差分プライバシーと個人情報保護法との関係

関原秀行（インハウスハブ東京法律事務所）、古川直裕（株式会社 ABEJA）、後藤大（晴海パートナーズ法律事務所）、木村菜生子（株式会社 Hacobu）、松村将生（株式会社エクサウィザーズ）、森田岳人（松田綜合法律事務所）、落合孝文（渥美坂井法律事務所・外国法共同事業）、柴山吉報（阿部・井窪・片山法律事務所）、堤雄輝（堤総合法律事務所）

本論文は、筆者全員が所属する AI 法研究会セキュリティ部会における検討内容に基づくものである。なお、主執筆者は関原秀行と古川直裕（本論文への貢献度順）である。本論文に関するご質問・ご意見等は、AI 法研究会のメールアドレス（aiandlaw.secretariat@gmail.com）にお送りいただきたい。

2022 年 4 月 1 日

第１　序論

１　本論稿の目的

差分プライバシー（differential privacy。以下「DP」という。）とは、識別不能性に基づくプライバシー保護技術であり、オリジナルのデータにノイズを付加すること等によってプライバシーを保護するものである。近年、企業等においてビッグデータの利活用が盛んであり、積極的な研究・活用が進められている。

DP は、プライバシーを保護するためのアプローチであることから、個々人に関連する情報、いわゆるパーソナルデータに対して用いられることが一般的である。もっとも、パーソナルデータの中核をなす「個人情報」の取扱いに関する規律を定める「個人情報の保護に関する法律」（平成 15 年法律第 57 号。以下「個人情報保護法」又は「法」という。）と DP との関係を整理した文献は我が国では多くはない[1]。特定の技術とそれに関連する法律との関係が不明瞭な場合、過剰な萎縮効果により適切な技術の導入が躊躇される可能性がある。本論稿では、DP に最も関連する法律である個人情報保護法との関係を検討した上、両者の関係を可能な限り明らかにすることによって、DP の導入における不要な萎縮効果を除去することを目的とするものである。

２　DP とは

⑴　DP の意義

DP とは、前述したとおり、プライバシー保護技術であり、オリジナルのデータに対して、ハッシュ化、サンプリング、ノイズの付加等のランダム化を施し、プライバシーを保護するためのものである[2]。

DP は、もともとはデータベースが開示する計算結果（統計量という。例えば、日本におけるがん患者のデータベースの場合、がん患者の合計数、がん患者の平均年齢、30 代のがん患者の合計数、がん患者の最高年齢などが統計量にあたる。）の公開におけるプライバシー保護を目指したものであった。つまり、このような統計量の公開であっても、攻撃者の前提知識等によってはプライバシー侵害が発生するため、開示する統計量に誤差を加えて不正確な値を開示するというわけである。例を出すと、日本におけるがん患者のデータベースが存在し、攻撃者は 30 代の人物 A 以外の全 30 代のがんの有無を知っているが A だけ知らないという場合、30 代のがん患者数の公開により攻撃者は A のがんの有無を知ることができてしまう。このため、真の 30 代のがん患者数が 10,000 である場合、これに一定の統計的

---

[1] 古川・渡邊編「Q&A AI の法務と倫理」（中央経済社・2021 年 5 月）は、法律実務家の立場から、差分プライバシーとプライバシー侵害との関係について考察されている（同 451-453 頁）。

[2] DP の理論的詳細については、Cynthia Dwork, Aaron Roth “The Algorithmic Foundations of Differential Privacy”(https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf)が詳しい。

なノイズを加えて、9,888 や 10,231 といった値を 30 代のがん患者数として公開するのである。

また、DP を応用して情報収集におけるプライバシー保護を図ることも可能である。例えば、スマートフォンのブラウザからあるサイトの訪問回数を収集する場合、訪問回数をいわば統計量とみて、スマートフォン側でノイズを加えた訪問回数を収集するということができる。

なお、DP に関する数学的説明については割愛する。主に、どの様なノイズを付加することができるか、どの程度のノイズを加えたら、どの程度の安全性が確保されるかに関する数学的な検討を行うものであるため、本稿には影響を与えないためである。

