平成25年3月21日 電子行政オープンデータ実務者会議資料

# オープンデータ流通推進コンソーシアムの取組と提言

平成25年3月21日 オープンデータ流通推進コンソーシアム 技術委員会主査 越塚 登 データガバナンス委員会主査 井上 由里子



# 1. 技術委員会の取組と提言

オープンデータ流通推進コンソーシアム 技術委員会主査 越塚 登





# (1) 技術委員会のミッションと実施内容・成果

#### ミッション

- (1) オープンデータを流通させるための技術仕様のあり方を検討
  - ▶ データ規格
    - □ オープンデータの表現モデル(オープンデータはメタデータとして表現される。それを表現するモデル)
    - □ オープンデータを表現するためのボキャブラリ(メタデータを理解するための標準的な辞書)
  - ▶ API規格
    - □ オープンデータを取得・交換するための手法に関する規定
- (2) 国際標準化のための作業検討
  - ▶ ITU-T, W3C等の適切な組織での国際標準化

#### 実施内容

- (1) 技術仕様のあり方に関する議論
  - ★ 仕様素案に関する意見交換
  - ▶ 類似する既存規格・取組の整理と整合性検討
  - ▶ 利活用・普及委員会などへの技術要件ヒアリングとその要件に対する技術仕様へのフィードバック検討
  - データガバナンスのライセンス表現や、それに基づくデータ処理に関する技術的な検討(2年目以降)
- (2) 国際標準化活動に関する意見交換
  - 1年目は、標準化の動向調査と体制に関するLアリング、意見交換

#### 今年度の成果

- (1) 技術ガイド案
- オープンデータ化する際の、データの形式に関する技術的な要求事項、およびそれを実現するための手順
- (2) コンソーシアム規格案
  - ▶ オープンデータ化のためのCSV形式データ規格案(平成24年度版)
  - 情報流通連携基盤システム外部仕様書案(平成24年度版)



# (2) 技術ガイド案 (概要)

- 幅広いアプリケーションやサービスが有効に利活用するために、国・自治体、企業等、さまざまな組織が保持するデータをオープンデータ化するための技術的な要求事項、およびそれを実現するための手順を示す。
- ▶ 技術的な内容や対象とする組織を限定するものではないが、扱う対象となる規定範囲は、主に国・自治体が公表するデータを念頭におく。
- ▶ 目的は、データによる情報伝達機能や、情報伝達の曖昧性排除ではなく、 即座にオープンデータ化できることを最重要視する。
  - データ作成の容易性に重点をおく。
  - オープンデータ化に関与する主体のデータリテラシー向上に応じて、技術ガイドの内容も見直す。



# (2) 技術ガイド案 ①表形式データ

#### ▶ テーブル全体に対する要件

- 1. 1つのデータシートには、1種類の表のみを含むべきである。
- 2. セルに、整形のためのスペース・改行、位取りのカンマを含めるべきでない。
- 3. 年の値には、西暦表記を備えるべきである。
- 4. 数値やタイトル・単位以外の情報を、セルに含めるべきではない。

#### セルに関する条件

- 5. すべてのセルが、他のセルと結合されているべきではない。
- 6. 値がない場合を除き、データセルが空白とすべきでない。

#### タイトルに関する条件

- 7. データの内容を示すタイトルは、1行で構成されているべきである。
- 8. データの単位が明記されているべきである。
- 9. データセルの内容・単位・記数単位を示すタイトルが、それぞれ別の行に記載されているべきである。
- 10. タイトルやデータ型は、一定の基準に従ったフォーマットで記述すべきである。

#### ▶ データの公開形式に関する要件

- 11. データセットは、オープンな標準データ形式で提供されるべきである。
- 12. データセットの属性や説明を表すメタデータを、XMLやRDFの形式を使ってフォーマルに記述すべきである。そのメタデータからデータセット本体へリンクし、たどれるようにすべきである。
- 13. データセットに含まれるデータ本体を、XMLやRDFの形式を使ってフォーマルに記述すべきである。



# (参考) 表形式データの技術ガイドを満たす形式 ①

	2010年の人口	2012年の面積
単位		km2
記数単位	1000	
北海道	5506	83457.06
青 森	1373	9644.7
岩手	1330	15278.89
宮城	2348	6862.12
秋田	1086	11636.3
山形	1169	6652.11
福島	2029	13782.76
茨城	2970	6095.72
栃木	2008	6408.28
群馬	2008	6362.33
埼玉	7195	3767.92
千葉	6216	5081.93
東立	13159	21 03.97

ガイド7: タイトルは1行 ガイド8: 単位を明記

ガイド9: 内容(1行目)・単位(2行目)・記数単位(3行目)を示すタイトルを別々の行に記載

ガイド11: オープンな標準フォーマット(CSV)で提供

CSV形式で出力

福岡	5072	4047.2
佐賀	850	2439.65
長崎	1427	41 05 .75
熊本	1817	7267.89
大分	1197	

宮崎

沖縄

鹿児島

ガイド1: 1つのデータシートに1種類の表

ガイド2: セルに整形のための改行・空白・カンマを含めない

ガイド4:数値やタイトル・単位以外の情報を、セルに含めない

ガイド5: すべてのセルが、他のセルと結合されない ガイド6: 値がない場合を除き、データセルが空白でない ,2010年の人口,2012年の面 積

単位,,km2

記数単位,1000,

北海道,5506,83457.06

青森,1373,9644.7

岩手,1330,15278.89

宮城,2348,6862.12

秋田,1086,11636.3

• • •

熊本,1817,7267.89

大分,1197,5099.58

宮崎,1135,6794.69

鹿児島,1706,9044.66

沖縄,1393,2276.64

レベル1を満たす形式



# (参考) 表形式データの技術ガイドを満たす形式②

ガイド7: タイトルは1行

ガイド9: 内容(1行目)・記数単位(2行目)を示すタイトルを別々の行に記載

年	年(和暦)	東京都の入口
記数単位	1 (10/0/	1000
1920	大正9年	3699
1925	大正14年	4485
1930	昭和5年	5409
1935	昭和10年	6370
1940	昭和15年	7355
1945	昭和20年	3488
1950	昭和25年	6278
1955	昭和30年	8037

ガイド11: オープンな標準フォーマット(CSV)で提供

CSV形式で出力

1985	昭和60年	11829
1990	平成2年	11856
1995	平成7年	11733
2000	平成12年	12064
2005	平成17年	12576
	1990 1995 2000	1985 昭和60年 1990 平成2年 1995 平成7年 2000 平成12年 2005 平成17年

