

# 第5回 国税庁保有行政記録情報の 整備に関する有識者検討会

国税庁 企画課データ活用推進室

# 資料内容

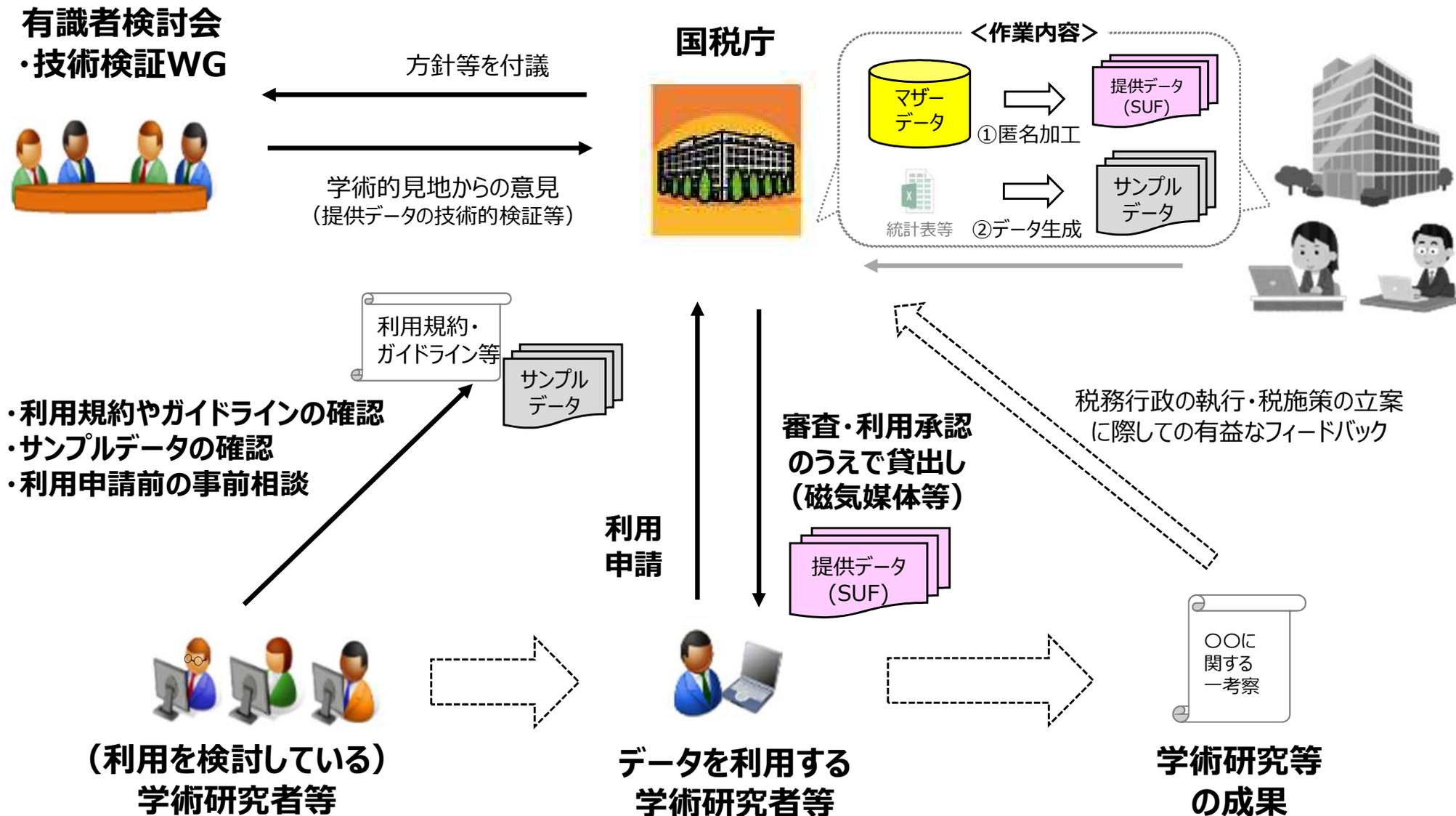
1. これまでの議論

2. 今後の方向性(案)

