



非特定情報のリスクの考察

2014.5.13

NTTセキュアプラットフォーム研究所

高橋克巳

- 準個人情報・個人特定性低減データを考えるために、非特定情報のリスクを考察する
 - 非特定情報：
誰であるのかわからない情報、無名情報、仮名情報とも
- リスクの類型化（マップ）
 - 個人特定
 - 個人への到達
 - データ分析
 - なりすまし
- 考察の前提
 - 狭義の個人情報以外のもののリスクを考えたため、対象は「個人情報を無名・仮名したもの」「もともと無名・仮名で収集したもの」「準個人情報かもしれないもの」などが混在している
 - リスクとして記述したもののすべてを一概に問題視するものではない

非特定情報のリスク 《個人特定 1/2》



1. 「対応表」が手に入り個人特定される
 - 非特定情報に対する氏名等への対応表があり、それが漏えいした場合
2. データが個人登録／個人と結びつけられて、それ以降個人特定される
 - 個人登録なしで使っていた端末IDを個人登録した場合
 - 顔認識データが個人と結びつけられた場合

非特定情報のリスク 《個人特定 2/2》



3. データが複数の異なった組織で取得可能で、そのデータが名寄せされ、その組織の一つになされた個人登録から特定される
 - あるサービスで個人登録なしで使っていた端末IDが、他のサービスで個人登録されていた
 - あるサービスで個人登録なしで集められた位置情報と、他の個人が登録されたサービス(アプリで)集められた位置情報とがマッチングされる
4. データ中のあるデータが(攻撃者の知識によって)個人特定される
 - データに特徴がある場合
 - 2m・150kg の人
 - 買い物履歴データと、ある人のブログの書き込みが一致している
 - データに識別性があり(母集団に対する)網羅性がある場合
 - 攻撃者がその母集団からそのデータに合致する人を知っている

6. データに含まれる情報から直接的な個人到達が起こる
 - 習慣的な移動履歴から、待ち伏せされる

7. データに含まれる情報から電子的な個人到達が起こる
 - メールアドレスにメールが送付される
 - 習慣的なウェブ閲覧履歴からブラウザにターゲティング広告される
 - ※一概に問題視されるものではない

8. 非特定の個人に対するデータが集積され、その結果個人特定される可能性が高くなる
 - 利用期間が長い場合(ある履歴の継続的な収集)
 - 利用範囲が広い場合(複数の履歴・情報の名寄せ)

9. 非特定の個人に対するデータから、特定の個人への風評等の被害が起きる
 - アドレスが「ブラックリスト」に登録されてしまう

10. 非特定の状態で「自分に」に関するデータが集積され、「自分に」に関する様々な推定が行われる
 - ※一概に問題視されるものではない

非特定情報のリスク 《なりすまし》



11. 非特定情報を用いて、不正アクセスが行われる
 - IDとパスワードを用いて不正アクセスが行われる
 - 生体情報を用いて不正アクセスが行われる

- 本類型に基づいて、詳細なケーススタディーが必要
 - リスクの詳細分析
 - 対策の分析
 - データ自身に対するもの(加工方法等)
 - 運用プロセスに関するもの(オプトアウト手段等)
 - 副作用の有無の分析
 - 他の事業者義務との関係(安全管理措置等)
- 準個人情報の方考え方
 - リスクの背後にある情報の性質を指標化し、各種のデータを準個人情報であるか否かフィルタリングする
- 個人特定性低減データの考え方
 - リスクを分析し、対応が必要なものに対して持つべき性質と加工方法を分析する
 - 低減データといっても、可能であれば個人識別性の低いデータへの加工(真面目な匿名化)を志すことは重要
 - 真面目な匿名化に応じて、提供者の報告等が楽になる方策があればよいと考える