

## 第2回 厚生労働省 データ利活用検討会 議事録

1 日 時 令和2年11月6日（金）15:00～16:20

2 場 所 厚生労働省 仮設第3会議室

3 出席者

### 【構成員】

阿部正浩（座長）、野口晴子（座長代理）、川口大司、川崎茂

### 【事務局（厚生労働省）】

鈴木統括官、武藤参事官、戸田企画官、大野審査解析室長

4 議 事

- 1 第1回利活用検討会における御意見について
- 2 アンケート集計結果について
- 3 オンライン拡張方針への対応について
- 4 その他

### [配布資料]

資料1 第1回利活用検討会における御意見について

- (1) ホームページの改善
- (2) 研究費の申請段階などでの広報周知

資料2-1 調査票情報の二次利用等の利用者アンケート集計結果

資料2-2 二次利用等利用者実態アンケート用紙

資料3 審査にかかる時間

資料4 オンライン施設への調査票情報の提供方針について（案）

資料5 行政記録情報の利活用に関する考え方（案）

5 議事録

○大野審査解析室長 それでは、定刻となりましたので、ただいまより第2回「厚生労働省データ利活用検討会」を開催いたします。

構成員の皆様方におかれましては、御多忙のところ、お集まりいただきまして、誠にありがとうございます。

私は、政策統括官（統計・情報政策担当）付審査解析室長の**大野**と申します。どうぞよろしく願いいたします。

「厚生労働省データ利活用検討会」の開催は、本日が第2回となります。本年5月22日に開催しました第1回は、新型コロナウイルス感染症の防止に資する観点から、書面開催といたしました。対面での開催は今回が初めてでございます。このため、構成員並びに事務局を御紹介いたします。

はじめに、中央大学経済学部教授の阿部構成員でございます。

○阿部構成員 阿部です。よろしくお願いいたします。

○大野審査解析室長 東京大学大学院経済学研究科教授の川口構成員でございます。

○川口構成員 川口です。どうぞよろしくお願いいたします。

○大野審査解析室長 早稲田大学政治経済学術院教授の野口構成員でございます。

○野口構成員 野口でございます。よろしくお願いいたします。

○大野審査解析室長 また、オブザーバーといたしまして、日本大学経済学部の川崎特任教授にも御出席いただいております。

○川崎構成員 川崎です。よろしくお願いいたします。

○大野審査解析室長 続きまして、事務局です。

政策統括官（統計・情報政策担当）の鈴木でございます。

○鈴木統括官 鈴木でございます。よろしくお願いいたします。

○大野審査解析室長 参事官の武藤でございます。

○武藤参事官 武藤でございます。よろしくお願いいたします。

○大野審査解析室長 政策企画官の戸田でございます。

○戸田企画官 戸田です。よろしくお願いいたします。

○大野審査解析室長 議事に先立ちまして、鈴木統括官から御挨拶を申し上げます。よろしくお願いいたします。

○鈴木統括官 構成員の皆様方、本日は厚生労働省の「データ利活用検討会」に御参集いただきまして、誠にありがとうございます。本来でしたら、第1回で御挨拶申し上げるところでありますけれども、先ほど御紹介ありましたように、第1回はコロナの影響で書面開催でございましたので、本日、改めて御挨拶申し上げたいと思っております。

さて、厚生労働省関係の統計につきましては、今色々統計改革をやっておる最中でございます。ガバナンスの強化でありますとか、業務の改善、処理システムの改善、EBPMの推進といったものを色々やっているところでございます。こういったことで、統計精度の強化のような形では色々進めておるところですけれども、統計精度を上げていくのは非常に重要なことではあります。統計というのはユーザーの皆様に使っていただいて、初めて意味があると考えております。そのために、併せまして、統計データをどういう形で国民の皆様提供し、使っていただき、分析いただくかが非常に重要なことと認識しております。

本検討会は、このデータの利活用につきまして御議論いただき、色々御示唆いただけたらと思っております。ぜひとも忌憚のない御意見をいただきまして、この会

が有意義になるようお願い申し上げます、冒頭の御挨拶とさせていただきます。

どうぞよろしくお願いいたします。

○大野審査解析室長 それでは、議事に移ります。

本検討会は、第1回開催時の互選により、阿部構成員に座長をお願いしております。

以後の進行については、座長をお願いいたします。

○阿部座長 それでは、よろしくお願いいたします。

初めに、本日の進め方ですが、議題ごとに事務局から説明を受けて、その後議論したいと思います。

早速ですが、資料1について事務局から説明をお願いいたします。

○大野審査解析室長 それでは、お手元の資料1を御覧ください。第1回「厚生労働省データ利活用検討会」では、ホームページ改善について貴重な御意見をいただいたところですが、そのうち、厚生労働省内で改善できる2点については、早速改善させていただいたところです。

1点目は、ホームページの改善です。1枚おめくりいただきまして、2ページを御覧ください。調査票情報等の利用のホームページのオンサイト施設の利用については、ページ左側にあるホームページから、右上のページへ、さらに右下のページへと、2回たどって、オンサイト利用が可能な調査票情報のページに到達してございましたけれども、それぞれのページについては、内容がほとんど変わらないにもかかわらず、ホームページが細分化されてしまっているという点への御指摘をいただきました。

また、総務省が作っていますマイクロデータ利用ポータルサイトであるmiripoというサイトがございます。こちらへのリンクについても、1回目のページからたどれなくはないのですが、かなりたどりにくい構成となっていたものです。miripoへのリンクを適切に貼る点につきましても、御意見をいただきました。

3ページを御覧ください。いただきました御意見を踏まえまして、調査票情報等の利用のページに、オンサイト施設の利用について、それぞれ利用可能な調査票情報の種類や年次も含めて記載するとともに、miripoへのリンクを適切にお貼りしている状況です。

次に、4ページを御覧ください。2点目ですがけれども、科学研究費の申請段階などにおける周知です。従来、厚生労働省の科学研究費のホームページには、調査票情報の二次利用についての説明がございませんでした。第1回「厚生労働省データ利活用検討会」にて、科学研究費の申請段階などで周知するなどの工夫が必要ではないかといった御意見をいただきましたので、今般、厚生労働省の科学研究費に関するページに、二次利用のホームページへのリンクを設置いたしました。これは、いずれも既に実現しております。

資料1にかかる説明は以上です。

○阿部座長 ありがとうございます。

皆様からいただいた御意見に関して、ホームページの改善等を行ったということですが、何か御質問あるいは御意見はございますか。

川崎構成員、どうぞ。

○川崎構成員 特に意見ということではないのですが、このあたりのこと、私も幾つか意見を申し上げておりました。早速対応していただき、細かなことではありますが、見やすくしておくというのは大事なことだと思いますので、いいことだと思います。

○阿部座長 ありがとうございます。

他には、よろしいですか。

では、次に資料2に移りたいと思います。事務局、お願いいたします。

○大野審査解析室長 それでは、お手元の資料2の1ページを御覧ください。厚生労働省における調査票情報の二次利用等にかかる利用者の実態を把握し、今後の改善につなげることを目的に、資料のとおりアンケートを行いました。第1回「厚生労働省データ利活用検討会」でいただいた御意見を踏まえて、報告を求める数ですが、当初より増やして100人程度に対して実施いたしました。

