

本手引きは、匿名データの利用を検討するに当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、「研究用匿名データの利用に係るガイドライン」もあわせてご参照ください。

資料5

# 申告情報の匿名データに係る 利用の手引き（案）

国税庁長官官房企画課 データ活用推進室

## はじめに

---

国税庁保有行政記録情報（以下「納税申告情報」という。）の提供は、各府省庁が保有するデータは全てオープンデータとして公開することを原則とする 「オープンデータ基本指針」を踏まえ、匿名加工した税務データを外部に提供する仕組みであり、共同研究の開始と並行して、令和3事務年度から継続的に検討されてきました。

国税庁では令和3年度に「国税庁保有行政記録情報の整備に関する有識者検討会」を設置し、データ利用に向けた「研究用匿名データの利用に係るガイドライン」の整備を行うとともに、納税申告情報等の第三者提供を令和7年度から本格実施しています。

納税申告情報は、1年間に生じた所得の金額と税額を計算するための情報であり、各種分析を想定した形式を備えておりません。したがって、データの利用に当たってはデータの特徴や限界を十分に把握しておく必要があります。

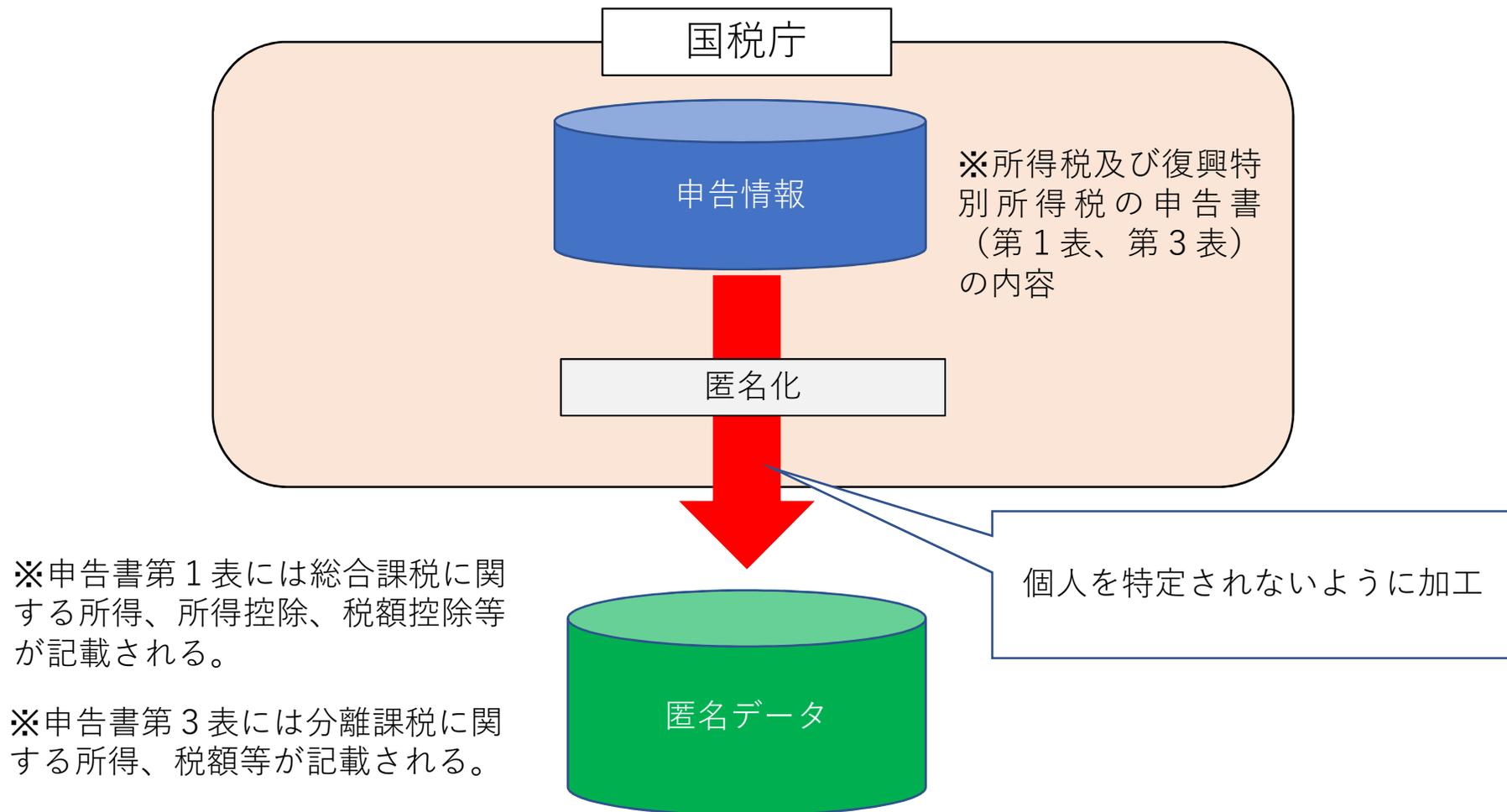
また、納税申告情報は納税者の個人情報で構成されているため、格納されているデータには匿名化処理を施してはいるものの、データ利用に当たっては提供申出者に対し一定の要件を備えるよう求めているとともに、申出に対しては、研究内容及び研究環境におけるセキュリティ等の観点から、国税庁において審査が行われることとなっております。