AI との関係に言及すると、DP を AI 開発に利用することもできる。DP を用いてオリジナルのデータに対してノイズを付加等することによって、統計データからオリジナルのデータに関する個々人に関する情報の推測、AI 学習済モデルから学習の元となったオリジナルのデータを推測すること等がより困難となり、オリジナルのデータを利用した場合と比較して、よりプライバシーを保護する形で、統計データや AI 学習済モデルを公開することが可能となる。

もっとも、小規模なデータセットに対して DP のアプローチを採用した場合、オリジナルのデータセットを利用した場合のアウトプットとの間の誤差が大きい可能性があるため留意が必要である。

⑵ DP の実例

近年、企業や行政機関等において DP のアプローチを採用している例が見受けられる。

例えば、Apple、Facebook のような IT 企業においては、データの収集に際して DP を用いていることを公開し、White paper 等を公表している事例が見受けられる[3] [4]。

また、米国の国勢調査局（US Bureau of Census）は、同国の国勢調査に DP を導入しており、DP におけるプライバシー保護の程度を示す数学的指標である ε 値も公表している。

⑶ DP 技術を適用するタイミング

DP のアプローチにおいては、オリジナルのデータに対してノイズを付加する等してランダム化の処理が行われる。当該処理を行うタイミングは、①クライアント側でランダム化の処理を行う場合と、②サーバ側でランダム化の処理を行う場合に大別される。

①は、スマートフォン側（クライアント側）においてオリジナルのデータに対してノイズを付加する等した上、企業が管理するサーバ側にノイズが付加されたデータを送信するケースである（このようにクライアント側でノイズを付加するケースを以下「ローカルモデル」

[3] https://www.apple.com/jp/privacy/features/

[4] https://privacytech.fb.com/differential-privacy/

という。)。

②は、スマートフォンからオリジナルのデータを企業が管理するサーバに送信した上、受領したオリジナルデータに対してサーバ側でノイズを付加するケースである（このようにサーバ側でノイズを付加するケースを以下「セントラルモデル」という。)。

3　個人情報保護法とは

個人情報保護法とは、「個人の権利・利益の保護」と「個人情報の有用性」とのバランスを図るための法律である（法 1 条)。

そして、同法は、企業等の民間事業者における「個人情報」の取扱いに関する規律を規定する[5]。

本論稿では、データの処理に当たり DP のアプローチを採用することが個人情報保護法の適用の有無に影響するかどうかを検討する。その上で、一般的なデータのライフサイクルに沿って、DP と個人情報保護法との関係を整理する。最後に、DP が契約関係に与える影響について考察する。具体的な検討項目は以下のとおりである。

① DP と個人情報保護法の適用関係
② DP データの取得
③ DP データの管理
④ DP データの利用
⑤ DP データの提供・公開
⑥ DP データに対する権利行使
⑦ DP における契約上の留意点

第 2　各論

1　DP と個人情報保護法の適用関係

まず、オリジナルのデータに対して DP のアプローチを採用してノイズの付加等を行うことが個人情報保護法の適用関係に影響するかどうかが問題となる。もっとも、個人情報保護法は、「個人情報」に限らず様々な情報の類型を定義し、情報ごとに規律を規定している。したがって、本論稿では、オリジナルデータにノイズの付加等を行うことが同法の主な規律対象である「個人情報」該当性に影響するかどうかを検討する。

結論から言えば、DP によるオリジナルデータに対するノイズの付加等の処理を行うことは、オリジナルのデータを利用する場合と比較してプライバシーが保護される可能性はあるものの、それ自体によって処理後のデータの「個人情報」該当性を左右するものではない。もっとも、ノイズの付加の方法等によっては、本来「個人情報」であったオリジナルデータ

---

[5] 2021 年 5 月「デジタル社会の形成を図るための関係法律の整備に関する法律」が成立し、同法の施行後は行政機関、地方公共団体等も個人情報保護法の適用対象となる。