1枚おめくりいただきまして、2ページを御覧ください。今回のアンケートの回収状況は、おおむね2割から3割程度の回収率となっております。

さらに1枚おめくりいただきまして、3ページを御覧ください。今回のアンケート回答者について、大学のホームページ等々を参考に、専門分野の情報を当方で取得、構成割合を示しております。「医療経済」以外の医療関係者については、「公衆衛生・保健・疫学」という分類とさせていただきます。利用者の半分程度は、この医療・公衆衛生等を専門分野とされている方々であり、以下、労働経済、公益社団法人・民間企業等の職員の方、労働経済及び医療経済以外の経済の方、社会学の方という順となっております。

さらに、1枚おめくりいただきまして、4ページを御覧ください。ここからがアンケートの回答結果でございます。厚生労働省ホームページ上の調査票情報の二次利用等に関する周知につきましては、「わかりやすい」と「ややわかりやすい」を合わせたものが35%、「ふつう」が44%、「わかりづらい」及び「ややわかりづらい」が21%となっております。

5ページ、6ページに実際のわかりやすさについての具体的な御意見を掲載しております。

以下も同様ですけれども、既に対応済みのものもございましたので、それらについては、二重の括弧に事務局のコメントとして付しております。例えば、5ページの下から2つ目は、ホームページ改正後には改善されたものなどです。

また、5ページの一番最後、「わかりやすくしたら利用者が増えるようなものではないので、わかりやすくする必要性は低い」という意見もいただいておりますけれども、当方としましては、利用者の利便性向上のためにもわかりやすくする必要があると考えております。

これらのいただいた御意見を大まかにまとめたものが7ページでございます。

内容につきましては、高評価していただいた方がいらしたほかには、利用区分については賛否が分かれました。また、「フロー図・手引き書・手順作成」を求める意見などがご

ございました。

場所につきましては、「利用者（利用ケース）別誘導の作成」といった意見をいただいております。

その他、利用者であるにもかかわらず、ホームページでの周知案内をそもそも「知らなかった」などの意見も3件ほどいただいているような状況です。

次に、1枚おめくりいただきまして、8ページを御覧ください。利用手続についての御意見です。ありがたいことに、かなり多岐にわたる御意見をいただきました。13ページまで御意見を記載しております。これらの御意見を大まかにまとめたものが14ページと15ページです。

まず、14ページを御覧ください。

申請書類について多かった意見は、内容の簡素化でした。同じ調査でも、調査年次によって調査項目が変わり、データレイアウトが変わることなどから、利用する調査項目を年次ごとに指定する作業部分についての指摘がかなり多くございました。また、集計結果の図や表をあらかじめ作成するのが困難であるという指摘もございました。後者については、オンサイト施設を利用できれば、図表を作成してから審査となりますために、対応可能となっているところでは、ほかに、電子化の要望と、使用可能な調査票情報の項目一覧表の公開といった御意見をいただいております。

次に、審査にかかる指摘につきましては、審査事項や手続の簡素化の要望があった一方で、利用者側の意見としては少々意外ですけれども、現行制度の審査自体の必要性を理解するといった意見もいただいております。また、使用しない調査項目を含む全調査項目の提供を求める声もいただいております。また、事後チェック対応となりますと、今のオンサイト施設利用のような形になると思われましても、オンサイト施設とは異なり、研究室に使用状況を管理するような設備を設置するわけにもいかないという問題もあることから、なかなか難しい面はございますけれども、要望はかなりございました。

15ページを御覧ください。承認・提供までの日数と費用でございます。「日数が長い」、「メールで速やかな返答を要望」といった御意見を多数いただいております。一方で、「短かった」、もしくは「以前より審査期間が短縮された」といった御意見もいただいております。

後ほど審査日数の議題でも触れますけれども、昨年度、審査日数が大幅に増加したことについてです。基本的には、申請順に処理しておりますので、お待ちいただいた方もいらっしゃると思いますが、一方で、申請者の申請書類への慣れにより、申請内容が前回と余り変わらない場合は審査が早くなることなども考えられます。もともと、令和元年5月に申請書について様式変更いたしましたので、その書類が増えております。変更前後に申請された場合は、申請内容によっては短くなるというのが、必ずしも全員に当てはまらないと思っております。

また、審査にかかる指摘同様、利用者側の回答としては少々意外ですけれども、承認・

提供までの日数、費用につきましても、「妥当である」、「秘匿性が高い統計利用の審査には時間が必要」という御意見もいただいております。

次に、16ページを御覧ください。提供物について、利用に当たっての御意見です。19ページまで御意見を記載しております。

これらの意見を大まかにまとめたものが20ページ及び21ページにあります。

20ページに参ります。データにつきましては、現在、テキスト形式及びCSV形式で提供しておりますけれども、それ以外の様式を求める意見もいただいております。

また、ドキュメントにつきましては、「データ定義書・符号表・符号情報等に対する不満」がございました。

22ページを御覧ください。次に、他省庁との比較です。そもそも複数の省庁に二次利用申請をされるケースが少なかったのか、回答数は10件ほどと少ないので、断定的なことは申し上げられませんが、厚生労働省が他省庁と比較して特に利用しづらいという結果にはなっておりませんでした。

23ページには具体的な御意見を記載しております。比較されるのは、主に「承認・提供までの日数」、「手続のわかりやすさ、指示」などについての点です。省庁間の違いではなくて、担当者の違いによるのではないかといった趣旨の指摘もございました。

次に、24ページを御覧ください。匿名データの利用意向です。「利用したい」とした者は4割程度、「学生に使わせてみたい」とした者は半数近くとなっています。「利用したい」とした者から見ると、その7割程度が「学生に使わせてみたい」となっております。

25ページから27ページまで御意見を記載しております。

これらの意見を大まかにまとめたものが28ページです。利用したい理由としては、「学生教育や実習が多い」という意見のほか、消極的な意見となりますけれども、「調査票情報の個票の利用のハードルが高いため」という意見がございました。

一方、利用したくない理由としては、リサンプリング等の匿名化処置をしておりますので、「個票データのほうがよい」や、「マッチングできないこと」などの意見が挙げられています。

29ページを御覧ください。今後利用してみたい匿名データとしては、既に匿名データが存在する「国民生活基礎調査」のほか、「国民健康・栄養調査」の回答が挙げられています。「国民生活基礎調査」は、複数の専門分野の方からなのに対し、「国民健康・栄養調査」は3人の方が回答されておりますけれども、いずれも公衆衛生系の専門分野の方からの回答でした。

次に、30ページを御覧ください。オンサイト施設について、わからない点として、メリットやルールなどが挙げられています。miripoなどのホームページに丁寧な記載もありますが、周知されていないのか、そういった回答となっております。

また、31ページにございますけれども、「初めて知った」という意見がかなり多い点が目立っております。周知不足等の御意見も含め、まだまだ周知されていないという結果に

なっております。

32ページを御覧ください。今後のオンサイト施設の利用意向でございますけれども、「利用したい」が17%なのに比べ、「利用したいとは思わない」が33%と高くなっております。

