2.4.3 データ・コンテンツのセキュリティー

（1）研究開発領域の定義

データを活用していくためには、その収集、流通、管理、解析などの過程においてセキュリティーやプライバシーの保護が必要となる。本研究開発領域では、個人情報や個人にかかわる情報であるパーソナルデータ[1]を利用するにあたり、データ活用と個人情報・プライバシー保護とを両立するための技術と、将来、データの保護に必要となる耐量子計算機暗号技術を扱う。

（2）キーワード

個人情報、パーソナルデータ、プライバシー保護、匿名化、秘匿計算、秘密計算、プライバシー保護データマイニング、差分プライバシー、局所差分プライバシー、準同型暗号、秘密分散、秘匿回路計算、連合学習、耐量子計算機暗号

（3）研究開発領域の概要

［本領域の意義］

「データは、インターネットにおける新しい石油である」と評されるように、データの活用は経済の発展における中心的役割を果たすことが期待されている。その中でも、「パーソナルデータ」は、さまざまな分野で、その活用による経済効果が期待されている。例えば、医療機関が保有する患者の医療情報を活用した創薬・臨床分野の発展や、カーナビなどから収集される走行位置情報を活用したより精緻な渋滞予測などにより、社会や産業の発展、人々の生活の質向上が期待できる。一方で、プライバシーは他人に知られたくない私事でありそれをコントロールする基本的人権であり、個人に関わるパーソナルデータを扱う際には適切なプライバシー保護が求められている。データ活用におけるプライバシー保護への要求は、欧州連合における一般データ保護規則（GDPR：General Data Protection Regulation）の適用開始を契機に高まっている。このような中で、データ活用と個人情報やプライバシー保護を両立するデータのプライバシー保護技術を確立することは、社会や産業の発展、人々の生活の質向上のために必要不可欠である。

暗号は、Webサイトの閲覧やVPN（Virtual Private Network）によるリモートアクセスなどで情報を安全に送受信するために用いられ、マイナンバーカードを使った各種サービスなどではデジタル署名で暗号が利用されており、現在のわれわれの生活の中で必要不可欠なものとなっている。一方で、近年、量子コンピューターの研究開発が活発に進められており、十分な性能を持つ量子コンピューターが実用化されると、現在利用されている暗号が解読されてしまう懸念がある。このため、量子コンピューターでも解読できない耐量子計算機暗号技術の確立が求められており、将来にわたり安全・安心なデジタル社会の維持・発展に必要不可欠である。

［研究開発の動向］

❶ これまでの研究開発の流れとトレンド

多くの企業が顧客の情報や購買履歴などを管理して、ビジネスに活用する動きが加速している。いわゆる

1 個人情報保護法において「個人情報」とは、生存する個人に関する情報で、氏名、生年月日、住所、顔写真などにより特定の個人を識別できる情報を指す（政府広報オンライン「「個人情報保護法」をわかりやすく解説：個人情報の取扱いルールとは？」内閣府大臣官房政府広報室）。
「パーソナルデータ」は、個人情報に加え、個人の属性情報、移動・行動・購買履歴、ウェアラブル機器から収集された情報など個人情報との境界が曖昧なものを含め個人と関係性が見出される広範囲の情報を指す（総務省「令和４年版情報通信白書 情報通信に関する現状報告（本編）」）。

ビッグデータと呼ばれる、大規模で機械的に収集される多量のデータが、あらゆる分野で注目されている。その一方で、データの活用から生じるセキュリティーやプライバシーの課題が浮き上がってきた。例えば、2018年にはソーシャルネットワーキングサービス（SNS）の個人情報が無断で政治広告のために不正利用されたり、2019年には2億6,700万人以上のユーザーID、電話番号、名前がパスワードやその他の認証なしにオンライン上で閲覧可能な状態に置かれていたとの報告があった[1)]。2013年には鉄道会社が利用履歴などのパーソナルデータを活用するためにデータを社外へ提供しようとして問題となった[2)]。一方、パーソナルデータの活用に関しては、個人が自らの意思でパーソナルデータを蓄積・管理するための仕組み（システム）であるパーソナルデータストア（PDS：Personal Data Store）や、個人の指示に基づき提供されたパーソナルデータをPDSなどで管理し、個人から提供されたパーソナルデータを使って事業を行う情報銀行が注目されている。2018年には「個人がパーソナルデータを自分自身のために使い、自分の意思で安全に共有できるようにする」という個人中心のMy Dataの考え方を世界に発信していくことを目指してMyData Globalが設立されている。日本でも、2019年に「公正で持続可能な社会を実現するため、パーソナルデータに関する個人中心のアプローチを推進し、個人をエンパワーする」ことを目的としてMyDataJapanが設立されている[2]。このように、データは価値の源泉であり、データは個人の意志に基づき個人情報やプライバシーを保護した上で活用していくことが大きな流れになっている。

