

# 第6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. 匿名データ提供に係る全体像

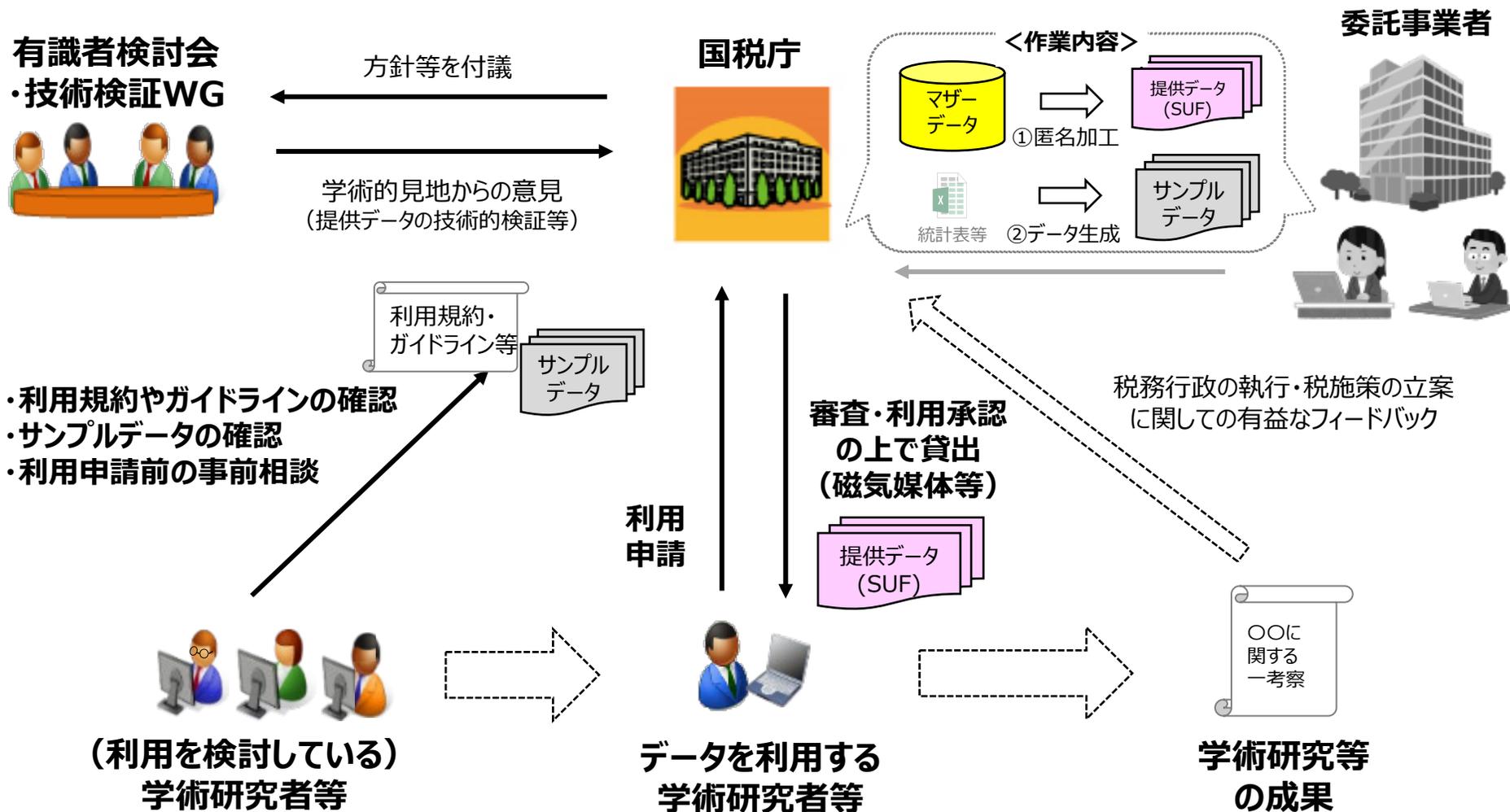
2. 匿名データに対する意見

3. 匿名データ提供の見直し(案)

