

第10回 ビッグデータ等の利活用推進に関する産官学協議のための連携会議 議事概要

1 日時 令和2年9月30(水) 10:00～12:00

2 場所 総務省第二庁舎 6階特別会議室

3 出席者

- ・ 構成員 高橋座長、庄司構成員、田原構成員、水野構成員
清家構成員（オブザーバー）
- ・ 審議協力者 横浜市立大学 佐藤教授
東京大学 竹内教授
JAXA 石田特任担当役
- ・ 事務局 総務省政策統括官（統計基準担当）付 統計委員会担当室

4 議題

- (1) 統計行政の最近の動きについて
- (2) ビッグデータ連携会議におけるこれまでの事例整理-公的統計作成への活用を中心に-
- (3) 地球観測データを活用したSDG指標の算出

5 配付資料

資料1 統計行政の最近の動きについて

資料2 ビッグデータ連携会議におけるこれまでの事例整理-公的統計作成への活用を中心に-
(参考: ビッグデータの公的統計への活用事例等について (多様化検討会資料1-3))

資料3 持続可能な開発目標 (SDGs) の進捗の測定に用いる指標について

資料4 SDG15.4.2(山地植生被覆指数) の試算について

資料5 観測データ利活用検証WG設置について

参考 第9回ビッグデータ等の利活用推進に関する産官学協議のための連携会議議事概要

6 議事概要

(1) 統計行政の最近の動きについて

- 総務省統計改革実行推進室の中村参事官より、資料1の説明が行われた

(2) ビッグデータ連携会議におけるこれまでの事例整理

概要は以下の通り。

- 総務省統計改革実行推進室の谷道企画官より、資料2の説明が行われた。
- 構成員からいただいた意見を整理し、議論した上で、今年中に報告書として発行する予定である旨を事務局から報告した。

主な質問・意見は次のとおり。

- ビッグデータは提供会社それぞれで精度の違いがある。また、最近ではフェイクデータであったという事例がでてきており、公的統計作成にそのまま使うというのは危険であると感じている。精度の検証だけではなく、ビッグデータの作成方法を把握し妥当性を検証する必要があると考えている。
- データ検証に関するポイントとしては、精度面のみでなくセレクションバイアスが生じていないかという代表性にも着目して記述するべきであるとする。また継続性という視点では継続的な提供のみでなく、今のデータと将来のデータが比較可能な状態で継続できるかということが重要である。
- データ作成プロセスが人的作業になっているところで問題が発生しているケースが見られるので、デジタル化を大きなテーマとして考えていくべきである。
- コロナ対応で厚生労働省がヤフーに検索情報の提供を依頼した際の事例も踏まえ、情報提供を依頼する手順や情報保護の考え方などの必要なコミュニケーションルールを整理しておくことが重要であると感じた。

(3) 地球観測データを活用したSDG指標の算出

概要は以下の通り。

- 総務省国際統計管理官室の津村管理官より資料3、JAXAの石田特任担当役より資料4、総務省統計改革実行推進室の谷道企画官より資料5の説明が行われた。
- 観測データ利活用検証WGの設置が承認された。

主な質問・意見は次のとおり。

- SDG 15.4.2 に関して、被覆データを変えた場合にどうなるかという各国の事例があれば教えて欲しい。
 - ▶ 日本と同様にFAOから各国に照会が行われているが、具体的な事例については提供されていない。

- P8でMGCI a 1 1を算出するためにKapos産地分類1～6の面積が必要だと思うが、これはFAOから提供されているか。
 - ▶ JAXA自身は自らが保有する面積データに基づいて計算をしているが、FAOから提供されているのは割合(%)情報のみであり、面積は提供されていない。

- Kapos産地分類6が一番大きいと予想されるが、分類5と6で何%ぐらいになるか。
 - ▶ おそらく30%～40%ぐらいだと記憶している。分類2が最も小さくて1.9k m²である。

- JAXAのデータでも、欧州のデータでも大体同じ結果が出ているという説明があったが、非植生全体が1%程度の規模の中で0.数%くらいずれるというのは割合的にかなり差があると見ることもできるが、どういうふうに解釈すればよいか。
 - ▶ JAXAの土地分類の精度が81.6%、ヨーロッパの土地被覆図データの分類精度は73.2%となっており、どちらもそれほど高い精度を持っているわけではない。ここで精度と言っているのは、国内の場合3万5,000地点の地上検証用データベースを保有しており、その正しい分類に対して、衛星データで分類した時の的中率を表している。

- JAXAのデータの特徴、精度の善し悪しの傾向は何か。
 - ▶ ヨーロッパの宇宙機関の検証方法をよく理解していないので、今回の取組の中で、ヨーロッパの土地被覆図データと比較しながら特徴を見いだしていきたいと考えているが、JAXAのデータは国内にフォーカスしているので、水田等の日本特有部分の分類の精度が高いと考えている。

- P12で局所起伏幅のアルゴリズムが不明のため反映できていないとあるが、国土地理院標高データ、Kapos山地分類データとも、Kaposのアルゴリズムを用いて処理していると考えて良いか。

- 地理院データはこの局所起伏幅の条件を除いて、計算している。
- 国土地理院の標高データの方が 10m 分解能と精度は高いので、K a p o s の標高データと差が出て来るのは当然である。富士山の北側のオレンジ色で見えているところが精度の差として表われているところだと想像している。
- 同じ意見であり、国土地理院のデータはK a p o s の局所起伏幅の条件を入れると、山地として評価され、結果的にK a p o s データのような図になると考えている。

以上