データ流通の将来像とそのセキュリティ技術

デジタルトランスフォーメーション（DX）の加速によりデータの価値は以前にも増して高まるとともに，セキュリティのリスクやプライバシに関する懸念が高まりつつあります．NTTセキュアプラットフォーム研究所では，この課題を暗号技術で解決し，個人・組織が所有するデータを目的に合わせて必要最小限で提供し，課題解決につなげる世界をつくっていきたいと考えています．本稿では，このような世界における安全なデータ流通を支えるNTT研究所の具体的な取り組みを紹介します．

みやざわ としゆき ふくなが としのり
宮澤 俊之 /福永 利徳
たかはし げん きくち りょう
高橋 元 /菊池 亮
たかはし せいじ はせがわ さとし
高橋 誠治 /長谷川 聡
NTTセキュアプラットフォーム研究所

データ流通の将来像

近年，さまざまな分野でデジタルトランスフォーメーション（DX）の加速により，企業や組織のヒト・モノ・プロセスなどのデータ化が進み，高度な分析処理により価値創造や業務効率化などの課題解決につながることが期待されています．これに伴いデータの価値が高まる一方で，セキュリティのリスクやプライバシの懸念も増しています．

企業活動のグローバル化やクラウド・IoT（Internet of Things）の普及によって，多種・多様なエンティティ（人・端末・組織など）が相互接続し，さまざまなデータの授受や共有が進むにつれ，企業や個人にかかわるデータの窃取・漏洩のリスクは高まっています．また，DXにおいて重要な役割を果たすAI（人工知能）や機械学習において個人情報や企業情報の利活用するためには法的な制約やプライバシ等に関する懸念があります．

このようなリスクや懸念を払拭するために，NTTセキュアプラットフォーム研究所では暗号技術を使って，個人・組織のモノ・ヒトに関するデータが目的に合わせて安全に相互流通し，課題解決につなげる世界を実現していきたいと考えています（図1）．本稿では，このような安全なデータ流通を支える技術として，あらゆる通信をエ

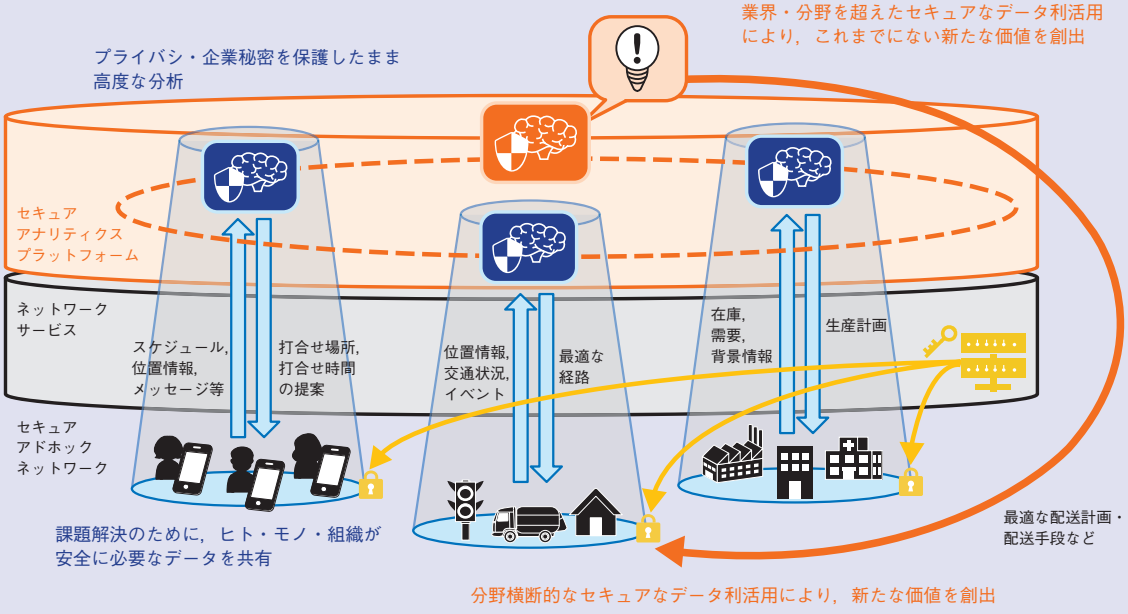


図1　セキュアなデータ流通・利活用のめざす世界

ンド・ツー・エンドで保護する「データの暗号化とその関連技術」，企業秘密やプライバシを保護しながら高度な統合分析を可能とする「秘密計算AI」，個人を特定できないようにパーソナルデータを加工し，これらの情報の利活用を促す「匿名加工技術」について説明します．