4. 今後のスケジュール

# 1. 匿名データ提供に係る全体像

- 令和7年4月より、まずは所得税について、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施した匿名データの提供を開始。
- 匿名データは、学術研究での利用を前提に、利用申請・承認のプロセスを経た上で、データを貸し出すことにより、提供。



## 2. 匿名データに対する意見

- 匿名データについて、有識者等からいただいたご意見は以下のとおり。

項番	改善点	内容
①	地域情報の提供	<b>地方財政を研究するには、細かい地域情報が必要</b> となるため、より詳細な地域情報を提供してほしい。
②	データのパネル化	<b>景気の変動などをより詳細に分析するには、個人のデータを年分ごとに比較できることが必要</b> となるため、データのパネル化を実施してほしい。
③	内訳削除処理の見直し	<b>分析に制限が生じるおそれ</b> があるので、当該処理を実施せずに提供してほしい。
④	トップ（ボトム）コーディング基準値の見直し	トップコーディング基準値を引き上げてほしい。例えば総所得金額だと3,000万円が基準値となっているが、 <b>高所得者層の情報を提供してほしい</b> ので、この基準値をもう少し引き上げることはできないか。
⑤	年齢情報の細分化	現状年齢情報は5歳刻みで提供されているが、 <b>1歳刻みで提供できないか</b> 。

## 2. 匿名データに対する意見

項番	改善点	内容
⑥	データ提供方法の見直し	利便性向上のため、提供方法を現状の磁気媒体（CD-R）を使った方法から、 <u>クラウドサービスを使用する方法に変更できないか。</u>
⑦	申込様式の簡略化	利用申請に必要な書類の様式について、 <u>記載すべき分量が多いので作成手続きを簡略化</u> できないか。
⑧	利用目的の拡充	利用目的の「 <u>税財政政策に関する</u> 」という制限について <u>緩和</u> できないか。 <u>教育目的</u> にも利用できないか。
⑨	周知・広報の実施	<u>記述統計を用いた分析が可能であることをより積極的にPR</u> してはどうか。

## 2. 匿名データに対する意見

### <第10回技術検証WG（技術的課題）（令和7年11月4日開催）の議事要旨>

匿名データに対する有識者等からいただいたご意見を踏まえ、意見交換を実施。税務データの匿名加工について以下の方向性で作業を進めることが適当との意見をいただいた。

#### ○ 地域情報の提供

- ・ 住所の粒度は、47都道府県とする
- ・ 特定リスクの高いレコードは、平均値に置き換えて提供することとし、当該平均値は、局所平均を取って提供する
- ・ 平均値に置き換えたレコードを判別できるようにフラグをつける
- ・ 抽出率は1%とする
- ・ 準識別子（年齢・住所・業種）のクロスごとの平均値・標準偏差を統計量として別途提供する
- ・ 上記を前提とした匿名データは、現行の匿名データから置き換えて提供する

#### ○ データのパネル化

- ・ まずは、安全性が高いと考えられる以下の方法により提供を開始し、提供後のニーズを踏まえ、必要に応じて有用性の向上を検討する
- ・ パネル化の情報範囲については、準識別子を年齢に絞った上で提供する
- ・ パネル化の年数は、連続した2年間とする
- ・ 連続した2年間のデータをパネルデータと捉え、1%標本抽出する
- ・ 年齢情報のグルーピング、定量情報のトップコーディングやラウンディング、内訳削除の処理を継続するが、コーディングについては、安全性の観点から0.5%から1%に基準を引き上げる
- ・ 都道府県レベルで提供する各年データとパネルデータは、分けて提供することとし、データのリンケージは認めないこととする

#### ○ その他

- ・ 同一のレコードが2年間のパネルデータに重複してサンプリングされても安全性には問題はない。サンプリングする際は、データの分布に歪みが生じないように留意すべき
- ・ パネルデータの統計的マッチング（Statistical Matching）は、個人を特定するためのリンケージには当たらないため、認めてもよいのではないか
- ・ 利用申請に必要な書類の様式について、作成手続きを簡略化するため、リサーチマップを利用する場合には、記載内容が正確であるかどうかを確認した方がよい
- ・ 利用目的の「税財政政策に関する」という制限については、将来的に緩和する方向で検討してはどうか