3. 今後のスケジュール

# 1. これまでの議論（匿名データ提供に関する全体像（イメージ））

- 学術研究での利用を前提に、利用者は事前に利用規約等・サンプルデータを確認し、利用申請が承認されれば、データを貸し出すことにより、提供。
- データ作成に当たっては、他の情報と照合しても個人等が特定されないよう**匿名加工**を実施。



# 1. これまでの議論（第9回技術検証WG（技術的課題））

## <第9回技術検証WG（技術的課題）（令和6年9月18日開催）の議事要旨>

### ○ 匿名データ（案）の作成と安全性・有用性の評価

- ・ 安全性については十分確保されており、匿名データと元データにおける照合・対応付けの可能性の観点からも問題はないと考える。
- ・ 匿名加工の処理を施しても、内訳から合計が類推されるレコード等については、有用性を保つために、レコード自体を削除するのではなく、内訳を削除して合計を残すという方法を取った方がよい。
- ・ 類推チェックのパターンについて、精査する必要があると考える。

### ○ 匿名データのグルーピングの一部修正等

- ・ パネル化及びグルーピングの一部修正については、安全性の観点から非常に難しく、議論にも時間を要す。また、税務大と共同研究の機会もあることから、将来の課題とするべきであると考える。
- ・ トップコーディングされたデータの統計数量を提供した方がよい。まずは、平均値、標準偏差を提供することでよい。

### ○ 利用申出・研究等の成果の公表に係る審査基準

- ・ 研究成果物の審査は、10の集計単位等を基準として行うことで問題ないと考える。実際の運用に当たっては、利用前の段階で、審査の際に問題が生じないよう審査基準等を説明しておくことが重要である。
- ・ 研究者は、早期に研究成果を出すことが求められているため、実際の運用を意識して、審査作業の体制を検討しておく方がよい。

### ○ 新規年分の提供時期

- ・ 新規年分については5年分まとめた提供で問題ないと考える。まずは、安全に運用し、おって提供早期化の議論を行えばよいと考える。

## 2. 今後の方向性（案）

### ◆ 匿名データに係る検討状況（1／2）

項目	検討結果
①利用者範囲	<ul style="list-style-type: none"><li>• <b>大学や公的機関等に所属する常勤の研究者</b> ※上記研究者の指導監督の下、博士研究員、大学院生（ただし、博士課程に限る）も補助者として利用可（計5名まで）</li></ul>
②利用目的	<ul style="list-style-type: none"><li>• <b>税・財政施策の改善・充実に係る統計的研究</b>に限る</li></ul>
③利用期間	<ul style="list-style-type: none"><li>• <b>原則2年間</b>（1年間の延長が可）</li></ul>
④対象データ	<ul style="list-style-type: none"><li>• <b>所得税申告書第一表、第三表に係る各項目</b>（提供開始時の年分は平成26～30年）</li></ul>
⑤提供方法	<ul style="list-style-type: none"><li>• CD-R等による<b>貸出し</b></li></ul>
⑥個人情報保護法上の取扱い	<ul style="list-style-type: none"><li>• 学術研究を目的とした<b>保有個人情報の提供</b>（個情法69②四）</li></ul>
⑦情報公開法上の不開示情報の該当性	<ul style="list-style-type: none"><li>• 目的外の利用をもって公にした場合や「適正管理措置」を講じないまま公にした場合、提供事務の遂行に支障を及ぼすおそれがある（情公法5六）ことから、<b>不開示情報に該当</b></li></ul>
⑧匿名加工の考え方	<ul style="list-style-type: none"><li>• <b>他の情報と照合しても個人等が特定されない匿名加工情報</b>（※）とする ※社会通念上、一般人の判断力や理解力をもって、具体的な人物と情報との間に同一性を認められない程度に匿名加工された情報</li></ul>