以下では、プライバシー保護対策のための代表的な技術として、①個人を識別不能にする匿名化技術、②プライベートなデータを暗号化したままで任意の計算を実行する秘密計算技術、③プライバシーを保護した上でデータマイニングを実施するプライバシー保護データマイニング技術、④抽出された知識からプライベート情報が漏えいしないように精度を落としたりノイズを加えたりする差分プライバシー技術について紹介する。

・匿名化技術

匿名化技術は、データとデータ主体（あるいは所有者）との間の相関を取りのぞく技術である。パーソナルデータの収集において、姓名などの識別子を削除しただけでは、上記の相関は完全には取りのぞけず、他の属性情報・履歴情報を束ねて見ることで個人が特定され得るリスクがある。このようなリスクを定式化し、低減するための考え方としてk-匿名性[3)]がよく知られている。具体的には、表形式データについて、パーソナルデータの属性値の組み合わせが同じであるデータが、パーソナルデータ集合中にk個以上存在している状態が、k-匿名性が成立した状態である。データの正確性は犠牲になるが、パーソナルデータを改変することで、k-匿名性を成立させ、個人特定を困難にする。その後、k-匿名性を基礎概念として、匿名化対象を表形式データからグラフや時系列データに拡張する研究や、k-匿名性モデルにおいて十分にプライバシーを保護できない状況下におけるより強力な匿名性定義の研究などが進められてきた（l-多様性、t-近似性など）。個人情報保護法による匿名加工情報の実装において実務上重要な技術である。

・秘密計算技術

秘密計算（マルチパーティー・コンピュテーション、MPC：Multi-Party Computation）技術は、複数のコンピューターが、開示できないデータを暗号化したまま処理して、計算結果以外の情報を一切開示することなく計算可能にする技術である。安全な秘密計算のためのプロトコルは、1980年代から研究が開始された。近年では、理論的には成熟しつつあり、実用的な時間で動作する秘密計算を実行するための汎用コンパイラーが開発され、専門家でなくても秘密計算を利用したシステム開発を行うことが可能になりつつある[4)]。秘密計算の研究は安全性などの理論も継続して研究されているが、ここ数年、実用のためのアル

2 MyDataJapanは、MyData Globalの日本拠点としての活動も行っている。

ゴリズム研究と応用研究が大きく増えている。秘密計算を実現する暗号要素技術にデータをシェアと呼ばれる断片に分割し機密性を守る秘密分散がある。秘密分散ベースの秘密計算の研究では、わが国が実用的なスループットや大規模データ処理に必須な機能であるセキュアソートで最高性能を打ち立てており、その汎用性や実用性が確認されている[5), 6)]。また、いわゆるデータ処理以外でも応用されており、セキュリティー用途、すなわちデータ暗号化時の秘密鍵の管理を秘密計算で行うことなどの成果がある。なお、秘密計算には、上述の秘密分散を用いたものの他に、準同型暗号（プライバシー保護データマイニングに記載）を用いたものなどがある。

・プライバシー保護データマイニング技術

プライバシー保護データマイニング（PPDM：Privacy Preserving Data Mining）技術は、利用者のプライバシーを保護しながらビッグデータの活用を実現する技術である。PPDM研究の原点は、2000年に発表された二つの同名の論文「Privacy Preserving Data Mining」である。一つは、公開鍵と秘密鍵のペアからなる公開鍵暗号方式を用いる暗号学的アプローチによるもの[7)]で、もう一方はランダムなデータを入力に加えてマイニング処理を行うランダム化アプローチによるもの[8)]であった。両論文のアプローチは異なるが、対象は両者ともプライバシー保護を考慮した決定木学習（与えられたデータから決定木と呼ばれる木構造のグラフを生成する手法）を実行するものであったことは、興味深い事実として知られる。この二つの論文を出発点として、PPDMに関して盛んに研究が行われるようになった。

PPDMの主な要素技術としては、暗号化したまま加算や乗算の演算が可能な準同型暗号があり、加算が可能なPaillier暗号[9)]や乗算が可能なRSA暗号[10)]が知られる。これらの要素技術の研究開発や安全性評価は2000年代にはほぼ完成していて実現可能性は確認されているが、暗号化にかかる計算コストが大きく、広い実用化のレベルには至っていない。この技術的な困難さを改良するために、加法、乗法の両方の演算が可能な完全準同型暗号などの暗号要素技術の改良が重ねられている。

・差分プライバシー技術

差分プライバシー技術は、データ収集者が信頼できる場合に、データ収集者が公開した統計情報から個人に関する情報が推測されることを防ぐ技術である[11)]。一方で、あらゆるデータ収集者が完全に信頼できるとは限らないため、個人がデータを提供する際にプライバシー保護処理を行い、その個人に関する情報が推測されることを防ぐことを保証する、局所差分プライバシー（LDP：Local Differential Privacy）が提案されるようになった[12)]。