33ページに具体的な意見がございます。利用したい理由としては、オンサイト施設利用のメリットである、実際のデータに触れながら分析できる点が回答されております。一方、利用したくない理由としては、「個票データを手元に置いて分析したい」、「アクセスの問題」、「夜間に使用したい」、「電算機を回しっ放しにしたい」等のニーズがあり、「利用時間に制限があること」などが挙げられております。

34ページ、一番下の部分になります。「今後、オンサイトで利用してみたい統計」を回答していただきました。回答は2件ございまして、そのひとつに業務統計であるNDB（National Data Bank = レセプト情報・特定健診等情報データベース）が挙げられております。

35ページの今後利用したい行政記録情報についてですけれども、36ページまでいただいた御意見に重複等々ありますけれども、色々な意見に分かれております。

38ページを御覧ください。マッチングキーのニーズについて御意見をいただいています。

39ページまで御意見を記載しておりますけれども、ニーズはあるとの意見が来ております。

いただいた意見で挙げたマッチングキーをまとめたものが40ページにございます。事業所番号や市区町村番号が多い状況ですけれども、利用できないマイナンバー等々を回答されたケースも、かなり多くなっています。皆様御存じのとおり、マイナンバーにつきましては、41ページ以降にございますとおり、提供については制限がございまして、現在利用できない状況になっております。

最後に、44ページを御覧ください。その他の意見として記載いただいたものになります。基本的には今までいただいた意見とかなり重複するものと、苦情的なものが多くなっております。

資料2に関わる説明については、以上です。

○阿部座長 ありがとうございます。

調査票情報の二次利用等を利用した方々に対しますアンケートを実施していただき、その結果を今、ざっと報告していただきました。委員の皆様から何か御質問あるいは御意見がございましたら御発言をお願いしたいと思います。

川口構成員、どうぞ。

○川口構成員 幾つかございます。まず8ページ目、手続、申請書類が大変だというところですが、恐らく「変数表」のようなものが送られてきて、「どの変数を使うかを選んでほしい」といったやり取りをさせていただくと、かなり手間が軽減できるのではないかなと思います。もしかしますと、既にそのようにしていただいているかもしれないですけれども。

あとは、一つ一つの変数について利用必要性の説明をつけないといけないというのが、実質的にはとても大変です。それは厚生労働省のルールというよりも、総務省で策定の統一的なガイドラインで、必要最低限の変数を提供する形になっていると理解しております。その部分を和らげる形にしないと、恐らく「申請するほう」も「申請されるほう」も大変なのは変わっていかないのでは、思います。

ですので、これは厚生労働省だけで対応できる問題ではないと思うのです。全体の枠組み自体も少し考え直していかないと、今まではマイクロデータを使った分析はそれほどされていなかったのですが、盛んになるに従って業務負担がどんどん増えていってしまうのかな、という印象を持っております。

それと、データの提供形式の部分、16ページです。これも厚生労働省だけで御対応いただくのは難しい部分があると思います。私どもの東京大学にCREPEというリサーチセンターがあるので、そちらのほうでレイアウト表を読み込むと、「Stata」にデータを読み込める。「Stata」の中に固定長のデータを読み込むファイルというのを、研究者の方はみんな作っているわけですが、恐らく同じものがたくさんリプリケートされているような状態だと思うのです。

そのレイアウト表から自動的に固定長のデータを読み取るようなファイルを、プログラムを作成するような自動化ソフトウェアを作って公開しているので、もしもリンクとかを貼っていただければ、研究者の方の負担も若干減らしていただけるのかなと思います。

それから、匿名データに関してです。私も総務省統計局が作っている統計で教育利用させていただいたことがあります。その際、匿名データを使うにしても、物理的な場所の確保とか、結構ハードルが高くなっています。すぐにダウンロードできる形になっているわけでもないですし、アンケート結果でも指摘があるように、秘匿処理がすごくされていて、都道府県がないとか、年齢が5歳刻みになっているといった問題があります。

これは先輩方が非常な御苦勞をされて、匿名化のデータを提供するというプロジェクトをやってこられたのだと思うのですが、これをそのまま進めていくのは結構難しいのではないかと、率直に言って思います。匿名化されていることによって、匿名化されている部分が結構厳しいので、研究利用がほとんどできないのです。では、教育での利用かといいますと、ダウンロードできるぐらいの気軽さで使える状態までなっていないと、教育利用でもなかなか利用は進まないのかなと思います。

そういったことで、結局匿名データはどちらにも使えないので、その利用件数が伸びないようになっている気がいたします。これから先、統計に割ける資源が限られている中で、匿名データに追加的な資源を投入すべきかどうかは、結構難しい判断ではないかなという印象を持っています。

次に、行政記録情報について、色々なことが記載してありまして、回答されている方々の御意見にとっても賛同するところが大きいのですが、特に私は、雇用保険のデータ

は非常に貴重だと思います。

例えば、労働力調査を使えば現在の就業の状態などが分かるのですが、どうしてもサンプルサイズが余り大きくないので、サンプリングのフレームが都道府県ではなく、全国の10地域になっており、場合によっては都道府県単位のケースもあって、細かい情報がなかなか分からない中で、雇用保険の情報があれば、失業された方とか再就職された方とか、あと、保険料が賃金水準にひも付いているので、賃金情報もありますし、パネルにもなっていますので、利用価値が非常にあると思います。

それで、労働市場のフローを把握する統計としては、雇用動向調査が既にあるのですが、雇用保険のデータがある程度地域レベルで集計し、個人の情報が暴露されない形で公開することを考えてもいいのではないかなという印象を持ちました。

最後に、マッチングキーで、マイナンバーが使えないのは当然かと思うのですが、労働統計ですと事業所系の統計がほとんどであって、法人番号は当然公開されていますので、法人番号を全ての統計に入れていただければ、マッチングキーとして使っていけるのではないかと思います。

以上です。

○阿部座長 ありがとうございます。

多岐にわたって御意見いただいたのですが、どうでしょう。他には何かありますか。

○大野審査解析室長 最初の1点目ですが、申請された件数の表については、データレイアウトをお渡しして、使用する項目を塗り潰していただくという方式で処理しております。つまり、年度によって調査方法が変わることもあって、カラムの位置がずれたり、表の項目の名前が変わることがありますので、申請手続においてはご苦勞おかけすることとなっているかもしれませんが、そうした形でやらせていただいております。

プログラムの自動化ですが、塗り潰したものが、古い調査データは紙媒体しか残っていないので1から作らなければいけません。新しいデータについては、一部塗ってあるところを読み込んで、どの項目が要求されているのかを感知するプログラムは少し作成しておりますけれども、まだ実用には追いついておりません。

他に、総務省と幾つか相談しなければいけない貴重な話題とご意見もいただきましたので、引き続き検討していきたいと考えています。

行政記録情報につきましては、業務統計のデータを基に、今やっているほかの調査の代替にならないかといった観点もございます。その辺も踏まえて検討していきたいと考えています。