を非個人情報として企業が取得することが可能となるケースは想定される。

⑴ 「個人情報」とは

個人情報保護法は、「個人情報」について、2つの類型を規定する（法2条1項)。具体的には、生存する個人に関する情報であり、以下の①又は②のいずれかに該当するものを「個人情報」と定義する。

① 個人識別符号が含まれるもの 例：顔認識データ、指紋認識データ、マイナンバー 等 ② 当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合でき、それにより特定の個人を識別することができるものを含む。）

したがって、DPによるノイズの付加等によって処理されたデータが「個人情報」に該当するか否かは、当該データが上記の「個人情報」の定義に該当するか否かによって左右されるものであり、DPによる処理を施したからといって当然に当該データが非個人情報となるものではない。

⑵ DPによる処理と「個人情報」該当性

特定のデータの「個人情報」該当性は、前述した「個人情報」の定義に該当するか否かによって左右される。そして、ローカルモデル、セントラルモデルのいずれの場合であっても、ランダム化によって特定個人を識別可能な情報が含まれるかどうかは、ランダム化の手法やサーバ側に送信するデータの種類に依存するものであり、ノイズの付加等の処理を行ったからといって当然に特定のデータが非個人情報となるわけではない。

すなわち、取得するデータが「個人情報」に該当するか否かは、取得するデータに個人識別符号（法2条2項）が含まれている場合を除き、クライアント側からサーバ側に送信されるデータが以下の①～③のいずれに該当するかによって定まるものと考えられる。

① サーバ側に送信されるデータ自体が特定個人識別性を有する場合 ② サーバ側に送信されるデータ自体は特定個人識別性を有しないがサーバ側において特定個人識別性を有することとなる場合 ③ サーバ側に送信されるデータ自体は特定個人識別性を有せずサーバ側で取得した後も特定個人識別性を有しない場合

①の場合、サーバ側で受領するデータ自体が特定個人識別性を有する情報であるため、企業がサーバ側で取得した時点で「個人情報」を取得したことになる。例えば、特定個人の氏名とともに購入履歴等の一定のデータをサーバ側に送信する場合、顔写真のように特定個人を識別可能な画像データをサーバ側に送信する場合がこれに該当する。

②は、例えば、氏名等の「個人情報」と紐づけ可能な識別子等の情報を購入履歴等の一定

のデータとともにサーバ側に送信する場合である。クライアント側で DP によるランダム化の処理を施した個人の購入履歴等のデータを、当該個人を一意に識別可能な識別子とともにサーバ側に送信する場合が想定される。送信された識別子とセットで氏名等の個人情報をサーバ側で保有している場合、上記の購入履歴等のデータは当該識別子によってサーバ側で取得した後は、氏名等の個人情報と紐づけが可能であり、その紐づけが容易であれば当該データの取得は「個人情報」の取得に該当する[6]。

③は、それ単体では特定個人を識別不能であり、かつサーバ側で取得した後も氏名等の個人情報と紐づけ不能なデータを送信する場合である。この場合、サーバ側で大量の行動履歴情報を蓄積すること等によって特定個人識別性を有するものと解されない限り[7]、当該データは非個人情報である。

なお、スマートフォンのようなクライアント側の端末内に個人情報が保存されており、それに対して DP による処理を行ったとしても、当該処理は企業の管理支配権の範囲外で行われているものである。したがって、サーバ側で取得する以前に企業の管理支配権の範囲外であるクライアント側で DP による処理を行うことは、一般的には当該企業における個人情報の取扱いには該当しないものと考えられる。

2　DP データの取得

クライアント側からサーバ側にデータを送信し、当該データを取得する場面における個人情報保護法上の主な論点は以下のとおりである。

①　DP を用いていることを利用目的として特定する必要があるか
②　DP によるデータ取得は適正取得義務と抵触するか

なお、前述のとおり、ローカルモデル、セントラルモデルのいずれの場合であっても、取得するデータの個人情報該当性には直接影響しない。ただし、ローカルモデルは、セントラルモデルと比較して、取得するデータと特定個人との結びつきが弱まることが一般的であるため、セントラルモデルよりもプライバシーに配慮した実装といえる。なお、後述するとおり、ローカルモデルの場合、取得したデータが真実を反映したデータかどうかを判断することが困難であるため本人からの訂正請求権の行使等の関係で影響が生じる可能性があると考えられる。

