

二次利用の促進のための府省のデータ公開に関する基本的考え方(ガイドライン)(案)

1 総論

(1) 電子行政オープンデータ推進の背景

インターネット等の普及により、データを低コストかつ即時に提供することが可能となるとともに、データを利用する企業・国民等においても、コンピュータの能力向上、端末の高度化・多様化等により、大量・多様なデータを処理・利用できるようになっている。

このような技術の進展を背景に、政府、独立行政法人、地方公共団体等が保有する多様で膨大な公共データについて、ビジネスや身近な公共サービスへの活用が期待されるようになってきている。

このような状況のもと、「電子行政オープンデータ戦略」(平成24年7月4日高度情報通信ネットワーク社会推進戦略本部(IT戦略本部)決定)は、公共データの活用を促進するための取組に速やかに着手し、それを広く展開することにより、国民生活の向上、企業活動の活性化等を図り、我が国の社会経済全体の発展に寄与することが重要であるとの考えを明らかにしている。

また、同戦略においては、我が国における公共データの活用の取組に当たり、①政府自ら積極的に公共データを公開すること、②機械判読可能な形式で公開すること、③営利目的、非営利目的を問わず活用を促進すること、④取組可能な公共データから速やかに公開等の具体的な取組に着手し、成果を確実に蓄積していくこと、という4つの基本原則を掲げている。

これらの社会・経済状況や政策方針を踏まえ、政府、独立行政法人、地方公共団体等が保有する公共データを、機械判読¹に適したデータ形式で、営利目的も含めた二次利用が可能な利用ルールで公開する「オープンデータ」の取組を推進することが重要である。

(2) オープンデータ推進の意義

これまでも政府は、各府省のホームページ等において保有するデータを公開してきており、情報提供という観点では一定の成果が出ている。

¹ ここでの「機械判読」とは、コンピュータプログラム(以下この注において、単に「コンピュータ」という。)が自動的にデータを再利用(加工、編集等)できるということである。人手をどれだけ要せずに、コンピュータがデータを再利用できるかにより、「機械判読に適した度合い」には、いくつかの段階がある。コンピュータが自動的にデータを再利用するためには、コンピュータが、当該データの論理的な構造を識別(判読)でき、構造中の値(表の中に入っている数値、テキスト等)が処理できるようになっている必要がある。

ただ、これまでのホームページ等による情報提供は、基本的に、人間が読む（画面上で又は印刷して）という利用形態を念頭に置いた形で行われており、検索も難しく、大量・多様なデータをコンピュータで高速に、横断的に又は組み合わせで処理・利用することが難しい。

大量・多様なデータをコンピュータで高速に、横断的に又は組み合わせで処理・利用できるようになれば、例えば次のようなことが可能となり、データの活用という観点から、人間が読むという利用形態だけでなく、機械判読という利用形態も考慮した情報提供も求められるようになってきている。

ア 経済の活性化、新事業の創出

データ収集や各種コードによるデータの横断的利用が機械で自動的に可能になることからコスト圧縮ができ、新しいサービスを提供するビジネスが可能となる。（例えば、気象、地質、交通その他の観測・調査データのような専門的データを収集・分析してビジネスに活用するなど）

イ 官民協働による公共サービス（防災・減災を含む。）の実現

複数の行政機関や民間のデータを組み合わせることで、民間からも、生活利便を高めるサービスや災害時に有用なサービスを提供できる。（例えば、子育て、教育、医療、福祉等の身近な公共サービスの内容、品質等を利用者に分かりやすく示す、災害時に迅速に複数の情報を組み合わせた情報発信が可能となるなど）

ウ 行政の透明性・信頼性の向上

政策・事業に関する計画、決定過程、決定内容、結果等について、横断的に検索・集計・比較することで、政策の変化・特徴の把握や、政策の妥当性の理解・評価ができる。（例えば、補助金や政府支出について、府省、分野、地域、支出先等別に分析するなど）

コンピュータで高速に、横断的に又は組み合わせで処理・利用することが期待できるデータは、統計等の数値データだけでなく、白書等の文書や地図等の図表も含まれる。

また、これまでは、理由が明確でないまま、各府省の判断でインターネットを通じた公開がされていないデータも多く存在している。

オープンデータの効果は、大量・多様なデータの横断的利用や組み合わせによりもたらされることから、横断的利用や組み合わせができるデータを増やすことが重要であり、これまでインターネットを通じて公開されていなかった情報にも、新たな利用が期待できるものがありうることから、公開可能な情報については、すべて公開するという理念の下、積極的にインターネットを通じて公開していくことが求

