



厚生労働省

ひと、くらし、みらいのために
Ministry of Health, Labour and Welfare

雇用保険・厚生年金データの学術利用の事業（案）について

厚生労働省 年金局

Ministry of Health, Labour and Welfare of Japan

IX. 新しい資本主義実現に向けた取組の確実な推進

2. フォローアップ等

(3) 行政保有データの利用制約の緩和

EBPMを推進するためには、行政保有データを行政機関の外部の大学等の研究者が学術研究のために利用することを円滑化し、その成果が政策立案に活用されることが不可欠である。

このため、税務・関税・レセプトデータを政策の検討のために学術研究に活用する取組も参考としつつ、行政が取得・保有する個人情報の学術研究目的の利用・提供方法のユースケースとして、雇用保険及び厚生年金のデータを学術利用できる環境を早期に整備する。

雇用保険・厚生年金データの学術利用について（案）

取組の趣旨・目的

- **EBPMを推進するため**には、行政保有データを行政機関の外部の大学等の研究者が学術研究のために利用することを円滑化し、その成果が政策立案に活用されることが不可欠
- このため、税務・関税・レセプトデータを政策の検討のために学術研究に活用する取組も参考としつつ、行政が取得・保有する個人情報の学術研究目的の利用・提供方法のユースケースとして、**雇用保険及び厚生年金のデータを学術利用できる環境を整備**

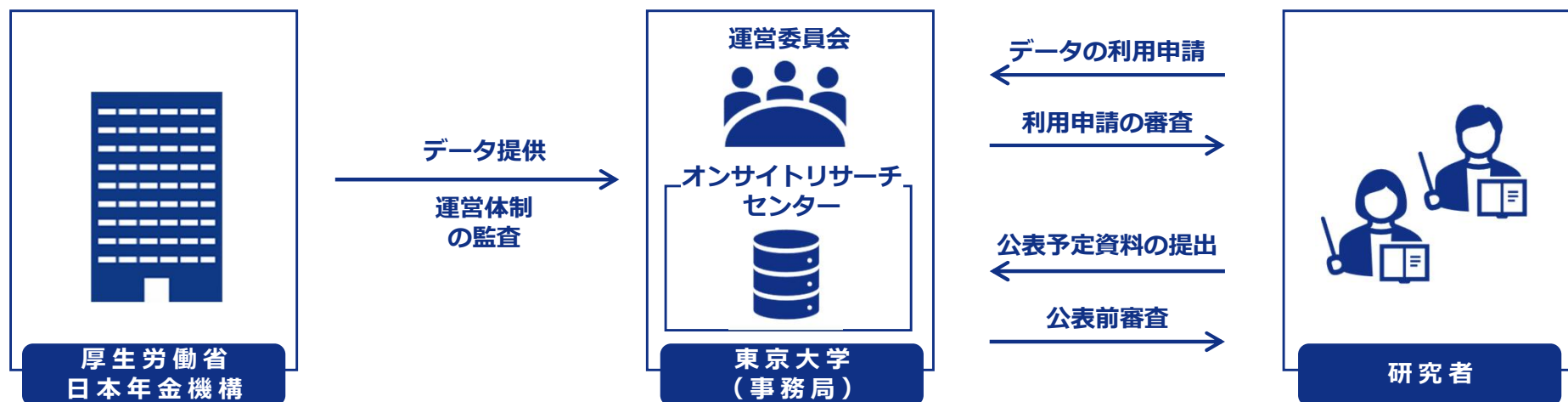
データ提供スキーム（案）

主な枠組み

- 厚生労働省・日本年金機構から、**雇用・年金に関する一定の項目に係る行政保有データを東京大学（事務局）に提供し、研究者が利用申請を行うことでオンサイトリサーチセンターのデータを研究目的で活用することができる**