ガイド1: 1つのデータシートに1種類の表

ガイド2: セルに整形のための改行・空白・カンマを含めない

ガイド3: 年の値に西暦表記を備える

ガイド4:数値やタイトル・単位以外の情報を、セルに含めない

ガイド5: すべてのセルが、他のセルと結合されない ガイド6: 値がない場合を除き、データセルが空白でない 1920,大正9年,3699 1925,大正14年,4485 1930,昭和5年,5409 1935,昭和10年,6370 … 1985,昭和60年,11829 1990,平成2年,11856 1995,平成7年,11774 2000,平成12年,12064

2005,平成17年,12576

年,年(和暦),東京都の人口

記数単位,,1000

レベル1を満たす形式

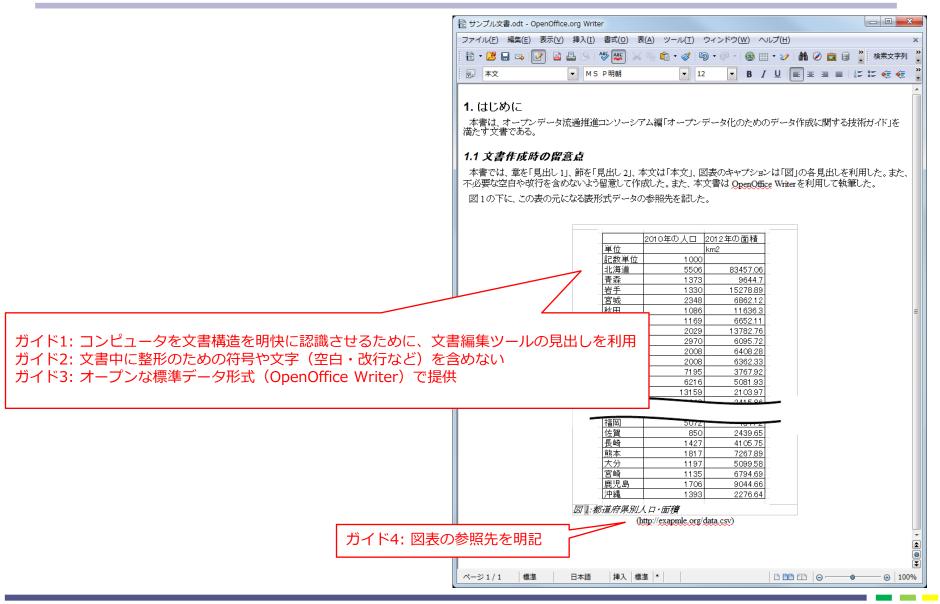


# (2) 技術ガイド案 ②文書形式データ

- 1. 文章に存在する部・章・節・図表などの構造が、コンピュータが明快に認識できる形で記述されているべきである。
- 2. 文章形式データに、整形のための符号や文字(空白・改行など)を含めるべきではない。
- 3. 文書形式データは、オープンな標準データ形式で提供されるべきである。
- 4. 文書形式データが図表を含む場合、それらを構成するガイド11までを満たす表形式 データが添付されているべきである。
- 5. 図表やグラフとそのキャプションの対応関係がわかるように、文書形式データを構成すべきである。
- 6. 文章形式データの属性や説明を表すメタデータを、XMLやRDF等の形式を使ってフォーマルに記述すべきである。そのメタデータは、文書形式データ本体へリンクすべきである。



# (参考) 文書形式データの技術ガイドを満たす形式





# (2) 技術ガイド案 ③地理空間データ・リアルタイムデータ

# ▶ 地理空間データの技術ガイド案

- 1. 地理空間データは、測地系を明記すべきである。
- 2. 地理空間データは、地理空間情報を記述するために広く使われているフォーマットで記述されるべきである。
- 3. 地理空間データの属性や説明を表すメタデータを、XMLやRDF等の形式を使ってフォーマルに記述すべきである。そのメタデータは地理空間データ本体へリンクするべきである。

# リアルタイムデータの技術ガイド案

- 表形式データのガイド11まで、地理空間データのガイド2までに準拠した形式のデータを、ファイルとして取得できるべきである。
- 2. リアルタイムデータの最新値・差分を取得する手法が提供されているべきである。
- 3. リアルタイムデータの最新値や差分を取得するための、メタデータ記述に対応したデータ取得規約が提供されているべきである。 または、メタデータ記述されたリアルタイムデータを取得する手法が提供されているべきである。



# (参考) 技術ガイド案のレベル

- 各ガイドの要求度に応じて、「レベル」を設ける。
  - ▶ レベル1
    - レベル1は、オープンデータが満たすことを強く推奨する要件である。
    - レベル1の要件は、データ本体の中身を修正したり手を加えたりすることなく、かつ容易にそのデータを扱うプログラムが書けることを目的とする。

#### ▶ レベル2

- レベル2は、オープンデータが満たすことを推奨する要件である。
- ▶ レベル2の要件は、データを取得した利用者が、そのデータの項目(タイトル)・値・単位、あるいは構造を正しく解釈した上で、データを扱うプログラムを書けることを目的とする。

#### ▶ レベル3

- レベル3は、オープンデータが満たすと望ましい要件である。
- レベル3の要件は、データ本体を修正したり手を加えたりすることなく、個別に作成された複数のデータを統合して利用(マッシュアップ)するプログラムが書けることを目的とする。
- そのために、データセットが含む個々のデータ本体の型や、データセット全体の構成や構造が、フォーマルに定義されていることを、レベル3の要件とする。
- レベル2・レベル3を満たすデータについては、アプリケーションが自動的にデータを取得・ 解読することにより幅広いアプリケーションでの利活用が期待される。

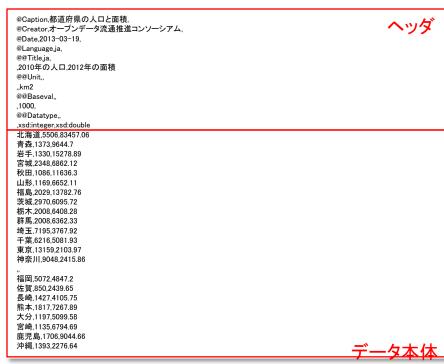


# (3) コンソーシアム規格案 [規格1] オープンデータ化のためのCSV形式データ規格案

- 表形式データのタイトルやデータ型をフォーマルに記述するための、フォーマット規格の1つ。
  - 本案ではCSVデータについて規定する。
  - 他の形式(TSV、Open Document Format、XML等)については、必要に応じて追って規定 する。
- 表形式データのキャプション・タイトル・単位などのメタデータを、データセットのヘッダとして記述する。

ヘッダ名	意味
@Caption	データセットのキャプション
@Creator	データセットの作成者
@Date	データセットの公開日
@Language	データセットの基本言語
@@Title	タイトル
@@Unit	カラムの単位(物理単位・貨幣単位)
@@Baseval	カラムの記数単位
@@Datatype	カラムのデータタイプ