差分プライバシーでは、個人が保持しているデータを個人から収集者へ提供する。収集者は収集した個人データに対して統計処理を行い、それを解析者へ公開する。この流れの中で、差分プライバシーにおいては、収集者が統計処理した結果を解析者に公開するときに個人データが漏えいしないようにプライバシー保護処理を行う。ただ、収集者は生の個人データを閲覧することができるため、個人が収集者を信頼できる必要がある。一方、局所差分プライバシーでは、個人がデータを提供する際にプライバシー保護処理を行い、個人のデータが漏えいしないようにする。従って、信頼できない収集者に対する個人データの漏えいも防ぐことができる。局所差分プライバシーでは、個人から収集者への提供データにノイズを加えて、元のデータが推測できないようにするとともに、収集者はノイズが入った提供データを用いて所望の統計処理を行う。従って、データセット全体で見たときには、差分プライバシーと比べて多くのノイズが加えられるため、実用性が低下しやすく、適用範囲が広いとはいえない。しかし、仕組みの単純さとプライバシー保証の強力さのために、多くのユーザーから情報を収集するGAFA（Google、Amazon、Facebook、Apple）を含むプラットフォーマーは、局所差分プライバシーを利用したデータ収集を取り入れ始めている。

以下では、暗号耐量子計算機暗号についての研究開発の流れとトレンドを紹介する。

現在、電子署名やWebサイトの閲覧、VPNによるリモートアクセスでは、公開鍵暗号であるRSA暗号

および楕円曲線暗号が広く利用されている。これらの暗号は、量子ゲート型の量子コンピューターとShorのアルゴリズムを利用することで、十分な性能を持つ量子コンピューターが登場した場合に、暗号が解読され安全性が危殆化することが懸念されている[13), 14)]。このため、十分な性能を持つ量子コンピューターが実用化されても安全性を保つことが可能な耐量子計算機暗号の研究開発と標準化が活発に進められている。

・耐量子計算機暗号技術

耐量子計算機暗号は、2022年時点では、公開鍵暗号を意味することが多い。その理由は、公開鍵暗号を除く、共通鍵暗号などの他の主要な暗号技術に対して量子コンピューターを用いた効率の良い解読方法がまだ提案されていないためである[15), 16)]。耐量子計算機暗号の代表的な候補として、格子に基づく暗号技術、符号に基づく暗号技術、多変数多項式に基づく暗号技術、同種写像に基づく暗号技術、ハッシュ関数に基づく署名技術が挙げられる。これらの公開鍵暗号は、整数の素因数分解や楕円曲線上の離散対数問題（ECDLP）とは異なる数学的な計算問題の計算困難性をその安全性の根拠としている[17)]。量子コンピューター実機を用いて素因数分解する計算や離散対数問題を解く計算の研究が進められており、その成果によって、RSA暗号や楕円曲線暗号に対する量子コンピューターの脅威を把握し、耐量子計算機暗号が必要となる時期を見積もることが期待されている。この研究に関する調査がCRYPTREC（Cryptography Research and Evaluation Committees）によって2020年度に実施され、2020年においてはRSA暗号および楕円曲線暗号に対する量子コンピューターの脅威は生じていないことが報告されている[18)]。しかし、量子コンピューターの開発が進み、RSA暗号などを解読するのに十分な性能を持つ量子コンピューターが登場したときに生じるリスクは極めて大きく、使用されている暗号を更新するためには長い年月が準備期間として必要となるため、早めに耐量子計算機暗号を使用できるように準備することが望ましい。そのため、CRYPTREC や米国・国立標準技術研究所（NIST：National Institute of Standards and Technology）などの世界各国の組織が耐量子計算機暗号の使用に向けた活動を実施している。

上記で紹介したプライバシー保護や耐量子計算機暗号に関する研究開発は、ACM Conference on Computer and Communications Security（ACM CCS）やIEEE Symposium on Security and Privacy（IEEE S&P）、USENIX Security Symposiumなどの国際学会や、国内の情報処理学会　コンピュータセキュリティシンポジウム（CSS）や電子情報通信学会　暗号と情報セキュリティシンポジウム（SCIS）で活発に議論されている。

❷ 海外・国内政策動向

個人データの取り扱いに関する研究は、欧州においては2018年から施行されたGDPRに大きく影響されたといえる。 GDPRの大きな特徴の一つは、IPアドレスやCookieなどのインターネットで利用される識別子を含む情報も、個人情報として取り扱うこととなったことにある。このことは、Web経由で個人のデータを暗黙的に収集してきた事業者に多くの影響を与えた。またGDPRは、個人情報を取り扱うサービスやシステムについて、設計段階でデータ保護が組み込まれ、利用者が明示的に設定しなくても、十分なプライバシー保護が初期状態で設定されていることを要求する（設計段階、および初期状態におけるプライバシー）。この設計思想は、プライバシー・バイ・デザインの影響を受けたものである。さらにGDPRは、プロファイリングを含む個人に対しての自動化された意思決定について、分析する側に透明性の確保（プロファイリングしている事実を知らせること、およびプロファイリングの方法やその影響について説明すること）などを求めるとともに、利用者はこのような自動化された意思決定を受けない権利を有するものとした。「プロファイリング」とは、「個人の特定の側面を評価するために個人データを自動的に処理すること」であり、特に個人の業務実績、経済状況、健康、個人的嗜好、興味、信頼、行動、所在、または移動など、個人について重要な判断を伴う分析・予測やそれを提供するシステムとそのロジックについて、透明性の確保と

