STRATEGIC BUSINESS INSIGHTS

scan™

2020 年 12 月

SoC1198

Automating Privacy in Data Use

By Rob Edmonds（Send us feedback）

データ利用におけるプライバシーの自動化

ビッグデータの利用が始まって間もない頃から、プライバシーをめぐる懸念によってデータ駆動型ソフトウェアの潜在力が制限されていた。たとえば 2005 年の「SoC083：プライバシー危機の蔓延」では、プライバシーへの脅威が急速に増えているようだと指摘されている。2013 年の「SoC655 ：ビッグデータ、大きな懸念」では、ビッグデータ・システムの開発を望む企業が、データ規制や消費者のプライバシーへの期待など、さまざまな問題に直面している状況が紹介されている。2017 年の「SoC933：スヌーピング技術」では、顧客サービスとプライバシーの侵害が、紙一重になりつつあると論じられた。2019 年初頭の「SoC1059：（プ）レビュー2018/2019：データとプライバシー」では、2018 年にScan™で採り上げられたプライバシー関連のトピックを総ざらいしている。2020 年の「SoC1171 ：コロナウイルスと市民の自由」では、Covid-19（2019 年型コロナウイルス感染症）パンデミック対策を目的としたデジタル・トラッキングの利用が、データ・プライバシーに真っ向から抵触する現状について解説している。しかし、データの大規模な収集・処理と、鉄壁の個人プライバシーとが互いに両立する未来が実現可能だとしたら、どうだろうか。このような未来が訪れたら、プライバシーに関する規制や懸念事項に縛られずに、ビッグデータの潜在力がようやく解き放たれるかもしれない。一般的なデータ駆動型コンピューティングはもとより、データ駆動の人工知能が大きく進歩する可能性が高い。医療、教育、消費者サービスなどの分野でブレイクスルーが起こり得る。プライバシーの問題からデジタル変革の目標を穏便なものにしていた企業も、意欲的なデジタル戦略に踏み切れるようになるだろう。不確実性は存在するが、現時点で見られるさまざまな道しるべから、自動化によってプライバシーの問題が基本的に取り払われた遠い未来が、説得力をもって浮かび上がっている。

プライバシー管理ツールを利用すると、法規制の順守に必要なコストが削減され、リスクが低下する。

自動プライバシー管理ツール、フェデレーテッド・マシン・ラーニングなどの新興のテクノロジーによって、企業が個人のプライバシーを保護しながら、急速に増大しつつあるデータを最大限に利用できるようになる可能性がある。

ヨーロッパの一般データ保護規則（GDPR）、カリフォルニア消費者プライバシー法など、プライバシー法制の順守を支援する半自動ソフトウェアは、さまざまなベンダーからすでに提供されている。この種のソフトウェアが投資家の多大な関心を集めている。たとえば米国の BigID は、1 億 5,000 万ドル近い資金を調達した。同社が提供するプラットフォームは、データ・アナリティクスと並行してコンプライアンス・ソフトウェアを使用することで、企業における個人データの識別とマッピングを支援する（これらのデータが、複数のオンプレミス・システムやクラウドベース・システムに分散されていても問題ない）。さらに、顧客や従業員からのデータ・アクセス要求の履行など、プライバシーに関連する機能が、このプラットフォームによって部分的に自動化される。また、米国の Securiti.ai は、設立後わずか 2 年強しか経っていないにも関わらず 8,100 万ドルの資金を調達している。同社も BigID と同様、機械学習を利用して複数の場所に分散した個人データを管理し、プライバシー・コンプライアンスのさまざまな側面を部分的に自動化する。他にも事例がある。2019 年 5 月に設立され現在までに 4,000 万ドルの資金を調達した米国の InCountry は、複数の司法権管轄区域に広がるデータ・ストレージの管理を通じて、企業におけるプライバシー・コンプライアンスを支援している。その他に注目すべきプロバイダーとしては、英国の Privitar、米国の TrustArc、そして 2020 年中頃に競争相手の Integris Software を買収した米国・英国の OneTrust がある。