以上です。

○川口構成員 どうもありがとうございました。

すみません、プログラムについて、説明が分かりにくくて非常に申し訳なかったです。我々が開発したプログラムは、レイアウト表で電子化されているものを読み込ませると、「Stata」というソフトウェアにデータを読み込ませるためのプログラムを吐き出してくれ

るソフトウェアを開発したので、もしよかったら使っていただければ、皆様の手間が省けるのではないかと思います。

○大野審査解析室長 ありがとうございます。

○阿部座長 今の川口構成員の御意見は、「Stata」だけで、ほかの「SPSS」や、アンケートに出ていたのは「R」と記載してありましたけれども、そういうものはない。

○川口構成員 ないです。「Stata」だけになります。「R」は多分開発してもいいのかもしれないですね。

○野口構成員 ただ、「Stata」に読み込ませれば変換ソフトはありますので。

○阿部座長 ありがとうございます。

そういうものは、変数名をどうつけるかというのは色々ありますから。

○川口構成員 レイアウト表に既存のものが入っているのです。それを読み込んで自動的に付ける形式になっています。総務省の担当者に確認したのですけれども、読み込みのファイルを研究者間で共有するのは問題ないそうです。ですので、「GitHub」とかで共有するみたいな仕組みを考えてもいいのかなと思いますね。自分たちが勝手にやればいい話ですけれどもね。

○阿部座長 ありがとうございます。

○野口構成員 変数名が決まっていないのですね。だから、「Stata」だとライブラリを作らなければいけないので、それがとても面倒です。ですから、変数名と定義を表みたいなものにあらかじめ示してもらえれば、もっと楽になりますね。それを読み込めば、一度に吐き出してくれるのですね。

○阿部座長 ほかに何かございますか。野口構成員、どうぞ。

○野口構成員 川口構成員がおっしゃったことに尽きるのですけれども、今、申請が混み合っている。それは素晴らしいことだと思います。先ほど川口構成員もおっしゃいましたが、皆様が個票を使って、きちんと国際的な質の高い研究をしつつ、それを日本の政策に資するようなインプリケーション（＝結果）が出てくるような研究をしようとされている。特に若い人は、大変皆さん優秀で、そういう人たちが増えてきたのはとても良いことですし、それによって申請が混み合っているのは、とても良いことなのです。

これは、厚生労働省に申し上げるのは何なのですけれども、トレードオフ（＝二律背反）だと思います。解決をするには2つしかなくて、審査の過程を簡易化するか、対応する人を増やすかのどちらかではないでしょうか。実は、私、自分のことで恐縮ですけれども、3年間の厚生労働科学研究費を使ってデータ利用の申請をして2年たっているのですけれども、まだデータが来ていないのです。ですから、残りの研究期間が3年間のうち1年ぐらいしかなくなっています。それぐらい今、承認申請が混み合っている。でも、人手が足りないのはよく分かっているのです、仕方がないなと思っています。

アンケートの御意見の中にも、それだけ重要な、非常に個人情報が入っているのだから、審査に1年ぐらいかけるべきだと、いうのはおっしゃるとおりで、アメリカなども本当に

凄惨な個人情報が入ったデータを申請したら、1年半ぐらい承認までは時間がかかります。医療の場合は特に承認までの審査に時間がかかるので、御意見として分かるのですけれども、トレードオフなので、処理人員を増やしていただけるか、川口構成員がおっしゃったように、例えばプロポーザル（＝提案）がしっかり記載してあれば、使うか使わないかに関わらず、一括してデータを提供するというのも、1つの案だと思います。

先ほど総務省の指針に従っているという話がありましたけれども、その指針を変えないと、いつまでも一個一個、変数をチェックするような審査のやり方をしていたら、手が足りないのです。ですので、これだけ皆さんがデータを利用するようになって来て、承認申請に時間がかかるようになってきますと、とにかくプロポーザルがきちんと記載してあれば、承認をして、全データを一括して渡すというやり方も良いのではないかと思うのです。

あと、川口構成員がおっしゃったことに尽きるのですけれども、匿名データとオンサイトは限界があると思っています。先ほど川口構成員がおっしゃったとおり、匿名データは研究には絶対に使えません。色々なところで細かい層といいますか、例えば都道府県とか市区町村とか企業とか、色々なところの固定効果と我々は呼んでいるのですけれども、そういった層が分析にとっても必要になってくるので、そういう層の情報が分からないと国際水準に達するような研究ができにくくなってきていますので、私も研究的なニーズとしては匿名データは先細りかなという気がします。

川口構成員がおっしゃったように、教育データとしては使い勝手が悪過ぎます。例えばアメリカで、いわゆるNDBの簡易版、アメリカですとメディケアのデータの簡易版になりますけれども、学生がいつでも使えるデータは、ネット上に存在するわけですね。それを学生がいつでも落としてきて授業とかレポートとかに使えるようになっています。ですので、匿名データに関しては、確かに限界があると思っています、もうやめた方がいいかなと私、正直思います。

オンサイトですけれども、私は第1回のときに意見を記載させていただいたのですけれども、オンサイトも限界があると思っています。去年、スウェーデンに行ったときにすごいなと思ったのは、統計省にサーバーがあるのです。統計省のサーバーに全部データが入っていて、研究者が申請すると、審査期間はそれなりに長いのですけれども、各研究者のオフィスで統計省のサーバーに繋がっていて全部分析ができるようになっていました。そのサーバーの中で必要な変数同士をつなぐ。向こうは、財務のデータも国民全員に対してすべて公表しているのですけれども、そのサーバーの中でつなげていて、研究者がデータを自分で持つことは一切ない様になっているわけです。要するに、クラウドになる訳です。研究者はそういう意味で比較的安心して使えるようになっています。

他方で、アンケートに罰則の話に記載されていらっしゃる方がいますけれども、海外では非常に罰則も厳しくなっています。医学者であっても、目的外で、自分がやっている手術のデータと統計データを突合して資料を作ったりすると、研究者でもあっても犯罪行為となってしまいます。ですから、非常に厳しいルールの中だけでも、研究者が非常に

安全に安心して使える環境が提供されているのです。

そのかわり、提供する側は使用する各機関からお金を取っているのです。例えばA大学の経済学部においてデータを使いたいという研究者がいれば、A大学経済学部が厚生労働省に年間契約をして、そんなに大きなお金ではないらしいですけれども、そのお金を統計省でマネジメントして、データを研究者に提供しているのです。ですから、今すぐにというのは無理でしょうけれども、将来的には研究者がデータを所持して管理するという世界から、省庁のクラウドあるいは省庁の持っているサーバーに研究者がアクセスして使うという環境になっていったほうがいいのではないかと、という気はします。

特に、今回、10月1日に法律が改正されて、NDB (National Data Bank = レセプト情報・特定健診等情報データベース) と介護保険のデータを全部結合したのです。これは素晴らしいことで、本当に世界に冠たる介護・医療データになったと思います。どこの国も持っていないデータだと思います。でも、そういったデータを今後、どの様に個人情報管理していくとか、研究者は使い勝手がよくて、アクセシビリティ (= 利用しやすさ) を高めるためには、発想の転換で、研究者にデータを渡すというのではなく、省庁が全部管理を行う。

