［安全なデータ活用を実現する秘密計算技術］

①秘密計算による安全なデータ共有
―秘密計算技術の概要と社会実装に向けた課題―

竹之内隆夫 | NEC セキュリティ研究所

データ共有の期待と秘密計算

データ共有による社会価値創出

AI や IoT 等の技術の進展に伴い，多くの企業・組織にとって，データの重要性はますます高まってきている．特に一部の企業では，戦略的にデータを収集・分析・活用し，企業価値を高めている．

従来のデータ活用は，組織内におけるデータ活用にとどまっていたが，今後は組織を跨いでのデータ活用が期待されている．たとえば，医療・ヘルスケア分野では，ゲノムバンクが保有するゲノム情報と，医療機関が保有するカルテ情報を結合して分析することで，ゲノムの特性に応じた効果的な投薬に関する研究などの医学研究が進むことが期待されている．このような組織間でのデータ共有は，さまざまな分野でも期待されており，たとえば，物流の最適化や不正な金融取引の検知などが挙げられる．企業・組織を跨いだ適切なデータ共有は，さまざまな社会的な価値を生む可能性がある．

データ共有の課題と秘密計算

しかし，このような複数組織でのデータ共有は，プライバシ侵害の恐れや，データの囲い込み思考があり，実際には十分には進んでいない．たとえば，個人情報をほかの組織に提供する場合は事前の本人同意が必要であるが，提供先の組織でデータが不適切に利用されプライバシが侵害される不安感などもあり，同意を得るのは容易ではない．また，個人情報に限らず企業秘密を組織間で共有する場合も，データが競合企業等に渡る恐れを懸念し，データを囲い込む傾向もある．

秘密計算技術は，このような課題を解決し，組織間での安全なデータ活用を実現し，より良い社会を実現するための基礎的な技術になると期待されつつある．

以降，まず秘密計算技術の概要を説明し，次に，そのユースケースの一例を示す．そして，秘密計算のさまざまな方式について整理した後，ほかの技術との連携の必要性を説明する．最後に実用化に向けた課題を述べる．

秘密計算とは

秘密計算とは，データを秘匿したまま処理できる技術である．この技術によって，複数の参加者（パーティ）が，各参加者が持つ秘密情報をほかの参加者に開示することなく，それらの秘密情報についてのあらかじめ決められた処理を行い，その処理結果だけを出力できる[1),2)]．このような処理をマルチパーティ計算とも呼ぶ．本特集では秘密計算技術とマルチパーティ計算技術を同じ技術として記載する．

図 -1 に秘密計算の概要を示す．この図のように，組織 A と組織 B が，それぞれ個人情報や企業秘密などの機密データ A，B を持つとする．このとき，秘密計算を用いることで，機密データ A，B を秘匿したまま結合して分析し，分析結果だけを開示することができる．つまり，秘密計算を用いれば，機密データ A，B は元データとしては外部に提供せずに，分析結果だけが外部に提供されることになる．なお，この図では，分析結果を組織 C が利用して

いるが，分析結果を組織 A や組織 B が利用してもよい．

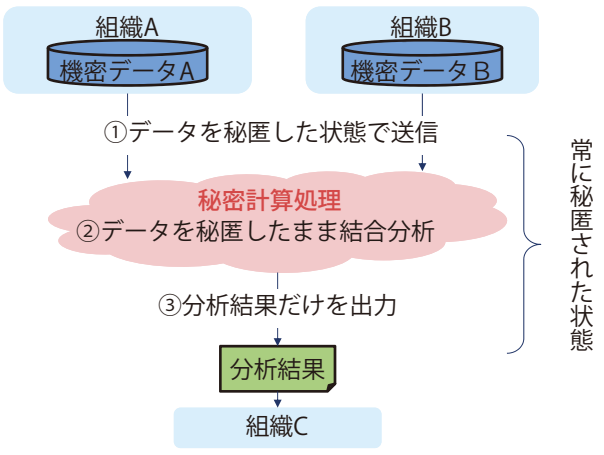
このように，秘密計算を用いれば，生データを組織外に開示することはないため，プライバシ侵害や競合企業にデータが渡る危険性を減らすことできる．これにより複数組織での安全なデータ活用が促進されることが期待される．

また，秘密計算は近年の研究によって大幅に高速化が進み，一部の方式は実用レベルに達したと考えられている．たとえば，単純な比較は難しいが，2004 年に提案された方式では秒間約 620 回の論理演算が可能であったが，2015 年に提案された方式は秒間約 1,300 万回の論理演算が可能である [2]．このように，秘密計算の処理能力は急速に向上しており，実用化が期待される．

