

視聴履歴等の取扱いに係る検討について (3. 匿名加工情報の取扱い)

平成29年5月11日

事務局

視聴履歴取扱指針等で定めることが望まれる事項

3. 匿名加工情報の取扱い

3-1 視聴履歴の匿名加工情報の作成方法

1. 視聴履歴の匿名加工情報の作成等

- ◆ 視聴履歴を取扱う匿名加工情報取扱事業者の義務については、改正放送分野ガイドライン(以下、「放送GL」という。)第29条の解説のとおり、「個人情報の保護に関する法律についてのガイドライン(匿名加工情報編)」のほか、個人情報保護委員会事務局レポート「匿名加工情報 パーソナルデータの利活用促進と消費者の信頼性確保の両立に向けて」(以下、「事務局レポート」という。)に準拠することとする。
- ◆ 視聴履歴の匿名加工情報の作成に当たって求められる加工について、以下の事項を明記。
 - ① 詳細な時刻情報を伴った視聴履歴は、他の情報と照合されることがあり、個人の特定につながるリスクがある。このため、詳細な時刻は丸めたり誤差を入れることが望ましい。【事務局レポート 4.1.5.2 (p27)】
 - ② 視聴履歴は個人の趣味・嗜好を表す傾向がある。一般的に趣味・嗜好から個人の特定に至ることは困難であるとしても、特に顕著な履歴に関しては適切な加工が望ましい。【事務局レポート 4.1.5.2 (p27)】
 - ③ 視聴履歴の期間が長いほど、その情報は一意となり得、その一意性から直ちに個人を特定することができないとしても個人が推測できる可能性がある。長期にわたる履歴を扱う場合、詳細な時刻や番組名などを適切に加工することが望ましい。【事務局レポート 4.2 (p28)】
- ◆ 視聴履歴を含む視聴関連個人情報を加工して匿名加工受信者情報を作成する場合について、一般的なデータの事例とユースケースと加工の方向性を記載する。具体的には、二次流通を前提としない契約者情報と視聴履歴が契約者IDで紐付けされている事例とそのユースケースを想定する。

<視聴履歴のユースケースにおけるデータのレイアウトイメージ>

受信者情報

契約者ID	機器ID	MACアドレス	氏名	性別	生年月日	電話番号	住所
53012602	XX-0001	XX-00-11-22-33-44	総務太郎	男性	1987年3月12日	03-222-XXXX	東京都 千代田区 霞ヶ関X-X-X
53597201	YY-0002	YY-11-22-33-44-55	情流花子	女性	1990年5月23日	090-444-YYYY	東京都 荒川区 荒川Y-Y-Y
81567824	ZZ-0003	ZZ-22-33-44-55-66	放政一郎	男性	1968年8月19日	03-123-ZZZZ	東京都 港区 六本木Z-Z-Z

システムの連結

視聴履歴

契約者ID	視聴日	視聴開始時刻	視聴終了時刻	視聴番組名	視聴チャンネル
53012602	2017年3月26日	20:16	20:45	サッカー・タイム	233
53012602	2017年3月28日	21:00	22:53	A刑事の事件簿	611
53012602	2017年4月2日	20:01	20:45	サッカー・タイム	233
53012602	2017年4月4日	21:05	22:45	A刑事の事件簿	611
53012602	2017年4月9日	05:20	05:29	〇〇教の時間	233

2. 視聴履歴のユースケースにおける加工の方向性

① 契約者情報に含まれる個人属性情報の加工の方向性【事務局レポート 4.1(p18)】

- ・ 契約者ID： 個人情報と視聴履歴を連結する符号に該当するため、仮IDに置き換える等によりすべてを削除する。
- ・ 機器ID/MACアドレス： 個人に密接し、かつ当該個人が容易に変更することができない外部から観察可能な符号であって、不変性が高いIDであるため(個人情報保護法施行規則第19条第5号)、再識別につながる可能性がある情報として削除する。
- ・ 生年月日： 少なくとも日については削除する一方で、同年同月に生まれた個人の人数が少ない場合は年月も削除の対象となる。曖昧化の程度については、住所の加工と合わせて検討する。
- ・ 電話番号： 多数の事業者で収集されており、個人の特定につながる可能性が高いため、削除する。ただし、住所の記述の曖昧化と平仄を揃える程度の情報であれば残すことは可能。
- ・ 住所： 個人の特定につながる可能性が高い情報であるが、有用性が高いことから、データセットの大きさや他の情報(生年月日等)の加工の程度を考慮して行う必要がある。町村以下の情報、郵便番号は原則として削除することが望ましい。

※ 上記のような、情報の項目それぞれについて一定程度曖昧化されるように部分的な削除や置換えを行う考え方のほか、住所・生年月日・性別等の複数の項目の組合せで一意にならないように各項目の加工レベルを調整する考え方も想定される。【事務局レポート4.1.1(p20)】

