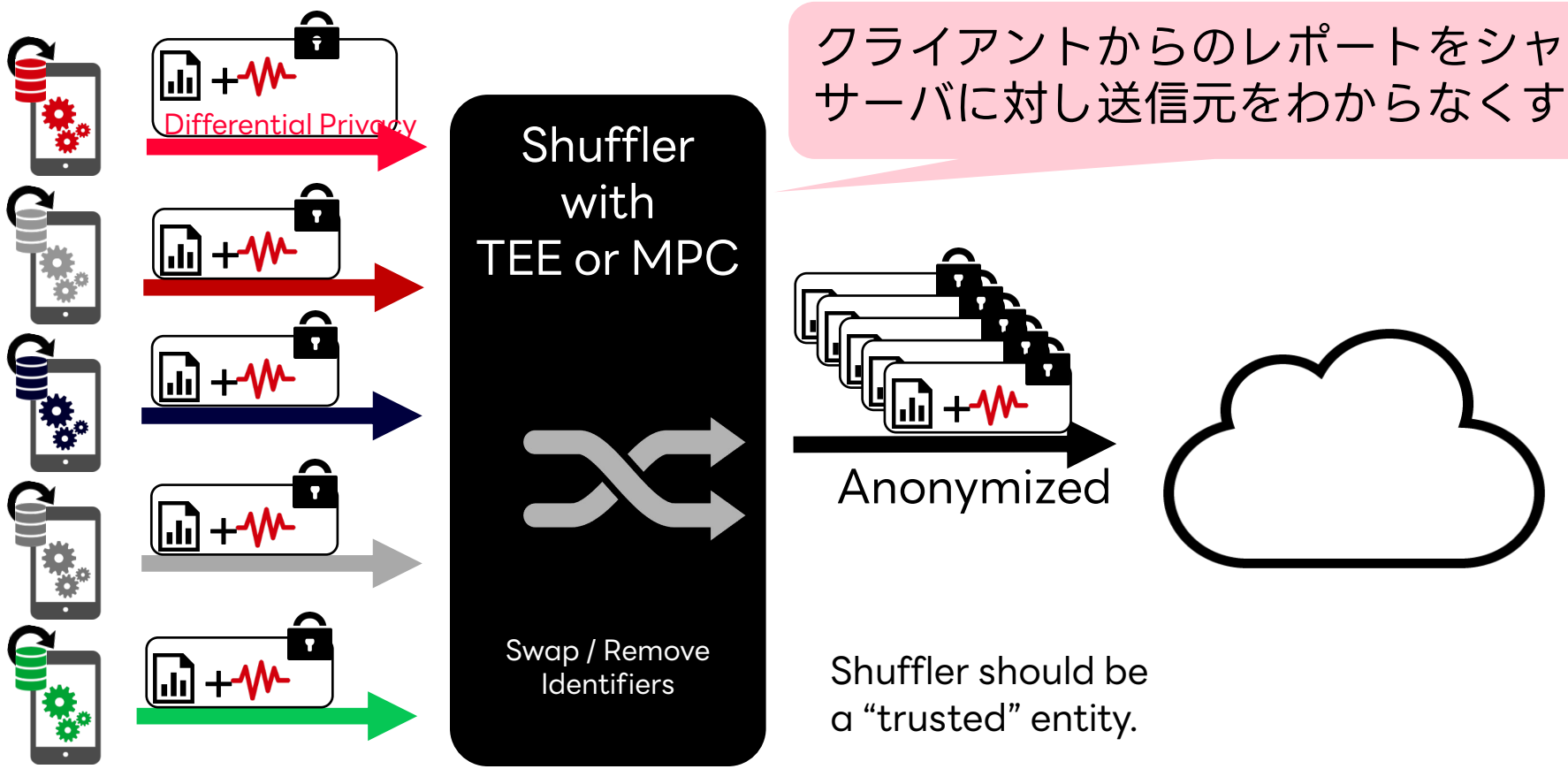
- MPCとTEEは機密性（データを秘匿したままの処理）を提供
- TEEは完全性（処理されているプログラムが不正改竄の検知）も提供