---

[6] 個人関連情報を個人データとして取得する場合、個人データの取得に該当するものとして確認記録義務の対象になることとされており、単体では特定個人識別性を有しない情報の取得も「個人情報」の取得に当たるものと解されている。

[7] 位置情報のように情報の蓄積によって個人情報となる場合がある（個人情報保護委員会「個人情報の保護に関する法律についてのガイドライン（通則編）」（平成 28 年 11 月・令和 3 年 8 月一部改正）・82 頁参照）。

(1) DP と利用目的の特定、通知・公表義務との関係

個人情報取扱事業者[8]は、「個人情報」を取り扱うに当たっては、その利用目的をできる限り特定しなければならず（法 17 条 1 項）、また、「個人情報」を取得した場合には、取得前に利用目的を公表している場合を除き、速やかに特定した利用目的を本人に通知し、又は公表しなければならない（法 21 条 1 項）。

「利用目的」を「特定」するとは、個々の取扱いプロセスごとにその目的を特定することを求める趣旨ではなく、あくまで個人情報取扱事業者が一連の取扱いにより最終的に達成しようとする目的を特定することを求めるものとされている[9]。

もっとも、「利用目的の特定」の趣旨は、個人情報を取り扱う者が、個人情報がどのような事業の用に供され、どのような目的で利用されるかについて明確な認識を持ち、できるだけ具体的に明確にすることにより、個人情報が取り扱われる範囲を確定するとともに、本人の予測を可能とすることであり、本人が、自らの個人情報がどのように取り扱われることとなるか、利用目的から合理的に予測・想定できないような場合は、この趣旨に沿ってできる限り利用目的を特定したことにはならず、例えば、本人から得た情報から、本人に関する行動・関心等の情報を分析する場合、個人情報取扱事業者は、どのような取扱いが行われているかを本人が予測・想定できる程度に利用目的を特定しなければならない[10]。個人情報保護法を所管する個人情報保護委員会が策定したガイドラインにおいては、この具体例として、ターゲティング広告の配信、信用スコアの算出・第三者提供を挙げている。

これらに鑑みると、法は最終的な利用目的を特定することが求めており、その内容は本人が予測・想定できる程度に特定しなければならないものではあるが、最終的な目的に至る過程で用いられる技術的手段までを特定することは求めていないものと解される[11]。

翻って DP について検討すると、DP はプライバシーを保護するための指標ないし技術的手段であり、取得したデータを用いて達成しようとしている最終的な利用目的そのものではない。したがって、DP を用いていること自体は、必ずしも個人情報保護法上は利用目的として特定する必要はないものと考えられる。

もっとも、透明性の観点から、データの取得又は取得したデータの処理の過程で DP を用いていることについて公表等を行うことは望ましい措置と考えられる。実際に、複数の企業がデータの処理に際して DP を用いていることを公表している（第 1、2(2)）。

---

[8] 個人情報データベース等を事業の用に供している者（法 16 条 2 項）

[9] 園部・藤原「個人情報保護法の解説《第二次改訂版》」（ぎょうせい・平成 30 年 2 月）・137 頁

[10] 個人情報保護委員会「個人情報の保護に関する法律についてのガイドライン（通則編）」（令和 4 年 4 月 1 日施行）（平成 28 年 11 月・令和 3 年 10 月一部改正）・31-32 頁

[11] 行政解釈上も「利用目的から合理的に予測・想定できる限り、必ずしも個別具体的な情報分析の技術的な手法まで含めて利用目的を特定する必要はない」とされている（令和 2 年改正 GL パブコメ・No.28）。