本マニュアルは、納税申告情報を使った研究を検討している方に、申告情報の匿名データの構造や、申出に当たっての要件、必要な手続等について、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です。「研究用匿名データの利用に係るガイドライン」とあわせて、匿名データの活用に合わせて参照いただきますよう、お願いいたします。

- 1.匿名データの構造・詳細
- 2.提供方法
- 3.利用要件
- 4.利用手続
- 5.利用審査
- 6.利用開始から終了の流れ

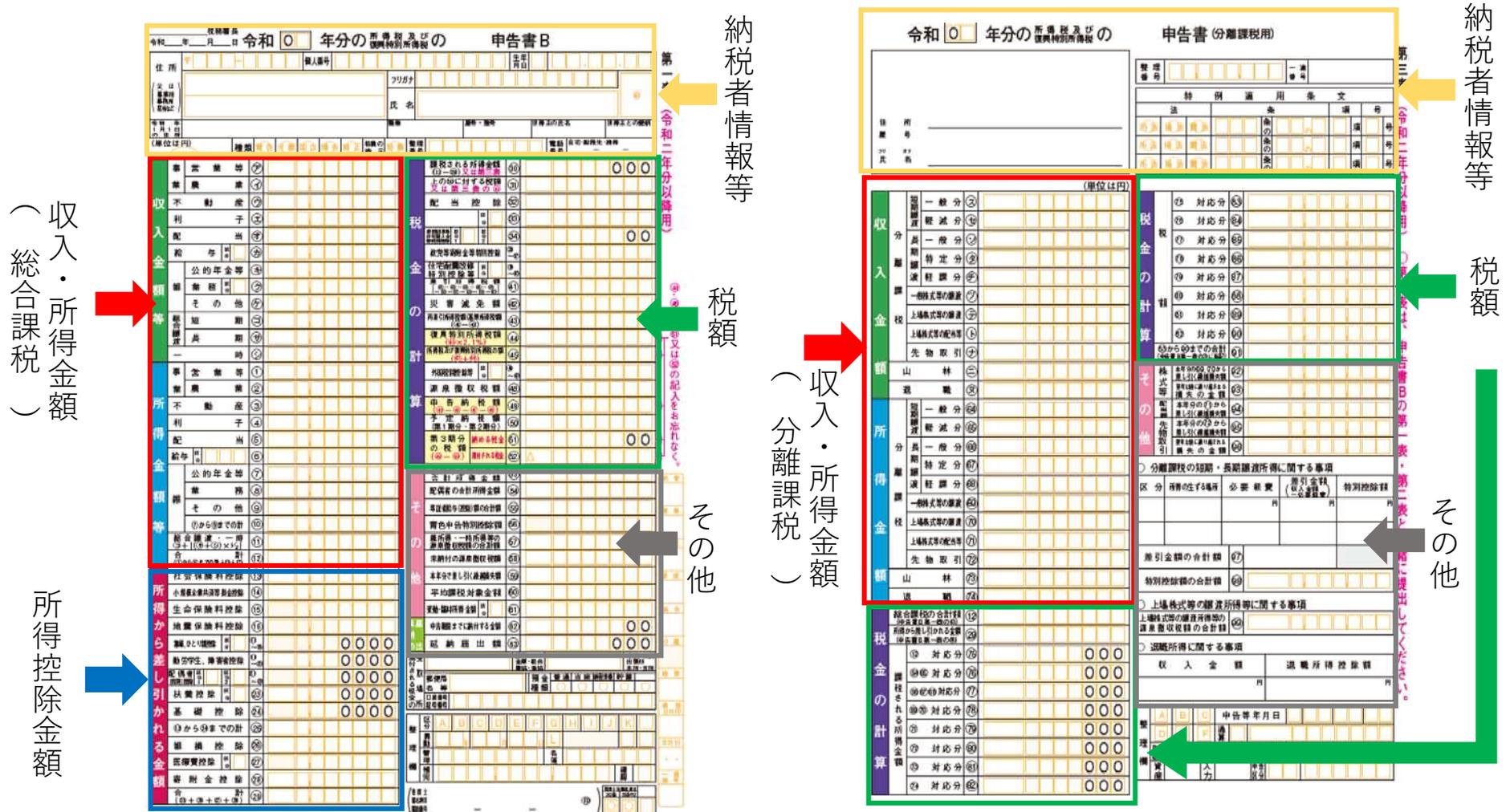
# 1. 匿名データの構造・詳細① (データのイメージ)