次に，まず秘密計算のユースケース例を示し，その後，秘密計算の具体的な方式について説明する．

秘密計算によって生まれる価値とユースケース

このような秘密計算であるが，まずは秘密計算によって生まれる価値を整理し，そして，それぞれのユースケースの一例を示す．



■図 -1　秘密計算の概要

秘密計算によって生まれる価値は，大きく以下の 2 つに整理できる．

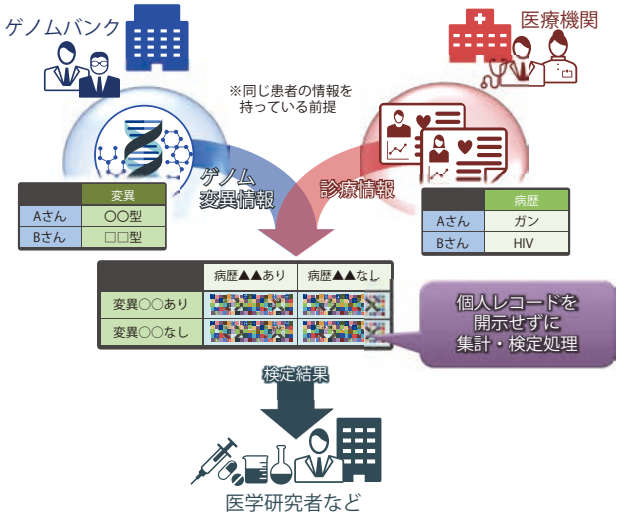
1. 組織間の安全なデータ共有による知識発見
2. 秘匿されたままの処理による安全性向上

1 つ目の価値は，これまで説明した通り，複数の組織間で秘密計算を利用する場合の価値である．2 つ目の価値は，複数事業者に限らず，単一の組織内で用いた場合にも得られる．以降に，ユースケースの一例を示す．

安全なデータ共有のユースケース

1 つ目の複数組織間での安全なデータ共有による新たな知識発見が生まれるという価値についての典型的なユースケースが，先ほど説明したゲノムやカルテなどの医療情報の組織間の共有分析である．医療情報は，個人情報の中でも特に機微性が高い反面，適切に活用することで医学研究が促進され社会価値が大きいため，データ共有が期待されている．

図 -2 に示した例のように，ゲノムバンクが持つゲノム情報と，医療機関が持つカルテ情報を，秘密計算を用いて秘匿しながら結合し，ゲノムの特性と投薬・疾病等の相関分析を行うことが可能である．この例では，ゲノム特性と疾病などの関係が統計的に優位な差があるか調べる検定処理を行い，その検



■図 -2　ユースケース例：ゲノム分析

定処理の結果だけを医学研究者に開示している．検定結果からは，元のゲノムやカルテ情報を推測することは困難であるため，プライバシ侵害や意図しない研究者にデータが不正にわたる心配も軽減できる．

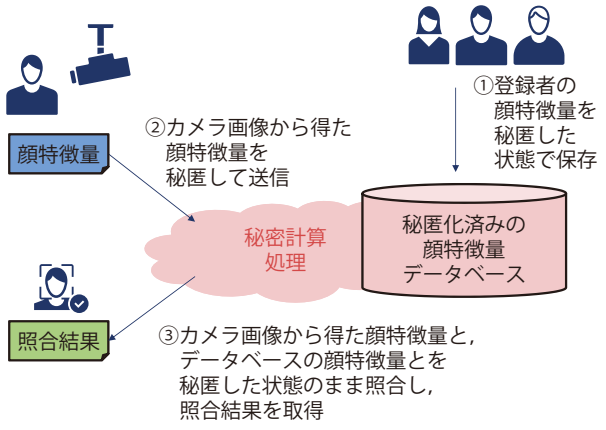
秘密計算を用いることで，医療情報のような機微性の高い情報を，従来よりも安全に結合分析できるため，医学研究の発展に寄与できると考えられている．このような秘密計算を用いたゲノム分析の実証は，いくつかの企業や大学等の秘密計算の研究者によって行われており，一定の成果が出ている．

安全性向上のユースケース

秘密計算のもう１つの価値は，安全性の向上である．従来の暗号技術では，処理を行うには，暗号化したデータを復号して元のデータにする必要があるため，元データに復号された際に，不正な攻撃者にデータを盗み見られる恐れがあった．それに対し，秘密計算では，秘匿したままの処理が可能であるため，データが漏洩するリスクを格段に減らせる．