⑵ DP と適正取得義務との関係

個人情報取扱事業者は、偽りその他不正の手段によって個人情報を取得することが禁止される（法 19 条 1 項)。

「不正の手段」は、「偽り」に限られず、不適法な又は適正性を欠く方法や手続が含まれ、具体的事案において当該規律に抵触するか否かは、事案ごとに個人情報保護法その他の法令の趣旨や社会通念に沿って判断される。例えば、本人に対して個人情報を収集しているという事実や収集する目的を偽って取得する場合、正当な権限なく他人が管理する個人情報を取得したり隠し撮りする場合、十分な判断能力を有していない子どもから親の個人情報を取得する場合等がこれに該当するとされている[12]。

この点、データの取得や利用に当たり DP を用いること自体は、個人情報保護法の他の規律に反せず、通常のデータ取得・利用と比較して本人のプライバシー保護に資するものであり、一般的に社会通念上適正性を欠く方法による取得には当たらない。したがって、DP によってランダム化したデータの取得は、不適正な個人情報の取得には該当しないと考える。これは、ランダムなノイズを加えた結果、本人に不利益に見えるデータが作出されたとしてもあてはまる。例えば、アダルトサイトへの訪問回数について、ランダム化したため実際の訪問回数よりも多い数字がサーバ側に記録されたることになったとしても（これが本人にとって不利益と言えるかは別の問題である。)、同様である。

ただし、本人を識別可能な情報とともにクライアント側でノイズを付加してデータを取得した場合において本人に不利益が生じるような場合には、例外的に不適正取得に該当することはあり得ると考えられる。

3 DP データの管理

ローカルモデル、セントラルモデルのいずれの場合であっても、サーバ側で取得したデータの管理に関する主な論点は以下のとおりである。

① DP は安全管理措置義務に影響を与えるか
② DP によるノイズの付加は正確性確保の努力義務に抵触するか

⑴ DP と安全管理措置義務との関係

ローカルモデル又はセントラルモデルのいずれを選択した場合であっても、サーバ側で取得したデータは「個人データ」に該当する可能性があり、その場合、個人情報保護法が規定する安全管理措置義務（法 23 条～25 条）の適用対象となり、当該データの安全管理のために必要かつ適切な措置を講じなければならない。

セントラルモデルの場合、サーバ内には、「オリジナルデータのデータベース」と「ノイズを付加した処理済データのデータベース」の 2 類型が存在する場合がある。このような

[12] 園部・藤原「個人情報保護法の解説《第二次改訂版》」（ぎょうせい・平成 30 年 2 月）・148 頁参照

場合、仮にオリジナルのデータが個人情報に該当し、ノイズを付加したデータが個人情報に該当しない場合には、両者が照合できない場合においては、「オリジナルデータのデータベース」は「個人情報データベース」に該当することから安全管理措置義務の対象となるが、「ノイズを付加した処理済データのデータベース」は「個人情報データベース」には該当せず、安全管理措置義務の対象外となる。

なお、講じるべき安全管理措置は、個人データが漏えい等をした場合に本人が被る権利利益の侵害の大きさを考慮し、事業の規模及び性質、個人データの取扱状況（取り扱う個人データの性質及び量を含む。）、個人データを記録した媒体の性質等に起因するリスクに応じて、必要かつ適切な内容としなければならないとされているところ[13]、DP によるランダム化後のデータが個人データであっても、オリジナルデータと比較して特定個人との関連性等が弱まっていたり、データの機微性が低減している場合には、講じるべき安全管理措置のレベルがオリジナルデータと比較して低い水準で足りるものと考えられる。

⑵ DP と正確性確保の努力義務との関係

個人情報取扱事業者は、「利用目的の達成に必要な範囲内」において、個人データを正確かつ最新の内容に保つ義務を負う（法 22 条）。

もっとも、常に正確なデータであることが求められているわけではない。個人情報を利用する際に当該情報に求められている内容については、その利用目的によって、最新時点における事実に関する情報が必要となる場合だけでなく、過去の一定時点の事実に関する情報が必要となる場合や、軽微な変更は重要でない場合などの様々な場合が想定されるため、法 22 条の正確性確保の努力義務は、個人情報取扱事業者がその保有する個人データをそれぞれの利用目的に応じて、その達成に必要な範囲内で正確性・最新性を確保するよう努めることで足りる[14]。つまり、当該規律において考慮される「利用目的の達成に必要な範囲」は、最終的な利用目的を達成するために必要な手段・過程を広く含む概念と考えられる。