説明責任を求めるとともに、そのような決定を受け入れない権利があることを定めている。

わが国においても、個人情報保護法が改正され、プライバシー保護の関連では、個人データを規則にのっとって加工することで利用目的の変更が緩和される仮名加工情報、利用目的に加え第三者提供が緩和される匿名加工情報の制度がそれぞれ2022年、2017年に導入された。また2018年、医療分野の研究開発に資するための匿名加工医療情報に関する法律「次世代医療基盤法」の施行や、同年に総務省および経済産業省がとりまとめを行った「情報信託機能の認定に係る指針」によって、認定された事業者によるデータ収集や利活用ができるようになってきた。さらに2022年には電気通信事業法が改正され、いわゆるサードパーティーCookieと呼ばれる仕組みによる利用者情報の外部送信に規制が入るようになった。今後の安心・安全なデータ利活用のため、データ・コンテンツに関するセキュリティー、およびプライバシー保護技術がますます重要になってきている[19)]。

［論文や特許の動向］

論文数、特許数ともに年々増加傾向にあり、成長率が高い研究領域である。論文動向に関しては、論文数では2018年に中国が1位となり、米国、欧州が続いている。 Top1%論文数では2021年に中国が米国を抜いて1位となっている。中国と欧州がTop1%論文数を増やしているのに対して、米国は近年減少している。日本の論文数は堅調に増加しているが論文シェアは低い状態が継続している。特許動向に関しては、特許ファミリー件数シェアでは2018年に中国が米国を抜き、以降、1位を維持している。Patent Asset Indexシェアでは米国が1位を維持しているが、中国が増加している。

（4）注目動向

［新展開・技術トピックス］

❶ 差別への配慮

データ解析に関わる個人情報の問題は、これまでは取得・収集データ（入力データ）の扱いにフォーカスされてきたが、AI技術の発展により、取得・収集された個人データを用いて学習したAIの出力データの扱いにも、配慮が必要となりつつある。例えば、AIシステムによる人種、性別、健康、宗教などによる差別の問題が挙げられる。 AIの入力データにこれらの情報が含まれる場合には、プライバシー・個人情報保護の問題となるが、AIによる出力や決定がこれらの情報と相関する場合には、差別の問題となる。このため、人工知能分野では公平性に配慮したAI（公平配慮型AI）の研究がホットトピックとなっている。

❷ 連合学習（Federated Learning）

従来、AIの学習は、拠点に分散している学習データを一カ所に集約していっていた。一方で、学習データを一カ所に集約するためには大量の学習データの送信や保存が必要となるとともに、パーソナルデータを扱う場合にはプライバシー保護が必要となる。連合学習では、学習データが拠点に分散したままでAIを学習させる学習方法であり、学習データそのものではなく学習の手がかりになる情報（学習モデルの勾配）だけを各拠点から収集してAIを学習させる。連合学習は、AIを使ったパーソナルデータの活用とプライバシー保護を両立する点でも注目されている。

［注目すべき国内外のプロジェクト］

❶ 戦略的創造研究推進事業におけるプロジェクト（JST）

・CREST「イノベーション創発に資する人工知能基盤技術の創出と統合化」

JSTの戦略的創造研究推進事業CREST「イノベーション創発に資する人工知能基盤技術の創出と統合化」研究領域においては、研究課題「プライバシー保護データ解析技術の社会実装」が実施されている。個人情報や企業の機密情報などのあらゆる機微情報を、安全性を保ったまま任意のデータ処理に適用可能とす

るプライバシー保護データ解析技術を創出することを目的としている。2016年度にスモールフェーズの研究を開始し、2019年度からは加速フェーズへと移行し社会実装に向けた研究が進められている。

・CREST「基礎理論とシステム基盤技術の融合によるSociety 5.0のための基盤ソフトウェアの創出」

2021年度、文部科学省において「Society 5.0時代の安心・安全・信頼を支える基盤ソフトウェア技術」という戦略目標が決定され、セキュリティーやプライバシーが組み込まれた基盤を開発するCREST「基礎理論とシステム基盤技術の融合によるSociety 5.0のための基盤ソフトウェアの創出」研究領域が設立されている。

❷ CRYPTREC（Cryptography Research and Evaluation Committees）（デジタル庁、総務省、経済産業省、NICT、IPA）

CRYPTRECは、デジタル庁、総務省、経済産業省、国立研究開発法人情報通信研究機構（NICT：National Institute of Information and Communications Technology）、独立行政法人情報処理推進機構（IPA：Information-technology Promotion Agency）が共同で運営する電子政府推奨暗号の安全性の評価監視などを実施するプロジェクトである。これまでに耐量子計算機暗号に関わる以下の調査や評価が実施され、報告書として公開されている。