匿名データは、所得税及び復興特別所得税の申告情報を匿名加工したものです。



# 1. 匿名データの構造・詳細② (申告書との対応関係)

提出された申告書の第1表・第3表の申告情報を匿名加工の上、匿名データとして提供します。



# 1. 匿名データの構造・詳細③（項目一覧1 ※平成26年分の提供項目）

## <納税者情報等（定性情報）>

課税年分  
申告年月日  
納税地  
納税地に係る自宅・事業所等の別  
年齢  
業種  
青色申告・白色申告の別  
申告の種類  
出国特例適用有無  
準確等適用有無

## <収入金額等（総合所得）>

営業収入金額  
農業収入金額  
不動産収入金額  
配当収入金額  
給与収入金額  
公的年金収入金額  
その他雑収入金額  
特後一時所得金額

## <所得金額（総合所得）>

営業所得金額  
農業所得金額  
不動産所得金額  
配当所得金額  
給与所得金額  
雑所得金額  
譲渡・一時所得金額  
総所得金額

## <所得控除金額>

社会保険料控除額  
小企業共済等控除額  
生命保険料控除額  
損害保険料控除額  
老年者寡婦寡夫控除額  
勤労学生障害者控除額  
配偶者控除額  
配偶者特別控除額  
扶養控除額  
基礎控除額  
年末調整控除合計額  
医療費控除額  
寄付金控除額  
所得控除合計額

※出現頻度の低い項目についてはデータの空欄化等の匿名加工を行っています。

# 1. 匿名データの構造・詳細④（項目一覧2 ※平成26年分の提供項目）

## < 税額等（総合） >

課税総所得金額  
算出合計税額  
配当控除額  
住宅取得特別控除額  
政治献金税額控除額  
差引所得税額  
復興基準所得税額  
復興特別所得税額  
復興合計所得税額  
源泉徴収税額  
申告納税額  
第1期予定納税金額  
第2期予定納税金額  
第3期納税額

## < 収入金額（分離所得） >

分離長期一般収入金額  
株式上場収入金額  
上場株式等配当収入金額  
商品先物取引収入金額  
退職収入金額

## < 所得金額（分離所得） >

分離長期一般所得金額  
株式上場所得金額  
上場株式等配当所得金額  
課税分長譲渡所得金額  
課税株式譲渡所得金額  
課税上場株式等配当所得金額

## < 税額等（分離） >

算出総所得税額  
算出分長譲渡税額  
算出株式譲渡税額  
算出上場株式等配当税額  
差引株式損失金額  
繰越株式損失金額  
差引配当損失金額  
税理士署名有無（定性情報）  
税務代理権限証書提出有無（定性情報）  
譲渡等特例適用有無（定性情報）  
第3表算出合計税額

## < その他 >

配偶者合計所得額  
専従者給与（控除）額  
青色申告特別控除額  
雑・一時所得源泉額  
繰越損失控除額  
延納届出額

※出現頻度の低い項目についてはデータの空欄化等の匿名加工を行っています。

## 1. 匿名データの構造・詳細⑤（匿名データの特徴等1）

---

### 主な特徴

- ✓ 申告された所得税及び復興特別所得税の第1表・第3表のデータに対し、個人等が特定されないよう匿名加工処理を実施したデータです。
- ✓ 課税年分ごとのデータとなっているため、個人別の時系列の分析はできません。

### 提供対象 課税年分

- ✓ 令和7年4月から、平成26～30の課税年分について提供します。

### 抽出方法 (サンプリング率)

- ✓ 申告件数全体を母集団とし、1%（約23万レコード）を抽出しています。
- ✓ 複数回申告を行っている方については、最終申告分のみを残しています。

## 1. 匿名データの構造・詳細⑥（匿名データの特徴等2）

---

### 定性情報の 匿名化処理

- ✓個人を特定できる項目（氏名、マイナンバー等）は削除しています。
- ✓一部の項目（年齢・住所・業種等）に対してグルーピングを実施しています。

### 定量情報の 匿名化処理

- ✓ラウンディングを実施しています。
- ✓各項目の99.5%タイルの値をもとにトップコーディング又はボトムコーディングを実施しています。
- ✓出現頻度の低い項目に値がある場合や、トップコーディング前の値が類推可能な場合には、一部の項目（総所得金額・申告納税額等）を除き空欄にする処理を行っています。