でも、マージはマイナンバーとかのデータでは無くて、匿名化されたキー番号として、その搭載しているサーバーやクラウドに研究者が外から入ってアクセスして、データはそこから外へは出さない。作成された研究結果は誰かがチェックして、それでメールなりを送ってくれるのです。そういうシステムに変えていったらいいのではないかな、と思いました。

○阿部座長 ありがとうございます。

私、野口構成員がおっしゃった、審査の簡素化といったときに、プロポーザルがうまく記載されていればいいという話ですけれども、そこを厚生労働省だけがやるのではなくて、例えば我々みたいな研究者がピアレビュー (= 査読) するとか、それによって「これは安心して出せます」という確認をやって、あとは事務的なことを厚生労働省が行うとか、分担をそれぞれでやっていけば、今後、利用申請は多分増えてくると思うのです。ですから、専門的なところは専門家がチェックする、事務的なところは事務方がチェックするという分担の在り方を模索したほうがいいかもしれません。野口構成員がおっしゃったことで、そんな風に思いました。

○野口構成員 実際、NDBとかは有識者会議を作っていますね。NDBも介護も作っていて、今度、両者が合体した有識者会議で議論することになっていますけれども、専門家が入っているので、そういうものを立ち上げてもいいかもしれません。

○阿部座長 あと、匿名データについて、私も思っているのですけれども、十数年前から始まったのですか。当時の時代の状況と今は全然違うと思うのです。当時としては画期的で、今まで全然使えなかったのが、非常に使えるようになった。ただ、今になってみると、もう使用に耐えうるレベルではなくなってきたので、発想を変えていかざるを得ないので

す。これは、厚生労働省の問題だけではないので、必ずしも本検討会で結論が出るわけではないですけれども、そういう意見があったことを伝えたほうがいいかもしれません。

ありがとうございました。

何かありますか。

○大野室長 NDBについてですけれども、申請は随時受け付けていて、適宜開催されてはいますが有識者会議を行うことから、審査には時間がかかるようです。

○野口構成員 提供まで1年少しかかります。でも、それは仕方ないかなと思いますね。

○阿部座長 川崎構成員、どうぞ。

○川崎構成員 今までの御意見、極めて共感を持って聞かせていただきました。そういう意味で、つけ加えることは余りたくさんはないのですが、申し上げます。

まず最初に申し上げたいのは、こういう利用者へのアンケートをきちんとされたのはいいことだと思います。こうやって見える形で整理していかないと、先に進まないところがありますので、私はこういった情報をぜひ総務省の担当にも共有しながら、今日、出たような議論も含めて、どうやって解決していくのが良いか。厚生労働省の担当として解決できる問題と、解決できない問題がありますので、そこは明らかにしながら、制度をどのよう変えていくかを議論したほうが良いと思います。

それから、先ほど阿部座長がおっしゃったことと極めて関係が深いのですが、この制度ができた当初、相当慎重に運用してきた面があると思います。というのは、かなりセンシティブな個人情報が含まれているということで、匿名データを提供したこと自体で調査への影響がないとか、そういう意味での相当な慎重さがあったと思うのです。しかし、これだけデータの利用をやり慣れてくると、まずは利用者側の要求水準はどんどん上がってきますし、それから、国民も発足当時に比べれば、EBPMと言ったら知っている人は相当増えてきているわけで、そういう中では、こういったマイクロデータをどんどん使わないといけないという理解は深まってきている気がするのです。

ですから、そういう意味でも、次の制度に向けてどうするかという議論は、もうそろそろ私は始めたほうが良いと思います。そういう意味では、これは大事な出発点ではないかと思います。

それから、もう一点、制度の問題とは別に、データのリンケージです。例えば、先ほどお話にあった、事業所ないし法人でリンクする。これは、次の利用の展望を開くためにも非常に重要なことだと思います。今まで、残念ながらリンクに対して非常に慎重にやってきたのも、リンクによって全て丸裸に把握されてしまう怖さを感じられている観点なので、相当慎重にやっていく面がありますが、研究者がそれぞれリンクしながら、1からやっていると、相当労力もかかりますね。

ですから、そういう意味では、もっとリンクしやすい仕組みをつくらなければいけないと思いますので、リンクのニーズをもう少し丁寧に把握していけたらと思います。それは、厚生労働省だけではなくて、もしかしたら総務省も一緒になってやってもらったほうがい

いかなと思います。

もう一点、先ほど野口構成員がおっしゃったことにも関連するのですが、リソースの問題の制約はすごく大きいですね。これにどのように対応するかを考えますと、各官庁ごとに対応するのはどうしても無理があると思います。それを解決するには、官学でもっと連携しないといけないのではないかと考えています。ですから、官庁もリソースを投入しなければいけないけれども、今、それだけでなく行政のリソースがかなり削られている中で、研究の基盤を作るところにリソースを投入し、もっと研究的な発想で体制整備をやっていくことを考えなければいけないのではないかと思います。

そういう意味でも、この議論が大変大事なスタートになるかと思っていますので、ぜひこれを深めていけたらと思います。以上です。

○阿部座長 ありがとうございます。

川口構成員、何かつけ加えることはありますか。

○川口構成員 どうもありがとうございます。

川崎構成員が今、おっしゃられたことは、本当にそのとおりだなと思います。例えば、厚生労働省だと厚生労働科学研究費というものがあって、統計整備のプロジェクトみたいなものを研究者が請け負うことがあってもいいのかなと思いました。例えば、業務データを集計して出すことにしても、それなりに工数がかかることで、どの様に出すと、使えるデータになるかの判断は、かなり研究的な要素も入ってくると思います。例えば科学研究費の枠組みの中でそういう作業をやっていくことがあってもいいのではないかなと思いました。

もう一つ、これは野口構成員が先ほど御指摘になったリモートアクセスの件です。オンサイトセンターというのは、実を言うと、もう既にリモートアクセスにはなっているのです。和歌山県にサーバーが置いてあって、オンサイトからアクセスするようになっていますので、運用の仕方、アクセスする元を研究室も含めていただければ、多分技術的には問題は解決すると考えております。実際、ニュージーランドもそうなっているのです。ニュージーランドのある大学に行って滞在したのですけれども、今、データを使っているので部屋に入らないで下さいという張り紙を研究室の入り口に貼っている人がおりました。その人は、統計局のサーバーに研究室からアクセスして使っていました。

前出のスウェーデンのパターンと恐らく全く同じで、レビューされた、集計された統計の結果とかグラフとかしか外には出せない仕組みになっているようで、そこが混み合っているとは言っていましたが、そういう仕組みに長い目で見たら変えていったほうがいいのかなと思いました。

○阿部座長 ありがとうございます。

○大野審査解析室長 最後のオンサイトの件ですけれども、日本のオンサイト施設ですと利用する部屋には監視カメラがついていまして、利用者は（自分の）カメラの持ち込み禁止になっており、写真撮影などをして個人情報を持っていかれないようになっています。

今のお話しのニュージーランドも、「ただいま立入禁止です」という張り紙をして利用しているということは、それ程厳しくないのかもしれませんが、ある程度の秘匿措置をニュージーランドとしては要求しているということですか。