・「格子問題等の困難性に関する調査」[20]（2015年3月）：
2013年度から2014年度に実施された耐量子計算機暗号の候補である格子に基づく暗号技術などの安全性に関する調査。

・「耐量子計算機暗号の研究動向調査報告書」[17]（2019年3月）：
2017年度から2018年度に実施された4種の耐量子計算機暗号の候補（格子に基づく暗号技術、符号に基づく暗号技術、多変数多項式に基づく暗号技術、同種写像に基づく暗号技術）の技術動向調査。

・「量子コンピュータが共通鍵暗号の安全性に及ぼす影響の調査及び評価」[15]（2010年1月）

・「暗号技術評価委員会報告」[16]（2020年3月）：
2019年度時点においては電子政府推奨暗号リストにある共通鍵暗号、暗号利用モード、ハッシュ関数に対する直近で現実的な脅威が生じる可能性は極めて低く、現状では CRYPTREC での具体的な対応は不要であると結論付けられている。

・「Shorのアルゴリズム実装動向調査」[18]（2021年6月）：
2020年度に実施されたShorのアルゴリズムと量子コンピューター実機を用いた、素因数分解およびおよび離散対数問題を解く数値実験の調査。

・「暗号技術評価委員会報告」[21]（2021年3月）：
2020年度時点においてRSA暗号や楕円曲線暗号の安全性に量子コンピューターの脅威は生じていないことが確認されている。

・「ハイブリッドモードの技術動向調査」[22]（2020年12月）：
耐量子計算機暗号とRSA暗号などの現在使用されている公開鍵暗号の双方を併用するハイブリッドモードに関する調査。

・「暗号技術評価委員会報告」[23]（2022年3月）：
2021年度から2022年度にかけて耐量子計算機暗号の技術動向調査が実施され、耐量子計算機暗号ガイドラインを2022年度中に作成する予定となっている。

・「CRYPTREC耐量子計算機暗号の研究動向調査報告書」[26]（2023年3月）、
「CRYPTREC暗号技術ガイドライン（耐量子計算機暗号）」[27]（2023年3月）：
耐量子計算機暗号の代表的な候補である5種類の分類（格子に基づく暗号技術、符号に基づく暗号技術、多変数多項式に基づく暗号技術、同種写像に基づく暗号技術、ハッシュ関数に基づく署名技術）につい

て調査し、主に2022年9月30日までの調査結果を報告書、およびガイドラインにまとめたものである。

❸ 耐量子計算機暗号の標準化（米国・国立標準技術研究所（NIST））

NIST（National Institute of Standards and Technology）は2016年より耐量子計算機暗号の標準化プロジェクトを実施している[24)]。標準化プロジェクトでは、鍵共有のための暗号方式（鍵共有暗号方式）と署名のための暗号方式（署名暗号方式）の標準化が進められている。2022年7月に3回目の評価ラウンドが終了し、鍵共有暗号方式として格子に基づく暗号1件が、署名暗号方式として格子に基づく暗号2件とハッシュ関数に基づく暗号1件が標準化されることとなっている。また、標準化に進まなかった4件について4回目の評価ラウンドを実施することと、署名暗号方式を新たに公募することが決定されている[25)]。

（5）科学技術的課題

ここでは、「（3）研究開発領域の概要、（4）注目動向」に関する科学技術的課題を紹介する。耐量子計算機暗号については、「2.7.5 計算理論」を参照いただきたい。

❶ AIとプライバシー・個人情報保護

人工知能の学習には大量の情報が必要であり、特に個人情報や個人の行動履歴などのパーソナルデータを入力とする場合には、大量のデータの適切な収集と管理にコストがかかる。さらに、GDPRをはじめとする法令上の規制から、個人情報の収集は必要最小限度にとどめることが求められており、個人情報の収集が可能であったとしてもなるべく収集量を少なくすることが必要である。そのため、個人情報の提供者である個人の手元に情報をとどめるなどの対策を取りつつ、AIを学習させる技術が注目されつつある。具体的には、既存のAIを少量の情報だけを用いて別の目的のAIに転換する転移学習、少量の情報をもとにその情報の特徴を踏まえ類似情報を大量に生成するGAN（Generative Adversarial Networks）、個人の手元に情報をとどめ情報そのものではなく学習の手がかりになる情報（学習モデルの勾配）だけを収集してAIを学習させる連合学習（Federated Learning）などである。これらの技術は本来個人情報やプライバシー保護とは無関係に機械学習技術として発展してきたが、GDPRの発足とともに、個人情報やプライバシー保護を目的とした利用技術としても発展してきており、これらの技術と局所差分プライバシーや秘密計算を併用した技術の研究開発も期待されている。

GDPRでは、自分の情報をAIがどのように使用するかの決定権を個人が持つことを保証するよう求めており、AIによる決定のロジックに透明性があることが必要とされている。深層学習を始めとしてAIによる決定は帰納的であり、決定のロジックが説明不可能であることが多い。 AIによる決定を、演繹的・説明可能にすることも課題となっており、そのための研究もここ数年盛んになってきている。