# 1. 匿名データの構造・詳細⑦ 匿名化処理（項目一部抜粋）

- ✓ ラウンディング・トップ／ボトムコーディングされた数値が提供されます
- ✓ 年齢は5歳階級（10区分）、住所は2区分、業種は12区分で提供されます。
- ✓ 納税者情報はコード化（P9）されて提供されます。

 ...コーディングされた値

サンプルNO	納税者情報等			収入金額等			所得金額				総所得金額	所得控除の合計額	課税所得金額等	申告納税額
	年齢	住所	業種	事業収入	給与収入	総合譲渡	事業所得	利子所得	給与所得	総合譲渡				
1	3	0	0	1,0980,000	0	3,500,000	3,840,000	0	0	1,750,000	5,590,000	1,300,000	4,280,000	430,000
2	0	1	9	0	3,880,000	3,500,000	0	0	2,660,000	1,750,000	4,410,000	900,000	3,510,000	△130,000
3	2	0	9	0	3,660,000	0	0	0	2,480,000	0	2,480,000	700,000	1,780,000	80,000
4	1	1	6	72,000,000	1,740,000		10,000,000	0	1,140,000	0	30,000,000	4,000,000	30,000,000	4,000,000
5	9	0	0								30,000,000	4,000,000	30,000,000	4,000,000

出現頻度の低い項目（0.5%未満）に値がある、またはトップコーディング前の値を推測可能（※）なデータは、内訳（赤枠内）を空欄化しています。  
 ※例えば、所得金額の内訳を合計することにより総所得金額の値が推測可能な場合を指します。

## 1. データの構造・詳細⑧ (納税者情報のコードの内容例)

項目名	項目説明												
年齢コード	<p>実年齢を基に階級化したコードが格納されます。 5歳刻み、最大は75歳以上です。</p> <p>【コード値】</p> <table> <tr> <td>0: 35歳未満</td> <td>5: 55-59歳</td> </tr> <tr> <td>1: 35-39歳</td> <td>6: 60-64歳</td> </tr> <tr> <td>2: 40-44歳</td> <td>7: 65-69歳</td> </tr> <tr> <td>3: 45-49歳</td> <td>8: 70-74歳</td> </tr> <tr> <td>4: 50-54歳</td> <td>9: 75歳以上</td> </tr> </table>	0: 35歳未満	5: 55-59歳	1: 35-39歳	6: 60-64歳	2: 40-44歳	7: 65-69歳	3: 45-49歳	8: 70-74歳	4: 50-54歳	9: 75歳以上		
0: 35歳未満	5: 55-59歳												
1: 35-39歳	6: 60-64歳												
2: 40-44歳	7: 65-69歳												
3: 45-49歳	8: 70-74歳												
4: 50-54歳	9: 75歳以上												
住所コード	<p>三大都市圏か否かでコード化しています。</p> <p>【コード値】</p> <p>0: 三大都市圏    1: 三大都市圏以外</p>												
業種コード	<p>業種ごとにコード化しています。</p> <p>【コード値】</p> <table> <tr> <td>0: 小売業</td> <td>4: 受託加工業</td> <td>8: その他の営業</td> </tr> <tr> <td>1: 卸売業</td> <td>5: 修理業</td> <td>9: その他の部</td> </tr> <tr> <td>2: 製造小売業</td> <td>6: サービス業</td> <td>10: 農業の部</td> </tr> <tr> <td>3: 製造卸売業</td> <td>7: 建設業</td> <td>11: その他事業の部</td> </tr> </table>	0: 小売業	4: 受託加工業	8: その他の営業	1: 卸売業	5: 修理業	9: その他の部	2: 製造小売業	6: サービス業	10: 農業の部	3: 製造卸売業	7: 建設業	11: その他事業の部
0: 小売業	4: 受託加工業	8: その他の営業											
1: 卸売業	5: 修理業	9: その他の部											
2: 製造小売業	6: サービス業	10: 農業の部											
3: 製造卸売業	7: 建設業	11: その他事業の部											

