匿名データ

総務省政策統括官(統計基準担当)

# ○地域情報など提供項目の情報が粗くニーズに合っていない

- ・現在の匿名化は、研究目的利用に耐えられるぎりぎりの水準の匿名化となっているところであるが、研究者からは、その匿名化でさえ地域情報が 粗い(「大都市」と「その他」の2区分のみ等)との不満が出ているところ
- ・しかし、データの厳密性を確保する必要のある学術研究等目的の利用において匿名性を担保するためには、ノイズ(匿名性を持たせるために意図的に混ぜる実際のデータとは異なる値のデータ)を加えるような匿名加工手法は採り難く、基本的には、地域情報や年齢区分等を一定程度まとめて区分する(コーディング)ような加工を行わざるを得ない状況にあり、<u>学術研究等を利用目的とするのであれば、加工の手法面で利便性を向上することは困難</u>

### ○提供されている匿名データの調査年次が古い

- ・匿名データの提供は、平成19年の法改正により導入された新たな仕組でもあったことから、万一のことのないよう、直近の調査結果を匿名化する ことは控えて来たところであるが、この点についても研究者からは不満の声
- ・しかし、学術研究目的の匿名データは、上記のとおりノイズを加えるような匿名加工手法が採り難いため、匿名処理の適否を判断するための膨大 なデータ検証等を実施する必要があり、<u>学術研究等を利用目的とする限り、調査結果の公表後、速やかにデータの作成・提供を行うことは困難</u>

### ○事業所・企業対象調査に対する利用ニーズに対応できない

- ・事業所・企業関係の研究を実施するに当たり、事業所・企業対象調査に係る匿名データに対する研究者のニーズは非常に強いところ
- ・しかし、事業所・企業対象調査に係るものは、どのように匿名化しても、学術研究等目的の利用に耐えるような匿名化では、所在地や業種、事業 規模のみで容易に調査対象が特定されるおそれがあり、研究上の有用性を確保しつつ匿名化措置を講じることが極めて難しいため、作成が困難

## オーダーメード集計

# ○オーダーした統計表の作成が必要になった都度、オーダーをし直す必要があり煩雑

- ・利用者が試行錯誤しながら研究を進める探索型研究を行う場合、オーダーした統計表の作成が必要になった都度、オーダーをし直す必要があり煩雑。 これに伴い発生する手数料も負担大
- ・しかし、調査票情報そのものは提供しないという制度の性格上、現在の仕組みを大きく変えることは困難



利用件数の低迷(年間利用件数は匿名データが30件台、オーダーメード集計が20件前後)