### 3. 匿名データ提供の見直し（案）

#### ① 現行の匿名データの変更

- 住所情報はこれまで三大都市圏か否かの2区分であったところ、47都道府県単位に変更
- 個人の特定リスクが高いレコードについては、別途の匿名化措置（平均値等への置換）を実施
- 出現頻度の低い項目やコーディング前の数値が類推可能な項目がある場合には、合計項目を残すのではなく、内訳項目を残す

#### ② パネルデータの試行提供

- 現行の匿名データとは別に、2年間のパネルデータを作成
- 試行的な提供と位置づけ、利便性向上に係る意見をいただきながらブラッシュアップを図る

##### 【現行の匿名データとの主な違い】

- 住所・業種情報は削除
- トップ（ボトム）コーディングの基準を上位（下位）1%

#### ③ 申込様式の簡略化

- これまで提出必須としていた職務経歴書について、researchmapで確認可能な場合には省略可とする等を見直し

### 3. 匿名データ提供の見直し（案）（匿名化の処理方法）

#### 単年データ

データ全体

- 全データのうち約**1%をサンプリング**（複数回申告を行っている者については**最終申告分のみをレコードとして残す**）
- 住所・生年月日・業種（準識別子）の組合せの件数が少ない場合には、**定性情報は最頻値、定量情報は平均値に置き換える**  
※サンプリングされたレコードと同じ準識別子の組合せである集団から、総所得金額の近いレコードを10件程度取り出したものを用いて最頻値・平均値を求める
- 住所・生年月日・業種（準識別子）の組合せごとに、定性情報の最頻値・定量情報の平均値を別途統計量として提供

定性情報

- **名前、マイナンバー、整理番号、性別情報は削除**
- **住所は47区分**（都道府県単位）、**生年月日は5歳階級**、**業種は12区分**でグルーピング

定量情報

- **収入金額・所得金額等の定量情報は上位0.5%等をトップ（ボトム）コーディング**  
※コーディング対象分について、**平均・標準偏差の統計量を提供**
- 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**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）に相当する項目を削除**
- 桁数に応じて**ラウンディング（切捨て）処理**