付与するヘッダ



本規格に基づくデータ記述例

# (3) コンソーシアム規格案



# [規格2] 情報流通連携基盤システム外部仕様書(第1版)(案)

# ▶ 概要

- 意味(セマンティクス)情報を持つオープンデータを扱う際の、データ表現形式とAPI に対する規格案
- 自分のデータをオープン化するときの手法やシステムを例示する意味を持つ。

# ▶ 規定範囲

- ▶ API規格
  - さまざまな利用シーンを対象とする。
    - □ データの直接取得・リアルタイムデータの取得・操作など。
    - □ 識別子を読み取ることをトリガとしてデータを取得するような利用法にも対応する。
  - ▶ SPARQLベースのAPIとRESTベースのAPIを提供する。
  - !利便性よりも、むしろ機能性や意味の正確性などに重点をおく
- データ規格
  - モデルはRDFに準拠する。
  - ▶ 広く使われているボキャブラリは、そのまま利用する。 → 次頁
  - 規定されていないが必要であるボキャブラリは、新規に定義する。

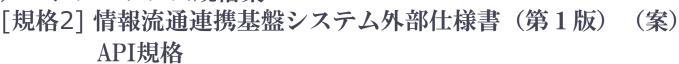
# (3) コンソーシアム規格案

# [規格2] 情報流通連携基盤システム外部仕様書(第1版)(案) ボキャブラリ規定(既存のボキャブラリ)



名称	規定範囲	ネームスペース	ボキャブラリ例
RDF基本構造	RDFでデータ構造を表現するための基本的なボキャブラリ。	http://www.w3.org/1999/02/22-rdf- syntax-ns#	rdf:subject(主語), rdf:predicate(述語)
RDFスキーマ	ボキャブラリを定義するためのボキャブ ラリ。	http://www.w3.org/2000/01/rdf-schema#	rdfs:subClassOf(サブクラス), rdf:range(値域), rdfs:subPropertyOf(サブプロパティ),
OWL	オントロジを記述するためのボキャブラリ。	http://www.w3.org/2002/07/owl#	owl:sameAs(同義), owl:inverseOf(反意)
ダブリンコア基本 要素	書誌情報を記述するためのボキャブラリセットであるが、Webリソースの属性を記述するために広く用いられている。 ISO 15836にて標準化。	http://purl.org/dc/elements/1.1/	dc:title(名前), dc:description(説明文) , dc:creator(作者), dc:format(メディアタイプ)
DCMI語彙	ダブリンコア基本要素を拡張し、その意味を細分化したボキャブラリ。	http://purl.org/dc/terms/	dcterms:alternative(代替タイトル), dcterms:audience(対象としている 利用者)
FoaF	人や組織に関する情報をRDFで記述するためのボキャブラリ。	http://xmlns.com/foaf/0.1/	foaf:familyName(姓), foaf:givenName(名), foaf:age(年 齢)
SKOS	シソーラス、分類体系、件名標目表、タクソノミー、フォークソノミー、およびその他の同種の統制語彙のような概念体系の基本構造や内容を表現するためのモデルを提供するボキャブラリ体系。	http://www.w3.org/2008/05/skos-xl#	skos:definition(ボキャブラリの定義文), skos:broader(広義である), skos:note(ボキャブラリ定義に関するノート)
geoSPARQL	位置や形状に関するボキャブラリや、空間演算を行うための関数ボキャブラリが 定義されている。	http://www.opengis.net/ont/geosparql# http://www.opengis.net/ont/sf# など	sf:wktLiteral(Well-Known Text規格の地理情報), sf:gmlLiteral(GML規格の地理情報)
DCAT	データセットを記述するためのボキャブラリが定義されている。	http://www.w3.org/ns/dcat#	dcat:theme(データセットのカテゴリ), dcat:accessURL(データにアクセス するためのリンク先情報)

# (3) コンソーシアム規格案





- ▶ RESTベースのAPIとSPARQLベースのAPIを提供する。
  - RESTべースのAPIでは、データ検索・取得コマンドのレスポンスにRDF/XML等を利用することにより、RDFモデルに基づくデータとの互換性を保つ。
  - Streams APIに対応することにより、リアルタイムデータの送受信にも対応する。

機能名	概要	提供する理由
SPARQLベースのAPI		
SPARQL-based Command	SPARQL 1.1準拠のデータ操作APIを提供する。	RDFモデルに基づくデータに対するアクセスを 提供するため。(既存の技術)
RESTベースのAPI		
Traceability/Realtime Data Management Command	トレースフォワード・トレースバックを含む、トレーサビリティに代表されるイベントを管理する機能。	対象応用分野の1つであるトレーサビリティで頻繁に発生する、トレーサビリティイベントを効率的に扱うため。
Geographical Data Management Command	GIS等地理情報処理を必要とするデータ検索・取得・操作機能。	実物や実環境を扱う上で、地理情報演算は重 要であるが、演算が複雑であるため。
Notification Management Command	データの登録・更新をNotificationとしてデータ利用者のシステムにコールバックする(Notification)仕組み。	センサ情報の共有を目指すcomsにも同様の機能が提供されているため。
Security Management Command	ユーザ・グループの管理と、データのアクセスルールに関する機能。	民間データでは課金処理を要するケースがあり、 ユーザ管理は必要である。comsやSODAにも ユーザ管理機能が提供されている。
Vocabulary Management Command	ボキャブラリ情報の登録・検索・取得に関する機能。	モバイル環境や組み込み機器に対応するため、 データの一部を登録・取得できる機能を提供す
Triple Management Command	RDFモデルの主語・述語・目的語からなる基本データの登録・検索・取得に関する機能。	<b>る</b> 。
Identification Resolution Command	IDをキーとしてデータを登録・検索する機能。	識別子を読み取ることをNotificationとしてデータを取得する利用法に対応するため。



# (4) 電子行政オープンデータ実務者会議への提言(技術関係)

- 技術委員会は、これまでの検討をもとに、以下を提言する。先行省庁での試行をふまえつつ、実務者会議におけるオープンデータ推進に向けたルール整備の参考としていただきたい。
  - 1. 表形式データや文書形式データ、地理情報データ、リアルタイムデータをオープンデータ化する際のデータ構造・形式に、技術委員会で検討した技術ガイド案、および規格書案を参考にしていただきたい。
  - 2. 政府がオープンデータを公開するAPIや高度なデータ形式について検討する際には、技術委員会で検討した平成24年度版「情報流通連携基盤の外部仕様書(第1版)(案)」を参考にしていただきたい。
  - 3. 政府の保有する情報をオープンデータ化するためには、フォーマットに関するルールのほかにデータ公開ポリシ作成や、データの信頼性確保のためのガイド制定が必要である。また、マニュアルやツール・テンプレートの整備、職員への研修などの周辺環境を整備する必要がある。これらの点についても並行して検討していただきたい。