## 2. 提供方法

### データ提供方法

✓匿名データは、CD-R等による貸出しにより提供を行います。

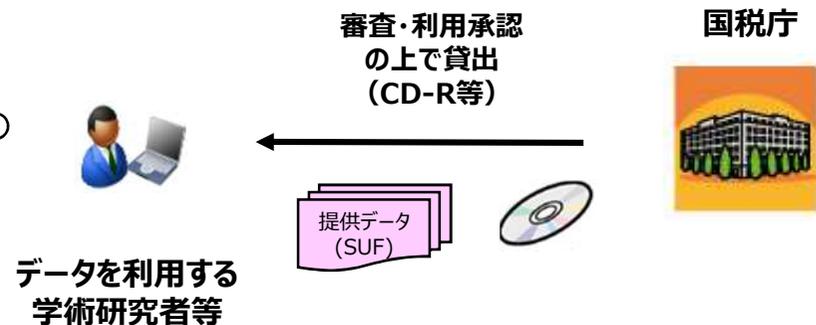
### 提供に当たっての 留意事項

✓貸出しは、一般書留での郵送（着払い）又は国税庁での直接受取りにより行います。

✓利用に当たっては、外部ネットワークから遮断された環境で行っていただき、施錠可能なキャビネットに厳重に保管していただきます。

✓使用後は、CD-R等の返却とともに、中間生成物等を削除いただきます。

- ✓利用場所は施錠可能な物理的な場所で、外部ネットワークから遮断された環境に限定
- ✓入退室管理を実施
- ✓施錠可能なキャビネット等に保管
- ✓管理簿の提出（国税庁の監査あり）



### 3. 利用要件①（利用者の要件）

---

①匿名データの利用者の範囲は、次のいずれかの機関に所属する常勤の研究者（当該機関において専ら研究に従事する者をいう。）です。

- ・ 公的機関
- ・ 独立行政法人（※1） ・ 地方独立行政法人（※2）
- ・ 大学（※3） ・ 大学共同利用機関（※4）

※1 独立行政法人通則法（平成11年法律第103号）第2条第1項に規定する独立行政法人をいう。

※2 地方独立行政法人法（平成15年法律第118号）第2条第1項に規定する地方独立行政法人をいう。

※3 学校教育法（昭和22年法律第26号）第1条に規定する大学（大学院を含む。）をいう。

※4 国立大学法人法（平成15年法律第112号）第2条第4項に規定する大学共同利用機関をいう。

②利用者の範囲は、上記①の利用者が所属する機関と同一の期間に在籍する、次の研究等の補助者を含みます。ただし、利用者が直接指導・監督する者で、次のいずれかの身分の者に限ります。

- ・ 博士研究員
- ・ 大学院生（博士課程後期相当に限る）

### 3. 利用要件②（利用者の要件）

---

利用目的が以下の両方に該当することが求められます。

#### 利用目的

- ✓ 学術研究の発展に資するものであること
- ✓ 税・財政施策の改善・充実に関する統計的研究であること

データの利用には以下についての留意が求められます。

#### 留意すべき事項

- ✓ 個体が第三者に識別されないように十分配慮し、個体を識別するために、他の情報と照合することは禁止します。
- ✓ 研究等の成果を公表する際、第三者が個体の識別ができないよう個票データを集計し分析した結果のみ公表可能とします。
- ✓ 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ことに配慮すること

### 3. 利用要件③（利用に当たっての制約事項）

---

#### 利用期間

- ✓ データの利用期間は原則として2年間を上限とします。
- ✓ 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、最長1年間を上限として、原則1回に限り延長することができます。

#### データの適正管理措置

- ✓ 匿名データの利用は、施錠可能な物理的な場所（国内）で、外部ネットワークから遮断された環境に限定します。
- ✓ 管理簿等を提出いただきます。その内容については、必要に応じて国税庁が監査します。