※これまでの議論を基に、今後の方向性を整理したもの（次ページも同様）。

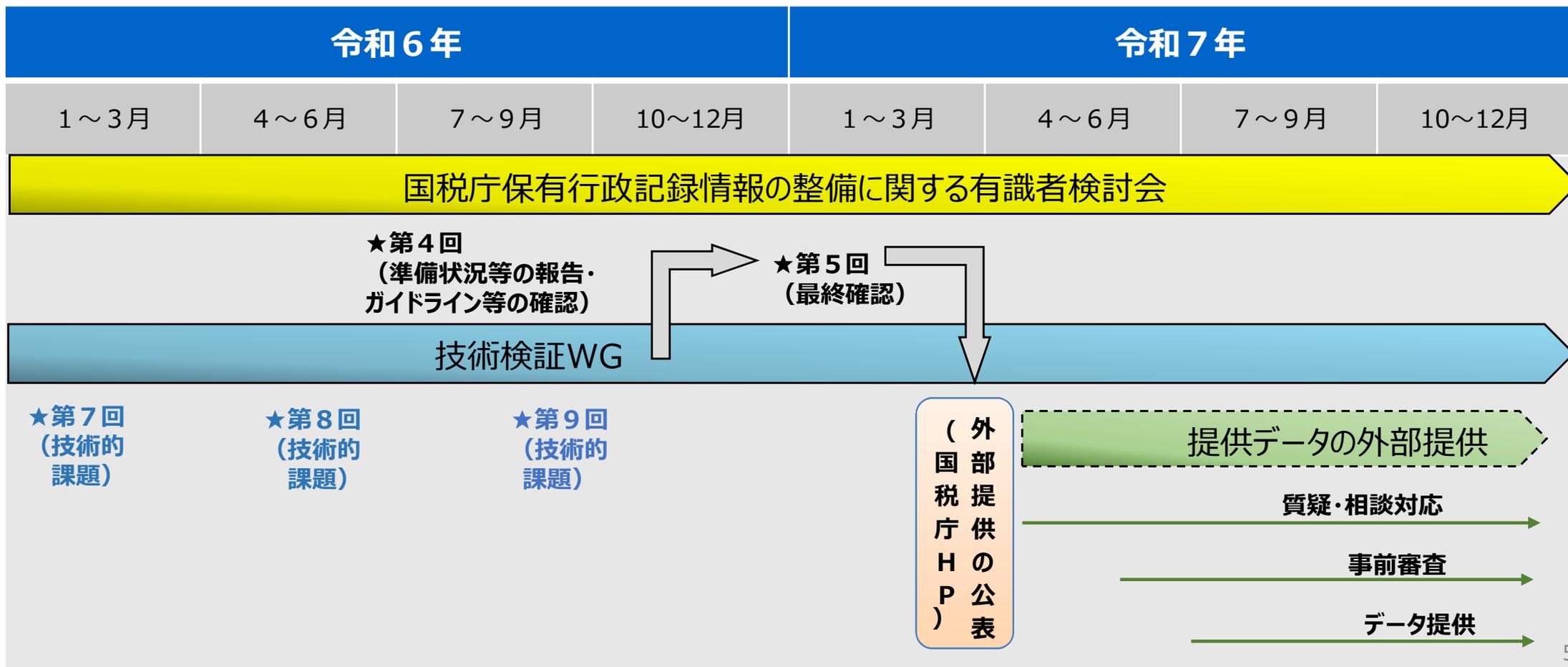
## 2. 今後の方向性（案）

### ◆ 匿名データに係る検討状況（2／2）

項目	検討結果				
⑨匿名加工の内容	<ul style="list-style-type: none"> <li>全データのうち約<b>1%をサンプリング</b>として抽出するほか、以下の加工を行う ※複数回申告を行っている者については<b>最終申告分のみをレコードとして残す</b></li> </ul> <table border="1" data-bbox="394 419 2188 815"> <tr> <td data-bbox="394 419 651 571">定性情報</td> <td data-bbox="651 419 2188 571"> <ul style="list-style-type: none"> <li><b>名前、マイナンバー、整理番号、性別情報は削除</b></li> <li><b>住所は2区分</b>（三大都市圏or否）、<b>生年月日は5歳階級、業種は12区分</b>でグルーピング</li> </ul> </td> </tr> <tr> <td data-bbox="394 571 651 815">定量情報</td> <td data-bbox="651 571 2188 815"> <ul style="list-style-type: none"> <li><b>収入金額・所得金額等の定量情報は上位0.5%等を基準としてトップ（ボトム）コーディング</b> ※コーディング対象分について、<b>平均・標準偏差の統計量を提供</b></li> <li>出現率の低い項目（0.5%未満）を含むレコードや、内訳等により上記<b>コーディング前の数値が類推されるレコードは、合計（総所得金額・申告納税額等）に相当する項目以外を削除</b></li> <li>桁数に応じて<b>ラウンディング（切捨て）処理</b></li> </ul> </td> </tr> </table>	定性情報	<ul style="list-style-type: none"> <li><b>名前、マイナンバー、整理番号、性別情報は削除</b></li> <li><b>住所は2区分</b>（三大都市圏or否）、<b>生年月日は5歳階級、業種は12区分</b>でグルーピング</li> </ul>	定量情報	<ul style="list-style-type: none"> <li><b>収入金額・所得金額等の定量情報は上位0.5%等を基準としてトップ（ボトム）コーディング</b> ※コーディング対象分について、<b>平均・標準偏差の統計量を提供</b></li> <li>出現率の低い項目（0.5%未満）を含むレコードや、内訳等により上記<b>コーディング前の数値が類推されるレコードは、合計（総所得金額・申告納税額等）に相当する項目以外を削除</b></li> <li>桁数に応じて<b>ラウンディング（切捨て）処理</b></li> </ul>
定性情報	<ul style="list-style-type: none"> <li><b>名前、マイナンバー、整理番号、性別情報は削除</b></li> <li><b>住所は2区分</b>（三大都市圏or否）、<b>生年月日は5歳階級、業種は12区分</b>でグルーピング</li> </ul>				