# 2. データガバナンス委員会の取組と提言

オープンデータ流通推進コンソーシアム データガバナンス委員会主査 井上 由里子





# (1) オープンデータ戦略推進のためのデータ利用ルールの在り方

- ▶ 国等が保有する公共データ(以下「公共データ」という。)を、広く国民が活用しやすくするためには、国等の著作権の取扱いを明確にし、誰もが自由に利用できるようにする環境を整備することが急務である。(なお、数値データ、法令等には著作権がないことに留意)
- 上記の検討に当たっては、公共データは税金で作成されたものであり、著作権の有無にかかわらず国民共有の財産であるという観点を十分に踏まえる必要がある。

課題解決の方向性	具体的内容と課題
(1)立法によるパブリック ドメイン化	<ul><li>○米国の立法例に倣い、国等が保有する公共データには著作権が発生しないよう著作権法を改正する。</li><li>※著作権法は、創作を奨励するためのインセンティブとして著作権という独占権を与える制度であり、国民の税金を用いて作成される公共データの創出プロセスに著作権がインセンティブとして働く余地はない。</li><li>[課題]著作権法の改正には長期間の検討が必要。</li></ul>
(2)国等の著作権の放棄	〇国等が著作権を有することを前提としつつ、国等が自らの著作権を放棄する。 [課題]著作権も国・地方公共団体の財産権を構成しうるものであり、国有財産法、 財政法、地方自治法、補助金等適正化法等との関係において、権利放棄 を行うことが適当かどうか検討が必要。
(3)二次利用促進のため のパブリック・ライセン スの活用	○欧州の一部や豪・ニュージーランドの例に倣い、国等が著作権を有することを前提としつつ、二次利用を促進するために著作権の不行使を内容とするパブリック・ライセンスを採用し、利用できる範囲を利用者にわかりやすく表示し、個別の交渉なしにオンラインで処理できるようにする。  [メリット]早期の実現が可能  個別交渉無しに簡便な権利処理が可能であり、実質的に(1)及び(2)と同等の結果を実現できる。

データガバナンス委員会では、オープンデータ戦略の早急な推進が求められていることを踏まえ、国に著作権のある公共データについて、(3)の二次利用促進のためのパブリック・ライセンスの活用を前提として検討を行った



# (2) 国内での採用が考えられるライセンスの検討

諸外国で利用されているパブリック・ライセンスを比較検討すると、以下のようになる。

ライセンスの特徴	Open Government Licence	Open License (LICENCE OUVERTE)	Open Data Commons License	Creative Commons License
二次利用促進のための制約の少ないラ イセンス	0	0	0	0
出典表示を義務としている	0	0	0	0
一部の条件の選択を認める柔軟性(※)	△ (商用のみ)	×	△ (改変時の承継の 有無のみ)	0
無保証に対応していること	0	0	0	0
他のパブリック・ライセンスとの互換性有り	0	0	0	0
	×	×	0	0
採用の実績	イギリス	フランス	パリ市、ドイツ	ドイツ、ニュージーランド、 オーストラリア



\* CC-BYを付与することができない場合でも、それ以外のクリーエイティブ・コモンズ・ライセンス(商用利用無しとするCC-BY-NC、改変利用無しとするCC-BY-ND等)の採用が可能になる。

- オープンデータ戦略を推進するためのパブリック・ライセンスに求められる条件としては、広く自由利用を認める内容であること、出典表示を義務づけることの2点を軸として、場合によって条件の一部の選択ができること、無保証に対応していること、互換性のあるライセンスであること等が挙げられる(諸外国での採用実績も考慮要素となる)。
- ▶ 上記の観点からは、クリエイティブ・コモンズ・ライセンスを採用し、その中で利用にあたって最も制約の少ないCC-BYライセンスを原則とするのが望ましいと考えられる。



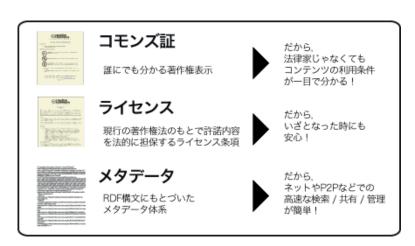
# 参考1. クリエイティブ・コモンズ・ライセンスの概要

## ▶ 概要

- クリエイティブ・コモンズとは、クリエイティブ・コモンズ・ライセンス(CCライセンス)を提供している国際的非営利組織とそのプロジェクトの総称。
- 2001年に組織が設立され、2002年にアメリカにおいて、ライセンスの最初のバージョンが公開されている。(日本では2004年に最初のバージョンが公開)
- ▶ CCライセンスはインターネット時代のための新しい著作権ルールの普及を目指し、様々な作品の作者が自ら「この条件を守れば自分の作品を自由に使って良い」という意思表示をするためのツール。
- CCライセンスを利用することで、作者は著作権を保持したまま作品を自由に流通させることができ、 受け手はライセンス条件の範囲内で再配布やリミックスなどをすることができる。
- ▶ 機械可読性があるので、CCライセンスの付された著作物を検索することは容易である。

### ライセンスの特徴

- CCライセンスは三つの要素によって構成されている。
  - ① 「コモンズ証」: 法律に詳しくない人でもライセンスの内容がすぐに理解できる簡潔な説明文
  - ② 「利用許諾」(ライセンス原文):同じ内容を法律の専門家が読むために法的に記述したもの
  - ③ 「メタデータ」:検索エンジンが利用するための、 作品そのもの(コンテンツ)に付された機械可 読性のあるデータ



【出典】 クリエイティブ・コモンズ・ジャパン ウェブサイト( http://creativecommons.jp/licenses/ )をもとにデータガバナンス委員会事務局作成