### 3. 利用要件④（利用に当たっての契約等）

---

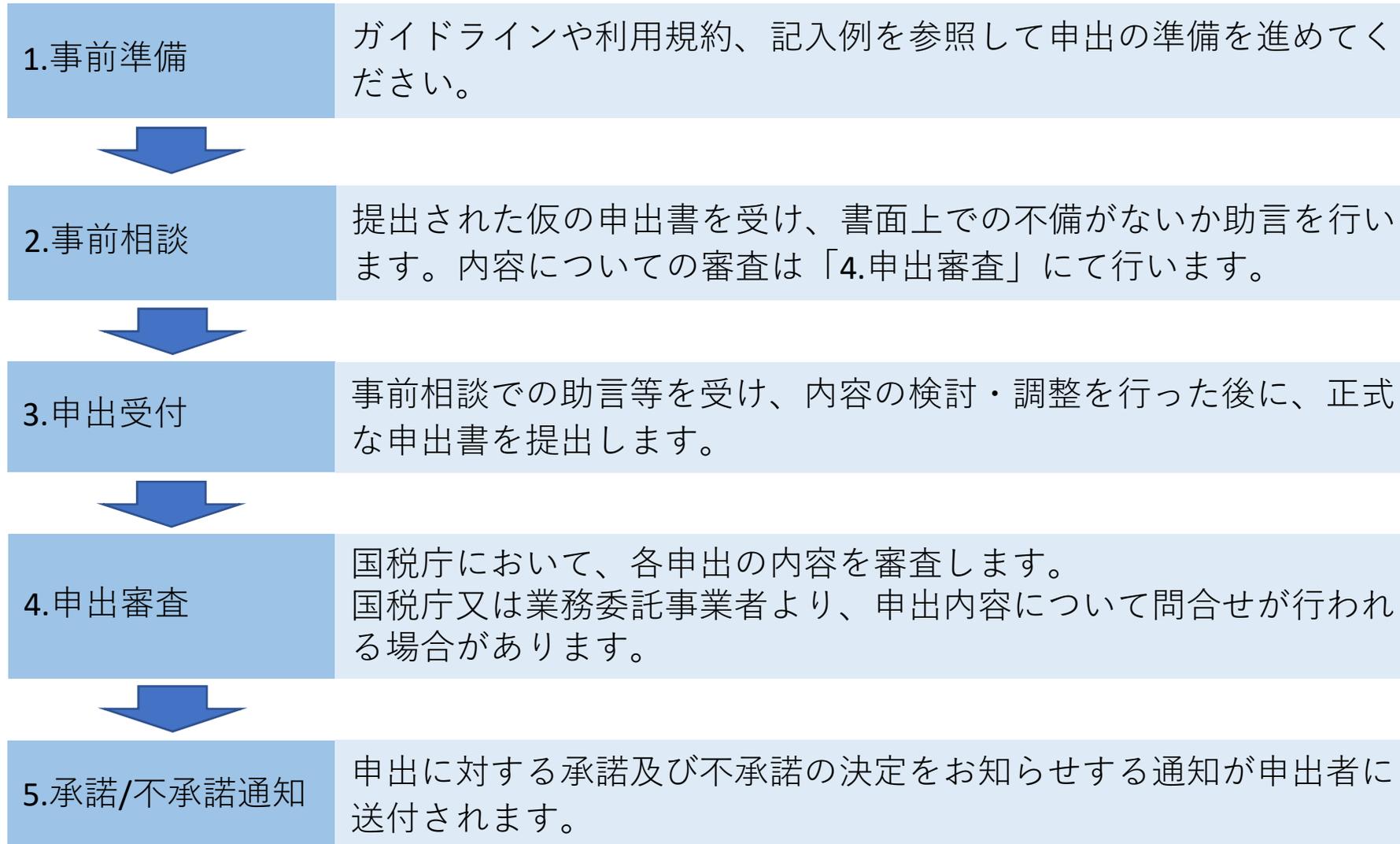
✓ 国税庁が定める利用条件（利用規約）に同意するとの誓約書を提出した上で、匿名データ等の利用を行っていただきます。

✓ 提供されたデータを紛失する、内容を漏えいする、承諾された目的以外に利用する等の事例は不適切利用としてみなし、事例に応じて匿名データに係るデータ提供を禁止すること、さらに、行為の態様に応じ利用者の氏名を公表すること等の措置を採ることとしています。