表：MPCとTEEの機密性と完全性の対応

	機密性	完全性	安全性の前提
MPC	対象	対象外	暗号学的な保証 （アルゴリズムの安全性証明あり）
TEE	対象	対象	ハードウエアの安全性に依存

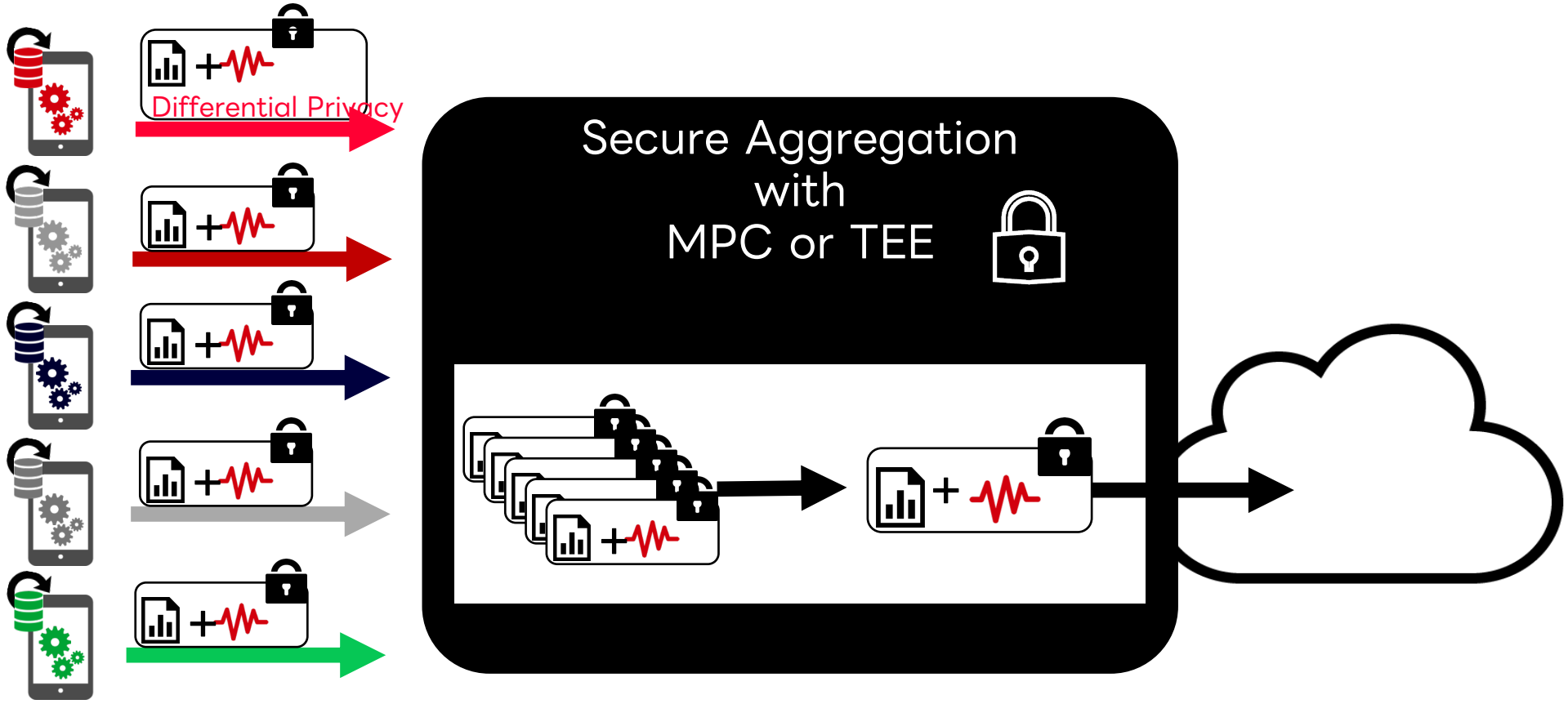
DPとTEE/MPCとの組み合わせ：シャッフリング

- クライアントが送るデータをシャッフルすることでプライバシー保護を強化
- シャッフル処理は、ある程度の信頼が必要であり、MPCやTEEが有効な手段の一つ



DPとTEE/MPCとの組み合わせ：Secure Aggregation

- クライアントが送るデータを、MPCやTEEで秘匿しながら集計することでプライバシー保護を強化



Contents

スタンプの自動推薦 (1/2)

- 意味的なタグをもとに推薦
  - Sticker suggestions based on semantic labels
- 入力文字からインクリメンタルに推薦
  - Incremental suggestions while text input, using pre-defined keywords associated with the each label

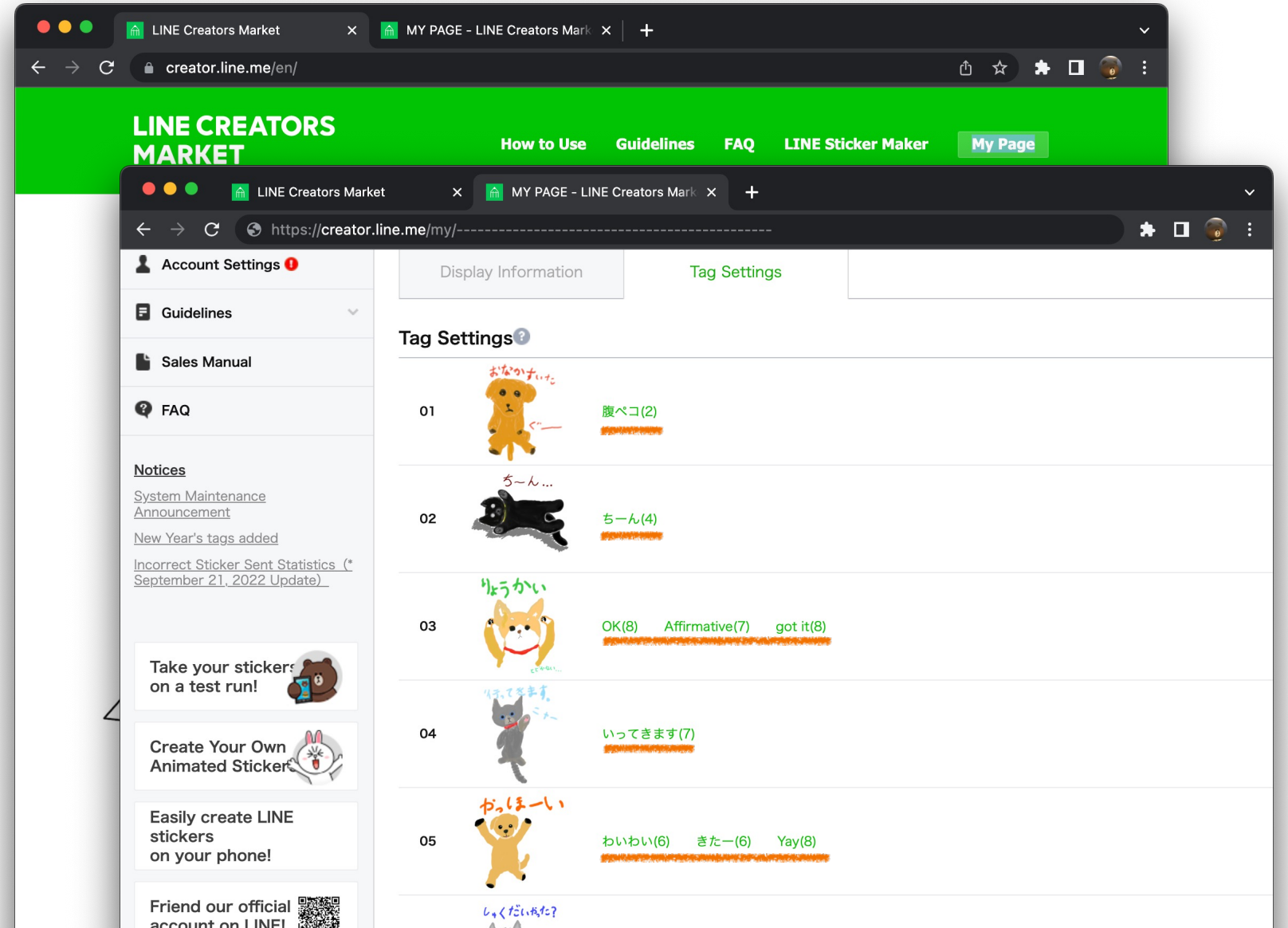
※ユーザの入力した文字は、対応する意味的なラベルを特定するためだけに利用されます。