たとえば，顔認証システムにおける登録者の顔特徴量情報の安全管理に，秘密計算を利用できる．図 -3 に示した例の場合，たとえば，登録者全員の顔特徴量情報を，秘密計算が可能な形で秘匿した状態でデータベース等に保存しておく．ここで，従来であれば，カメラから取得した顔特徴量とデータベースに格納された登録者の顔特徴量との照合を行う際には，データベースに格納された顔特徴量を元データに復元してから照合を行う必要があった．しかし，秘密計算を用いれば，秘匿したままの照合処理が可能である．そのため，登録者の顔特徴量情報は常に秘匿された状態であるため，顔特徴量が漏洩する危険性は格段に減る．

このように秘密計算は，顔特徴量のような機微性が高い情報の安全管理にも用いられることが期待される．また，この場合，複数組織間というよりは，単一組織内（もしくは処理の委託先の組織内やクラウド事業者内）での利用が想定される．



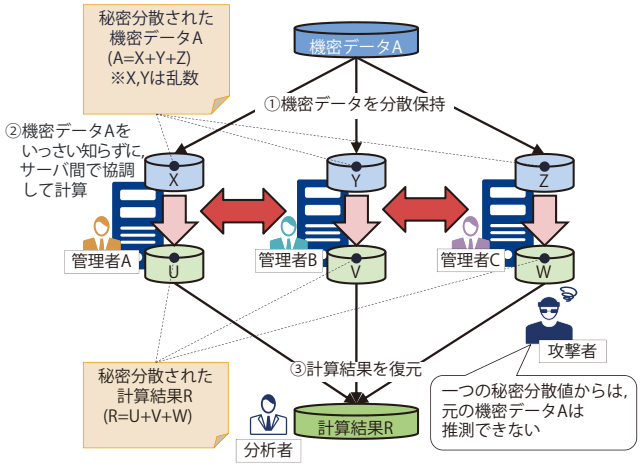
■図 -3　顔特徴量を秘匿したまま照合

秘密計算のさまざまな方式

秘密計算には，さまざまな方式が存在する．ここでは，代表的な方式である「秘密分散を用いた方式」「準同型暗号を用いた方式」「Garbled circuit（秘匿回路）を用いた方式」の３つを説明する．

秘密分散を用いた方式

図 -4 は，秘密分散を用いた秘密計算の一例である．ある機密にしたいデータを，複数の値に分け，そのうちの一定数が集まると復元できるような形式で値に分ける（秘密分散）．この個々の値（秘密分



■図 -4　秘密分散型の秘密計算の一例

散値，シェア）は，単体では乱数にしか見えず，元の値は推測できない．この値を，管理者の異なる複数のサーバに配置すれば，管理者が結託しなければ元データに戻ることはないので安全といえる．

秘密分散を用いた秘密計算では，この秘密分散した値に対して，秘密計算を行う処理手順に従って各サーバ間で情報をやりとりし，元の値に戻すことなく処理を行う．そして，処理結果の秘密分散値が生成され，それを最後に集めて復元すると，処理結果を得ることができる．つまり，常に秘密分散された状態で処理が行われる．

この処理を，図 -5 に示したように複数の組織が持つ機密データに対して行う場合を考える．すると，この図で示したように，組織 A と B の機密データを常に秘密分散された状態のまま，結合分析し，分析結果だけを出力する処理が実現できる．

なお，秘密分散は，暗号のような計算量に依存する安全性（計算量的安全性）よりも強い安全性（情報量的安全性）となる．また，秘密分散の安全な方式は，いくつか提案されており，一部は ISO/IEC 19592-2:2017 で標準化されている．

また，この方式は，AND や XOR のような論理ゲートを処理可能であるため，その組合せによって理論上は任意の処理が可能である．また，近年高速化を実現し，文献 2）で述べられている通り，実用レベルの処理速度を実現できる方式として期待されている．

準同型暗号を用いた方式

準同型暗号とは，暗号化したまま加法や乗法が計算可能な暗号方式である．たとえば，Paillier 暗号は加法が可能な準同型暗号であり，以下のように暗号化した a と暗号化した b を掛け算すると，暗号化した a＋b が得られる．