データの暗号化とその関連技術

データ流通で行われる重要なデータの授受や共有は，それを行う複数のエンティティの間でのみでき，それ以外には一切情報が漏れないことが重要です．特に最近では，サービス提供者からの情報漏洩リスクも考慮し，データ流通サービスを提供する企業やそのシステム管理者に対してさえもデータを秘匿したいというニーズが高まっています．これを実現するために，信頼するエンティティ間で暗号の秘密鍵を共有し，授受や共有するデータをその秘密鍵で暗号化すること，またそのうえで共有データを検索できることが好ましいですが，大きな技術課題が2つあります．

1番目の課題は，複数エンティティでの効率的な鍵の共有です．データ流通サービスでは，多くエンティティでの鍵の共有が必要ですが，多くのシステムで使われている2者間で鍵を共有するプロトコルを繰り返し実行することは大変非効率で非現実的です．

これに対しNTTでは，サービス提供者が設置する鍵仲介サーバを介して複数のエンティティ間で効率的に鍵の共有ができる技術の研究に取り組んでいます．このとき，鍵仲介サーバでは共有される鍵が復元できないことが原理的に保証されています．現在では，エンティティの数によらずに一定の時間で効率的に鍵の共有ができる方式を発明しています．これにより「その時点で通信にかかわる任意の数のエンティティのみによる，情報流通サービス提供者からも秘匿されたデータの授受・共有」が実現されます．

2番目の課題は，暗号化された共有データのデータ流通サービス提供者の計算機資源を用いた検索です．データ流通サービスはクラウド型で提供されるケースが多く想定されるため，共有データの検索をデータ流通サービス提供者の計算機資源を用いて行うことが望ましいですが，検索のために暗号化データを復号してしまうと，サービス提供者へのデータ秘匿ができません．そこでNTTでは，データとは別に検索インデックスを暗号化し，それを暗号化したまま検索することでこれを実現する方法を研究しています．上述したとおり，共有データはエンティティの追加や削除のたびに頻繁に再暗号化されるため，処理が複雑になるのですが，これを効率的に行う方式を考案しています．

上述した2つの技術の関連技術として，エンティティの追加や削除のたびに，共有する鍵の更新や，それに伴い暗号化された共有データを復号せずに新しい鍵で再暗号化を効率的に行う技術も考案しています．これらの技術を組み合わせ，サービス提供者に対してもデータの内容を漏らさない，電話・チャットなどのコミュニケーションサービスをすでに商用化しています[(1)]．

秘密計算AI

通常，データを利活用するためには，通信時や保管時に暗号化していたとしても，処理を行う際には元データに戻して処理する必要があります．このことは，データ所有者からすると情報漏洩のリスクを感じることから，企業秘密や個人のプライバシにかかわるデータの利活用に抵抗感を持つユーザや組織が少なくありません．特に所有者から他者，または同一組織内であっても，データを提供して積極的に利活用したい場合には，このことは大きな障害だと考えられます．

NTTはそのような要因の解消に貢献するため，データを暗号化したまま処理ができる秘密計算技術の研究開発を世界に先駆けて取り組んでいます．NTTが取り組む秘密計算技術は，ISO国際標準である秘密分散技術を利用して暗号化されたデータを，一度も元データに戻さずに分析を行うため，企業の秘密情報や個人のプライバシにかかわる情報などの情報を安全に，安心して提供し利活用できる社会の実現に貢献すると期待されています．これまでに，統計分析を行う秘密計算技術は実用段階に達しています．

現在，NTTではさらに高度な分析ができる秘密計算技術の研究開発を進めており，最近では，AIの中でも活用が進み始めているディープラーニングの標準的なアルゴリズムを，暗号化したまま一度も元データに戻さずに処理できる技術を世界で初めて実現しました[(2)]．これは，ディープラーニングでのデータ活用に必要な①データ提

供，②データの保管，③学習処理，④予測処理，のすべてのステップを暗号化した状態で行うことができることを意味します（**図2**）．この技術によって，AIを用いてデータ活用する際に，データ所有者が安心してデータを提供でき，データの量や種類の増加や，これに伴う精度向上・高度分析の実現につながると考えています．例えば個人の位置情報やスケジュールを暗号化したまま，天気や企業のイベント情報などと併せて学習することで，最適な飲食店の仕入れや人員リソースの配備を予測することや，レントゲン写真，MRI，CTスキャン，顕微鏡写真などの医療データを秘匿しつつ学習し，検査結果に悪性腫瘍があるかなどを高速かつ精度良く判定することが可能になると期待されます．