スタンプの自動推薦 (2/2)

- 意味的なタグをもとに推薦
  - Sticker suggestions based on semantic labels
- 入力文字からインクリメンタルに推薦
  - Incremental suggestions while text input, using pre-defined keywords associated with the each label

※ユーザの入力した文字は、対応する意味的なラベルを特定するためだけに利用されます。

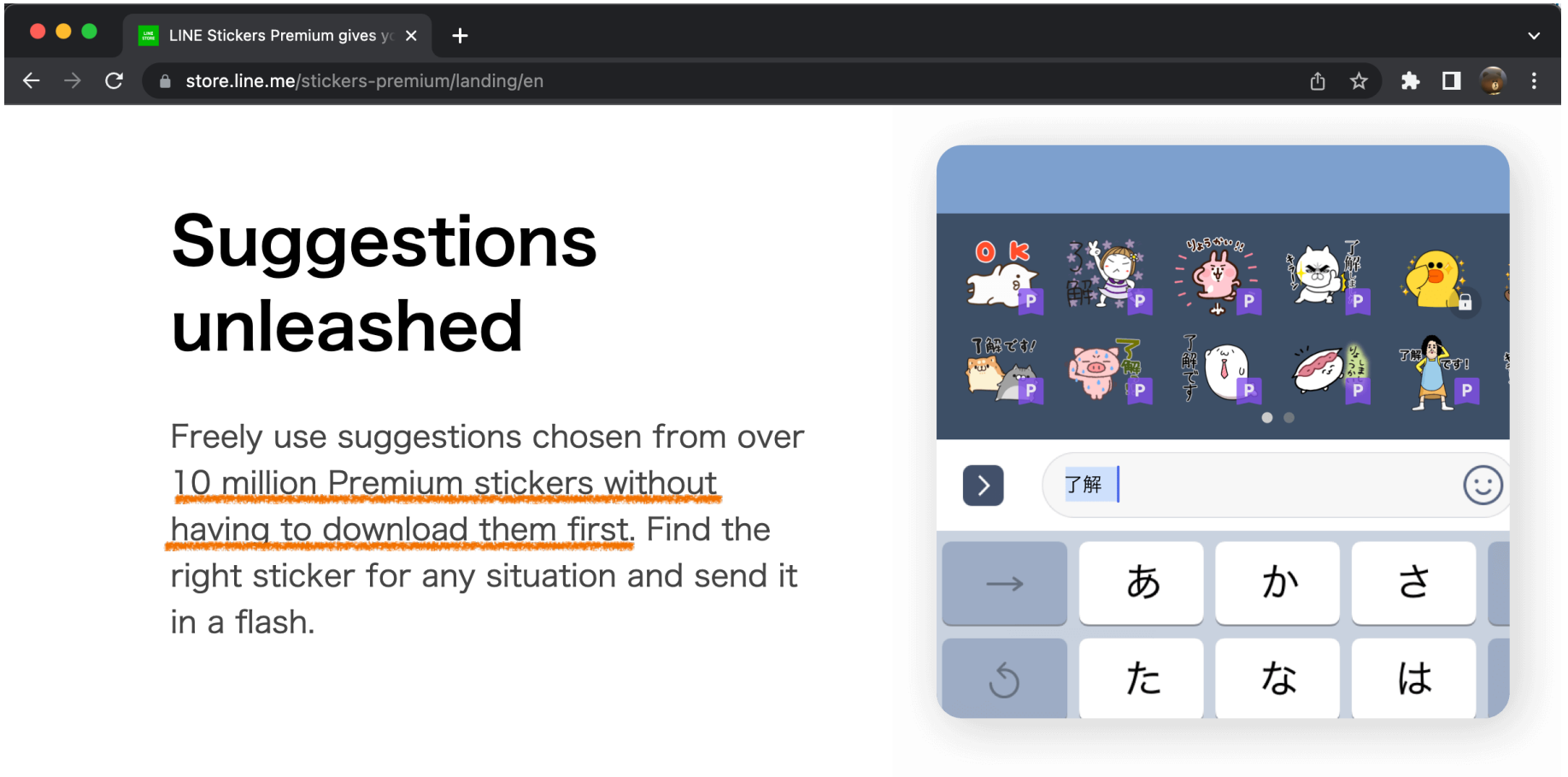
スタンプの意味的なタグ（キーワード）



- 各スタンプに意味的なタグが付与
- 日本語の場合は500を超えるタグが存在

https://creator.line.me

「LINEスタンプ プレミアム」サービスとは



- 1000万以上の
スタンプが使い放題

⇨ 推薦が重要

- 事前ダウンロードが
不要な機能に
FL+DPを適用

https://store.line.me/stickers-premium/landing/en

Federated Learning(FL)とDifferential Privacy(DP)の適用

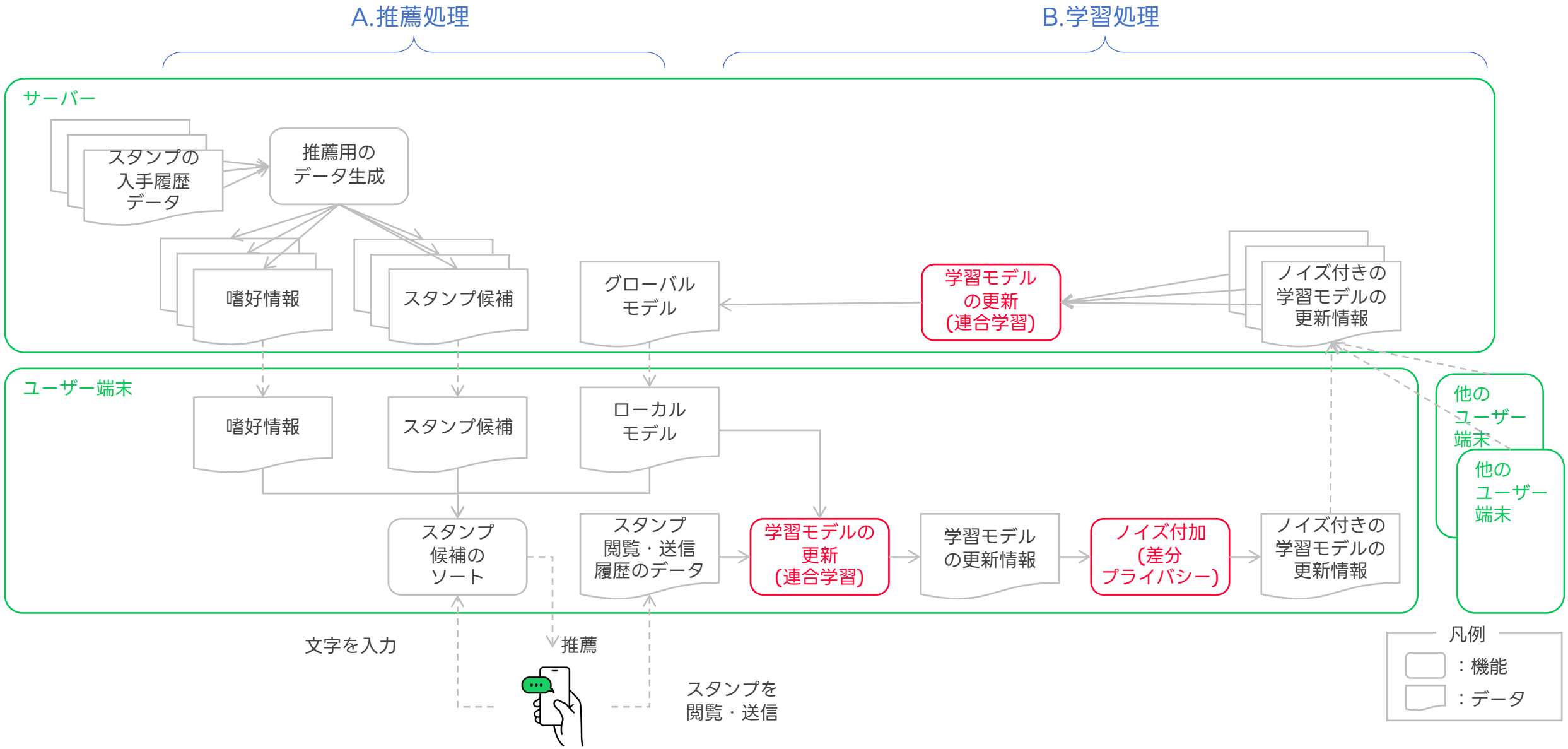
- スタンプの推薦処理を2段階に分け、特にプライバシーの観点で取り扱いに注意が必要なデータを用いる処理に、FL+DPを適用

	(第１段階) 推薦候補の生成	(第2段階) 推薦候補の並べ替え
スタンプの個数	1,000,000 → 100	100
学習データ	スタンプの入手履歴データ （購入や無料ダウンロード等）	トークルーム等での スタンプ閲覧・送信履歴のデータ
推論処理	サーバー	クライアント端末
学習処理	サーバー	主にクライアント端末

FL+DPを適用

※ 図出典：https://linecorp.com/ja/security/article/429

処理の概要



参考：技術詳細を示したホワイトペーパーを公開(2023年9月)

LINE

連合学習向けの差分プライバシー技術に関するホワイトペーパーの公表

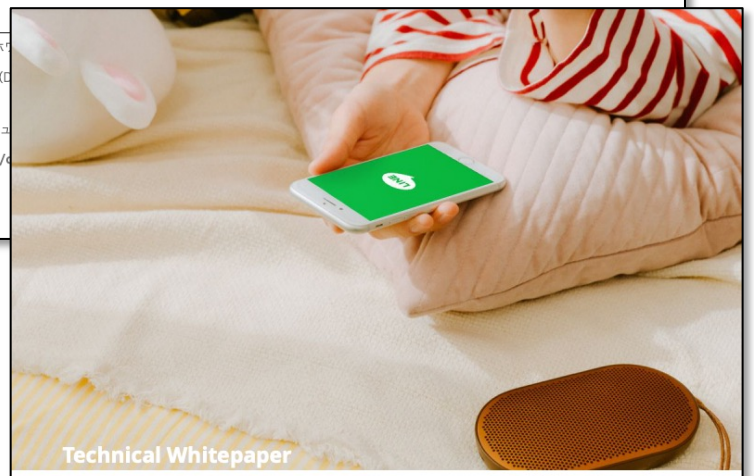
2023.09.29

LINE株式会社は本日、技術ホ

LINEは、差分プライバシー（D
に導入しています。

本ホワイトペーパーは、セキュ

ホワイトペーパー： https://



Technical Whitepaper

Differential Privacy in LINE Federated Learning

v1.0, September 2023

Contents

Further Extension

Note the description here is not included in the first release of LINE's federated learning platform. We are now studying the feasibility of the following extensions to reinforce the usability of federated learning under rigorous privacy guarantees.

One of the most important extensions will be introducing a trusted shuffler[4,6,12,14,21] to amplify local privacy through anonymizing the identity of clients. To introduce the shuffle model, we also need to consider how to securely implement the shuffler. Our current idea is to use TEEs (Trusted Execution Environments). This will improve the efficacy of federated learning while securely preserving the privacy of users under differential privacy with such trusted entities.

Secure aggregation[5,17] with secure computation is also another option for us. The secure aggregation will increase the efficacy under the securely implemented private federated learning with a combination of secure multi-party computation, homomorphic enryption, TEE, and differential privacy.