提供するデータ項目

- オンサイトでの提供を行うことを前提として、**データ提供は全数を対象に行う**
- 提供項目は、**法人や個人を識別できないよう加工（ハッシュ化等）**



- データ利用申請に対する事前審査基準
公表物における基準の制定
- 個人情報の漏洩防止等の措置の確認

- データ利用申請に対する事前審査、
公表物における審査を実施
- 個人情報の漏洩防止等の措置を実施

- 東京大学に対し、データの利用申請
公表予定資料の提出を実施
- データ利用はオンサイトのみ

雇用保険・厚生年金データの学術利用の事業概要（案）

1 事業に関係する主体の役割と関係規定

1.1 オンサイト環境での提供について

- 本事業については、閣議決定「新しい資本主義のグランドデザイン及び実行計画 2025 年改訂版」で「税務・関税・レセプトデータを政策の検討のために学術研究に活用する取組も参考としつつ、行政が取得・保有する個人情報の学術研究目的の利用・提供方法のユースケースとして、雇用保険及び厚生年金のデータを学術利用できる環境を早期に整備する」こととされている。
- この閣議決定に関し、東京大学から、厚生労働省が保有する雇用保険と厚生年金データの提供を受けて、学術利用環境を研究者へ提供する事業を実施したい旨の申し出があったこと、また、東京大学には匿名医療保険等関連情報データベース（NDB（National Database of Health Insurance Claims））データの第三者提供に係るオンサイト環境の運用実績があることから、厚生労働省から東京大学にデータを提供した上で、東京大学において本事業を実施することを検討している。
（参考）NDB データの第三者提供の形態は、①媒体による提供、②オンサイト環境を通じた提供、③HIC※環境を通じたりモートアクセスによる提供、といった方法で実施。このうち、②オンサイト環境を通じた提供については、東京大学のオンサイトリサーチセンターなどにより実施されている。
※HIC（Healthcare Intelligence Cloud）：医療・介護等データ解析基盤。クラウド上に構築されたプラットフォーム。
- 閣議決定の内容や、オンサイト環境で行政保有データを研究者の利用に供する点で類似の事業となっていることから、NDB データの第三者提供を参考として事業を検討する。
なお、NDB データの第三者提供は、国（厚生労働省）が第三者提供の申出者にそれぞれデータを提供するスキームであることに対して、本事業は、国（厚生労働省）からデータの提供を受けた東京大学がデータの学術利用環境の提供主体である点に留意が必要である。

1.2 厚生労働省と東京大学の関係

- 厚生労働省は閣議決定に基づき、雇用保険及び厚生年金のデータを学術利用で

きる環境を早期に整備するため、東京大学に雇用保険と厚生年金のデータ（以下「保険年金原データ」という。）を提供する。

- 東京大学においては、厚生労働省から提供されたデータを活用して、EBPM（証拠に基づく政策立案）推進に資する研究が行われるよう、オンサイトでデータ利用する環境（オンサイト環境）を構築（東京大学オンサイトリサーチセンター）し、研究者に提供する。
- データには個人が識別できる可能性がある情報が含まれており、本事業において個人情報に特定・漏えいされることのないよう、事業運営を行う。
- 東京大学においては、このデータを活用した研究成果が得られるよう本事業を実施する。

1.3 厚生労働省、東京大学、利用者に係る関係規程

- 「合意文書」においては、データを提供する厚生労働省と、データ提供事業の実施主体となる東京大学における基本的な事項を定めている。
- さらに、「運用管理規程」では、合意文書に規定する安全管理措置を具体的に定めるものとして、厚生労働省から匿名化したデータの提供を受ける東京大学が、東京大学オンサイトリサーチセンターの運用において個人情報保護と情報セキュリティの観点から遵守すべき事項を定めている。
- 一方で、利用者との関係については、データを提供する東京大学とデータを利用する利用者の間のルールを定めるものとして「ガイドライン」を定める。

2 データ関係

2.1 データ提供の前提

- 今回提供するデータについては、個人情報保護法及び日本年金機構法の規定に基づき、学術研究目的の提供を行うものである。厚生労働省は、提供を受ける東京大学に対し、データの利用方法について制限を設ける等、データの適切な管理のために必要な措置を講ずることを求めるものとする。
- これらの法律においては、本人又は第三者の権利利益を不当に侵害するおそれがあると認められるときには、情報の提供ができないこととされており、権利利

益の侵害がないよう、提供データの一部項目について、ハッシュ化¹する等の処理をする。

- ハッシュ化等の処理をすることにより、個別の項目ごとにみれば直接個人識別できないものになるが、提供項目の組合せや雇用保険・厚生年金の連結によって個人が特定される可能性があることから、本事業の実施に当たっては、データ利用に際し、必要な安全管理措置の実施を求めるものとする。

