

行政データにおける機械可読性に関するルール（概要）

- データ整形にかかる負担とコストを減らし、データ利活用を推進するため、行政データの機械可読性を担保するルールを整備
- 国の行政機関が今後作成するテーブルデータ（オープンにするデータ）を対象
- 各府省の取組を促すに際し、3段階の水準を設定

レベル	基準	具体的なルール	
レベル1	閲覧・転記可能 表形式データとして最低限の機械処理が可能	<ul style="list-style-type: none"> ・ ファイル形式はExcelかCSVとする ・ 1シート（ファイル）に複数の表を掲載しない ・ データが分断されていないようにする ・ データ本体と無関係な情報は含まない ・ すべての列に意味が推測できる項目名を入れる ・ スペースや改行等で体裁を整えない ・ 1セル1データとする ・ 機種依存文字を使用しない 	<ul style="list-style-type: none"> ・ 空白とゼロが明確に区別されている ・ オブジェクトを使用しない ・ 書式でデータの違いを表現しない ・ セルの結合をしない ・ 不要な行や列は非表示のまま残さない ・ 1行1データで表現する ・ 文字列にカンマが含まれているフィールドの値をダブルコーテーション（"）で囲む
レベル2	集計・分析可能 統計的な集計・分析作業が前処理なしで即座に実施可能	<ul style="list-style-type: none"> ・ 数値データは数値属性とし、文字列を含まない ・ データ内での項目名等の省略をしない ・ 各列が一意に識別可能な項目名を持つこと 	<ul style="list-style-type: none"> ・ 選択肢回答を標準化する ・ 選択肢列と「その他」の詳細記入を分離する ・ 数式を使用している場合は数値データに修正する
レベル3	連携・自動化可能 異なるデータセット間での結合・比較分析、プログラムによる自動処理が可能	<ul style="list-style-type: none"> ・ 項目名行から始まり、次行からデータ入力とする ・ 回答のコード表は別添とする ・ 数値データの同一列内に特殊記号（秘匿等）を使用する場合はその定義を明記する ・ データの単位を記載する ・ 時間軸の表記は標準化する 	<ul style="list-style-type: none"> ・ 地域コード又は正式な地域名称を表記する ・ 年によりフォーマットが著しく異なるようにする ・ データの定義や更新履歴を記載する ・ データは縦持ち形式とする

今後の予定

- 3月 各府省庁DX推進連絡会議・デジタル社会推進会議幹事会決定
4月～ 各府省において運用開始