Conclusion

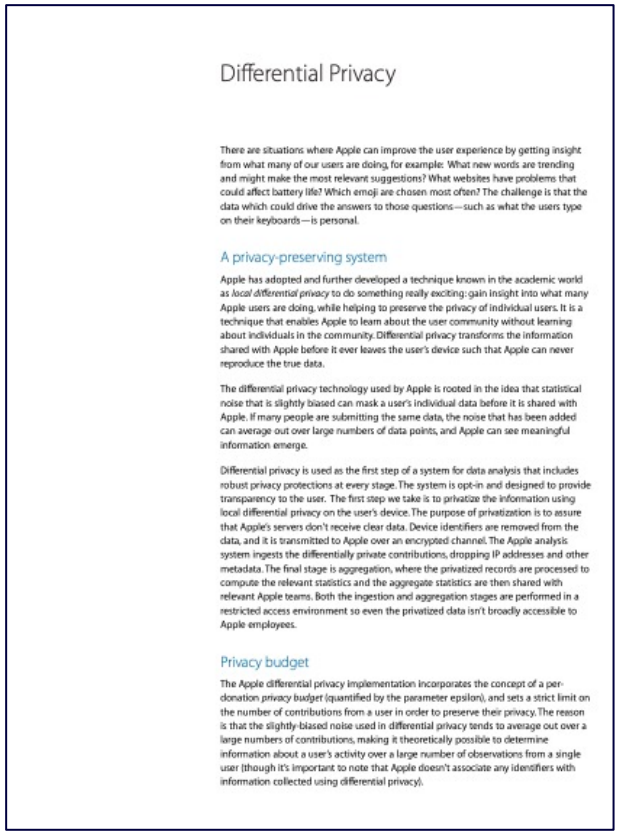
This white paper describes how federated learning on LINE Apps preserves the privacy of users with the differential privacy mechanisms. We always strive to seek better implementation and hyper-parameters that achieve higher efficacy as well as preserving sufficient privacy.

出典：https://linecorp.com/ja/security/article/460

参考：他社の情報公開の例

- いくつかの企業は技術詳細を公開（技術の透明性） → 信頼感・ブランド

Appleの例：ホワイトペーパーを開示



Differential Privacy

There are situations where Apple can improve the user experience by getting insight from what many of our users are doing, for example: What new words are trending and might make the most relevant suggestions? What websites have problems that could affect battery life? Which emoji are chosen most often? The challenge is that the data which could drive the answers to those questions—such as what the users type on their keyboards—is personal.

A privacy-preserving system

Apple has adopted and further developed a technique known in the academic world as *local differential privacy* to do something really exciting: gain insight into what many Apple users are doing, while helping to preserve the privacy of individual users. It is a technique that enables Apple to learn about the user community without learning about individuals in the community. Differential privacy transforms the information shared with Apple before it ever leaves the user's device such that Apple can never reproduce the true data.

The differential privacy technology used by Apple is rooted in the idea that statistical noise that is slightly biased can mask a user's individual data before it is shared with Apple. If many people are submitting the same data, the noise that has been added can average out over large numbers of data points, and Apple can see meaningful information emerge.

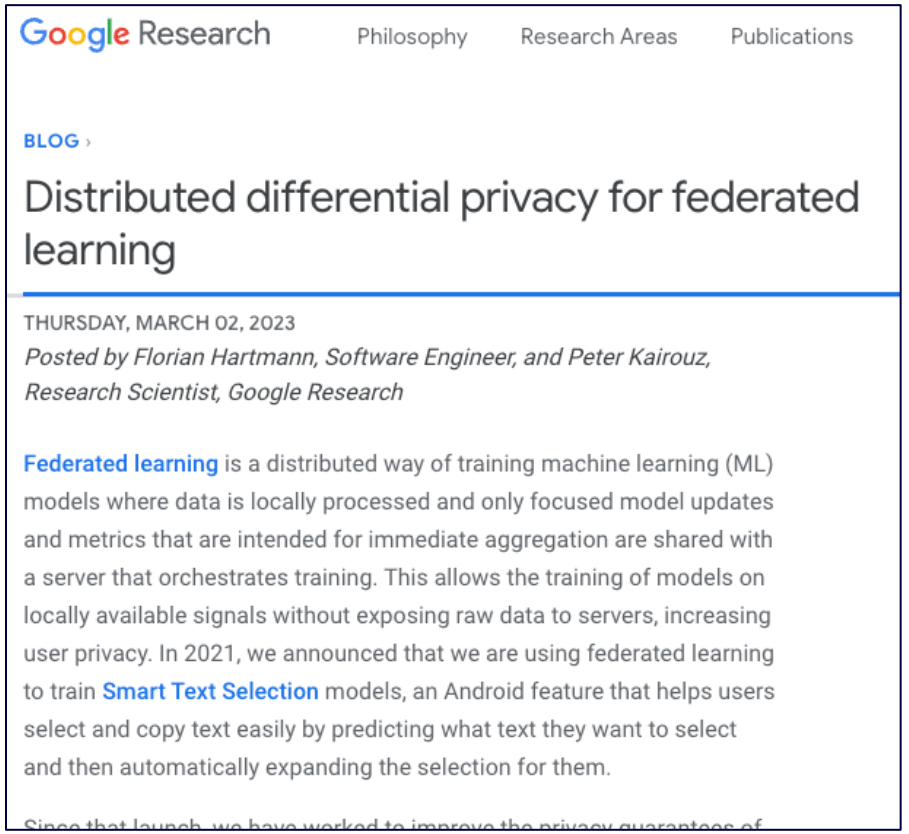
Differential privacy is used as the first step of a system for data analysis that includes robust privacy protections at every stage. The system is opt-in and designed to provide transparency to the user. The first step we take is to privatize the information using local differential privacy on the user's device. The purpose of privatization is to assure that Apple's servers don't receive clear data. Device identifiers are removed from the data, and it is transmitted to Apple over an encrypted channel. The Apple analysis system ingests the differentially private contributions, dropping IP addresses and other metadata. The final stage is aggregation, where the privatized records are processed to compute the relevant statistics and the aggregate statistics are then shared with relevant Apple teams. Both the ingestion and aggregation stages are performed in a restricted access environment so even the privatized data isn't broadly accessible to Apple employees.

Privacy budget

The Apple differential privacy implementation incorporates the concept of a per-donation *privacy budget* (quantified by the parameter epsilon), and sets a strict limit on the number of contributions from a user in order to preserve their privacy. The reason is that the slightly-biased noise used in differential privacy tends to average out over a large numbers of contributions, making it theoretically possible to determine information about a user's activity over a large number of observations from a single user (though it's important to note that Apple doesn't associate any identifiers with information collected using differential privacy).