<データ項目毎のリスクと加工方法>

データ項目	想定されるリスク	望ましい加工方法の例
契約者ID	<ul style="list-style-type: none"> ・ 契約者を特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。(項目削除)
機器ID	<ul style="list-style-type: none"> ・ 契約者を特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。(項目削除)
MACアドレス	<ul style="list-style-type: none"> ・ 他の事業者でも収集している可能性があり、それと照合して個人が特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。(項目削除)
氏名	<ul style="list-style-type: none"> ・ 単体で個人を特定できる。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。(項目削除)
性別	<ul style="list-style-type: none"> ・ 他の情報との組合せによって、個人が特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 加工しない。他のデータ項目を加工することで個人が特定されるリスクを下げることで対応する。
生年月日	<ul style="list-style-type: none"> ・ 他の情報との組合せによって、個人が特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 20歳未満、20代、30代、40代、50代、60代、70歳以上の7つの年代に置き換える。(丸め)
電話番号	<ul style="list-style-type: none"> ・ 他の事業者でも収集している可能性があるため、他のデータと照合して個人が特定されるリスクがある。 ・ また、本人にアクセスされるリスクがある。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。(項目削除)
住所	<ul style="list-style-type: none"> ・ 他の情報との組合せによって、個人が特定されるリスクがある。 ・ 他の事業者でも収集している可能性があるため、他のデータと照合して個人が特定されるリスクがある。 ・ また、本人にアクセスされるリスクがある。 	<ul style="list-style-type: none"> ・ 市区郡単位より細かい情報を削除する。(丸め) ・ 郵便番号は、市区郡より細かい情報であるため削除する。(項目削除)

3-2 視聴履歴の加工の方向性 ②視聴履歴の加工（長期間の履歴）

② 視聴履歴の加工の方向性

- ◆ 視聴履歴の加工に当たっては、利用したいデータ項目や利用目的を踏まえ、適切な加工手法を選択する必要がある。

例 1)： 番組横断的に長期間連続する視聴履歴を匿名加工して提供する場合

例 2)： 短期間のみ、特定のチャンネルのみ、特定の番組のみの視聴履歴を匿名加工して提供する場合

1) 番組横断的に長期間連続する視聴履歴を匿名加工して提供する場合【事務局レポート 4. 1(p18)】

- ・ 契約者ID： 個人情報と視聴履歴を連結する符号に該当するため、仮IDに置き換える等により全てを削除する。
- ・ 視聴開始時刻/視聴終了時刻： 詳細な時刻情報を含むデータベースは、視聴番組や視聴チャンネルを表す情報とセットになることで、異なるデータセット間における共通の識別子として機能し得るため、削除する。
- ・ 番組名： 番組横断で蓄積された履歴は、一意性を持ったデータとなり得るため、番組のジャンルなどに置き換える。

<データ項目毎のリスクと加工方法>

データ項目	想定されるリスク	望ましい加工方法の例
契約者ID	<ul style="list-style-type: none"> ・ 契約者を特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。（項目削除）
氏名	<ul style="list-style-type: none"> ・ 単体で個人を特定できる。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。（項目削除）
視聴日	<ul style="list-style-type: none"> ・ 詳細な時刻情報は、他の情報と照合されることがあり、個人の特定につながるリスクがある。 ✓ 提供先が、一部重複する視聴履歴を保有している場合のほか、詳細な時刻情報が付与されている非特定視聴履歴を保有している場合等も、照合されるリスクが高まる。 	<ul style="list-style-type: none"> ・ 「視聴日」のみを残し、「視聴開始時刻」と「視聴終了時刻」を削除することで、個人が特定されるリスクを下げることで対応する（丸め）。
視聴開始時刻		<ul style="list-style-type: none"> ・ 削除する。（項目削除）
視聴終了時刻		<ul style="list-style-type: none"> ・ 削除する。（項目削除）
番組名		<ul style="list-style-type: none"> ・ 番組のジャンル等のより上位の概念に置き換える。（一般化）
視聴チャンネル	<ul style="list-style-type: none"> ・ 視聴履歴の期間が長いほど、その情報は一意となり得、その一意性から直ちに個人を特定することができないとしても配慮が望ましい 	<ul style="list-style-type: none"> ・ 加工しない。
録画再生回数	<ul style="list-style-type: none"> ・ 再生回数が極めて多いなどの特異な記述は特定の個人の識別又は下の個人情報の復元につながるリスクがある。 	<ul style="list-style-type: none"> ・ 再生回数を「〇回以上」という情報に置き換える。（トップコーディング）