# 参考2. クリエイティブ・コモンズ・ライセンスの種類・評価

1.1.5*	メージ ライセンス名称		公共データに適		
イメーン		出典表示	商業利用	改変	用する上での当 委員会の評価
© <u>()</u>	表示 2.1 日本 (CC-BY 2.1 Japan)	必須 (タイトル、全ての著 作者、URLを表示)	許可	改変を許可する(※)	最も利用範囲が広い ので、推奨。
© S BY NC	表示-非営利 2.1 日本 (CC-BY-NC 2.1 Japan)	必須 (タイトル、全ての著 作者、URLを表示)	許可しない (改変されたもの の商業利用も許可 しない)	改変を許可する(※)	電子行政オープンデー タ戦略では、「営利目 的・非営利目的を問わ ず」としている。
© O B	表示-改変禁止 2.1 日本 (CC-BY-ND 2.1 Japan)	必須 (タイトル、全ての著 作者、URLを表示)	許可	許可しない	改変(二次利用)を行う ことができない。
BY NC ND	表示-非営利-改変禁止 2.1 日本 (CC-BY-NC-ND 2.1 Japan)	必須 (タイトル、全ての著 作者、URLを表示)	許可しない	許可しない	電子行政オープンデー タ戦略では、「営利目 的・非営利目的を問わ ず」としている。
BY SA	表示-継承 2.1 日本 (CC-BY-SA 2.1 Japan)	必須 (タイトル、全ての著 作者、URLを表示)	許可	改変を許可するが、改変されてできた二次的著作物は、このライセンスと同一のライセンスを採用すること。(※)	同一ライセンス同士で なくては結合できない ため、利用しづらい。
EY NC SR	表示-非営利-継承 2.1 日本 (CC-NC-SA 2.1 Japan)	必須 (タイトル、全ての著 作者、URLを表示)	許可しない (改変されたもの の商業利用も許可 しない)	改変を許可するが、改変されてできた二次的著作物は、このライセンスと同一のライセンスを採用すること。(※)	同一ライセンス同士で なくては結合できない ため、利用しづらい。

※ 著作者の人格権を侵害する改変は許可しない

【出典】 クリエイティブ・コモンズ・ジャパン ウェブサイト(http://creativecommons.jp/licenses/)等をもとにデータガバナンス委員会事務局作成



# (3) 利用ルールを策定する上での留意点(例)①

▶ CCライセンス(CC-BY)を活用した利用ルールを策定するに当たっては、以下のような点について留意する必要がある。

#### ○ 著作権がない数値データや簡単なグラフ・表の取扱い

- ▶ 数値データや簡単なグラフ・表には著作物性がないが、CCライセンスは、ライセンスの対象が著作物であることを前提にしている。
- 著作物性のない部分を含む公共データにライセンスを付与した場合、以下のような問題がある。
  - 本来は著作権がないものであるにも関わらず著作権があるかのような誤解を与えるおそれがある( 負のラベリング効果)。
  - ▶ クリエイティブ・コモンズのCC-BYを用いると、出典の明示が法的義務であるかのような誤解を与える恐れがある。

#### 【考えられる対応方法(例)】

- (1) 著作権のない部分を個別に特定した上で、当該部分に著作権のない旨の表示を付す。
  - ・限界事例では著作物性の判定は容易ではないという問題がある。
- (2)個々の情報の著作物性の判定はせず、著作権が無い部分についてはもともと自由に利用できるものであり出典明示は提供者側の希望であることを利用ルールに説明する。



# (3) 利用ルールを策定する上での留意点(例)②

#### 〇 第三者の権利の取扱い

- ▶ 公共データに、第三者が著作権その他の権利(肖像権、商標権等)を有する部分が含まれている場合 (例:白書)、以下のような課題がある。
  - ▶ 第三者が権利を有する部分については、国が利用を許諾することをできるわけではない。
  - ▶ 公共データの利用者にとって、第三者が権利を有する部分がわかりづらい。

#### 【考えられる対応方法(例)】

- (1)利用者が第三者の権利との抵触を回避することができるよう、公開にあたり、第三者の権利について調査・確認を行い、当該部分を削除したり、当該部分を明示してCCライセンスの適用除外となる 旨を明記する。
  - 特に過去の公共データについては、権利関係の確認には多大な労力がかかる
  - ・新たに作成される公共データについては、たとえば、調査研究を委託する際、作成される報告書の 二次利用を可能とする内容の契約として、予め権利の集約・明確化を図ることが可能である。
- (2)第三者の権利との抵触の回避は利用者の判断と責任に委ね、公開に際して、権利関係の調査・確認は行わない。
  - ・過去の公共データについては、利用の二一ズ等を踏まえ、費用対効果の観点から取扱いを決める ことが妥当であると考えられる。
- ※ データを提供した第三者との契約によって国の再利用の許諾権原が制限されている場合の取扱いに ついては、当該契約の内容を精査する必要がある。

#### ○ 個別法による制約のある公共データの扱い (例:気象業務法、測量法等)

著作権法以外の個別法に二次利用を制約する根拠がある場合には、利用ルールにわかりやすく説明 することが必要となる。



# (4) ケーススタディの検討フロー ①全体像

- ▶ データガバナンス委員会では、情報通信白書(ウェブ版)を対象に、①CCライセンス(CC-BY)を採用する場合の課題の洗い出し、②第三者の権利関係の確認等にかかる労力等の調査を目的として、ケーススタディを実施した。
- ▶ 情報通信白書(ウェブ版)の全体をチェックし、①著作物性の有無、第三者が権利を有する可能性のある部分等の要検討箇所の抽出・分類、②第三者の許諾の確認をした上で、③CC-BYが適用されない部分を明示して公開することを予定している。

# 公開することを予定している。 情報通信白書 (ウェブ版) ※作業 手順 ・作業手順に則り、該当箇所を抽出・分類。

# 3 CC-BYが適用されない部分を明示の上、公開・②の結果に基づき、CC-BYが適用されない部分を明示し、CC-BYを前提とした利用規約とともに公開する。

#### ※作業シートに記載

要核	討箇所	抽出		

#### ② 第三者の許諾の確認(各担当)

- ・①で抽出・分類した要検討箇所について、以下の作業を行う。
- a. 照会先の特定
- b. 第三者による二次利用の可否の確認
- c. (必要があれば)著作権者等への二次利用許諾依頼
- d. 確認結果及び依頼結果の整理
- ・作業負荷が大きい場合(特に過去の資料について)、要 検討箇所全てをCC-BY適用外とすることも考えられる。

※二次利用許諾確認書の送付

※作業シートに結果を記載

·	·		確認	土里た
			整	理



# (4) ケーススタディの検討フロー ②フルバージョン

- 情報通信白書を構成する素材を検討した結果、具体的な作業手順としては、以下のようなものになる。
- ▶「○」はCC-BYにより自由に利用できる部分、「☆」は第三者の権利に関する確認が必要である部分、「★」は 第三者の権利に関して許諾が得られない、または許諾の確認がとれない部分である。なお「一」は著作権がないため理論的にはCC-BYを付すことができないが、自由に利用が可能な部分である。