❷ 局所差分プライバシーによるデータ活用

個人にかかわる情報であるパーソナルデータを活用するに際には、プライバシー保護が必要となる。局所差分プライバシーには、個人データを提供するユーザーやデータ収集者の間のやりとりの制限を定めた対話可能性という概念が理論解析において必要となる。ユーザーが一斉にデータをランダム化し、それらの処理済みデータを収集者がいったん収集してから統計処理を行うモデルを非対話的モデル、ユーザーがデータをランダム化する際にユーザーと収集者全員に共有された乱数を活用できるモデルを公開コインモデルと呼ぶ。公開コインモデルはユーザー負担の増加が少なく、非対話モデルに比べて統計処理で大きな精度の向上が見られる場合があり、前に紹介したGoogleやAppleの事例で利用されている。ユーザー一人ずつ逐次的にデータの収集を行う逐次的対話モデルや、同じユーザーに対して何回もデータ収集を行うことが可能な完全対話モデルは、プライバシーに配慮した機械学習を行うための対話モデルとして盛んに研究が行われている。加えて局所差分プライバシーにおいては、極めて多くのユーザーからのデータ収集がプロ

セスに含まれること、ユーザーはスマートフォンなど限られた計算能力と限られた通信帯域しか持たないデバイスを通じてデータ提供を行うことなどの事情から、サンプル複雑度に加えて、ユーザーサイドにおける送信データ生成に要する時間やデータ提供時の通信量なども議論の対象となる。スマートフォンやIoTなどデータ収集に利用されるデバイスやインフラに合わせたデータ収集スキームと理論解析は未解決課題である。

❸ AI学習モデルのプライバシー保護

AIでパーソナルデータを利用する際には、学習モデルのプライバシー保護も必要となる。深層学習は、画像認識や機械翻訳などの多くの分野で大きなブレークスルーを達成しているが、高精度なモデルを作成するためには、大規模なデータセットによるモデル学習が必要である。この際、データ収集に伴うプライバシーの問題に加えて、学習済みのモデルの公開に伴うプライバシー情報漏えいの問題にも対処することが必要である。これに対して、差分プライバシーを保証した深層学習が活発に研究されている。特に、ローカルにデータを所有する多数のクライアントから独立にモデルの更新情報だけを収集し、グローバルな深層モデルを学習する連合学習においては、クライアントからのデータ収集におけるプライバシー保護や、各クライアントがモデルの更新情報を計算するために配布されるグローバルモデルからのプライバシー情報漏えいのリスクがあることから、差分プライバシーや局所差分プライバシーの活用が模索されている。差分プライバシーの適用においては、データを信頼できる中央サーバーが収集し、そのデータを用いて中央サーバーがモデル学習を行うことによって、学習されたモデルの公開におけるプライバシーを保護しようとしている。その実現においては、モデル学習に用いる目的関数をランダム化した上でモデルを学習する方法や、学習の結果得られたモデルをランダム化する方法、学習に用いるデータをランダム化する方法などが提案されている。局所差分プライバシーの適用においては、中央サーバーが信頼できないケースを想定し、各クライアントが中央サーバーに提供するデータを提供前にランダム化することによって、データ提供におけるプライバシーを保護しようとしている。深層学習モデルの学習に必要なデータが大規模化するにつれて、こういったプライバシーを保護したデータ収集・活用の技術の必要性も高まると考えられる。

（6）その他の課題

❶ 法規制

わが国の個人情報保護法は、入力データとしての個人情報を保護するために必要な措置や、その措置を緩和するための手続き（匿名加工情報・仮名加工情報）を定めているが、急速に進展する人工知能などのデータ利用技術や秘密計算技術にキャッチアップできているとは言いにくい。データの活用と個人情報・プライバシー保護に関して、法制度は「だれもが理解できる範囲」の技術しか想定していない。世界的なAI開発競争の波に乗り遅れないためにも、最先端の技術の利用を促進するための工夫が必要である。例えば、用途や範囲を限定した上で、既存の規制にとらわれることなく新たな技術の実証を行える場を導入することなども考えられる。日本政府からは、データ活用の在り方とAI技術活用の在り方についての「データ戦略タスクフォース　第一次とりまとめ」（2020年12月）や、その具体的な取り組みの方向性となる「包括的データ戦略」（2021年6月）、「AI戦略2022」（2022年4月）が公表されている。

❷ 産学連携

産学連携は一昔前に比べれば活発になり、特に企業が所持するデータを利用した研究は盛んになった。一方で産学官の人材の行き来は欧米・中国に比べ活発ではなく、産は産、学は学、あるいは産から学への一方通行に限られる。クロスアポイントメント制度や時限付きでアカデミアの人材が積極的にインダストリーの中で活躍できるような事例が増加してゆけば良い効果が生まれる可能性がある。