出典： https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf

Googleの例：技術Blogにて発信



Google Research Philosophy Research Areas Publications

BLOG ›

Distributed differential privacy for federated learning

THURSDAY, MARCH 02, 2023

Posted by Florian Hartmann, Software Engineer, and Peter Kairouz, Research Scientist, Google Research

Federated learning is a distributed way of training machine learning (ML) models where data is locally processed and only focused model updates and metrics that are intended for immediate aggregation are shared with a server that orchestrates training. This allows the training of models on locally available signals without exposing raw data to servers, increasing user privacy. In 2021, we announced that we are using federated learning to train Smart Text Selection models, an Android feature that helps users select and copy text easily by predicting what text they want to select and then automatically expanding the selection for them.

Since that launch, we have worked to improve the privacy guarantees of

出典： https://blog.research.google/2023/03/distributed-differential-privacy-for.html

参考：ユーザへの技術の説明について

ACM CCS 2022参加レポート

Takao Takenouchi 2023-01-11
ML Privacy Team, Senior Privacy Evangelist

(論文3) "Am I Private and If So, how Many?" - Communicating Privacy Guarantees of Differential Privacy with Risk Communication Formats

技術の専門家ではない一般利用者へ、差分プライバシーによるプライバシー保護の効果をどのように説明すると効果的かを調査した論文です。

調査の結果、機能そのものを説明するよりも、個人のデータが推測される確率などを数値で説明する方が良いという結論ですが、特に個人的には以下の点も興味深かったです。

- 数値表現を入れた説明と入れない説明で比較した結果、数値表現を入れると逆に警戒される恐れがある（参考：論文の6.2章「the quantitative notification cause a more cautious reaction」）
- 個人個人に応じて(技術や数値表現への詳しさ等に応じて)、説明を変えることが有効な可能性がある（参考：論文の6.3章）

技術を実用化する際には、どのように技術を説明するかも重要であり、その観点で大変参考になる内容でした。

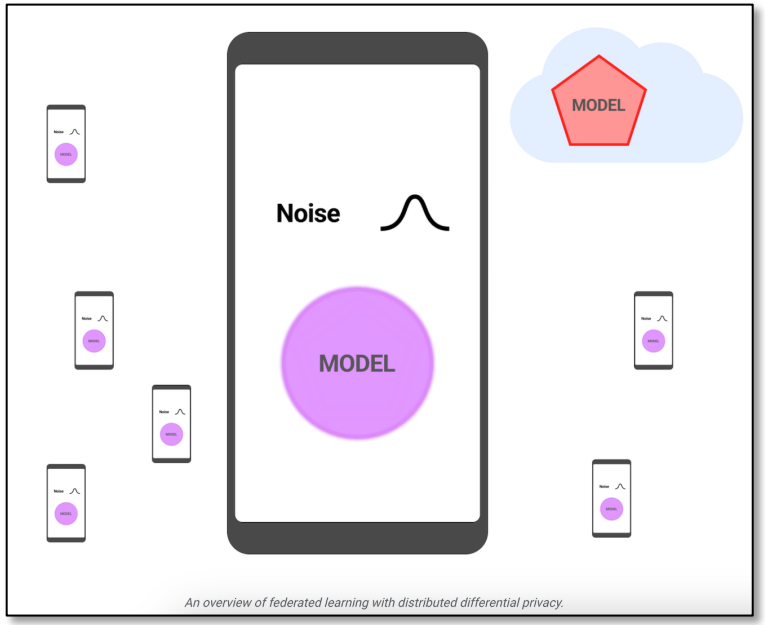
出典：https://engineering.linecorp.com/ja/blog/ACM-CCS-2022-report

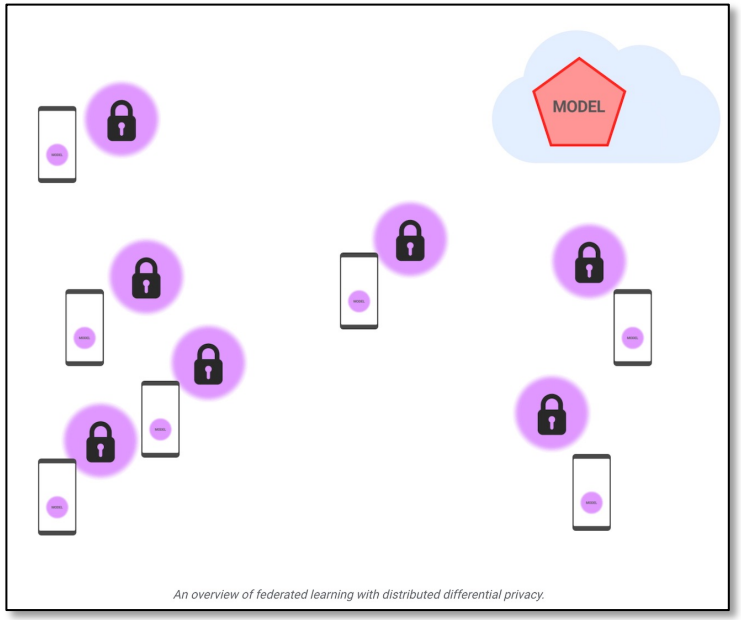
Contents

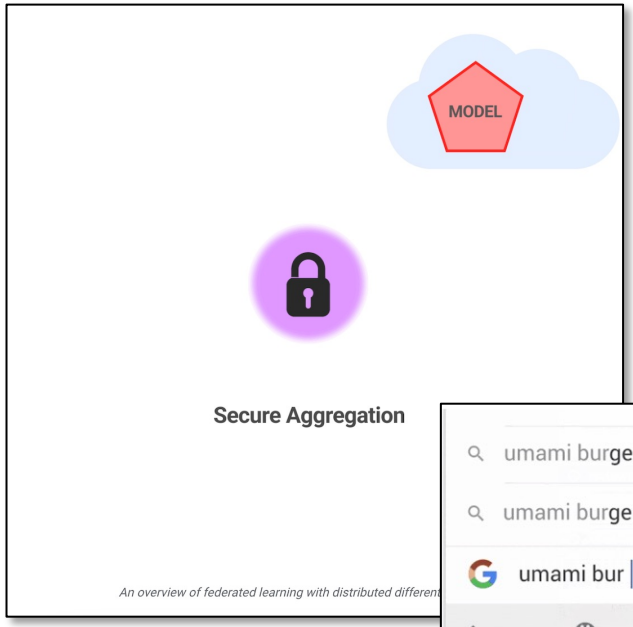
事例①：DP+FL+MPCの事例(Google)