この点、DP によるランダム化は、事実と異なるデータを取得し、又はデータに対して事実と異なる修正を加える等するものであり、その目的はプライバシーに配慮しながら統計データの作成・公開や AI 学習済モデルを開発・提供するといったものである。その目的を達成するためには DP によるランダム化処理を行う必要がある。そのため、差分プライバシーによるランダム化の処理は、利用目的の達成に必要な過程・手段であり、そのような処理を行ったとしても正確性確保の努力義務の規律には抵触しないものと考えられる。

4 DP データの利用

---

[13] 個人情報保護委員会「個人情報の保護に関する法律についてのガイドライン（通則編）」（平成 28 年 11 月・令和 3 年 8 月一部改正）・45 頁

[14] 園部・藤原「個人情報保護法の解説《第二次改訂版》」（ぎょうせい・平成 30 年 2 月）・161 頁参照

個人情報取扱事業者は、本人の同意なく、法 17 条 1 項により特定された利用目的の達成に必要な範囲を超えて「個人情報」を取り扱うことが禁止される（法 18 条 1 項）。

ランダム化したデータが「個人情報」に該当する場合、利用目的の制限規定（法 18 条 1 項）の対象となるものの、前述したとおり DP は個人情報を利用する最終的な目的ではなく、その目的を達成するための手段に過ぎない。そのため、DP によるランダム化を施したデータの最終的な利用目的が特定されている限り、その目的を達成する過程においてデータのランダム化の処理を行ったとしても、当該処理は利用目的の達成に必要な範囲内であり、DP を利用していること自体は利用目的として特定する必要はなく、当該データは利用可能と考える。

なお、ランダム化後のデータのみが存在するデータベース内に保管されているデータが「個人情報」に該当しない場合や DP により処理されたデータの利用結果としてのアウトプットである統計データ自体（「個人情報」に該当しない。）は、個人情報に該当しない以上、そもそも利用目的の制限規定の規律の対象外である。

5　DP データの提供・開示

ランダム化されたデータ自体が「個人データ」に該当する場合、法が定める個人データの第三者提供規制（法 27 条 1 項）の対象となり、当該データを第三者に提供するためには原則として本人の同意が必要となるが、DP の目的がランダム化したデータを利用した統計データの作成や AI モデルの作成にあることから、実際はこのようなケースは基本的には想定し難いと思われる。

本来の利用用途であるランダム化したデータを用いた統計データの作成の場合、アウトプットされる統計データは「個人情報」には該当しないため[15]、オリジナルデータの本人の同意を要せずに、公開を含む第三者提供を行うことが可能である。

また、ランダム化したデータを用いて作成された AI モデルには学習済のパラメータが含まれているが、当該パラメータは特定の出力を行うために調整された処理・計算用の係数に過ぎず、複数人から収集してランダム化した学習用データを用いて生成された学習済パラメータは一般的に特定個人との対応関係が排斥されているため「個人情報」には該当せず、元データの本人の同意なく[16]、公開を含む提供が可能である。

なお、後述するとおり、ランダム化したデータを利用して作成された統計データや学習済モデルは、オリジナルのデータを用いて作成された統計データや学習済モデルとは異なるものであり、提供先の第三者との関係における契約関係等において留意が必要となる場合

---

[15] 個人情報保護委員会「個人情報の保護に関する法律についてのガイドライン（通則編）」（平成 28 年 11 月・令和 3 年 8 月一部改正）・82 頁

[16] 個人情報保護委員会『「個人情報の保護に関する法律についてのガイドライン」及び「個人データの漏えい等の事案が発生した場合等の対応について」に関する Q&A』（平成 29 年 2 月 16 日・令和 3 年 6 月 30 日更新）・Q1-7-2 参照