✓ 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（昭和37年法律第160号）の対象外となります。

✓ ガイドラインの規定は、変更される可能性があります。

## 4. 利用手続①（申出～審査～提供に至る具体的な手順について）

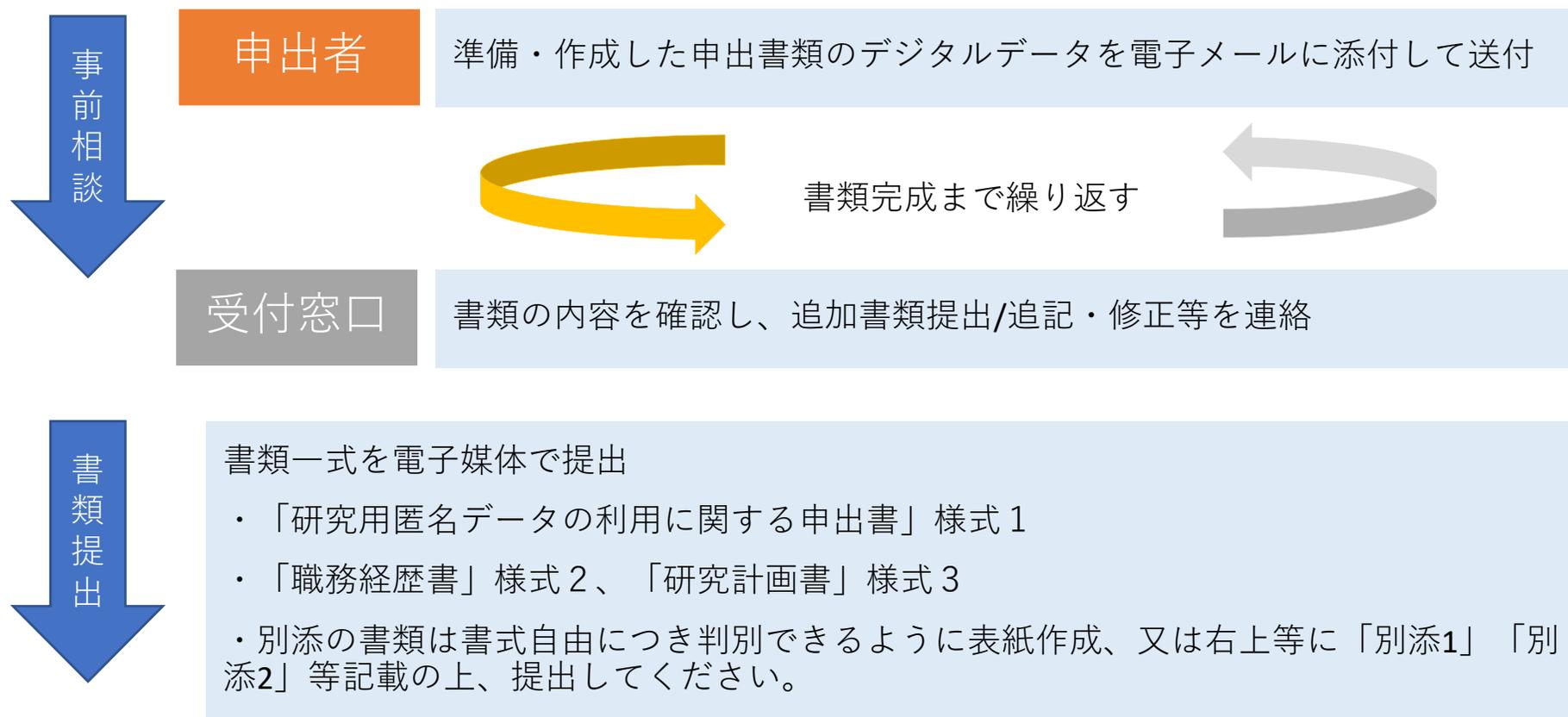


承諾された研究については、利用者名・職名・研究の名称等が公表されます。

承諾された場合には、事務局より誓約書等の提出をお願いする等、手続を進めていきます。

## 4. 利用手続② (事前準備/事前相談について)

✓ガイドライン・ホームページ・利用規約をご覧ください。書類一式を準備したのち受付窓口までご提出ください。



## 4. 利用手続③（準備すべき書類について）

- ✓ ガイドライン第5の7に規定する書類として申出書、職務経歴書、研究計画書（様式1～3）に加えて、以下の別添書類（別添1～4）をご準備ください。別添書類は書式自由につき判別できるように表紙、又は右上等に「別添1」「別添2」等、記載をしてください。

凡例（◎：提出必須 ○：該当時提出必須 △：提出任意）

申出に必要な提出書類一式				
様式/別添		書類題目	備考	
指定様式	様式1	研究用匿名データの利用に関する申出書	必ずホームページから最新版様式をダウンロードの上、ご利用ください。	◎
	様式2	職務経歴書		
	様式3	研究計画書		
様式自由	別添1	外部研究資金に関する書類	外部研究資金の制度名等を記載した場合、その事実が確認できる書類を添付してください。（科学研究費助成事業に係る資金の場合、研究者番号）	○
	別添2	研究内容に関連する資料、申出者の関連論文及び著作物一覧	必要に応じ、添付してください。	△
	別添3	過去の研究成果に関する書類	必要に応じ、実績を証明する書類を添付してください。	△
	別添4	その他必要な書類	その他参考となる書類がある場合は添付してください。	△