○川口構成員 端末が置いてある場所に関して、物理的な制約がかかっているということだと思います。

○大野審査解析室長 ありがとうございます。

○阿部座長 では、また何かあれば、再度御発言いただくことにして、資料3に移りたいと思います。資料3は4と一緒にですね。資料3、4、併せて説明をお願いいたします。

○大野審査解析室長 それでは、お手元の資料3を御覧ください。様式変更やコロナ対応などに伴い、平成31年度、令和元年度につきましては、前年度対比で審査日数が47に対して99日と、倍以上、非常に審査日数がかかっています。

1枚おめくりいただきまして、平成元年度の審査日数の分布を見たものが2ページですが、中央値で見ていただきますと84日となっておりますが、平均所得などと同様に、高いケースが平均を引き上げています。

さらに、3ページに審査にかかる時間を種類別に見たものがございます。統計法第33条第1項第1号と第2号を比べますと、第1号88日に対して第2号が119日と多くなっておりまして、主に研究者が使われているのは第2号のほうですので、研究者の方の日数のほうが内容が複雑なこともあり、長めになっている状況です。

それから、その下の種類別の第2欄ですけれども、通常新規ものであれば平均審査日数は132日。それから、調査事項の追加、目的変更、年次追加等々のものであれば79日。それから、利用期間の変更、利用者の追加、利用場所の変更等、軽微なものであれば33日程度かかっているという結果となっております。一度審査されたことがある場合には、審査時間は相対的に短くなるという傾向はございます。

引き続き、資料4について、戸田のほうから説明させていただきます。

○戸田企画官 それでは、お手元の資料4を御覧ください。

先ほどアンケートの中でもオンサイト施設の話にございましたとおり、統計法第33条に基づく調査票情報の二次利用に関しては、審査時間がかなり長くなっているということで、利用者の皆様に多大な御迷惑をおかけしているところでございます。オンサイト施設を活用したいという方、したくないという方、それぞれいらっしゃるということは承知しておりますが、一部活用したいという方がいらっしゃるというところ。あとは、オンサイト施設に関しては、平均的ではございますけれども、事前審査での時間が短縮されるというところ。あとは、利用者の選択肢を増やすという観点から、オンサイト施設への調査票情報の提供方針を定めていきたいと考えております。

方針を定めるにあたり、アンケートの結果にもございますように、研究者の方々、複数年次を活用して作業されることがございますので、こちらといたしましても、多くの年次に関してオンサイト施設で利用可能としたいところがございますが、オンサイト施設を運

営する統計センターでの事務作業がございまして、その調整を踏まえて、以下のように、利用可能となる統計調査や、その年次を徐々に増やすことを検討しております。

主な方針案の内容ですが、基本は、基幹統計調査につきましては、令和3年度中に10年分提供する。例えば、令和2年度までオンサイト施設で利用可能となっていない調査については、3年度中に10年分提供する。あとは、既にオンサイト施設で利用可能になっている調査がございまして、そうしたものについては、3年度中に10年分提供されるところと考えております。

例えば、基幹統計調査の中では、国民生活基礎調査がございまして、こちらは御案内のように、大規模年と中間年がございまして、大規模年と中間年を含めて10年分ということを考えております。

また、医療施設調査は、静態と動態という2種類の調査がございましてけれども、こちらはその両方を含めて10年分。

3年に一度実施されている患者調査もございまして、こちらは3回分を考えています。

また、一般統計調査につきましては、利用者のニーズの高い一般統計調査を検討し、該当する統計調査に関しては、令和3年度中に直近年次1年分を提供することを考えております。ただ、一般統計調査の中でも縦断調査も検討しておりますが、こちらは複数の年次を組み合わせて利用することがありますので、上記の提供年数にかかわらず、調査開始から直近年次まで提供することを考えております。

私からは以上です。

○阿部座長 ありがとうございます。

資料3が審査にかかる時間、資料4がオンサイト施設への調査票情報の提供方針ですが、何か御質問、御意見ございましてでしょうか。

オンサイト施設への調査票情報の提供ですけれども、最低10年分というのは、何故10年なのかがよく分からないのですが、提供できるものは全て出してもいいのではないかなと思うのですが、10年に区切ったり、あるいは一般統計ですと1年と区切っていますけれども、何か理由があるのですか。

○戸田企画官 基本としては、可能な限り利用可能としたいのですが、統計センターにおいて実際にオンサイト施設の中で利用可能とするためには、実際に実施府省からデータの提供を受けた際に、中身のチェックですとか、統計センターで使えるように諸々の事務作業があり、1年間でできる業務量には限界がございまして、その調整の中で、基幹統計は10年分、一般統計は1年分ぐらいが妥当という話になっているところです。

○阿部座長 例えば令和3年は10年分やって、それはデータとして置いてあるわけですね。令和4年になったら、そこから5年とかプラス1年とか、つけ足していくことは想定できるのですか。

○戸田企画官 そうですね。4年度以降もまだ方針としては定めていないのですが。

○阿部座長 もし順番にやっていくとしたら、そういうふうになるのですかね。分かりまし

た。とりあえず、事務的な作業量の観点から、基幹統計は10年分、一般統計は1年分ですべてやっていくということですね。

どうぞ、野口構成員。

○野口構成員 ひとつ質問してよろしいですか。私自身がこのオンサイト施設を利用したことがないので、イメージがつかないのですけれども、例えば私が国民生活基礎調査を使いたいとして、オンサイトのある研究機関に行って使うとき、固定票が出てくるのですか。どういうイメージなのでしょう。要するに、我々研究者がデータをもらおうと、固定票を切って、自分たちでデータクリーンアップするわけですね。オンサイト施設に行ったら、それが例えば「Stataファイル」となって置いてあるのでしょうか、それとも自分で1からその施設内で作るのでしょうか。

○戸田企画官 承認されたオンサイト施設で利用するとなった場合には、その施設でCSVのファイルが置いてあるので、そこで御自身の統計ソフトで読み込んで分析する。基本は、「R」とか「SAS」ですとか、幾つかの統計ソフトは既にインストールされているパソコンです。「Stata」とかはインストールされていないのですが、統計センターにあらかじめ依頼すれば、自分自身のライセンスを持ち込んでインストールすることは可能です。

○野口構成員 なるほど。ということは、そのオンサイトセンターのサイトにあるPCというか、物理的な中にデータが入っているのですか。

○戸田企画官 実際は、サーバー経由でシンクライアントを介されていますけれども、そのオンサイト施設の中で見ることができます。

○野口構成員 分かりました。それは、我々研究者が通常二次利用申請で借り受けるデータと遜色ないデータが入っているということですね。

○戸田企画官 第33条の現行の場合ですと、借り受けられるのは申請において必要な変数だけですけれども、オンサイト施設に関しては、基本、全てのデータが搭載されています。

○川崎構成員 これについて、いい悪いという意見ではなく、どんどんやっていただければいいので、特に意見があるわけではないのですが、どんどん推進していただくと、行く行くはこれを積み重ねていくと、いわば厚生労働省も含めた全府省のマイクロデータの巨大なアーカイブが誕生することになるはず。そうしますと、データ間のリンケージもしやすくなるわけですので、オンサイトにおけるデータの蓄積をきちんと積み重ねていただくのは、先ほど野口構成員がおっしゃったような、例えばスウェーデンが実際に対応しているのも、行く行くは夢ではなくなりますので、その布石としては非常に大事な動きだと思いますから、是非これは着実に進めていただけたらと思います。