3-2 視聴履歴の加工の方向性 ③視聴履歴の加工（限定された履歴）

2) 短期間のみ、特定のチャンネルのみ、特定の番組のみの視聴履歴を匿名加工して提供する場合【事務局レポート4.1(p18)、4.3(p31)】

- ・レコード一部抽出：複数データの連結による再識別ができないように、①短期間の番組横断的な視聴履歴については、各日毎に異なるサンプル、②特定チャンネルにかかる視聴履歴の場合、チャンネル毎に異なるサンプル、③特定番組のみの視聴履歴の場合、1番組（例えば1つのドラマの1話から最終話まで）を超えて作成する場合は番組毎に異なるサンプルを（確率的に）抽出する。
- ・契約者ID：個人情報と視聴履歴を連結する符号に該当するため、仮IDに置き換える等により全てを削除する。
- ・視聴開始時刻/視聴終了時刻：詳細な時刻情報を含むデータベースは、視聴番組や視聴チャンネルを表す情報とセットになることで、異なるデータセット間における共通の識別子として機能し得るため、①加工対象とする期間を十分な短期間のみ限定する、②特定のチャンネルのみに限定する、③特定の番組に限定することにより、個人が特定されるリスクを下げるのが望ましい。
さらに、具体的な加工の手法として、丸め（ラウンディング）や、ノイズ（誤差）付加を行う。
- ・チャンネル名、番組名：当該番組を視聴した者が極めて少ないチャンネルや番組については、匿名加工情報の利用目的を考慮して適切な情報の粒度とすることや、匿名加工情報の作成の是非を検討する。また、録画した放送番組の一定期間における再生回数が極めて高いなど、その視聴の希少性の高さ等から個人の特定につながる可能性があるものは、再生回数の曖昧化を行うことが望ましい。

<データ項目毎のリスクと加工方法>

データ項目	想定されるリスク	望ましい加工方法の例
契約者ID	<ul style="list-style-type: none"> ・ 契約者を特定されるリスクがある。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。（項目削除）
氏名	<ul style="list-style-type: none"> ・ 単体で個人を特定できる。 	<ul style="list-style-type: none"> ・ 全て削除する、又は仮IDに置き換える。（項目削除）
視聴日	<ul style="list-style-type: none"> ・ 詳細な時刻情報は、他の情報と照合されることがあり、個人の特定につながるリスクがある。 <ul style="list-style-type: none"> ✓ 提供先が、一部重複する視聴履歴を保有している場合のほか、詳細な時刻情報が付与されている非特定視聴履歴を保有している場合等も、照合されるリスクが高まる。 	<ul style="list-style-type: none"> ・ 加工しない。短期間のみ、特定のチャンネルのみ、特定の番組のみに限定することで、個人が特定されるリスクを下げることで対応する。
視聴開始時刻		<ul style="list-style-type: none"> ・ 丸める（ラウンディング）、又は、一定の分布に従った乱数的な数値等を付加するノイズ（誤差）付加を行うことで対応する。
視聴終了時刻		<ul style="list-style-type: none"> ・ 同上
番組名		<ul style="list-style-type: none"> ・ 加工しない。
視聴チャンネル		<ul style="list-style-type: none"> ・ 加工しない。
録画再生回数		<ul style="list-style-type: none"> ・ 再生回数が極めて多いなどの特異な記述は特定の個人の識別又は下の個人情報の復元につながるリスクがある。

3. 視聴履歴の匿名加工情報の提供について

◆ 視聴履歴を加工した匿名加工情報の提供について、以下の事項を明記。

- ① 同一の個人の蓄積された視聴履歴は、一意性を持ったデータとなり得るため、それを識別子にして識別につながることもあり得ることから、詳細な時刻や番組名などを適切に加工するとともに、同一の事業者に視聴履歴の匿名加工情報を提供する場合は、都度、異なるサンプルを抽出する、視聴データの期間が重複しないように提供する等の配慮が必要である。【事務局レポート 4. 2(p28)】
- ② 視聴履歴を匿名加工情報にする際には、匿名加工受信者情報のデータの流通範囲を検討することが望ましい。一次流通のみに限定し、二次流通を禁止する等して特定の事業者¹に提供する場合(スライド4、5で検討した加工の方向性)に比べて、転々流通を容認する場合には、より強度の曖昧化を検討する必要がある。【事務局レポート 4. 2(p28)】

4. 非特定視聴履歴の取扱いについて

◆ 非特定視聴履歴は個人情報に関する義務は生じないが、特定個人の識別リスクの観点から、以下の事項を明記。

- ① 非特定視聴履歴は、視聴履歴を有する事業者等に対して提供される場合、特定の個人が識別されてしまう可能性があることに留意しなければならない。
- ② 視聴履歴を有する事業者等において、あらかじめ非特定視聴履歴の取得を通じた個人情報の取得に同意を得ていない限り、非特定視聴履歴を取得し、特定の個人を識別、又は、容易に照合できる状態にすることは、個人情報の不適正な取得となり、削除が必要となるので留意しなければならない。【GL第7条第1項、第35条第1項】

※事前の同意については、利用規約などによる包括同意でも構わないが、本人が同意に係る判断を行うために必要な情報として、「〇〇社のテレビ受信機により取得される視聴データ」等、社名と製品名を明示しなければならない。【GL第35条第1項解説】