定量情報	<ul style="list-style-type: none"> <li><b>収入金額・所得金額等の定量情報は上位0.5%等を基準としてトップ（ボトム）コーディング</b> ※コーディング対象分について、<b>平均・標準偏差の統計量を提供</b></li> <li>出現率の低い項目（0.5%未満）を含むレコードや、内訳等により上記<b>コーディング前の数値が類推されるレコードは、合計（総所得金額・申告納税額等）に相当する項目以外を削除</b></li> <li>桁数に応じて<b>ラウンディング（切捨て）処理</b></li> </ul>				
⑩利用申出の審査	<ul style="list-style-type: none"> <li>利用者が国税庁に利用の申出を行った後、<b>国税庁において研究目的・内容の妥当性や、実行可能性等を審査した上で、匿名データの提供を行う</b></li> </ul>				
⑪研究成果物の審査	<ul style="list-style-type: none"> <li><b>成果物の公表前に、国税庁において学術研究に利用されたか、研究成果に個体を識別できる情報が含まれていないか等を審査</b></li> <li><b>10の集計単位等を基準として、個体を識別できる情報が含まれていないかを確認</b></li> </ul>				
⑫利用者による管理	<ul style="list-style-type: none"> <li><b>匿名データは、施錠可能な物理的な場所（国内）に限定して利用可能であり、外部ネットワークから遮断された環境でのみ利用可能</b></li> <li><b>管理簿等を提出させた上で、必要に応じて国税庁が監査することが可能</b></li> </ul>				
⑬利用者に対する守秘義務	<ul style="list-style-type: none"> <li>利用者に匿名データを利用させる際、当該データの開示や漏えい、目的外の利用が無いよう、<b>利用者に誓約書を提出させ、守秘義務を負わせる</b></li> <li>情報漏えいなど<b>契約違反があった場合の措置も併せて規定</b>（違反者及び代表者等の氏名等を公表、今後の匿名データの利用禁止など）</li> </ul>				
⑭新規年分の追加	<ul style="list-style-type: none"> <li>提供開始後は、5年後に、5年分をまとめて提供</li> <li>おって、提供早期化について検討</li> </ul>				

### 3. 今後のスケジュール

- 令和6年度中に、国税庁HPにおいて外部提供に係る詳細を公表し、令和7年度から外部提供を開始予定。
- データの外部提供後は、当面は提供体制の整備・充実化を図ることとし、その後、提供対象税目（所得税以外の税目）の拡大等も併せて検討する。



## 参考

# ガイドライン(概要)

## 目的

- 匿名データの利用手続や利用申出に係る審査基準及び必要な事務処理等を明確化し、国税庁がこれらの事務を適切かつ円滑に実施できるようにすることを目的として定めるもの（ガイドライン第1）

## 個票データ等の利用に際しての基本原則

- 国税庁は、学術研究の発展に資するものであって、税・財政施策の改善・充実に関する統計的研究であることを確認する（ガイドライン第3の1（1））
- 利用者は、税務データに関する秘密の保護及び公の秩序又は善良の風俗に反しないように十分配慮する必要がある、個体を識別するために当該データを他の情報と照合してはならない（ガイドライン第3の1（2））
- 匿名データの秘密の保護及び適正管理の確保に当たっては、ガイドライン及び利用規約に従うものとする。また、匿名データの利用は、適正管理措置が講じられた場合においてのみ可能とする（ガイドライン第3の2）