2.2 提供するデータの項目

- 厚生年金と雇用保険で提供するデータの主な項目は以下のとおり。

厚生年金 （※はハッシュ化して提供する）

- ・ 現存又は喪失時の被保険者属性（性、生年月、都道府県、取得年月、喪失年月、喪失原因、氏名※、生年月日※）
- ・ 基礎年金番号※
- ・ 標準報酬月額（8.8万円～65万円）、標準賞与額（150万円以下）の履歴情報
- ・ 被保険者が属する事業所情報（適用年月、全喪年月、全喪原因、都道府県、業態、内外国、本支店、法人番号※、郵便番号※、電話番号※、等）
- ・ 属性や履歴等の結合に用いる被保険者コード※、事業所コード※

雇用保険 （※はハッシュ化して提供する）

- ・ 現存又は離職時の被保険者属性（性、生年月、取得年月、離職年月、離職理由、氏名※、生年月日※、等）
- ・ 被保険者が属する事業所情報（設置年月、廃止年月、産業分類、法人番号※、郵便番号※、電話番号※、等）
- ・ 事業所の労働保険料申告時の情報（常時使用労働者数、雇用保険被保険者数）
- ・ 被保険者の介護休業開始時の情報（介護休業開始年月、賃金月額）
- ・ 被保険者の育児休業開始時の情報（育児休業開始年月、賃金月額）
- ・ 被保険者の教育訓練給付金の支給申請に関する情報（教育訓練講座名、受講開始年月、受講終了年月、教育訓練経費、等）
- ・ 属性等の結合に用いる被保険者番号※、事業所番号※

¹ ハッシュ化とは、ハッシュ関数と呼ばれる特殊な計算手順により、元のデータからハッシュ値と呼ばれる規則性のない固定長の値を求め、その値によって元のデータを置き換える方法のこと。同じデータからは常に同じハッシュ値が得られる一方で、少しでもデータが異なるとまったく類似しない別のハッシュ値が生成され、また、ハッシュ値から元のデータを割り出したり、同じハッシュ値を持つ別のデータを生成したりすることは極めて難しいとされている。

2.3 提供するデータに係る個人情報の配慮

- ハッシュ化の仕様については、厚生労働省で管理し、東京大学には提供しない。
- 基礎年金番号、雇用保険被保険者番号は個人情報保護法における個人識別符号であり、当然にハッシュ化の対象とする。
- さらに事業所を識別する符号についても、行政保有データでは当該事業所に属する被保険者が1名である事業所が数多く存在し、その勤務先事業所が特定されることは属する被保険者を特定することにつながりかねないことから、ハッシュ化の対象とする。

2.4 提供するデータの厚生年金と雇用保険のデータの連結

- 本事業の実施に当たり、厚生年金と雇用保険のデータを連結させることはデータの利用価値を上げるために有効。
- 厚生年金と雇用保険では届書情報の仕様・規格が統一されていないという制約があることから、比較的一致することが期待される半角英数、半角カナの項目を主なキーとして連結させることが考えられる。
- 当初は、生年月日、被保険者カナ氏名、事業所法人番号、事業所郵便番号、事業所電話番号の5つのハッシュ化した項目を用いる。
- 連結の状況等を踏まえ、来年度以降の事業において、連結のキーをさらに増やすことができないか検討する。

3 東京大学における事業実施体制

3.1 運用チーム

- 運用チームは、運用管理規程に定める役割や権限に従って、東京大学オンサイトリサーチセンターにおけるデータの保管及び利用者への二次提供に係る運用管理（5.6 のオンサイト環境からの持ち出し審査、5.8 の成果物の公表前審査など）を実施する。

3.2 運営委員会

- 運営委員会は、申請者からのデータ利用申請の審査を行う。また、必要に応じて、オンサイト環境からの持ち出し審査、成果物の公表前審査を行う。

- メンバー構成は申請者と関係を有する委員は審査に関与しないことに留意し、複数の外部委員を含むメンバーを予定。

4 東京大学における安全管理措置等

- 保険年金データ（保険年金原データを東京大学が利用者提供用に加工したデータをいう。以下同じ。）及び保険年金データを用いて生成する生成物（以下「保険年金データ等」という。）の管理に係る「実施場所」、「保管場所」、「利用場所」を合わせた区域を「取扱区域」とする。
- 保険年金データ等の取扱区域を特定し、立ち入り管理及び制限するための措置を講じる。
 - ※ 「取扱区域」から持ち出し可能な情報は、最小集計単位の原則を満たすなど、個人・事業所が第三者により特定されないことを運用チームとして確認しているものに限る。
- 「実施場所」（東京大学内のサーバ管理端末室）では、閉域 LAN により保管場所にあるサーバと接続されたサーバ管理端末から、保険年金データの入出力や安全管理措置等に係るサーバの操作を行う。
- 「保管場所」（東京大学のサーバ室）では、保険年金データ等をサーバで保存する。
- 「利用場所」（東京大学内の利用者端末室）では、閉域 LAN を用いた仮想デスクトップ基盤（VDI）²により接続された利用者端末により、利用者が統計解析ソフトウェアの操作を行う。なお、利用者端末から保険年金データには直接アクセスできず、利用者端末にはデータを保存できない等の仕様とする。
- 情報セキュリティ事故が発生した場合には、東京大学の情報セキュリティポリシー等に従って、必要な措置を講ずるとともに、速やかに厚生労働省に報告する。