今後はAIの知見を持つパートナーと連携して実証実験等を行うことで，秘密計算を使ったディープラーニングの効果を実証していきたいと考えています．

匿名加工技術

近年，パーソナルデータの利活用に注目が集まり，市場が本格的に活性化しようとしています．NTTでは，パーソナルデータの安全な利活用を促すデータ加工技術の研究を進めています．

2017年施行の改正個人情報保護法によって，個人情報を特定の個人を識別することができないように加工し，当該個人情報を復元できないようにした「匿名加工情報」は，本人の同意なしでも目的外利用，第三者提供が可能となりました．例えば，小売事業者が持つ購買履歴データを匿名加工情報にすることで，製造事業者が消費者属性と購買傾向に基づいた新製品開発を行うことができます．

匿名加工では，匿名性だけを高めようとすると，元のデータの特徴が大幅に損なわれ有用性の低いデータとなってしまいます．そのため，データ保有者の個人識別のリスクを低減したいというニーズと，データ活用者の元データの特徴が保持されたデータを入手したいというニーズの両方を満たす最適な匿名加工が必要となります．NTTは，匿名性と有用性のバランスの取れた匿名加工情報の作成を支援する「匿名加工情報作成ソフトウェア」を開発しました．本ソフトウェアは2018年よりNTTテクノクロスから提供可能となっており，医療・金融分野を中心に市場展開が進められています．

本ソフトウェアは，個人情報保護委員会が規定した匿名加工情報の加工基準1号～5号に対応するための多彩な匿名化技法，評価技法を備えています．特徴的な技法としては，撹乱的な手法によりデータの粒度を変えず高い有用性を確保するNTT独自の匿名化技法であるPk-匿名化を備えています．従来技術であるk-匿名化は「33歳」を「30代」，「東京都千代田区」を「東京都」にするなど，データを抽象化することによってk-匿名性*を担保する手法ですが，情報損失が課題の1つでした．これに対し，データを撹乱させるPk-匿名化を導入することにより情報損失がなく，より正確で幅広い分析が可能となります（**図3**）．

NTTは，データがより活発に利活用される世界の実現をめざし，個人識別のリスクを低減する研究だけでなく，個人の属性の推測リスクを低減する研究にも取り組んでいます．データ利活用において統計情報のような計算結果を用いる場合，例えば2人分のテストの平均点を用いると，1人分の点数を知る人はもう1人分の点数を推測できてしまうという問題があります．このような計算結果に対するプラ