プライバシー管理ソリューションの出現によって、コンプライアンスにかなりの金額を投資できる企業と、そうでない企業との間で競争条件が公平化される可能性がある。米国の Google（Alphabet）では、GDPR の施行開始に先立って、GDPR コンプライアンスを確保する作業に延べ数百年に相当する人的時間が費やされたとみられる。プライバシー管理ツールを利用すると、法規制の順守に必要なコストが削減され、リスクが低下する。テクノロジー業界以外の企業でも、国際市場向けの新しいデータ・サービスの開発が促進される可能性がある。

データ・プライバシーの自動化を目的とするソフトウェアのもう 1 つの系統が、フェデレーテッド・ラーニング・システムによる AI である。機械学習 AI は、データ・セット（通常、非常に大容量）に依存してインテリジェントな挙動を学習する。ところが多くの場合、プライバシー法やプライバシーをめぐる懸念のため、機械学習に必要とされる十分なデータに開発者がアクセスすることができない。フェデレーテッド・ラーニング・システムは、AI トレーニング・タスクを個別のモジュールに分割し、これらのモジュールをローカル・デバイスに分散することにより、そのような懸念に対処するシステムである。たとえば、AI トレーニングに必要なデータが個人の携帯電話に含まれている場合、モジュール化した機械学習トレーニング・システムは、携帯電話それ自体に存在するデータを利用することができ、そのデータをクラウドにアップロードする必要はない（トレーニング済みの AI モジュールだけがアップロードされる）。フェデレーテッド機械学習ソフトウェアを開発している企業としては、米国の Google、IBM、Intel、Micosoft などがある。

プライバシーの自動化は、さらなる前進が期待できる。目立ったところでは、IBM が最近、完全準同型暗号（FHE）のフィールド試験を完了した。FHE は、コンピューターが暗号化データを復号化せず、暗号化されたままの状態で演算や論理演算を実行することを可能にする暗号化技術である。基本的に FHE を使用すると、暗号化されていないデータに対して実行できる操作は何でも、暗号化された機密データに対して（少なくとも理論上は）実行できる。IBM によるこの前進は、将来への有望なステップになり得るが、FHE を使用するには、演算能力とメモリをかなり増強しなければならない。たとえば、FHE で暗号化された機械学習モデルは、非暗号化モデルと比べて、同じタスクを実行するために必要な演算能力は 40～50 倍、必要なメモリは 10～20 倍であることが IBM の試験で明らかになっている。このようなパフォーマンス・トレードオフがあるため、金融サービス、医療、官庁などの業種への応用は、今のところ狭い範囲に限られそうだ。それでも演算速度が向上し FHE 技法が進化するにつれ、トレードオフが軽減される可能性がある。

自動プライバシー・ツールは今のところ限られており、ほとんどの場合、プライバシー関連のタスクを全面的に自動化するものではない。それでも、企業がプライバシー・コンプライアンス・タスクを自動化できるようになり、個人データの処理によってプライバシーが危険にさらされることのない未来への道筋を指し示している。このような未来の実現に向けて望みが持てそうになりつつあるが、それは少なくとも 10 年は先の話であり、かなりの不確実性も存在する。技術的な課題はまだ顕著であり（たとえば FHE は依然として開発の初期段階である）、プライバシー保護ソフトウェアへの信頼も育てる必要がある。喜ばしいことに、このような未来に至る道筋には、さまざまな可能性が存在する。たとえば、すでに出回っている市販のツールを利用して、プライバシー・コンプライアンスを能率化することはすぐにでも可能である。

SoC1198

本トピックスに関連する Signals of Change		関連する Patterns	
SoC1171	コロナウイルスと市民の自由	P1273	監視技術はもろ刃の剣
SoC1059	(プ)レビュー2018/2019:データとプライバシー	P0843	プライバシーとセキュリティ:難しいバランス
SoC843	(プ)レビュー2015/2016:データの負の側面	P0661	2014 年における『1984 年』