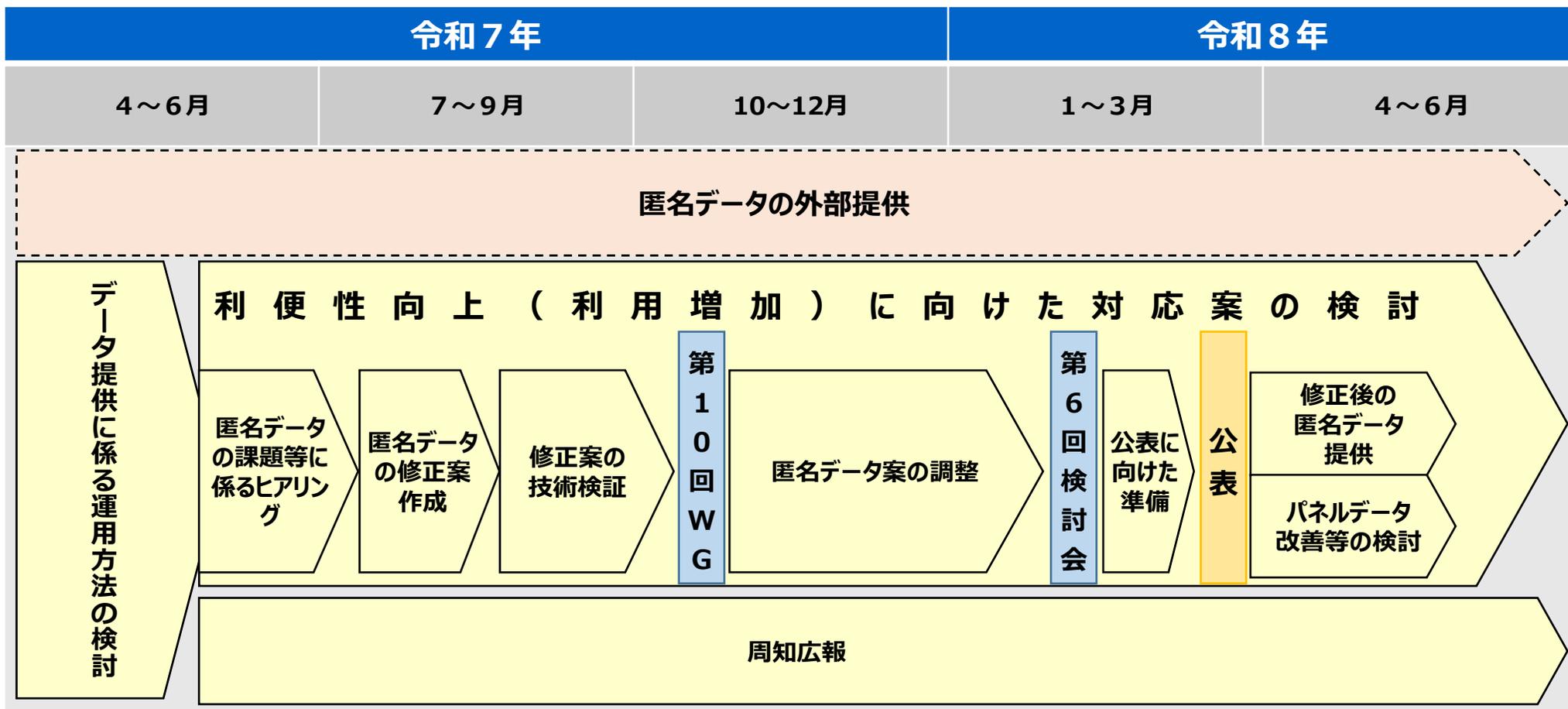
### 3. 匿名データ提供の見直し（案）（匿名化の処理方法）

#### パネルデータ

- | パネルデータ |   |
|--------|---|
| データ全体  | <ul style="list-style-type: none"><li>2年間のパネルデータ</li><li>全データのうち約<u>1%をサンプリング</u>（複数回申告を行っている者については<u>最終申告分のみをレコードとして残す</u>）</li></ul>  |
| 定性情報   | <ul style="list-style-type: none"><li><u>名前、マイナンバー、整理番号、性別情報、住所、業種は削除</u></li><li><u>生年月日は5歳階級でグルーピング</u></li></ul>   |
| 定量情報   | <ul style="list-style-type: none"><li><u>収入金額・所得金額等の定量情報は上位1%等をトップ（ボトム）コーディング</u><br/>※コーディング対象分について、<u>平均・標準偏差の統計量を提供</u></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<u>コーディング前の</u></li><li><u>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）に相当する項目を削除</u></li><li>桁数に応じて<u>ラウンディング（切捨て）処理</u></li></ul> |

## 4. 今後のスケジュール

- 令和8年4月を目途として、見直し後の匿名データ提供開始を目指す。
- 上記の提供開始後は、パネルデータに対する利便性向上に係る意見をいただき、パネルデータのブラッシュアップを図るほか、教育利用や提供対象税目（所得税以外の税目）の拡大を検討する。



# 參考資料

# 1. ガイドライン及び様式の改正について

○ 申込様式の簡略化に係るご意見を踏まえ、以下①～③の事項についてガイドライン・様式の見直しを行いたい

- ① researchmapの会員IDを利用することによる職務経歴書の省略
- ② 外部研究資金獲得状況の記載省略
- ③ 研究計画書の文字数制限(2,000～3,000字)の撤廃