## 匿名データの利用期間

- 利用期間は、原則として2年間を上限とする（ガイドライン第4）
- 管理責任者がやむを得ない理由により利用期間の延長を希望する場合、国税庁は、最長1年間を上限として、原則1回に限り延長を認めることができる（ガイドライン第9の3）

# ガイドライン(概要)

## 利用申出手続

- 申出者は申出書の記入例を参照して、仮の申出書を作成した上で提出し、国税庁はそれを基に必要な助言を行う。仮の申出書に基づき、申出内容の検討・調整を行った後、正式な申出書を提出することとする（ガイドライン第5の1）
- 匿名データの利用は契約に基づくものであり、行政不服審査法の対象外である（ガイドライン第5の2）

## 利用申出に対する審査・決定

- 匿名データの利用申出に関する審査は、ガイドラインで定める審査基準に基づき、国税庁において行う（ガイドライン第6）

## 利用後の措置等

- 利用者は、申出書の利用期間に記載された返却期限までに、中間生成物等について、復元できないように消去又は適切に破棄した上で、匿名データの提供媒体を国税庁に返却し、併せて、データ措置報告書を用いて報告しなければならない（ガイドライン第10の1）

# 利用規約(概要)

## 目的

- 匿名データの利用に関する利用者と国税庁の契約の内容及び本契約に関連する申出者又は利用者との基本的事項を定めるもの（第1条）

## 検査等

- 国税庁が匿名データの利用状況及び管理状況について利用者に対して検査を行う場合、利用者は、これを拒まないものとする（第10条）

## 利用者の保証等

- 利用者は、申出書、報告書その他匿名データの利用に当たって国税庁に提出した書類の記載内容を確認し、その内容が真実であることを表明し、保証する（第14条）

## 契約に違反した場合の措置

- 国税庁は、利用者が本契約に違反し、又は、措置要件に規定する行為を行ったと認められた場合は、本契約の解除の有無にかかわらず、事前に当該利用者から漏えい等に至る経緯及び意見を聴取した上、その内容に応じた措置をとることができる（第19条）

# 様式一覧

手続き	様式		ガイドライン上の記載箇所
①申出時	様式 1	研究用匿名データの利用に関する申出書	第 3 の 2 (1)①、第 5 の 7 (1)(2)
	様式 2	職務経歴書	第 5 の 7 (2)①、第 6 の 2 (5)①
	様式 3	研究計画書	第 5 の 7 (2)②、第 6 の 2 (5)②
②承諾後	様式 4	研究用匿名データの利用に関する誓約書	第 8 の 1 (2)
	様式 5	研究用匿名データの依頼書	第 8 の 2
	様式 6	研究用匿名データの受領書	第 8 の 4、6 (1)(2)
③利用時	様式 7	研究用匿名データに係る管理簿	第 8 の 7 (2)①②、8 (1)①、(5)②、9、第10の 2
	様式 8	研究用匿名データ管理状況報告書	第 8 の 7 (2)②
④変更時	様式 9	所属等変更届出書	第 9 の 1 (1)、2 (1)、3 (1)
	様式10	記載事項変更依頼申出書	第 9 の 1 (2)、2 (2)、3 (1)(2)
⑤終了後	様式11	データ措置報告書	第10の 1
	様式12	利用実績報告書	第10の 2・3

# 利用の手引き(概要)

税務データを使った研究を検討している方に、データの構造や、申出に当たっての要件、必要な手続き等について、あらかじめ知っておいていただきたいことを簡潔にまとめたもの。

本手引きは、国税庁保有研究用匿名データの利用を検討するにあたりご参照いただく目的で作成されたものです。  
利用の申出を行う際は、国税庁保有研究用匿名行政記録情報の利用に係るガイドラインもあわせてご参照ください。

## 国税庁保有研究用匿名行政記録情報 利用の手引き

国税庁長官官房企画課 データ活用推進室

### <目次>

- データ構造・詳細
- 提供方法
- 利用要件
- 利用手続
- 利用審査
- 利用開始から終了の流れ