



# ビッグデータとシステム監査

株式会社 SCC  
内桶 孝雄

# ビッグデータとは

- 「市販されているデータベース管理ツールや従来のデータ処理アプリケーションで処理することが困難なほど巨大で複雑なデータ集合の集積物を表す用語」  
(Wikipedia)

# パーソナルデータの利活用

- ビッグデータの中でも、パーソナルデータの活用について注目が集まっている。
- パーソナルデータは、「個人に関する情報」をいい、特定の個人を識別できる場合は「個人情報」となる。

# JR東日本の事例

- J社が記名式SUICAの乗降データをH社に販売し、H社はそれを統計データとして加工し販売する予定だった
- 利用者やマスコミから批判や反発を受けた。

# 大阪駅ビルの事例

- J研究機構は4月から2年間予定していた実証事件の実施を延期すると発表した。
- 実験の内容は、「大阪ステーションシティ」という商業ビルに90台のカメラを設置し、人の顔に当たる部分を自動的に識別して切り出し、その人物が施設内でどれくらい滞留しているか、どのように移動しているかなどを把握するもので、防災に役立てるものとしていた。
- JR大阪駅や同商業ビルの利用者は、そこを通らないことでしか実証実験への参加は拒否できず、プライバシー侵害に関する懸念の声が上がっていた。

# システム監査実施において

- **コンプライアンスの側面から**
  - 個人情報の提供にあたっては、原則として本人の同意が必要
- **技術的な視点から**
  - 「個人情報」に該当しないようにするための匿名化

# コンプライアンスの側面から (個人情報か)

- 個人情報に該当するか
  - 特定の個人を識別できるかどうか境界線であり、住所や氏名が含まれているかどうかは判断の基準とはならない
- 提供するデータが個人情報でないこと

# コンプライアンスの側面から (取得・提供)

- **取得時のコントロール**
  - 適切に取得が行われたかどうか
  - 個人情報の取得時に、利用目的が本人に通知または公表されているかどうか
- **提供時のコントロール**
  - データが個人情報に該当する場合、第三者に当該データを提供する際、本人の同意が得られているか、あるいはオプトアウトの仕組みがあるか



# 技術的な視点から (照合容易性)

- どのようなこと
  - 「他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるもの」は個人情報である
  - 社内においてこのような状況にあるデータは個人情報である
- コントロール
  - システム的にできない、ということだけではなく、社内全体でできないことを保証することが必要
  - そのためのコントロールがあるか

# 容易に照合

- Q:事業者の取扱部門ごとにデータベースがあり、他の取扱部門のデータベースへのアクセスが、規程上・運用上厳格に禁止されている場合、「容易に照合することができ」（法第2条第1）るといえますか。
- 
- A:他の取扱部門のデータベースへのアクセスが規程上・運用上厳格に禁止されている場合であっても、双方の取扱部門を統括すべき立場の者等が双方のデータベースにアクセス可能な場合は、当該事業者にとって「容易に照合することができ」る状態にあると考えられます。ただし、経営者、データベースのシステム担当者などを含め社内の誰もが規程上・運用上、双方のデータベースへのアクセスを厳格に禁止されている状態であれば、「容易に照合することができ」とはいえないものと考えられます。（2007.3.30）
- （「個人情報の保護に関する法律についての経済産業分野を対象とするガイドライン」等に関するQ&A）

# 技術的な視点から（匿名化）

- どのようなことか
  - 特定の個人を識別できないよう、データを加工する
  - 「完全な匿名化」は困難と言われている
- コントロール
  - 十分な匿名化が図られているかどうか
    - 住所や氏名が含まれていない場合でも、特定個人識別性が残る場合がある

# 個人情報に該当する可能性がある例

氏名	年齢	家族1	家族2	家族3
A田 一郎	42	妻(36歳)	長女(10歳)	長男(8歳)
B川 二郎	47	長女(21歳)	次女(16歳)	長男(11歳)
C山 三郎	35			

A田 一郎	0601_01
B川 二郎	0601_22
C山 三郎	0601_48



変換表は操作が完了したら消去する

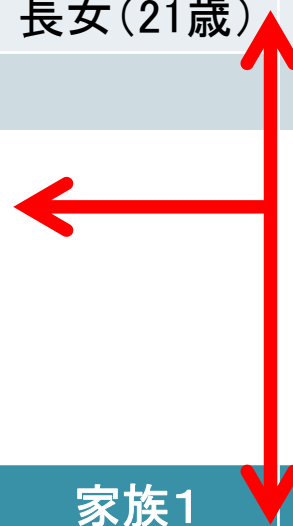
氏名	年齢	家族1	家族2	家族3
0601_01	42	妻(36歳)	長女(10歳)	長男(8歳)
0601_22	47	長女(21歳)	次女(16歳)	長男(11歳)
0601_48	35			

# 個人情報に該当する可能性がある例

氏名	年齢	家族1	家族2	家族3
A田 一郎	42	妻(36歳)	長女(10歳)	長男(8歳)
B川 二郎	47	長女(21歳)	次女(16歳)	長男(11歳)
C山 三郎	35			

A田 一郎	0601_01
B川 二郎	0601_22
C山 三郎	0601_48

氏名	年齢	家族1	家族2	家族3
0601_01	42	妻(36歳)	長女(10歳)	長男(8歳)
0601_22	47	長女(21歳)	次女(16歳)	長男(11歳)
0601_48	35			



# 個人情報に該当しない例

氏名	年齢	家族1	家族2	家族3
A田 一郎	42	妻(36歳)	長女(10歳)	長男(8歳)
B川 二郎	47	長女(21歳)	次女(16歳)	長男(11歳)
C山 三郎	35			

A田 一郎	0601_01
B川 二郎	0601_22
C山 三郎	0601_48



変換表は操作が完了したら消去する

氏名	年代	家族
0601_01	40~	3人
0601_22	40~	3人
0601_48	30~	0人

# 個人情報に該当しない例

氏名	年齢	家族1	家族2	家族3
A田 一郎	42	妻(36歳)	長女(10歳)	長男(8歳)
B川 二郎	47	長女(21歳)	次女(16歳)	長男(11歳)
C山 三郎	35			

A田 一郎	0601_01
B川 二郎	0601_22
C山 三郎	0601_48

氏名	年代	家族
0601_01	40~	3人
0601_22	40~	3人
0601_48	30~	0人

# システム監査上の留意点

- **照合容易性がないことを証明できるか**
  - システム的にできない、ということではなく、全社において、容易に照合できない、といった状況となっているか
- **提供先での再識別可能化への考慮**
  - 再識別できないよう加工したパーソナルデータであっても、提供先において特定の個人が識別できるよう加工される恐れもある



# パーソナルデータの利活用に関する制度見直し

- 2014年6月：大綱決定・公表
- ↓パブリックコメント
- ↓法案作成
- 2015年1月：通常国会に法案提出
  
- パーソナルデータに関する検討会
  - <http://www.kantei.go.jp/jp/singi/it2/pd/>