凡例: A = 総務省が独自に作成したデータ D = 引用ルールに従って掲載・利用したデータ

B=総務省委託調査で作成したデータ E=著作権対象外のデータ

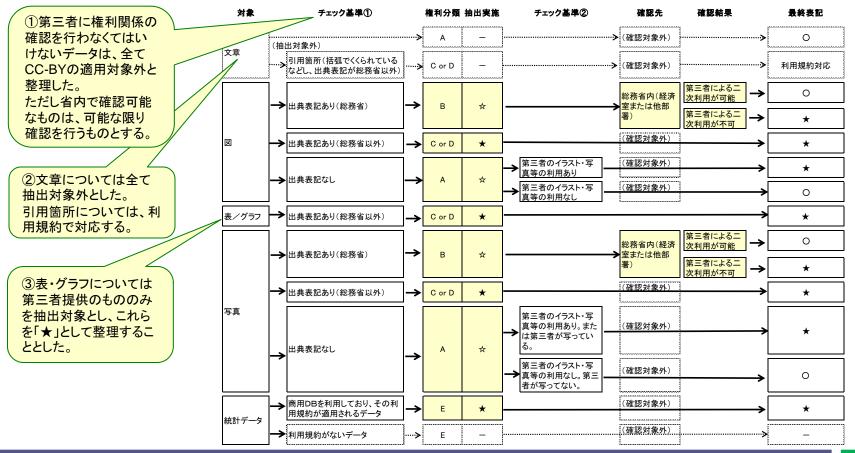
C = 第三者から掲載の許諾を受けて利用しているデータ

白書の担当部署)で担当者に確認(契約内容等)



# (4) ケーススタディの検討フロー ③簡略版

- ▶ データガバナンス委員会の委員からは、前掲の手順を簡略化した方が望ましいのではないかという意見が出ており、その意見を反映させた手順が以下になる。
- ①第三者が権利を有する部分またはその可能性のある部分の取扱いについては、省内で確認可能な素材は、担当者に第三者の許諾をとる必要があるかどうかを確認して可能な限り自由な利用を許諾できるよう努めるが、省外に確認が必要な素材や、省外に確認の必要があると判明した素材は、「★」印を付すという整理をする、②文章については抽出対象外として利用規約で対応する、③表・グラフについては第三者が提供したもののみを抽出対象として「★」印を付すことで、簡略化を図っている。





# (5) ケーススタディの結果

- 平成24年版の情報通信白書について、実際に確認が 必要な素材を洗い出すと、665件の確認事項が生じた。
  - 総務省内で確認可能な素材が574件、第三者への確認を要する素材が91件の見込み。
- ▶ 簡略版でのチェックを行ったところ、125件の確認事項となった。
  - ▶ 全て総務省内で確認可能な素材の見込み。
- ▶ 簡略版への移行で確認をしなくて良くなったものは、以下の通りである。
  - ▶ CC-BY適用不可候補 540減(665→125)
    - 文章33件(引用部分について利用規約でまとめて対応のため)
    - 表・グラフ441件(著作権はないものと整理。また 第三者の権利を有する可能性がないため)
    - ▶ 図 22件(第三者が権利を有する可能性があるものは確認せずともCC-BY不可としたため)
    - ▶ 写真 4件(第三者が権利を有する可能性がある ものは確認せずともCC-BY不可としたため)
    - ▶ 表・グラフ40件(第三者が権利を有する可能性があるものは確認せずともCC-BY不可としたため)
- ▶ 平成23年度以前の年度についても調査を行っているが おおむね数百件の確認事項が生じる。
- 近年のものについては確認先がわかるものが多いが、 過去のものになると確認先が不明になるものも多いと考えられる。

#### (1) フルバージョンで実施した場合の集計

A(総務省	が独自に作成しているデータ) 	380
	文章	
	図	4
	写真	
	表・グラフ	32
B(総務省	の委託調査で作成したデータ)	154
	文章	
	図	6
	写真	
	表・グラフ	8
AまたはB	(AとBの区別がつかないもの)	40
	文章	
	図	1
	写真	(
	表・グラフ	2
C(第三者	から掲載の許諾を受けて引用) 	3
	<b>図</b>	
	写真 キャンニー	
	表・グラフ	1
D(著作権	法上認められた引用ルールに従って掲載・利用) 	7
	文章	
	写真	
	表・グラフ	:
CまたはD	(CとDの区別が付かないもの)	81
	文章	2
	☑	1:
	写真	
	表・グラフ	3
E(数値デ	・一タや法令など、著作権の対象外のデータ)	(
	文章 (法律の引用)	(
	表・グラフ	(

#### (2) 簡易版で実施した場合の集計

経済室内の確認が必要なもの(総務省内他部署の物である可能性もある)	58
図	5
写真	
総務省の他部署への確認が必要なもの	67
***************************************	
図	6
写真	6



# (6) 利用規約案①(情報通信白書を例として)(たたき台)

- ▶ 二次利用を促進するために利用規約案を、ケーススタディを行った情報通信白書を対象に作成した。(なお、本利用規約案は、府省のホームページ全体を対象としたものではなく、「情報通信白書」掲載ページを対象としたものである。)
- ▶ 冒頭部分に、原則として自由に二次利用が行えることを平易に表現した上で、詳しい解説を設ける構成とした。 また、著作物性のある部分については、海外との互換性や機械判読可能性を担保するため、クリエイティブ・コモンズ・ライセンスを採用し、その「表示ライセンス(CC-BY)」により利用可能な点にも言及した。

#### ○ 情報通信白書は、原則として、自由にご利用いただけます

- ・情報通信白書(ウェブ版)は、<u>図表リスト</u>で★印が付いている箇所及び第三者の出典が表示されている文章等を除き、どなたでも自 由に、複製・改変・頒布・公衆送信等のあらゆる利用ができます。商用利用も可能です。
- 利用する際には、出典の表示をお願いします。(→出典表示の記載例)

#### 〇 詳しい利用方法については、以下を御覧ください

#### 【★印が付いている箇所及び第三者の出典が表示されている文章】

- ・<u>図表リスト</u>で★印が付いている箇所または第三者の出典が表示されている文章は、第三者が著作権その他の権利(例:写真につき肖像権・パブリシティ権など)を有している可能性があります。利用にあたっては第三者の権利を侵害することのないよう注意してください。
- ・第三者が著作権を有している情報であっても、著作権法上、引用など、著作権者の許諾無く利用できる場合があります。(→<u>著作権</u> 者の許諾が不要とされている利用方法)

#### 【★印が付いていない箇所】

- ・数値データ、簡単な表・グラフ等には著作権はありませんので、自由にご利用いただけるものですが、出典表示をお願いしています。
- ・著作物性のある文章や図などの著作権は、国が保有し、総務省が管理していますが、自由な利用を認める「<u>クリエイティブ・コモン</u> <u>ズ・ライセンス 表示 2.1 日本</u>」により利用を許諾しています。