められる。

また、同様の観点から、独立行政法人や地方公共団体、民間企業等においても、本ガイドラインによる政府の取組を参考に、オープンデータを推進することが期待される。

(3) 本ガイドラインの対象

「電子行政オープンデータ戦略」において、オープンデータの取組については、政府が保有するデータについて率先して取組を推進するとされていることを踏まえ、本ガイドラインは、基本的には、これまでの電子行政オープンデータ実務者会議の議論、先行的な取組を実施している府省の取組等をもとに、早急に取り組むべき事項として、各府省の保有するデータの公開に関する基本的考え方を整理したものである。

(4) 本ガイドラインの改定及び取組の留意事項

本ガイドラインの内容については、今後の電子行政オープンデータ実務者会議の議論の進展や関連技術の進展等を踏まえ、随時改定していくことが必要である。

また、同戦略において、オープンデータの取組について、独立行政法人、地方公共団体、公益企業等の取組に波及させていくとされており、各府省の取組に当たっては、民間、地方公共団体等と十分に連携を図り、民間、地方公共団体等に円滑かつ速やかにオープンデータの取組が普及できるように留意する必要がある。

さらに、具体的に取組を行う際には、十分な情報セキュリティを確保した上で、推進していくことが重要である。

2 二次利用を促進する利用ルールの在り方

(1) 各府省がインターネットを通じて公開するデータの著作権等の位置づけ

各府省がインターネットを通じて公開するデータに関し、著作権等の関係で留意すべき事項としては、次のような事項が挙げられる。

- ・ 単なる事実や数値データは、それ自体としては、著作物とはならず、著作権の保護対象にはならない。編集著作物やデータベースの著作物と認められる場合も、素材・数値データが著作物でない場合は、素材・数値データそのものを利用することは著作権法の観点からは制限されない。
- ・ 著作権者は、あらかじめ著作物の利用に係る考えを表示しておくことができるので、国が著作権者である著作物について、インターネットを通じて公開するに当たり、どのような条件で利用を認めるかは、著作権法の範囲内で、国が判断し、表示することができる。

なお、各府省がインターネットを通じて著作物を公開することについては、著

著作物が国有財産法第2条に規定する国有財産に該当しないため、国有財産法の適用はない。また、国有財産法は、インターネットを通じて公開されている著作物が二次利用されることに対し何ら制約を加えるものではない。

- ・ 国が著作権者となる著作物の中にも、第三者が著作権者である著作物が含まれる場合があり、そのような著作物をどのような条件で利用を認めるかについては、当該第三者（著作権者）の判断による。

(2) 各府省がインターネットを通じて公開するデータの利用ルールの在り方

各府省がインターネットを通じて公開するデータの二次利用を促進する観点から、公開データの利用ルールについては、以下の考え方によるものとする。

- ・ 著作物でないデータについては、著作権の保護対象外である（著作権を理由とした二次利用の制限はできない）ことを明確にする。
- ・ 国が著作権者である著作物については、国において、どのような利用条件で公開するかを決定できることから、広く二次利用を認める（著作権以外の具体的かつ合理的な根拠に基づき二次利用を制限する場合を除き、制約なく二次利用を認める）形で、あらかじめ著作物の利用に係る考えを表示する。当該表示については、できるだけ分かりやすく統一的なものとする。
- ・ 著作権を根拠に公開データの一部について二次利用の制限を行う場合には、例えば、二次利用の制限をする部分の著作物について第三者が著作権者であること、既に作成・保有している著作物について著作権者が明確でないこと等、二次利用を制限する理由とともに、二次利用を制限する部分を明確に表示する。
- ・ 本ガイドライン策定後、各府省が新たに作成・入手するデータについては、各府省がインターネットを通じて公開した場合に当該データの二次利用を認めることができるよう、事前に関係者との間で合意をとるよう努める。このため、本ガイドライン策定後の委託・請負契約の検討・締結等に当たっては、それを念頭に置いた対応（例えば、委託調査の契約の内容を、成果物である報告書を府省がインターネットを通じて公開する場合、当該公開データの二次利用を認めることの支障とならないようなものとする等）が求められる。
- ・ 個別法の規定等、著作権以外の具体的かつ合理的な根拠に基づき公開データの二次利用を制限する場合は、制限の範囲を必要最小限に限定し、その内容及び根拠を明確に表示する。当該表示については、できるだけ分かりやすく統一的なものとする。
- ・ 各府省がインターネットを通じて公開しているデータを第三者が二次利用し、当該二次利用されたデータを利用した者に損害が生じた場合も、各府省は責任を負わない旨を明確にする。

3 機械判読に適したデータ形式による公開の拡大の考え方

(1) 目指すべきデータの構造やデータ形式

各府省がインターネットを通じて公開するデータについては、それをコンピュータで機械的に読み取り、処理するといった利用を考慮して、データの構造（タグの付け方、表の形式等）を整えておくことが重要である。また、ある内容を示す用語や同じ用語の定義が組織（各府省、独立行政法人、地方公共団体等）により異なっているとデータを横断的に又は組み合わせで処理・利用することが困難となるため、用語やその定義の標準化が望ましい。