❸ 人材育成

日本における本分野のトップ国際会議での存在感は非常に小さい。トップ国際会議での発表には粘り強く精密な実験と精緻な議論を行う必要があるが、そもそも博士課程を目指す学生が減少する中、アカデミアでは目先の成果を追い求め、チームで息の長い研究を行う体力が失われている。また産業界では、研究成果を広くオープンにするなど人材を引き寄せ発展を促す戦略を取っていない場合も多い。研究者を目指す学生を手厚く支援し、キャリアプランを充実化させ、研究開発に取り組みたいと思う若い研究者を地道に増やすこと、また流行の分野に大型予算を配分するだけでなく、基礎的な成果にも分け隔てなく継続的に中規模の予算を多方面に配分することが必要である。

❹ 経済安全保障

本領域は、特定重要技術領域のひとつである「サイバーセキュリティ技術」に関係しており、経済安全保障の視点が重要になってきている。セキュリティーは、製品の安全性確保のためには必須であり、社会の中のさまざまなシステムで利用されるため公共性が高い。また、さまざまな用途のシステムに適用できるため汎用性が高く、防衛用途にも使うことが可能なデュアルユース技術である。本領域の基礎研究、応用研究・開発では、米国、欧州、中国がリードしており、経済安全保障の観点から、日本の技術力の強化が望まれる。サプライチェーンでは、セキュリティー機能を含むハードウェアやソフトウェアを他国に依存している場合もあり、自国あるいは同志国から調達できる体制を整備することが望ましい。

（7）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	○	→	・暗号理論の基礎研究に従事する研究者は多く、論文も多く出ているが、統計的プライバシー、AIセキュリティー・プライバシーについては、取り組む研究者の数も少なく存在感が薄い。
	応用研究・開発	○	→	・企業による秘密計算実装の提供などが行われているが、応用分野における先進的なプロジェクトは少ない。
米国	基礎研究	◎	↗	・多くの学術論文が発表されている。いずれの研究領域においても、コアとなる理論的アイデアはほとんど米国の大学・企業の研究者から提案されている。
	応用研究・開発	◎	↗	・局所差分プライバシーなど理論成果の実サービスへの導入が進んでいる。産学の人材交流も活発である。
欧州	基礎研究	○	→	・GDPR施行もあって、データ利活用とプライバシーを見据えた基礎的な研究が活発である。
	応用研究・開発	◎	↗	・エストニアにおける秘密計算の実用化など、実用を見据えた動きは活発である。
中国	基礎研究	◎	↗	・中国本土の大学・企業でも、分野問わずトップ国際会議における論文数は年々増加している。 ・2018年に論文数で1位となっている。
	応用研究・開発	○	↗	・民間企業において、秘密計算などの実用例が出始めている。
韓国	基礎研究	○	→	・各種の暗号アルゴリズムの基礎的な研究を行い、国際標準に提案活動を行っている。
	応用研究・開発	△	→	・特に目立った活動は見られない。

（註1）フェーズ

基礎研究：大学・国研などでの基礎研究の範囲

応用研究・開発：技術開発（プロトタイプの開発含む）の範囲

（註2）現状　※日本の現状を基準にした評価ではなく、CRDS の調査・見解による評価

◎：特に顕著な活動・成果が見えている　　○：顕著な活動・成果が見えている

△：顕著な活動・成果が見えていない　　×：特筆すべき活動・成果が見えていない

（註3）トレンド　※ここ1～2年の研究開発水準の変化

↗：上昇傾向、→：現状維持、↘：下降傾向

参考文献

1）独立行政法人情報処理推進機構（IPA）「情報セキュリティ白書2020」https://www.ipa.go.jp/files/000087025.pdf,（2023年2月24日アクセス）.

2）総務省「平成29年版情報通信白書　データ主導経済と社会変革（本編）」https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h29/pdf/29honpen.pdf, https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h29/pdf/29honpen.pdf,（2023年2月24日アクセス）.

3）Latanya Sweeney, “k-Anonymity: a model for protecting privacy,” International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10, no. 5（2002）: 557-570., https://doi.org/10.1142/S0218488502001648.

4）Marcella Hastings, et al., “SoK: General Purpose Compilers for Secure Multi-Party Computation,” in 2019 IEEE Symposium on Security and Privacy（SP）（IEEE, 2019）, 1220-1237., https://doi.org/10.1109/ SP.2019.00028.

5）Toshinori Araki, et al., “High-Throughput Semi-Honest Secure Three-Party Computation with an Honest Majority,” in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security（New York: Association for Computing Machinery, 2016）, 805-817., https://doi.org/10.1145/2976749.2978331.

6）Gilad Asharov, et al., “Efficient Secure Three-Party Sorting with Applications to Data Analysis and Heavy Hitters,” in Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security（New York: Association for Computing Machinery, 2022）, 125-138., https://doi.org/10.1145/3548606.3560691.

7）Yehuda Lindell and Benny Pinkas, “Privacy Preserving Data Mining,” in Advances in Cryptology - CRYPTO 2000, ed. Mihir Bellare, Lecture Notes in Computer Science 1880（Berlin, Heidelberg: Springer, 2000）, 36-54., https://doi.org/10.1007/3-540-44598-6_3.