#### 〇 免責事項

- ・掲載されている情報の正確さについては万全を期しておりますが、万が一、誤りなどありましたら下記までご連絡ください。
- ・なお、情報通信白書に掲載している情報を用いたことで、利用者に損失等が発生した場合でも、総務省は責任を負いかねます。

#### ○ 情報通信白書に関するお問合せ先 (以下の掲載内容は仮です)

·総務省 情報通信国際戦略局 情報通信政策課 情報通信経済室



# (7) 利用規約案②(過去データ等の場合) (たたき台)

- ▶ 過去の公共データ等については、権利関係を調査し第三者の利用許諾を確認するのに多大な労力がかかる一方、利用ニーズが小さいこともある。その場合、費用対効果を考慮し、権利関係の確認を行わずに公表して、第三者の権利に係る問題の処理は、利用者の判断と責任に任せることが考えられる。下記は、そのような場合のための規約案である。
- ▶ 利用者が第三者の権利のある情報であるか否かを判断するにあたって手がかりとなる出典表記等は、権利関係について一応の目安になるにすぎない。第三者の権利に抵触したことによる責任は利用者が負うこととなる。
- ※ 個別法による制約がある場合は、「〇ご利用にあたって」の欄に、禁止されている利用形態と該当する個別法名を具体的に記載する。この利用規約案は情報通信白書を例としているため、個別法による制約は記載していない。

#### 〇 ご利用にあたって

- ・情報通信白書に掲載している情報は、第三者の権利に抵触しないかぎり、どなたでも自由に、複製・改変・頒布・公衆送信等のあら ゆる利用ができます。商用利用も可能です。
- 利用する際には、出典の表示をお願いします。(→出典表示の記載例)

#### 〇 詳しい利用方法については、以下をご覧ください

- (1)第三者の権利を侵害しないようご注意ください。
  - ・第三者が著作権を有している箇所や、第三者が著作権以外の権利(例:写真につき肖像権・パブリシティ権等)を有している情報 については、利用者の責任で、個別に許諾を得て利用してください。
  - ・なお、数値データ、簡単な表・グラフ等には著作権はありません。また、第三者が著作権を有している情報であっても、著作権法 上、引用など、著作権者の許諾無く利用できる場合があります。(→<u>著作権者の許諾が不要とされている利用方法</u>)
- (2) クリエイティブ・コモンズ・ライセンスについて
- ・掲載されている著作物で、国が著作権を保有するものは、自由な利用を認める「<u>クリエイティブ・コモンズ・ライセンス 表示 2.1 日</u>本」により利用を許諾しています。
- ・数値データ、簡単な表・グラフ等には著作権はありませんので、自由にご利用いただけるものですが、出典表示をお願いしています。

#### 〇 免責事項

- 掲載されている情報の正確さについては万全を期しておりますが、万が一、誤りなどありましたら下記までご連絡ください。
- 掲載している情報を用いたことで、利用者に損失等が発生した場合でも、国は責任を負いかねます。
- 第三者の権利に抵触したことによる損害について、国は責任を負いかねます。

#### 〇 お問合せ先 (略)



# (8) 利用規約案③(今後作成するデータ等の場合) (たたき台)

- ト 情報通信白書のケーススタディで用いた★マークは、自由に二次利用できない箇所を明示する試みであるが、二次利用できない理由が、第三者の著作権が存するからなのか、著作権以外の肖像権等が存するからなのかは明確にしていない。これは過去のデータにおいて、確認作業を行う作業負担等を考慮したものである。
- ▶ 今後作成する公共データについては、自由に二次利用ができない理由を白書の素材の出典とともに表記しておくことで、個別の権利処理をしようとする利用者の便宜をできる限り図るようにすることができる。下記は、そのような場合の規約案である。
- 2)の【】内の言葉が表記されている素材については、利用する際に制約があるという整理である。また総務省名で出典表記をする場合は確実に権利処理を行ったものに限る。権利処理ができない場合は、2)【】の表記を利用する。
- ※ 個別法による制約がある場合は、「〇ご利用にあたって」の欄に、禁止されている利用形態と該当する個別法名を具体的に記載する。この利用規約案は情報通信白書を例としているため、個別法による制約は記載していない。

#### 〇 ご利用にあたって

- ・情報通信白書に掲載している情報は、第三者の権利に抵触しないかぎり、どなたでも自由に、複製・改変・頒布・公衆送信等のあら ゆる利用ができます。商用利用も可能です。
- ・利用する際には、出典の表示をお願いします。(→<u>出典表示の記載例</u>)

#### 〇 詳しい利用方法については、以下をご覧ください

- (1)第三者の権利を侵害しないようご注意ください。
- ・国以外の第三者が著作権を保有している箇所については、当該第三者にお問い合わせいただくか、著作権者の許諾が不要とされている利用の範囲内でご利用ください。(→著作権者の許諾が不要とされている利用方法)
  - ※ 第三者が著作権を保有している箇所は、以下のように表記されています。

【表記例:「出典:〇〇調査(〇〇社)」】(第三者により二次利用が許諾されている場合はその旨を記載してあります)

- ・著作権以外の第三者の権利(例:写真につき肖像権・パブリシティ権など)がある情報を利用する場合は、個別に権利者にお問い合わせください。
  - ※ 第三者が権利を有する箇所は、以下のように表記されています。

【表記例:「注:上記の写真には肖像権があります」等】

- (2) クリエイティブ・コモンズ・ライセンスについて
- ・掲載されている著作物で、国が著作権を保有するものは、自由な利用を認める「<u>クリエイティブ・コモンズ・ライセンス 表示 2.1 日</u>本」により利用を許諾しています。)
- ・数値データ、簡単な表・グラフ等には著作権はありませんので、自由にご利用いただけるものですが、出典表示をお願いしています。

#### ○ 免責事項、お問い合わせ先 (略)



# (9) 契約書に盛り込むべき条文案(たたき台)

- ▶ 今後作成される公共データについては、第三者による二次利用を可能にすることを前提として作成することが必要である。
- 今後作成する公共データのうち、事業者等に委託して作成する場合には、委託契約書の条文に以下の3点を盛り込むことが望ましい。
  - ▶ ①新たに作成した著作物の著作権は総務省に譲渡する。 (譲渡ができない場合には)新たに作成した著作物について、第三者が二次利用することを含めて利用を許諾する。
  - ▶ ②新たに作成した著作物について、総務省及び総務省以外の第三者が利用する場合に著作者人格権を行使しない。
  - ※二次利用に制約がある既存著作物が含まれている場合には、その部分についても可能な限り二次利用が可能となるような契約 内容とし、利用許諾を得る。二次利用を可能とする許諾を得られたものと、許諾の得られなかったものが区別できるようにする。
- ▶ 以下に、総務省及びコンソーシアム事務局企業における現在の標準的な契約書を踏まえつつ作成した条文案を例として示す。