機械判読に適したデータ形式については、特定のアプリケーションに依存しないデータ形式であることを要件とし、可能なところから、順次より高度な利用が可能なデータ形式での公開を拡大していく。

統計データについては、「統計調査等業務の業務・システム最適化計画」（平成18年3月31日各府省情報化統括責任者(CIO)連絡会議決定。平成24年9月7日最終改定）に基づき、従来から、統計情報の電子的提供の推進を含む取組が行われているところである。同計画においては、統計表を表計算ソフトで利用可能なスプレッドシート又はCSV形式ファイルにより作成し、提供するとともに、統計情報データベースを通じてデータ提供を行うことが記載されており、これを着実に実施することは、オープンデータの観点からも重要と考えられる。

なお、統計情報データベースについては、オープンデータの観点から、地理情報を活用した統計データの拡充を行うとともに、統計データの機械からのアクセス性等の利便性の向上を図る。

(2) 数値（表）、文章、地理空間情報のデータ作成に当たっての留意事項

統計情報データベースを通じて提供される統計データ（最適化計画に基づき統計情報データベースを通じた提供を推進している統計表管理システムの統計表を含む。）以外に、各府省がインターネットを通じて公開するデータは多種多様であるが、その主要なコンテンツであり、掲載後の更新頻度が比較的少ない（比較的長期にわたって掲載される）数値（表）、文章、地理空間情報（空間上の特定の地点又は区域の位置を示す情報（当該情報に係る時点に関する情報を含む。）及び当該情報に関連付けられた情報（地理空間情報活用推進基本法第2条第1項））について、そのデータの作成に当たっての留意事項を別添のとおり整理した。

本ガイドライン策定後、各府省が新たに作成し、インターネットを通じて公開する数値（表）、文章、地理空間情報については、人間が読む、印刷することを念頭に置いた従来のデータ形式（代表的なものとして pdf）のほか、別添の留意事項に示す事項を踏まえて作成した（構造が整った）データを、機械判読に適した、特定のアプリケーションに依存しないデータ形式でも公開することに努めるものとする。

特に、重点分野（白書、防災・減災情報、地理空間情報、人の移動に関する情報（交通、旅行、観光、引越、出入国等に関する情報）、予算・決算・調達情報）については、優先的に取り組むこととし、具体的に別添の留意事項に従って作成・公開すべきデータについて、費用対効果を踏まえつつ、実務者会議で検討する。

また、上記実務者会議における議論を踏まえ、機械判読に適したデータ形式のデータも納入させるための委託・請負契約の方法等について検討を行う。

4 インターネットを通じて公開するデータの拡大についての考え方

(1) 原則公開の理念

公開できない理由が明確なものを除き、保有するデータはすべて公開するという理念の下、具体的な取組としては、以下のとおり進めていくこととする。

(2) 現在インターネットを通じて公開している情報のデータ形式の整備

オープンデータの取組に当たっては、まずは、現在インターネットを通じて公開している内容の情報について、今後新たなデータを公開するに当たり、機械判読に適したデータ形式のデータも公開する取組からスタートする。特に、重点分野（白書、防災・減災関連情報、地理空間情報、人の移動に関する情報、予算・決算・調達関連情報）については、優先的に取り組むこととする。

（注）従来からインターネットを通じて公開されていたデータで、機械判読に適したデータ形式でないもの（代表的なものとして pdf）について、遡って機械判読に適したデータ形式でのデータの公開も行うことについては、既に公開しているデータより機械判読に適したデータ形式の基データの探索やデータの整形、メタデータ付与などの作業が必要となるため、それを行うことが適当な範囲について、実務者会議で検討することとする。

(3) 重点分野に関する公開データの拡大

上記の取組と並行して、上記(2)の重点分野について、従来インターネットを通じて公開されていないが公開可能な情報のうち、オープンデータ化（二次利用可能で機械判読に適したデータ形式でのデータ公開）することが適当なものの公開を進める。

具体的にどのような情報のオープンデータ化が適当かについては、実務者会議で検討することとする。

なお、上記情報のデータ形式が機械判読に適したデータ形式でなく、データ形式の変換に多くのコストを要する場合には、当面、従来のデータ形式で公開すればよいこととする。

(4) その他の公開データの拡大

重点分野について実務者会議で検討する情報以外の情報に関しては、新規にインターネットを通じて公開するためのコストが小さいデータや、利用者のニーズ（要望）の強いデータは、公開できない（行政機関の保有する情報の公開に関する法律第5条の不開示情報に該当する等）ものを除き、オープンデータ化していくこととする。

新規にインターネットを通じて公開するためのコストの考え方や利用者のニーズ（要望）を把握する仕組みについては、実務者会議で検討することとする。

なお、上記情報のデータ形式が機械判読に適したデータ形式でなく、データ形式の変換に多くのコストを要する場合には、当面、従来のデータ形式で公開すればよいこととする。

（以上）