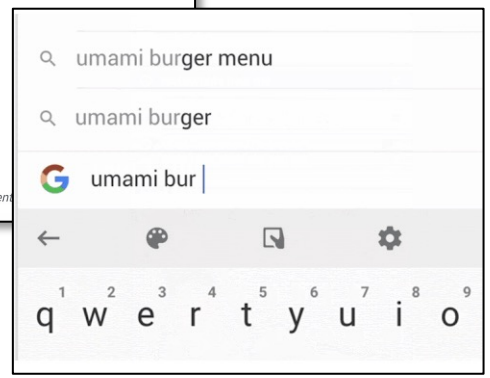
GoogleはDP+FL+MPC(Secure Aggregation)の技術をGboardに適用

クライアント端末で差分プライバシーのノイズを付与後に、MPC（Secure Aggregation）で集約後にサーバへ









Source：https://blog.research.google/2023/03/distributed-differential-privacy-for.html

Source: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

事例②：連合学習+秘密計算の金融不正検知(NICT)

- 複数の金融機関（千葉銀行、三菱UFJ銀行、中国銀行、三井住友信託銀行、伊予銀行）と連携して不正送金等を自動検知
- 連合学習と秘密計算（準同型暗号）を組み合わせた技術

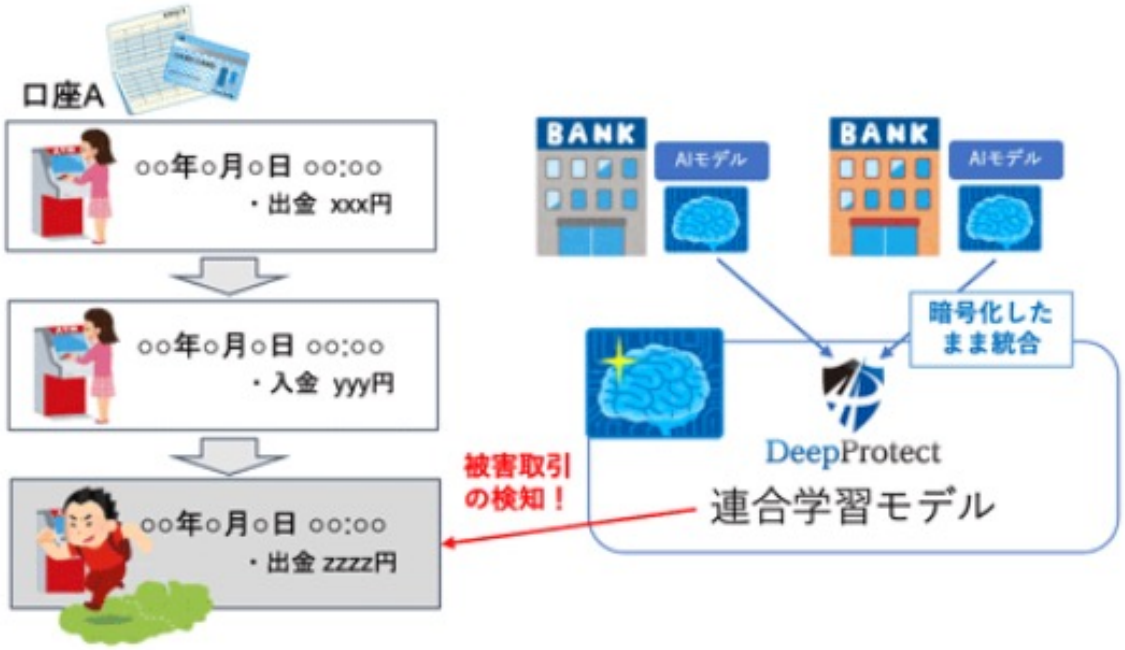


図1 被害取引の検知

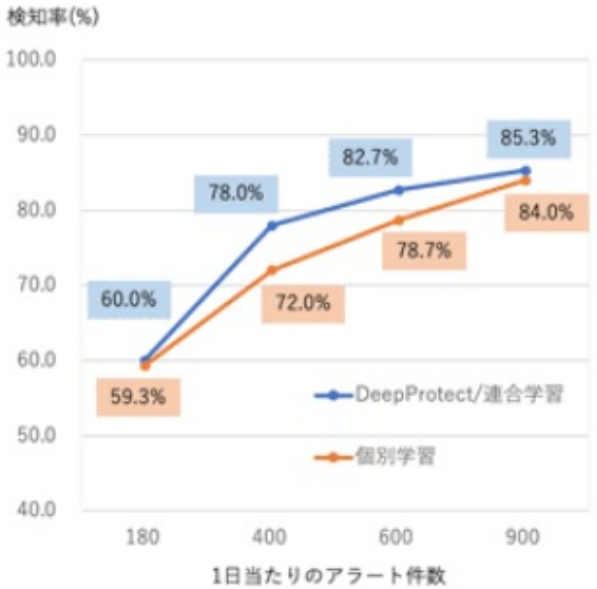
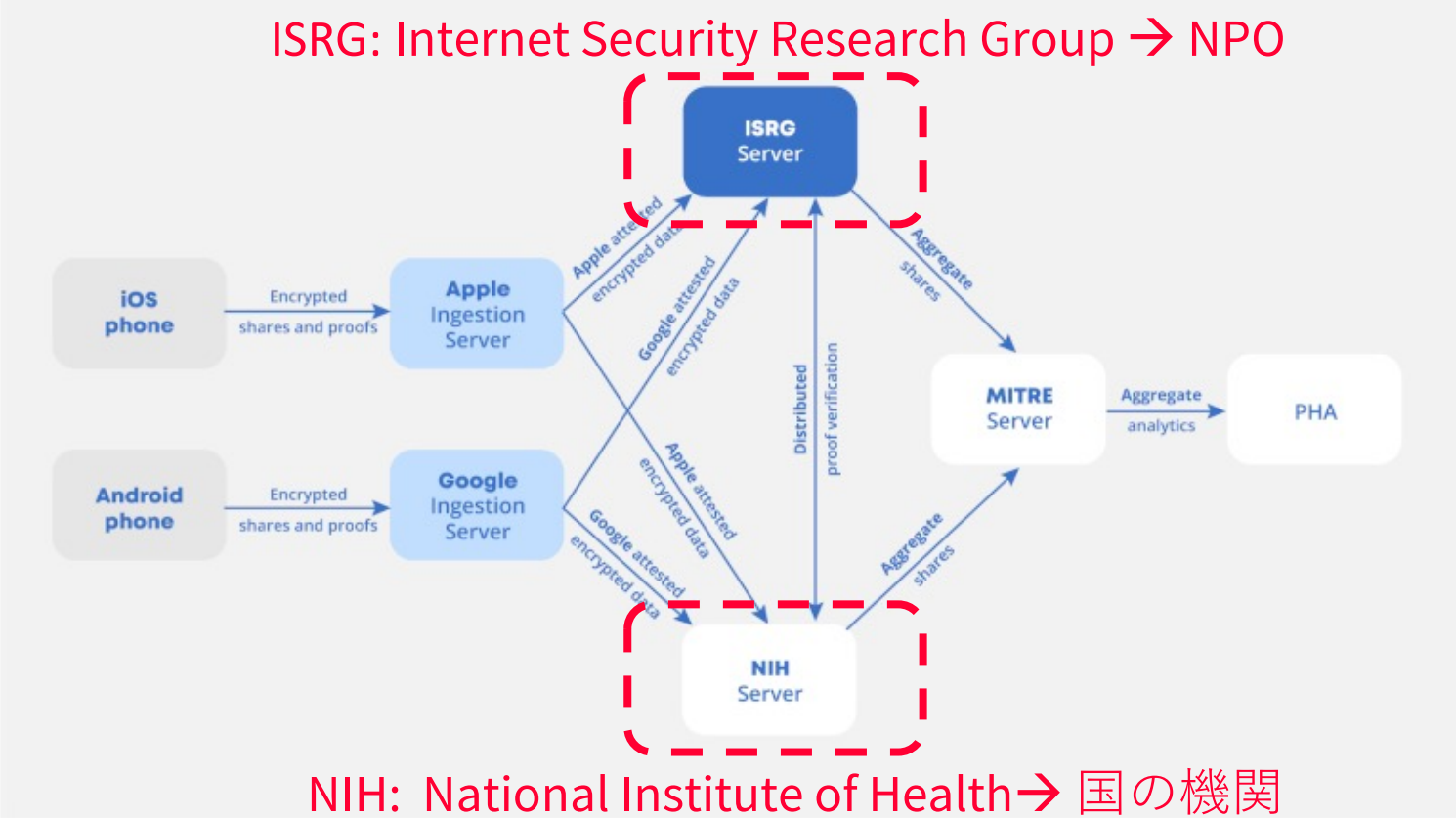


図2 被害検知グループの検知率（一例）
（実データを基にした検知結果を銀行の許可を得て表示）

図出典：”プライバシー保護連合学習技術を活用した不正送金検知の実証実験を実施～被害取引の検知精度向上や不正口座の早期検知を確認～”, 国立研究開発法人情報通信研究機構, 国立大学法人神戸大学, 株式会社エルテス, https://www2.nict.go.jp/oihq/seeds/detail/0024.html

事例②：MPC+DPのCOVID-19関係の集計(GoogleとApple)

- AppleとGoogleが連携し、MPCとDPを組み合わせCOVID-19関係の集計を実施
- MPCのサーバは独立管理が望ましいため、NPOや国の機関が管理



MPC: Multi-party computation
DP: Differential Privacy

図出典：" Exposure Notifications Private Analytics: Lessons Learned From Running Secure MPC at Scale", https://divviup.org/blog/lessons-from-running-mpc-at-scale/

参考：Mozilla(Firefox)は同様な技術を用いてブラウザの挙動の安全に集計

- Mozilla(Firefox)は、同様な技術を用いてブラウザの挙動の安全な集計を推進

FIREFOX

Built for privacy: Partnering to deploy Oblivious HTTP and Prio in Firefox

OCTOBER 12, 2023 BOBBY HOLLEY

Protecting user privacy is a core element of Mozilla's vision for the web and the internet at large. In pursuit of this vision, we're pleased to announce new partnerships with Fastly and Divvi Up to deploy privacy-preserving technology in Firefox.