がある。

6　DP データに対する権利行使

サーバ側で取得したデータ及び当該データをランダム化したデータに対する本人からの権利行使に関する主な論点は以下のとおりである。

①　ランダム化したデータに対する開示請求権の行使への対応
②　ランダム化したデータに対する訂正請求権の行使への対応

⑴　ランダム化したデータに対する開示請求権の行使への対応

個人情報取扱事業者は、データの主体である本人から、本人が識別可能な保有個人データに対して開示請求を受けた場合、原則として遅滞なく当該保有個人データを開示しなければならない（法 33 条 1 項、2 項)。

DP によりランダム化されたデータについても請求主体である本人を識別可能であり個人情報保護法が規定する「保有個人データ」の定義に該当する場合、個人情報取扱事業者は請求を受けたデータを開示する必要がある。

なお、ランダム化したデータが令和 2 年改正で新設された「仮名加工情報」(法 2 条 5 項)に該当する場合、開示請求の対象外となる。

⑵　ランダム化したデータに対する訂正請求権の行使への対応

個人情報取扱事業者は、データの主体である本人から、当該本人が識別可能な保有個人データの内容が事実でないものとして、当該保有個人データの内容の訂正、追加又は削除を請求された場合には、利用目的の達成に必要な範囲内において、遅滞なく必要な調査を行い、その結果に基づき、当該保有個人データの内容の訂正等を行わなければならない（法 34 条 1 項、2 項)。

訂正請求等は、「利用目的の達成に必要な範囲内において」対応すれば足り、例えば、現時点のデータとしてではなく単に履歴データとして保存しておく必要があったり、一定時点のあるいは一定の状況下におけるデータとして記録しておく必要がある場合には対応が不要とされており[17]、同条項の「利用目的の達成に必要な範囲」には最終的な利用目的のみならずそれを達成するための手段も含まれるものと考えられる。

DP によるランダム化によって保有個人データの内容が事実と異なる場合であっても、原則として、プライバシーを保護しつつ統計データや AI モデルを作成する目的を達成するための手段であるため訂正請求への対応を行わなかったとしても適法と考えられる。

なお、令和 2 年改正により、ランダム化したデータが仮名加工情報に該当する場合には、訂正請求等の対象外となる。

---

[17] 園部・藤原「個人情報保護法の解説《第二次改訂版》」（ぎょうせい・平成 30 年 2 月）・241 頁参照

7　DP における契約上の留意点

差分プライバシーによるノイズの付加によって生成されるデータは、オリジナルのデータとは異なるデータである。したがって、ノイズを付加したデータを用いて作成された統計データや AI の学習済モデルは、オリジナルデータを用いて作成された統計データや AI 学習済モデルとは異なるアウトプットとなる。データが大規模になればなるほどアウトプットはオリジナルデータによるアウトプットに近づくが、それでも異なるデータである以上、オリジナルデータと同じアウトプットとはならない。

そして、企業は、それらのアウトプットを外部に公開・提供する場合がある。例えば、ノイズを加えたデータを用いて作成した統計データを企業のサービス利用者数として公開する場合、ノイズを加えたデータを用いて開発した AI モデルを公開・提供する場合等が想定される。

このようにオリジナルデータと異なるノイズが付加されたデータを元に作成された統計データは、外部の第三者から見た場合、オリジナルデータを元に算出・作成されたものとの誤解が生じる可能性があるため、当該統計データの公開に当たっては DP によるノイズを付加したデータを元に作成されたものであることを明記することが望ましい。

ノイズを付加したデータを元に料金、報酬等の金額が算出されることはオリジナルデータを用いた場合の金額と異なる金額となり得るため基本的にはこのような金額の算出にノイズを付加したデータを利用することは避けるべきと考えるが、仮にノイズを付加したデータによって金額を算出する場合、契約内容に金額の算出の元データに DP によりノイズを付加されたデータを用いること、オリジナルデータによって算出される金額とは異なる金額が算出されそれが料金、報酬等となること等を盛り込むことがトラブルの回避につながるものと考える。

また、AI 学習済モデルの開発の受託の場面において、学習用データとしてノイズが付加されたデータを用いる場合には、その旨を契約書に盛り込むべきである。

以上