8）Rakesh Agrawal and Ramakrishnan Srikant, “Privacy-preserving data mining,” in Proceedings of the 2000 ACM SIGMOD international conference on Management of data（New York: Association for Computing Machinery, 2000）, 439-450., https://doi.org/10.1145/342009.335438.

9）Pascal Paillier, “Public-Key Cryptosystems Based on Composite Degree Residuosity Classes,” in Advances in Cryptology - EUROCRYPT ’99, ed. Jacques Stern, Lecture Notes in Computer Science 1592（Berlin, Heidelberg: Springer, 1999）, 223-238., https://doi.org/10.1007/3-540-48910-X_16.

10）Ronald Linn Rivest, Adi Shamir and Leonard Max Adleman, “A method for obtaining digital signatures and public-key cryptosystems,” Communications of the ACM 21, no. 2（1978）: 120-126., https://doi.org/10.1145/359340.359342.

11）Cynthia Dwork, et al., “Calibrating Noise to Sensitivity in Private Data Analysis,” Journal of Privacy and Confidentiality 7, no. 3（2016）: 17-51., https://doi.org/10.29012/jpc.v7i3.405.
12）Shiva Prasad Kasiviswanathan, et al., “What Can We Learn Privately?” SIAM Journal on Computing 40, no. 3（2011）: 793-826., https://doi.org/10.1137/090756090.
13）Peter W. Shor, “Algorithms for quantum computation: discrete logarithms and factoring,” in Proceedings 35th Annual Symposium on Foundations of Computer Science（IEEE, 1994）, 124-134., https://doi.org/10.1109/SFCS.1994.365700.
14）Peter W. Shor,“Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer,” SIAM Journal on Computing 26, no. 5（1997）: 1484-1509., https://doi.org/10.1137/S0097539795293172.
15）細山田光倫「量子コンピュータが共通鍵暗号の安全性に及ぼす影響の調査及び評価（2020年1月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/exreport/cryptrec-ex-2901-2019.pdf,（2023年2月24日アクセス）.
16）国立研究開発法人情報通信研究機構, 独立行政法人情報処理推進機構「CRYPTREC Report 2019：暗号技術評価委員会報告」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/report/cryptrec-rp-2000-2019r1.pdf,（2023年2月24日アクセス）.
17）国立研究開発法人情報通信研究機構, 独立行政法人情報処理推進機構「耐量子計算機暗号の研究動向調査報告書（2019年3月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/report/cryptrec-tr-2001-2018.pdf,（2023年2月24日アクセス）.
18）高安敦「Shorのアルゴリズム実装動向調査（2021年6月3日）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/exreport/cryptrec-ex-3005-2020.pdf,（2023年2月24日アクセス）.
19）国立研究開発法人科学技術振興機構研究開発戦略センター「戦略プロポーザル Society 5.0 時代の安心・安全・信頼を支える基盤ソフトウエア技術」https://www.jst.go.jp/crds/pdf/2020/SP/CRDS-FY2020-SP-06.pdf,（2023年2月24日アクセス）.
20）暗号技術調査（暗号解析評価）ワーキンググループ「格子問題等の困難性に関する調査（2015年3月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/exreport/cryptrec-ex-2404-2014.pdf,（2023年2月24日アクセス）.
21）国立研究開発法人情報通信研究機構, 独立行政法人情報処理推進機構「CRYPTREC Report 2020：暗号技術評価委員会報告（令和3年3月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/report/cryptrec-rp-2000-2020.pdf,（2023年2月24日アクセス）.
22）株式会社レピダム「ハイブリッドモードの技術動向調査」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/exreport/cryptrec-ex-3004-2020.pdf,（2023年2月24日アクセス）.
23）国立研究開発法人情報通信研究機構, 独立行政法人情報処理推進機構「CRYPTREC Report 2021：暗号技術評価委員会報告（令和4年3月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/report/cryptrec-rp-2000-2021.pdf,（2023年2月24日アクセス）.
24）Computer Security Resource Center, “Post-Quantum Cryptography,” National Institute of Standards and Technology（NIST）, https://csrc.nist.gov/projects/post-quantum-cryptography,（2023年2月24日アクセス）.
25）Gorjan Alagic, et al., “NIST IR 8413-upd1: Status Report on the Third Round of the NIST

Post-Quantum Cryptography Standardization Process," National Institute of Standards and Technology（NIST）, https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8413-upd1.pdf,（2023年2月24日アクセス）.
26）「耐量子計算機暗号の研究動向調査報告書（2023年3月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/report/cryptrec-tr-2001-2022.pdf,（2024年7月18日アクセス）.
27）「暗号技術ガイドライン（耐量子計算機暗号）（2023年3月）」Cryptography Research and Evaluation Committees（CRYPTREC）, https://www.cryptrec.go.jp/report/cryptrec-gl-2004-2022.pdf,（2024年7月18日アクセス）.