Mozilla builds a number of tools that help people defend their privacy online, but the need for these tools reflects a world where companies view invasive data collection as necessary for building good products and making money. A zero-sum game between privacy and business interests is not a healthy state of affairs. Therefore, we dedicate considerable effort to developing and advancing new technologies that enable businesses to achieve their goals without compromising peoples' privacy. This is a focus of our work on web standards, as well as in how we build Firefox itself.

出典：https://blog.mozilla.org/en/products/firefox/partnership-ohttp-prio/

- 関連技術は、W3CやIETFにて標準化検討されつつある
- 広告効果の分析にもユースケースの一つである

1. *Health statistics*: The COVID-19 Exposure Notification system developed jointly by Apple and Google includes a Private Analytics system that informs health authorities about how effectively the system is being used [11].
2. *Identifying malicious origins*: Mozilla's Origin Telemetry project helps browser vendors to identify malicious web pages through aggregate measurements without exposing users' browsing history [25].
3. *Advertising measurement*: Meta's private ads measurement products allow a Publisher and Advertiser to privately compute statistics about the effectiveness of advertising campaigns [5, 7]. The system has been used in production for queries with up to 1 billion records.

出典：
Benjamin Case, et al., "Interoperable Private Attribution: A Distributed Attribution and Aggregation Protocol", https://eprint.iacr.org/2023/437

まとめ（再掲）

- プライバシー保護技術の重要性
  - データ活用にプライバシー保護が必要
  - 継続的な新技術の適用が必要
- トレンドなプライバシー保護技術
  - 差分プライバシー、連合学習、MPC/TEE(”秘密計算”)
  - 組み合わせた事例が増加

- 今後の競争軸：「連携」
  - 組み合わせが重要 → 異なる技術分野の連携
  - 技術だけでない　→ 技術者と非技術者の連携