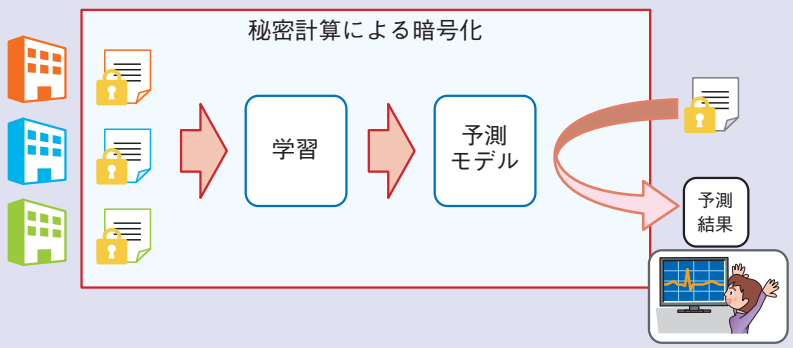


図2　秘密計算AI

＊k-匿名性：加工後のデータから対応する個人を1/k以上の確率で識別できない特性．

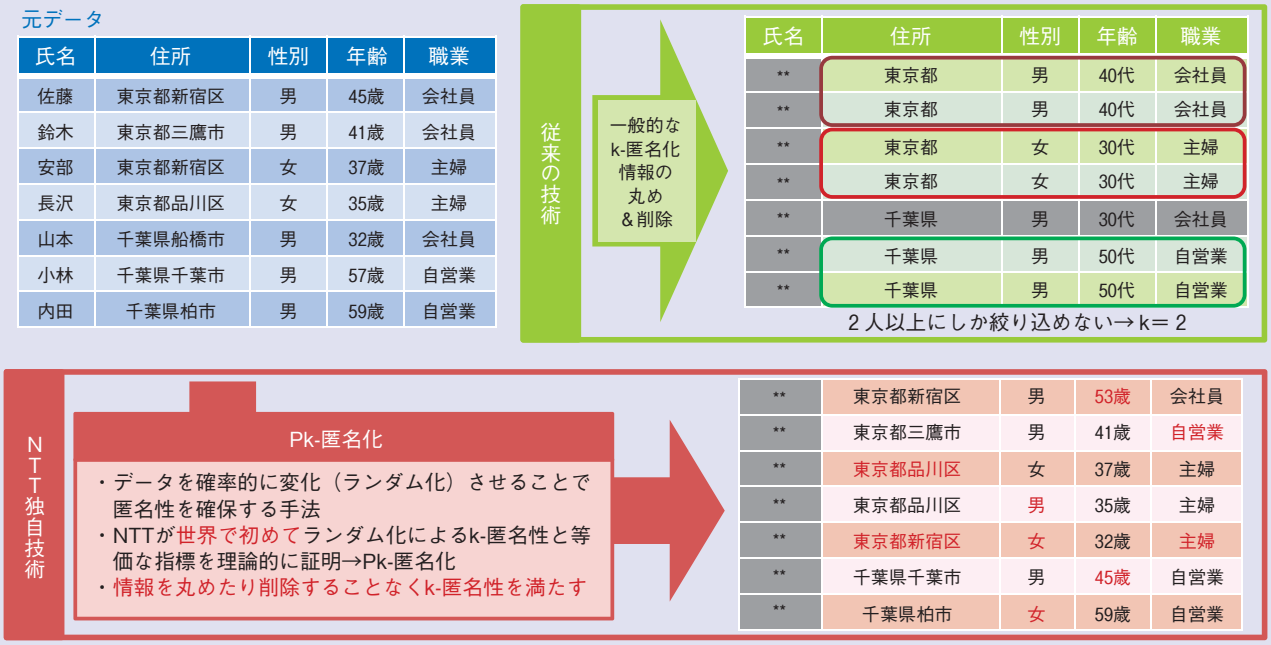


図3　Pk-匿名化の概要

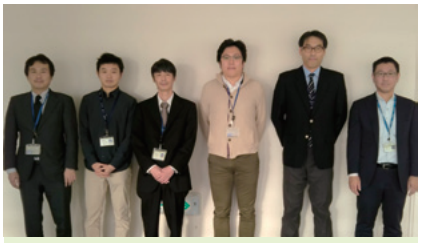
イバシは，出力プライバシと呼ばれ統計分野で広く研究が行われており，NTTでは特に，データ分析技術として有望な機械学習の出力プライバシに注目し，リスク分析技術，保護技術の研究を進めています．

今後に向けて

冒頭で紹介した安全なデータ流通を実現するために，NTTでは本稿で紹介した技術以外にもさまざまな暗号技術の設計・開発に取り組んでいます．例えば，研究開発が加速している量子計算機の実現を見据えた暗号方式の開発やその安全性評価や，プログラムの処理内容を暗号学的に解析不可能にし，安全なプログラムの流通を可能とする暗号学的プログラム難読化などがその一例です[3]．暗号理論・技術の専門性を基に，NTTグループのお客さまや社会的課題の解決に貢献すべく，データセキュリティの研究開発に取り組んでいきます．

■参考文献

⑴ 吉田・岡野・奥山・小林：“サーバからの漏洩・盗聴を防ぐ暗号ビジネスチャット”，NTT技術ジャーナル，Vol.29，No.2，pp.18-22，2017.
⑵ https://www.ntt.co.jp/news2019/1909/190902a.html
⑶ 草川：“耐量子暗号技術の研究動向，”NTT技術ジャーナル，Vol.31，No.2，pp.23-26，2019.



（左から）宮澤　俊之／長谷川　聡／
高橋　誠治／菊池　亮／
福永　利徳／高橋　元

私たちは，最先端の暗号理論研究から次世代の暗号通信・システム方式の研究まで幅広くデータセキュリティの研究に取り組んでいます．増加するデータの安全性を確保しつつ，活用可能とする技術の実現によって新ビジネスの創出と課題解決に貢献していきます．

◆問い合わせ先
NTTセキュアプラットフォーム研究所
企画担当
E-mail　scpflab@hco.ntt.co.jp