※ 改正部分はアンダーラインの箇所である。

項番	改正前	改正後
①	第2 用語の定義	第2 用語の定義 <u>8 researchmap</u> 本ガイドラインにおいて「researchmap」とは、国立研究開発法人科学技術振興機構（JST）が運営しているデータベース型研究者総覧をいう。  <u>9 researchmap会員ID</u> 本ガイドラインにおいて「researchmap会員ID」とは、「researchmap」に登録する際に設定するIDをいう。

# 1. ガイドライン及び様式の改正について

## <ガイドライン>

※ 改正部分はアンダーラインの箇所である。

項番	改正前	改正後
①②	<p>第5 研究用匿名データの利用申出手続 7 申出書の記載事項及び添付書類 代表者は、国税庁が別に定める様式に従い次の(1)①から⑪までの事項を記載した申出書を作成し提出するものとする。また、申出書には、次の(2)①から⑥までの事項を記載した書類を添付するものとする。</p> <p>(1) 研究用匿名データの利用に関する申出書 ① 各申出者及び補助者の氏名、所属機関名、職名、住所、連絡先（電話番号、電子メールアドレスを含む。以下同じ。）、補助者においては監督を行う申出者の氏名並びに申出者のうち代表者及び管理責任者の表示</p> <p>② 各申出者及び補助者の所属機関の所在地、連絡先 ③ 各申出者が、申出時点から起算して過去3年以内に獲得した外部研究資金の制度名、採択年度及び研究課題名</p> <p>④ 各申出者の研究等の分野での過去の実績（公的統計の調査票情報を含むマイクロデータの利用経験の有無を含む） ⑤ 各申出者及び補助者の研究用匿名データの利用場所及び保管場所 ⑥ 各申出者及び補助者の本ガイドライン、利用規約の了承の有無 ⑦ 研究の名称 ⑧ 研究の概要（研究の内容、研究用匿名データの利用方法、作成する資料等の内容等） ⑨ 研究の成果の公表時期及び方法並びに発表予定の内容 ⑩ 利用を希望する研究用匿名データの種類、年分及び利用期間 ⑪ その他必要な事項</p>	<p>第5 研究用匿名データの利用申出手続 7 申出書の記載事項及び添付書類 代表者は、国税庁が別に定める様式に従い次の(1)①から⑪までの事項を記載した申出書を作成し提出するものとする。また、申出書には、次の(2)①から⑥までの事項を記載した書類を添付するものとする。<u>ただし、申出書にresearchmap会員IDを記載した場合は、職務経歴書の提出を省略しても差し支えないが、researchmapの登録情報が最新情報に更新されていることを確認すること。</u></p> <p>(1) 研究用匿名データの利用に関する申出書 ① 各申出者及び補助者の氏名、<u>通称名、生年月日</u>、所属機関名、職名、住所、連絡先（電話番号、電子メールアドレスを含む。以下同じ。）、補助者においては監督を行う申出者の氏名<u>及び補助者の職名</u>並びに申出者のうち代表者及び管理責任者の表示</p> <p>② 各申出者及び補助者の所属機関の所在地、連絡先 <u>（削除）</u></p> <p>③ <u>各申出者の賞罰</u> ④ 各申出者の研究等の分野での過去の実績（公的統計の調査票情報を含むマイクロデータの利用経験の有無を含む） ⑤ 各申出者及び補助者の研究用匿名データの利用場所及び保管場所 ⑥ 各申出者及び補助者の本ガイドライン、利用規約の了承の有無 ⑦ 研究の名称 ⑧ 研究の概要（研究の内容、研究用匿名データの利用方法、作成する資料等の内容等） ⑨ 研究の成果の公表時期及び方法並びに発表予定の内容 ⑩ 利用を希望する研究用匿名データの種類、年分及び利用期間 ⑪ その他必要な事項</p>

# 1. ガイドライン及び様式の改正について

## <様式>

項番

改正前

改正後

様式 1

### 研究用匿名データの利用に関する申出書

以下、①から⑩については、各申出者が記載し、職務経歴書（様式2）を添付して提出すること。

①利用者  代表者、 申出者、 補助者

(氏名)