○阿部座長 川口構成員、どうぞ。

○川口構成員 私は、阿部座長が先ほど質問されたことをとても心配しています。10年分提供して、それで終わりになってしまったら困るなど率直に言って思ったのです。ですけれども、戸田企画官からの回答で、それはあくまでも事務的な制約がかかっている、1年間に対応できるものが10年分までだという御回答だったので、また次の年になったら追加

的に10年分提供されたデータが出てくることが期待できると思いましたが、安心しました。

○阿部座長 この話とは違いますけれども、過去の調査結果の個票というのは、どこまで保存してあるのですか。ずっと保存してあるのですか。

○大野室長 現在はデータ化されたものは永年保存となっています。紙の調査票自体は捨てています。

○阿部座長 データは残っているのですね。

○川口構成員 miripoに行くと、各調査がどこまで電子媒体で残っているかは分かります。

○野口構成員 ちなみに、電子媒体で残っていないものは、もう保存されていないですね。

○大野室長 大昔、コンピューターがない時代のものは多分残っておりません。人口動態調査も昭和47年ぐらいからしかありません。

○野口構成員 その前の調査票は全部廃棄されたということですね。

○大野審査解析室長 電子という形では残っていないです。

○阿部座長 ありがとうございます。

あとは、資料3のほうですけれども、平均審査日数は、要因として調査票情報の変更ですか。それと、コロナウイルスと申請様式の大幅な変更が要因で、ほぼ倍になったという御説明でしたけれども。

○大野審査解析室長 平成30年に統計法改正がなされまして、その後、平成31年、元年にガイドラインの改正がございました関係で、申請時につける書類の数が増え、様式もかなり変わりました。それによりまして、前年まで、例えば地方自治体では毎年同じようなものを自分のところの業務用に作っておられたものや、毎年同じ申請で担当者の名前を少し変えるぐらいで済んでいたものについては、大幅に書き直しになってしまいます。加えて、そもそもつける書類自体が増えたということもございますので、かなり大きな影響があると考えています。

○阿部座長 利用申請が増えたから、1件当たりの審査日数が増えたという要因は、それほど影響していないのですか。

○大野審査解析室長 審査件数が増えたということで、さらにこれにより1件当たりの審査日数が増えてしまうともものすごい滞留が起きることになりますけれども、実際、滞留自体は起きているのですけれども、それほど長期の滞留になっておりません。申請のあった順番に処理していますので、審査件数が増えたことで1件当たりの審査日数が増えているという影響は余りないかと思っています。

○阿部座長 分かりました。

ほかに何かありますか。

では、よろしければ、資料5に移りたいと思います。資料5の御説明をお願いします。

○戸田企画官 では、資料5を御覧ください。「行政記録情報の利活用に関する考え方」でございまして、本検討会において、調査票情報の二次利用に係る審査に関する項目だけ

ではなく、行政記録情報の利活用の方針に関しても検討したく、利活用に関する考え方について御議論していただければと思います。資料5につきましては、現状の取組と今後の課題をまとめたものでございます。

行政記録情報と申しましても、定義が幅広いものですので、統計的活用になじむもの、なじまないものがある中で、統計活用になじむものに関しては業務統計として公表しているところです。

下側の表に、統計ユーザーとしての利活用という観点で、各種審議会・白書などで業務統計が活用されているといったところです。

今後の方針といたしましては、EBPMの取組を政府全体で進めているところですが、1つの例として、EBPMの推進に係る若手・中堅プロジェクトチームなど、EBPMの取組の中において業務統計等の活用を検討することを考えています。

以上が統計ユーザーとしての活用の考え方ですが、統計の部門に関しては、それだけではなくて、統計メーカーとして活用方針というところもございまして、そちらについても御説明いたします。

これまでメーカーとしても、母集団情報の整備や推計精度の向上、調査項目の代替といった観点で検討し、既に活用を進めているところです。

例えば、調査項目の代替に関しては、先ほどお話ございましたように、レセプトデータであるNDBが整備されたことに伴い、社会医療診療行為別統計は、一般統計調査から業務統計に変更したという経緯があります。

また、ここには記載しておりませんが、介護保険総合データベースが蓄積されたことに伴い、介護給付費等実態調査を業務統計として継続されることになっています。

その他、調査項目の代替につきまして、具体的な統計ですとか御意見、お考え等がございましたら、ぜひ御教示のほどよろしくお願ひしたいと考えております。

こちらから以上です。

○阿部座長 ありがとうございます。

これは、具体的に「こういうことができないか」ということを言ってもらってもいいということですね。

○戸田企画官 はい。

○川口構成員 よろしいですか。先ほど少し触れさせていただいたのですが、雇用保険のデータを集計していただいて出していただくと、完全に代替になるかどうかは分からないのですが、雇用動向調査と同種の情報が得られるのかなど。もちろん、雇用保険から外れていらっしゃる方もいらっしゃるのですが、そういう方をどの様に捕捉するかという問題は残ると思いますけれども、かなり精度が高い統計ができるのではないかと。既に、職安統計に関しては、統計として利用できるように集計されたものが公表されていますので、雇用保険に関しても同様のものがあるといいのではないのでしょうか。

アメリカの統計で、クォーターリー・センサス・オブ・エンプロイメント・アンド・ウェ

イジズという統計があるのですけれども、これはカウンティ（行政区分）のレベルで四半期ごとに雇用保険のデータを集計して、平均的な賃金と雇用の創出と喪失が分かる様になっています。例えば、日本でも市町村のレベルで四半期毎にデータを出すことをすると、市町村だけで1500あって、四半期ごとに6000のNがあるようなデータができますので、それだけでも実証分析みたいなものができると思うのです。

市町村の単位で集計されていて、さらに年齢や性別で区切っていくわけですがけれども、これはどんどん切っていくと、また秘匿の問題が出てくるのですけれども、そのバランスをうまく取りながら、余りにも小さなセルは落とすという作業をしながら集計したものを出していくと、かなり使い度があって、かつ個人情報暴露しないようなデータができると思います。

○阿部座長 ありがとうございます。

雇用保険に加入していればしっ皆ですから、週20時間以上の労働者については把握できる。

○野口構成員 去年、毎月勤労統計調査の不正問題の話がありましたけれども、毎月勤労統計調査に代わって全体的な動向が見られる調査は何かという話になった際に、雇用保険情報を神林博史（東北学院大学教養学部人間科学科）教授が使ったらいいのではないかとおっしゃっていて、それで比べてみようと言う話になっていました。実際はそこまで検討会では実証しないまま終わってしまったのです。しっ皆データがあるというのは強いですね。