(甲:総務省 乙:受託者)

#### 第〇条 著作権及び著作者人格権

- 1 乙は、乙が本業務を行うにあたり新たに作成した著作物(以下「新規著作物」という)の著作権 法第27条及び第28条に定める権利を含むすべての著作権を甲に無償で譲渡する。
- [1 乙は、乙が本業務を行うにあたり新たに作成した著作物(以下「新規著作物」という)の著作権 法第27条及び第28条に定める権利を含むすべての<u>著作権の権利を留保するが、甲が第三者</u> に二次利用を許諾することを含めて、無償で利用を許諾する。]
- 2 乙は、甲及び<u>新規著作物と乙が従来より有している著作物(以下「既存著作物」という)を利用</u> する第三者に対し、一切の著作者人格権を行使しない。
- 3 新規著作物の中に既存著作物が含まれている場合、その著作権は乙に留保されるが、可能な限り、甲が第三者に二次利用することを許諾することを含めて、無償で既存著作物の利用を許諾する。また第三者の著作物が含まれている場合、その著作権は第三者に留保されるが、乙は可能な限り、甲が第三者に二次利用することを許諾することを含めて、第三者から利用許諾を取得する。成果物納品の際には、第三者が二次利用できる箇所とできない箇所の区別がつくように留意し、第三者が二次利用をできない箇所についてはその理由についても付するものとする。

① 著作権を甲に譲渡する 旨を記述

- ①'著作権を甲に譲渡せず、 利用許諾のみをする場合 の記述
- ② 第三者が二次利用する 場合にも著作人格権を行使 しない旨を記述
- ③可能な限り既存著作物からも利用許諾を得た上で、 二次利用に制約がある既存 著作物が区別できるように する旨を記述



# (10) その他の検討すべき事項

- 利用規約案や委託契約書のほかにも留意すべき事項として、今後オープンデータ化を行うに当たって、以下の5つの事項について検討を行う必要がある。
  - ① マニュアルの作成等
    - ▶ 実際に公共データを公開する際の手順等を記載したマニュアルを作成し、各府省で共有する。
    - ▶ 公共データの公開を前提として業務が実施されるよう、外部委託方法等の業務フローを見直す。
    - ▶ 各府省の担当者個人の判断により異なる取扱いがなされることのないよう、諸外国で用意されている ような公共データ公開の支援ツールを用意する。(例: AusGOAL等)
  - ② 職員向けの研修
    - ▶ 公共データ公開の手順について、職員向けの研修を実施する。
  - ③ 利用者向けのヘルプデスク・府省の担当者向けのヘルプデスク
    - 現場職員に疑問が生じた場合に問い合わせができる職員向けヘルプデスクを設置する。
    - 利用者が疑問を感じたときに問い合わせができる利用者向けヘルプデスクを設置する。
    - ▶ 多く寄せられた問い合わせ内容については、④に示すFAQに掲載する。
  - ④ 利用者向けFAQ・職員向けFAQの作成
    - ▶ よくある問い合わせ等については職員向けのポータルサイト等にFAQを作成・掲載する。
    - 利用者向けの問い合わせについてもFAQを作成する。
  - ⑤ リスク対策とノウハウの蓄積
    - ▶ 公開した公共データに関するクレーム等が生じた場合に、担当者個人の責任問題とならないような対策を講ずる。
    - 類似のクレームが複数発生した場合には、対応のノウハウを蓄積して、各府省で共有する仕組みを 構築する。



# (11) 電子行政オープンデータ実務者会議への提言①

- データガバナンス委員会は、公共データの二次利用促進のための利用ルールのあり方について、以下の事項を提言する。
  - 1. 国が保有する公共データは、利用を制限すべき実質的かつ合理的理由のない限り、誰もが自由に利用できるようにすべきであり、国の著作権を根拠とした利用制限を認めるのは妥当でない。したがって、国の保有する公共データは、その著作物性の有無にかかわらず、二次利用を自由に認めるのが原則であることを全府省に周知すべき。
  - 2. 著作物性のある公共データについては、立法により国の著作権を否定することも考えられるが、オープンデータ戦略を早急に推進する観点から、国が自らの著作権を行使せず二次利用を認める内容の利用ルールを策定・導入することを検討すべき。
  - 3. 利用ルールの策定に当たっては、以下のような点に留意すべき。
    - 利用ルールは、様々なデータを機械で重ね合わせて利用(マッシュアップ)する際の便宜のため、できる限り統一し、また、社会で広く使われているライセンスとの互換性についても明記し、機械判読可能性にも考慮すること。
    - ② 国の保有する公共データの中に第三者が著作権等の権利を有する部分が含まれている場合には、 第三者の許諾が必要となりうることから、その部分をできるかぎり明確に示して、利用者の注意を喚 起すること(その際、利用者に過剰な委縮効果が生じないよう、数値データや簡単な表・グラフ等は 著作権の保護対象でないこと、引用など著作権者の許諾なく利用できる場合があることなど関連法 制度のポイントを付記することが望ましい。)
    - ③ 公共データの二次利用を国が規制すべき旨を定めた個別法が存在する場合には、その内容をわか りやすく表示すること。
    - ※ データガバナンス委員会ではこれらの点に配慮した利用規約案(たたき台)を作成しているので参考とされたい(27頁~29頁参照)。



# (11) 電子行政オープンデータ実務者会議への提言②

- 4. 今後新たに作成する公共データと過去の公共データは、取り扱いを区別し、それぞれ以下のような取扱いとする方向で検討すべき。
  - ① 今後新たに作成する公共データについては、たとえば調査研究等を外部に委託する際に、その報告書の二次利用を可能にする内容を委託契約書に盛り込むなど、二次利用を前提として予め権利関係の集約化・明確化を図ること。
    - ※ データガバナンス委員会では、委託契約書の例(たたき台)を作成しているので参考とされたい(30頁参照)。
  - ② 一方、過去の公共データについては、権利関係の確認作業等の負担が大きく、費用対効果の観点から見合わないと判断される場合には、権利関係の確認は利用者の責任に委ねることとし、その旨を利用ルールに明示すれば足りるとすること。
    - ※ データガバナンス委員会では、各府省が権利関係の調査及び第三者等の許諾の可否の確認を行う場合の手順 (23頁~25頁参照)及び利用規約案(たたき台)(28頁参照)を作成しているので参考とされたい。
- 5. 上記と並行して、関連マニュアルの作成、職員向け研修の実施、ヘルプデスクやFAQの整備、リスク対策等、様々な環境整備を行ってほしい。