(所属機関名・職名)

(住所) 〒

(電話番号)

(電子メールアドレス)

(補助者の監督を行う申出者の氏名)

⑧研究用匿名データの利用場所（保管場所を含む。）

※レイアウト図を添付してください。

⑨代表者の管理責任  代表者が利用開始後に管理責任者も兼ねることに了承します。

⑩ガイドライン等の了承の有無  本申出書は、研究用匿名データの利用に係るガイドライン、研究用匿名データの利用規約及び研究用匿名データの利用に関し国税庁が事前に説明・確認した内容（ウェブサイト等で周知した内容を含む。）を了承した上で提出するものです。

⑪所属機関

(所属機関名)

(所在地) 〒

(電話番号)

(電子メールアドレス)

(代表者又は管理者の氏名)

⑫申出日から過去3年以内に獲得した外部研究資金の制度名、採択年度、研究課題名

※公的機関によるものだけでなく、獲得している外部研究資金を全て記載すること。

制度名	採択年度	研究課題名
<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>

⑬申出者の本申出書に記載された研究等の分野での過去の実績

※公的統計の調査票情報を含むマイクロデータの利用経験の有無についても記載すること。

様式 1

### 研究用匿名データの利用に関する申出書

以下、①から⑩については、各申出者が記載し、職務経歴書（様式2）を添付して提出すること。

※research\_chose 会員ID を記載した場合、職務経歴書（様式2）の提出を省略しても差し支えない。

①利用者  代表者、 申出者、 補助者

(氏名)

(通称名)

(生年月日)  年 月 日 ( 歳)

(research\_chose 会員 ID)

本申出書は research\_chose の登録情報が申出書提出日において最新情報に更新されていることを確認した上で提出するものです。

(所属機関名・職名)

(自宅住所) 〒

(電話番号) (自宅)

(携帯)

(電子メールアドレス)

住所 〒

氏名

利用者との関係

電話番号 (自宅)

(携帯)

(補助者の監督を行う申出者の氏名)

(補助者に該当する場合はチェック)  博士研究員（ポスドク）  大学院生（博士課程後期相当）

⑧研究用匿名データの利用場所（保管場所を含む。）

※レイアウト図を添付してください。

⑨代表者の管理責任  代表者が利用開始後に管理責任者も兼ねることに了承します。

⑩ガイドライン等の了承の有無  本申出書は、研究用匿名データの利用に係るガイドライン、研究用匿名データの利用規約及び研究用匿名データの利用に関し国税庁が事前に説明・確認した内容（ウェブサイト等で周知した内容を含む。）を了承した上で提出するものです。

⑪所属機関

(所属機関名)

(所在地) 〒

(電話番号)

(電子メールアドレス)

(代表者又は管理者の氏名)

⑬賞罰

種類	備考
年 月 日	<input type="text"/>

⑭自由記載欄

※通常の研究業績、公的統計の調査票情報を含むマイクロデータの活用経験があれば記載してください。

①②



# 1. ガイドライン及び様式の改正について

## <様式>

項番

改正前

改正後

①

年 月～ 年 月		
年 月～ 年 月		
年 月～ 年 月		
年 月～ 現 在		

所属学会及び社会における活動等	
現在所属している学会名	
年 月	
年 月	
年 月	
年 月	
年 月	

賞 罰		
年 月 日	種 類	備 考
年 月 日		
年 月 日		
年 月 日		
年 月 日		

現在の職務の状況					
勤務先	職名	学部等又は 所属部局の名 称	勤務状況 (担当科目等)	常勤又は 非常勤の別	備考

# 1. ガイドライン及び様式の改正について

## <様式>

項番

改正前

改正後

研究業績について				
著書、学術論文等の名称	単著 共著 の別	発行又は 発表の年月日	発行所、発表雑誌 等 又は発表学会等の 名称、掲載頁	概要
1. 著書（学術書、教科書等）				
2. 学術論文				
3. その他（学会発表、翻訳等）				

上記のとおり相違ありません。

年 月 日

氏名

①