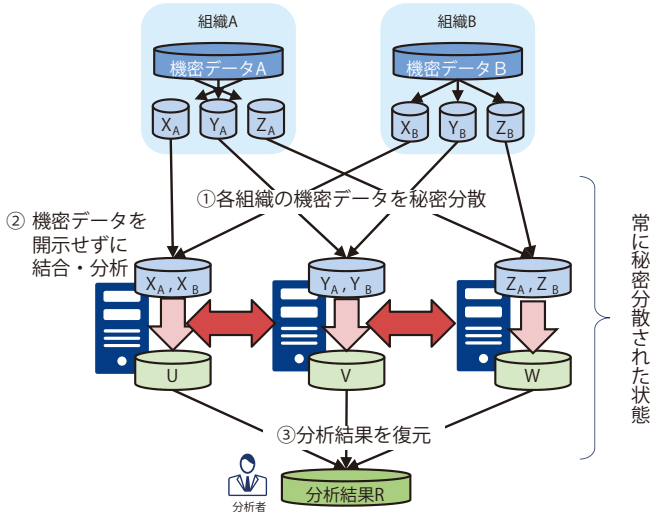
$$\mathrm{Enc}(a) \times \mathrm{Enc}(b) = \mathrm{Enc}(a+b)$$

また，Enc (a × b) のような結果が得られる乗法が可能な準同型暗号も存在する．

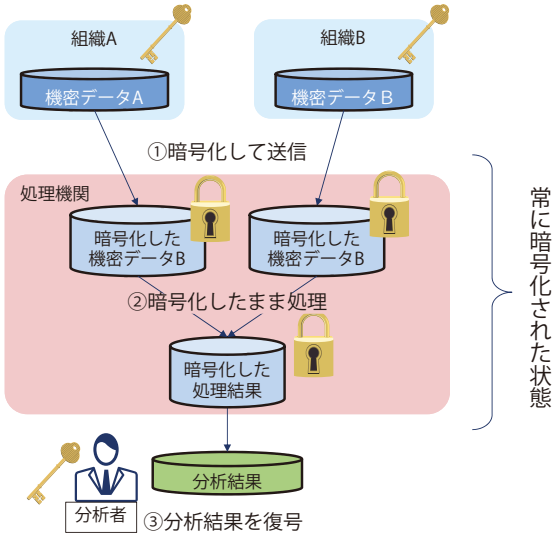
さらに，2009 年に Gentry らによって示された，加法も乗法も可能な完全準同型暗号も存在する．

このような準同型暗号を用いることで，図 -6 に示したように，組織 A，B が持つ機密データを同じ鍵で暗号化し，暗号化したまま結合分析して，分析者に暗号化した分析結果を提供し，復号するということが可能である．

この方式では，鍵をどのように管理するかが安全性を確保する上で重要である．鍵を不正に入手できないような仕組みが必要となる．



■図 -5　複数組織の機密データを秘密計算



■図 -6　準同型暗号を用いた秘密計算

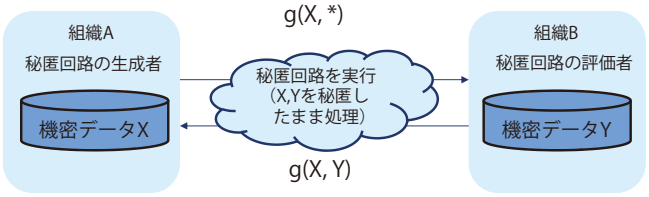
また，この方式は，特定の処理に特化するのであれば，ある程度高速に処理できるため，すでに実用的な場面で使われている．さらに，新たな暗号方式やそれらを組み合わせた処理に関する研究も活発で，実証実験もいくつか行われている．しかし，さまざまな処理を行える完全準同型暗号は，まだ処理速度等に課題があり，実用化のためには今後のさらなる研究が期待される．

Garbled circuit を用いた方式

Garbled circuit を用いた秘密計算とは，1986 年に Yao によって基礎的な技術が提案された論理回路を秘匿して処理する方式である．この方式は，図 -7 に示したように，2 つの組織間の片方で Garbled circuit を生成し，その回路に対して双方の機密情報を秘匿しながら回路を実行し，結果を得る処理を行うものである．この方式は，通信量は大きいが，通信回数を少なくできるなどの特徴がある．

そのほかの方式

また，上記以外の方式も存在する．たとえば，近年ではハードウエアを用いた方式として，Intel SGX（Software Guard eXtentions）のような，TEE（Trusted Execute Environment）という特別なハードウェアを用いた方式がある．これは，ハードウェア内に鍵が格納されており，ハードウェア内で復号して処理し，結果を出力する処理が可能である．ハードウェア内で処理するため，非常に高速である．しかし，たとえば Intel SGX では 10 回程度の実行で鍵を抽出できるようなソフトウェアレベルでのサイドチャネル攻撃が見つかるなど，いくつかの安全性の課題が見つかっている[2), 3)]．そのため，現在では本方式は機密データを保護した処理方式としては不十分と考えられ，さらなる研究が期待されている．