5 研究グループ（利用者）への対応

5.1 研究グループの範囲

- 研究グループは、日本学術振興会が実施する科学研究費助成事業の対象となる研究機関、厚生労働省の施設等機関である国立社会保障・人口問題研究所もしく

² 仮想デスクトップ基盤（VDI（Virtual Desktop Infrastructure））は、OS やアプリケーションなどのデスクトップ環境をサーバ上に仮想化し、ネットワークを通じて利用者端末からアクセスできる仕組みのこと。

は厚生労働省が所管する独立行政法人労働政策研究・研修機構に所属する研究者及び研究補助者、海外の同等の研究機関に所属する研究者及び研究補助者、又はこれらに準ずると認められる研究機関に所属する研究者及び研究補助者（以下「研究者等」という。）から構成されるものとする。

- 申請は研究者等からなるグループを単位とし、研究代表者・運用管理責任者をおく。査読者が海外に拠点を有する場合も含め、申請研究グループにデータ利用者を追加する手続きのもと、査読者がオンサイト環境のデータを利用することが可能となる。
- 安全管理の観点から利用者を特定する必要があるため、匿名の査読者はオンサイト環境のデータを利用することはできないが、学術雑誌等から、査読者に代わって補助者によるオンサイト環境の利用を求められた場合は、当該補助者の利用者追加の手続きを行う。

5.2 研究グループからの申請

- 研究グループから、データの利用目的、研究計画（研究の背景・目的、研究デザイン、スケジュール）等を記載した利用申請書等を提出させる。
- 研究グループの研究者等の所属機関から保険年金データを利用した研究を行うことの承認を証明する書類を添付する。
- 研究グループの研究者等から、法令及びガイドラインを遵守すること等を記載した誓約書を提出させる。

5.3 申請の審査

- 運営委員会において、データの利用目的、安全管理対策、結果の公表予定等の観点で審査を行う。
- 審査においては、国際的な査読雑誌での公表を優先採択基準とする。

5.4 研究グループから手数料の支払い

- 手数料については、初年度の試行実施の状況等も踏まえ、審査や提供にかかる実費を勘案して設定する。

5.5 研究グループのデータの利用

- 研究グループは、オンサイト環境においてデータの利用を行う。

5.6 研究グループによるデータの持ち出し

- データの持ち出しについては、個人の特定のおそれがあるデータの持ち出しを認めないこととし、特定の個人・事業所が第三者に識別されないことを、運用チームの審査により確認した上で持ち出しの許可をする。

※ 運用チームのメンバーが研究グループの構成員となる場合の審査は、運営委員会の委員が行うこととするなど、利益相反に配慮した運営とする。

- 具体的には、以下の（１）から（５）までの基準に基づく。

（１）最小集計単位の原則

i) 個人等の数の場合

原則として、成果物において個人等の数が 10 未満になる集計単位が含まれていないこと（個人数が10未満の場合はマスキングを行うこととするが、「0」の場合を除く）。

なお、個人等の数の集計単位が雇用保険事業所の郵便番号 3 桁による地域区分の場合には、地域区分内の被保険者数等を基準として秘匿性を高める。

※具体的な基準は、試行期間の実施状況を踏まえ、設定。

ii) 事業所数 3 未満の場合

事業所数の属性情報による集計数が 3 未満となる場合は、マスキングを行うこと（ただし個人等の数が「0」を除く）。

※ マスキングにあたっては、マスキングを行ったセルの値が他のセルの値や行計・列計等から算出できる場合、これを算出できないように当該セル以外のセルもマスキングを行う。