## 5. 利用審査（利用前の審査）

---

- ✓ 「事前相談」で形式を整えていただいた上で、受付窓口で申出を受付いたします。
- ✓ 申出内容は、受付窓口で確認を行います。その際、内容が不明瞭な場合には、申出者に対して内容の確認を行う場合があります。

### 審査基準

- ✓ 税務行政の目的に沿ったデータ利用であるか
- ✓ 個体を識別するような分析を行っていないか
- ✓ データ利用が研究内容等から判断して必要最小限であって、**公表時期が整合的**であるか
- ✓ 申出者の**過去の研究等の実績を勘案**して、実行可能であるか
- ✓ 研究公表予定日が利用期間と比較して整合的であり、**公表内容が適切**であるか
- ✓ **申出書及び添付書類の記載事項**（申出者の氏名・連絡先、提供データ内容・利用期間、利用場所・保管場所、提供方法等）**に問題ないか**

## 6. 利用開始から終了の流れ1 (貸出し後～成果発表前審査～公表)

### 1. データの受取

誓約書等の書類の提出を確認後、データの提供を行います。



### 2. データの利用・公表準備

データの分析を行い、申出書に記載された公表時期及び公表方法に基づいて、公表準備を行います。



### 3. 審査受付

事前に公表する内容について、国税庁に報告します。



### 4. 内容審査

国税庁において、研究成果の公表内容を事前に審査します。  
審査事項を満たさない場合は、公表を禁止する場合があります。



### 5. 審査結果通知

事務局内で議論を踏まえた、公表の可否に関する決定をお知らせする通知が利用者に送付されます。

成果の公表  
データの返却

公表が許可された研究については、研究概要・成果等が国税庁HPに公表されます。

## 6. 利用開始から終了の流れ2 (管理簿の作成について)

- ✓ データを受領後3か月以内に、管理簿を作成し、国税庁に提出します。(2か所以上で使用する場合は、管理状況報告書も提出します)
- ✓ データの利用期間が1年を超える場合は、定期的(1年経過の都度、1か月以内)にデータ管理状況報告書及び管理簿を国税庁に提出します。別途、国税庁が報告を求める場合もあります。
- ✓ 管理簿には、データに関する事項、管理責任者、利用者の範囲、利用場所、利用状況等を記載します。
- ✓ 管理状況報告書には、検査年月日、検査実施者、検査場所、検査状況等を記載します。

### < 管理簿様式 >

1. 提供を受けた研究用匿名データの内容						
研究用匿名データの名称	年次	ファイル数	受領年月日	提供媒体	媒体枚数	返却期限
					枚	

2. 利用者の範囲								3. 利用状況					備考 (その他具体的な 業務内容等)			
利用者の氏名 (管理責任者は◎)	所属・職名	研究用匿名データに関する権限						保有する 複写媒体	業務の種類	利用年月日	データの利 用場所	データの保管場 所		利用終了日	データの 廃棄/削除日	中間生成物の 廃棄/削除日
		利用	保管	複写	複写回数											

## 6. 利用開始から終了の流れ3（研究成果の公表に当たっての審査）

- ✓ 研究成果の公表に際しては、公表前に国税庁で審査を行います。
- ✓ 審査基準は以下の通りです。
- ✓ 公表物によっては、審査に一定程度の期間を要する場合があります。
- ✓ 公表期日があるものについては、公表期日の1カ月前までに公表を予定する内容を国税庁に提出してください。
- ✓ ただし、公表期日が利用期間終了日以降の場合は、利用期間終了の1カ月前までに、公表を予定する内容を国税庁に提出してください。

### 審査基準

- ✓ 当該研究等の成果とあらかじめ承諾された申出書の内容が整合的であるか。（ガイドライン第13の3の(1)）
- ✓ 研究成果に第三者が個体を識別できる内容が含まれていないか。（ガイドライン第13の4）
  - (1) **個人等の数が原則として10未満となる集計単位が含まれていないこと。**
  - (2) **個人等の数が10以上であっても、集計単位の比較等により、個人が特定される可能性のある集計単位が含まれていないこと。**

等

## 6. 利用開始から終了の流れ4 (データの返却・削除について)

- ✓ 研究終了後は、利用期間満了までに、匿名データを国税庁に返却します。
- ✓ 返却は、一般書留による郵送返却（送料は利用者負担）又は受付窓口での直接返却のいずれかの方法により行います。
- ✓ 返却とあわせて、研究での中間生成物について、復元できないように消去又は適切に破棄した上で、国税庁にデータ措置報告書を用いて報告します。
- ✓ また、研究等の成果、利用実績報告書及び管理簿を作成の上で、提出します。