■図 -7　Garbled circuit を用いた方式

以上のように，秘密計算にはさまざまな方式が存在する．各方式の詳細や適用事例等については，本特集の以降の記事で説明しているため，参照いただきたい．

実際の利用のためのほかの技術との組合せ

秘密計算を実際に利用する際には秘密計算以外の技術との組合せも必要となる場合がある．以降の節で，代表的な技術を挙げる．

処理結果からの情報漏洩防止

秘密計算で実行する処理は，処理結果から元の機密情報が推測できないようにすべきである．たとえば，先ほどのゲノム分析の例では，分析結果として，検定処理の結果だけを出力している．検定結果からは，元データの推測は困難であるため，ある程度の安全性は確保できる．しかし，たとえば，条件を変えて複数回分析処理を行うことも考えると，分析結果の差から何らかの元データを推測できる可能性も出てくる．

秘密計算技術は入力を秘匿したまま出力を得られる技術であるが，出力からの情報漏洩に関する安全性は保障していない．そのため，秘密計算を実際に適用する際には，出力からの情報漏洩を防止する他の技術と組み合わせることを検討する必要がある．たとえば，さまざまな処理クエリに対応するようなサービス形態であれば，処理クエリを制限するようなクエリ制御技術や，分析結果にノイズを加えるなどの出力プライバシ保護などの技術との組合せも有用であろう．

異なるデータベースのレコード結合

秘密計算を用いて複数組織のデータを結合するには，レコードの紐づけが必要となる．先ほどのゲノム分析の例では，患者を識別する共通の ID が存在する前提であるが，実際には，どのようにレコードの一致判定を行うかも検討する必要がある．レコードの紐づけは，完全な一致による判定だけでなく確率的に一致を判定する手法もある[4]．実用化のためには，このようなレコードの一致判定を，秘密計算を用いて秘匿しながら処理することも検討する必要がある．

まとめと実用化に向けた課題

本稿では，秘密計算の概要やユースケースの一例や，秘密計算の代表的な方式について説明した．秘密計算は，新たな社会的な価値を生む重要な技術となる可能性があると考えている．本稿などを参考に，実サービスへの適用可能性の検討が進めば幸いである．

また今後，秘密計算を実用化し，社会実装を進めていく上ではいくつか課題がある．以降に主な課題を述べる．

制度整備

秘密計算は，従来の暗号とは異なり，秘匿したままの処理が可能である．そのため，従来の暗号化を前提とした制度・ガイドラインとは異なる新たな制度も必要と考える．

また，さまざまな方式があり安全性の前提も異なる．たとえば，秘密分散方式では結託を防止する必要があり，準同型暗号方式は安全な鍵管理の基準も必要となるだろう．今後，このような安全性の基準を明確にするとともに，新たな制度・ガイドラインを整備していくべきと考える．

今後の制度整備に向け，社会受容性や技術の検証のための実証実験や，技術者だけでなく政府関係者や法学者も含めた議論が必要と考える．

技術の浸透・使いやすさ向上

秘密計算は 1980 年代からある技術であるが，近年高速され，実用化・普及が見えてきている．しかし，秘密計算にはさまざまな方式が存在し，一般の技術者への理解が十分浸透しているとは言えないのが現状である．

たとえば，秘密分散方式の秘密計算は，さまざまな処理に対応可能であり，かつ，実用レベルの速度が出ている方式の 1 つである．今後は，このような実用レベルとなってきている方式を先鋒として，技術の普及や，秘密計算アプリケーションを作るためのツール類の整備など，実用化に向けたさまざまな取り組みを進めていくべきと考える．そして，秘密計算が広く普及し，さまざまな社会的な価値を生んでいくことを期待する．

参考文献

1) 佐久間淳：データ解析におけるプライバシー保護，機械学習プロフェッショナルシリーズ，講談社 (2016).
2) 荒木俊則，五十嵐大，高橋克巳，竹之内隆夫，ティブシ・メディ，花岡悟一郎，古川　潤：秘密計算の実用可能性，SCIS2018.
3) Moghimi, A., Irazoqui, G. and Thomas, E : Cachezoom : How SGX Amplifies the Power of Cache Attacks, CHES 2017.
4) Harron, K. : Introduction to Data Linkage, The Administrative Data Research Network (ANDR) Publication (2016).

(2018 年 7 月 4 日受付)

竹之内隆夫（正会員） takenouchi@bu.jp.nec.com

2005 年 NEC 入社．現在，セキュリティ研究所主任研究員．2013 年電気通信大学大学院情報システム学研究科博士後期課程修了．博士（工学）．主としてプライバシ保護技術，秘密計算技術の研究開発に従事．