（２）年齢区分

100 歳以上については、同一のグループとすること。

（３）個票の取扱い

個人、事業所等の個票を含まないこと。

（４）識別子の取扱い

ハッシュ化したものを含む基礎年金番号、事業所識別コード、法人番号、その他

の識別子又は連結キーを含むデータは含まないこと。

(5) グラフの取扱い

グラフのバックデータとなる集計表が(1)から(4)の基準を満たすことに加えて、当該グラフの作成に用いる項目に限ること。

グラフのバックデータとなる集計表が基準を満たす場合に当該グラフの持ち出しを認めるため、利用者は散布図等のグラフの持ち出しを申請する場合、バックデータとなる集計表をあわせて窓口へ申請すること。

5.7 持ち出したデータの保管・共有

- 運用チームの審査を経た上で持ち出した結果（最終生成物）について、個人が特定されないことを、すでに運用チームが確認していることから、当該結果（最終生成物）のみ利用する最終生成物利用者との共有のため、当該結果（最終生成物）の、クラウド等での保管を認める。

5.8 研究成果の公表前の審査

- 研究成果は公表前に、運用チームにおいて審査を行う。必要に応じて運営委員会も審査に参加する。
 - 公表物においては、持ち出し審査後のデータをもとにしつつも、データに対する記述が加わることで個人が特定されることがないか、個人が特定されるとの誤解を招かない適切なものとなっているか等について、運用チームにおいて審査を行う。
- ※ 運用チームのメンバーが研究グループの構成員となる場合の審査は、運営委員会の委員が行うこととするなど、利益相反に配慮した運営とする。

5.9 研究成果の公表

- 研究グループにおいて、学会誌等において研究成果の公表を行う。

5.10 終了時の処理

- 研究グループは、保険年金データ等の利用を終了したときは、当該データ等を消去して、廃棄したデータ、廃棄対象外データ（公表した成果物及び東京大学が許可したもの）、廃棄方法等を東京大学へ報告する。

6 基礎資料の作成、試行実施、本格実施について

6.1 東京大学による基礎資料の作成

- 研究グループがどのような研究を行えるかについて、運用チームが基礎資料（保険年金データの仕様書、保険年金データの各変数の記述統計等）を作成する。

6.2 東京大学による試行実施（公募によるデータ提供の開始前）

- 本格実施に向けた体制整備の観点から、東京大学において試行実施を行う。
- 試行実施においては、オンサイト環境整備及び基礎集計を行う。
- 試行実施においては、研究者等は東大グループ、実施期間は令和8年度下半期を予定している。
- 試行実施において実施内容等に変更が生じた場合は、厚生労働省と協議のうえ合意書等を見直す。

6.3 本格実施（公募によるデータ提供の開始後）

- 本格実施は令和9年度から実施することを予定している。
- 本格実施においては、同時にオンサイト環境で研究できるグループは3程度を想定している。
- 1グループの研究利用期間は、基本3年であるが、本事業については、査読を経て論文公表する予定を採択の観点とする事情があることから、1年の延長を2回できる。

7 合意文書の解除・合意文書に違反した場合の措置等について

- 東京大学に、データの取扱いに当たり重大な過失がある等、合意の解除に当たる事由が認められた場合は、厚生労働省は合意文書を解除できるものとする。また、東京大学に合意に違反した不適切な取扱いが認められた場合、厚生労働省が東京大学にとるべき措置について、合意文書に定める。なお、これらの措置等により、研究者等のデータ利用が中断されることがあり得る。
- 合意に違反した場合の措置については、違反行為の態様に応じて、段階的に定めることとする。具体的には、次のとおりとする。
 - ・ データに関して、特定の個人を識別することを試みたり、厚生労働省から提供

されたデータを紛失したりする等の行為が認められる場合は、速やかにデータの返却や消去を求める。

- ・ 運用管理規程に反するセキュリティ上の危険を伴うデータの利用や、オンラインサーチセンターの利用にかかる情報（アカウント情報等）の漏えいが認められた場合は、一定期間、データの利用や提供を禁止する。
- 厚生労働省は、保険年金原データ及び保険年金データ等の利用環境について東京大学の実地監査を行うことができる。また、東京大学は、保険年金原データ及び保険年金データ等の管理・利用状況について、定期的に厚生労働省に報告するものとする。厚生労働省は、実地監査や東京大学からの報告等を踏まえ、東京大学に対し、必要な改善を求めることができる。

8 事業実施の財源

- 本事業の実施に当たっては、東京大学が、文部科学省の教育研究活動充実分（組織改革分相当）の補助を受けて実施する。
- 合意書の履行に要する費用負担については、厚生労働省と東京大学で協議の上、定める。