

推奨データセットについて



令和2年12月22日
内閣官房情報通信技術（IT）総合戦略室

推奨データセットとは

- 「推奨データセット」は、オープンデータの公開とその利活用を促進することを目的とし、政府として公開を推奨するデータと、公開するデータの作成にあたり準拠すべきルールやフォーマット等を取りまとめたものです。
基本編と応用編から構成されます。
 - (i) **基本編**：推奨データセットの対象データの中でも、特にオープンデータに取り組み始める地方公共団体の参考となるようなデータを基本編として位置付けています。
 - (ii) **応用編**：推奨データセットの対象データの中で、基本編以外のデータを応用編として位置付けています。応用編では、地方公共団体に限らず、民間事業者等の保有するデータについても対象とします。
- 必ずしも最初から全てのデータセット公開に取り組まなければならないというのではなく、本データセットを参考に、各団体において公開可能なデータセットから公開を進めていただくことを期待するものです。
- また、既に推奨データセットと同様のデータセットを公開している場合、フォーマットの共通化による利用者の利便性向上の観点から、推奨データセットとデータ項目等を合わせることが望ましいですが、必ずしも対応しなければならないものではありません。

推奨データセットの検討について

(1) データセットの選定

- 公開当初の推奨データセットは、世界最先端IT国家創造宣言・官民データ活用推進基本計画における重点分野や「地方公共団体アンケート」（平成28年12月実施）におけるニーズの高い分野を中心に先進地方公共団体の公開済データ等を参考にしつつ、地方公共団体がオープンデータの取組を開始するにあたって、公開することが推奨されるデータセットを選定。
- その後は、各施策の検討が具体化していく中で、公開すべきと判断されたデータや公開における標準的な形式を周知すべきと判断されたデータ、有効なオープンデータ活用事例等で活用されているデータセットなど、必要に応じて、データセットを追加する方針。
- 令和2年9月末に次の3つの推奨データセットを公開。
 - 学校給食献立情報
 - 小中学校通学区域情報
 - 支援制度情報

(2) データ項目の検討

- 選定したデータセットについて、先進地方公共団体の公開済データ項目等から、オープンデータフォーマット標準例のデータ項目を洗い出し、**共通語彙基盤**等を参考に、共通的な項目やオープンデータ利活用等の観点から必要と思われるデータ項目を絞り込んだ上で、区分（必須、任意など）やデータの形式等を設定。

推奨データセットの準拠について

- **データ項目定義書の必須項目に値が設定されている**
- **データ項目定義書の項目と並び順等の構成が同じであり、かつ、項目が削除されていない**
- **推奨データセットを拡張して使用する（データ項目定義書にない項目を追加する）場合は、最終列の項目として追加する**
- **データ項目定義書の各シートの注意事項、データ項目特記事項シートおよびファイル命名規則シートの内容に従って作成されている**
- **各項目が共通語彙基盤の値型とあっていることが望ましい**

直近の取組

第5回 オープンデータ官民ラウンドテーブルの実施

<オープンデータ官民ラウンドテーブルとは>

データの公開・活用要望に関する公募を行い、データの公開・活用を希望する方々と、データを保有する府省庁等が直接対話する場を設けることにより、よりニーズにマッチしたオープンデータ化やデータ利活用を促進し、データの価値向上と多様なサービスの創出に貢献するもの。

●今回は「健康・医療・介護・子育て」分野のテーマとして次の内容を検討

- ①介護サービス事業所・障害福祉サービス事業所に関するデータ
- ②放課後児童クラブ・児童館に関するデータ

●議論の結果

- ①は、推奨データセット「介護サービス事業所一覧」の形式にて厚労省よりデータを公開
- ②は、推奨データセット「子育て施設一覧」の対象に放課後児童クラブ・児童館を加えることで対応