○阿部座長 ただ、雇用保険だと労働時間が取れていないですね。

○川口構成員 収入は分かります。

○阿部座長 保険料率がわからないのでは。

○川口構成員 ですので、多分、毎月勤労統計調査は事業所レベルではパネルデータになっていて、それを共通事業所でインデックスを作っているんじゃないですかね。仕方がないので、事業所の中の労働者はどんどん入れ代わっていきますので、そうすると、同じ労働者の賃金の動き方は分からないです。ですから、雇用保険のデータがあれば、個人の労働者の賃金がどれくらい変動しているかが分かっていくので、チェーンにしていくと賃金上昇率がどのくらいかが、ある種、品質調整済みの価格指数みたいなものができるようになりますので、マクロ政策とかを考えると非常に重要だと思うのです。

日本銀行の展望レポートは、金融政策決定会合で配られる資料ですがけれども、毎回、毎月勤労統計調査の賃金率が載っていますね。明らかに金融政策決定のために使われているのですけれども、労働者の構成が変わっていることがアジャスト（＝調整）できていないので、その部分でできたデータが入っていくと、より質の高い政策運営ができるようになっていくと思います。

○阿部座長 毎月勤労統計調査の話を出してききましたけれども、難しいですね。労働者がどういう構成になっているか、事前に分からないで調査に充てていくので、結果とし

てどうなるか分からない中でやっていた。

○野口構成員 雇用保険情報には事業所番号が入っていないのですか。

○阿部座長 私、詳しくは分かりませんが、毎月勤労統計調査のような事業所単位みたいにはなっていません。ですから、企業がどういうふうに事業所ですと出してくるかに関わってくるのではないかと思います。

○鈴木統括官 一括して保険料を払いたいところはまとまっています。ですから、そもそもの設計が統計処理を前提に作っていませんので、そこが一番問題になっています。

○川口構成員 事業所単位でまとめてあるわけですね。

○阿部座長 ですから、企業があつて、事業所が幾つかあるような場合に、全部で出すところと分けて出すところがあるわけですね。出す際の手間とかに関連して。

○鈴木統括官 雇用保険情報は結局、保険のシステムですので、保険料を頂く、給与を払うというシステムになっているのです。ですから、保険料をまとめて企業で払いたいという方でしたら、その場合はまとまってしまいますし、事業所単位でやって払っているところも当然あります。そういう事情がありますので、統計処理的に使おうとすると、やや難しい部分が出てくるのです。

○野口構成員 さきほどの労働時間とか、どういう形態で働いているかというときに、保険情報なのに、もし事業所番号があれば、あるいは結合することが可能ではないかという話を去年の議論の中でしているのですが、かなり厳しいですね。

○阿部座長 それは難しいでしょうね。例えば、ある大学は、後樂園にキャンパスがあつて、多摩にもキャンパスがあるのですけれども、多分保険料は多摩で（まとめて）払っている。そうしますと、後樂園で働いている人たちも多摩で働いていることになりかねない。

○川口構成員 なるほど、所在地の問題ですね。

○阿部座長 川口構成員が最初におっしゃったように、市町村で適切に判ればいいのですけれども、大企業の事業所が、もしかしたらその市町村にはないともありえます。

○川口構成員 個人の番号というのは、雇用保険についているのですね。

○阿部座長 雇用保険番号についています。

○川口構成員 そうすると、個人の賃金がどのように動いているかは分かるということですね。

○鈴木統括官 ただし、短時間過ぎると、パートは入ってきませんので、抜けてしまいますから、バイアスとなってしまいます。

○阿部座長 マイナンバーとかが雇用保険でも使えればいいのですけれども、使えないですからね。ああいうものが使えますと、色々なことができそうだなと思います。フランスなどはそれが非常にうまくできていて、行政側もそれで非常に楽に行政運営しているところがあるのです。例えば、兼業しているような人が1つの仕事で失業したと言ってハローワークに届けに来ると、すぐに失業認定するのです。日本だと、多分本当に失業したかどうかを、かなり念入りにチェックするのですけれども、ほとんどチェックしないで、すぐ

に。

○野口構成員 分かってしまうのですね。

○阿部座長 ええ。なぜかという、マイナンバーがついているので、その人が本当に事業所から抜けたか、あるいは別のところで仕事に就いたか、全部分かってしまう。監視しているところがあって、失業しているのに就業していると分かると、モニターで赤い点減するようになっていきます。この届けはおかしいみたいなものまで見えるような仕組みになっています。統計においても、日本で言うところの事業所統計と個人統計のマッチングが全部できる、エンプロイ、エンプロイヤーデータになっていますし、色々なことが全部分かるようになっていきます。日本では、マイナンバーの制約が大き過ぎますね。

○野口構成員 スウェーデンも全く一緒です。全部つながっています。国民は、なかなか嘘がつけない仕組みとなっているのです。

○阿部座長 違う話になって、申し訳ありません。

ですから、どこまで使うかは、もう少し検討していかないといけないかもしれません。実際、どのように調査されているのかとか、行政側がどのようにデータを把握しているのかとか、そのあたりも少しこの場に出してもらって見ていった方がいいのではないのでしょうか。もしかしたら、具体的に「こういう行政記録は何とかならないか」という提案が、厚生労働省側にあれば、それを議論するのはあるかもしれません。ユーザー側としては、使えたらうれしいというのは、色々出てくると思います。

○川口構成員 そうですね。

恐らく先ほどの事業所と本社の話とかも、事業所・企業統計とかを使って、今、もうかなりフレームはしっかりしていますので、ほかの統計を使いながら接合していくと、事業所別に分けたりできないのかなとか、色々考えることはありますけれどもね。

○阿部座長 何か具体的にこれを検討しようとか、出してもらったほうがいいかもしれません。

○川口構成員 そんな感じがします。例えば、事業所と企業の話で、確かに混ざっているかもしれないですけども、例えば企業のレベルで出しているところが5%であれば、それを落として集計するということがあるし、程度問題だと思うのです。50%、50%だったら難しいかなという話になるでしょうし、その辺は私たちには資料が全然ないので、検討できないですね。

○戸田企画官 具体的には今ご覧いただける資料はないのですが、雇用保険の話は、担当課とも相談して、次回、この場で情報提供させていただくか、もしくは具体的に「ここまで統計的に活用できる」といったような話を詰めることができれば、また御報告するような形で進めたいと思います。

○川口構成員 ありがとうございます。

○阿部座長 分かりました。

かなり早いペースで議論が進みましたので、事務局が準備した資料に関しては、これで

全てと言うことでよろしいのですね。

全体を通して、何か皆様からご発言はありますか。特にないでしょうか。それでは、予定よりもかなり早いですけれども、このあたりで本日の議論は終えたいと思います。色々と御意見いただきまして、ありがとうございました。もし後日、何か言い足りないことがございましたら、事務局にお伝えいただければと思います。

それでは、今後の進め方について、事務局からお願いいたします。

○大野審査解析室長 本検討会につきましては、今年度中に2回の開催を予定しております。次回は、本日御指摘いただいた御質問、御意見に対する回答及び報告書の骨子案の検討を行います。次々回をもって最終回と考えております。次々回は、報告書骨子案に基づいた報告書案の検討を議題としたいと考えております。

日程等につきましては、追って御連絡いたします。

○阿部座長 それでは、以上をもちまして第2回「厚生労働省データ利活用検討会」を終了したいと思います。

お忙しい中、ありがとうございました。