

官民連携による
公共データの戦略的利活用に関する提言
(Ver 1.0)

平成24年 2月 27 日



g コンテンツ流通推進協議会

<http://www.g-contents.jp/>

はじめに

gコンテンツ流通推進協議会(以下、「当協議会」)では、平成16年3月の設立以来、地理空間情報(位置情報、時間情報を伴うコンテンツ。gコンテンツと称す)の流通を通じて、異種の事業者のマッシュアップや、市場の拡大に向けて、SVG(Scalable Vector Graphics)の規格化など技術面の検討、新サービス・新産業を創出するために必要な措置などの制度面の検討を続けている。

一方、昨年3月の東日本大震災において、安否情報や避難所情報などの公開にあたり、当協議会会員企業もボランティアな支援を活発に行った。その際に、大きな課題となったものの一つに、国や行政機関等が所有する多くの一次情報が、機械判読可能な形になっていないため、その利用に大変な手間がかかると言うことであった。

これまで、当協議会では、地理空間情報を利用したサービスを実施するためには、その網羅性と鮮度の観点から、行政機関などが所有する情報の利用について検討し、関係府省へ提案等を出してきたところであるが、今般の災害対応を通じて、これらの情報が利用できる環境を整えることは喫緊の課題であると共に、行政情報は営利・非営利を問わずに、積極的に利活用し、価値を生み出して行くべきではないか、との問題意識を協議会内で合意形成したところである

当協議会は、地理空間情報の利用によって、更に高度なサービスや、新産業の創出が促進するために、国・行政機関等が所有する情報について機械判読可能な形で公開されることを期待し、ここに提言するものである。

2012年2月27日

gコンテンツ流通推進協議会

提言の内容

◎国・行政機関等が所有する地理空間情報等公共データを機械判読可能な形式で公開する仕組みの整備が必要である。

【本提言の観点】

1. 戦略性

行政情報は営利・非営利を問わず、積極的に利用し、新たな価値を生み出していくべきである、特に、その価値を最大化するための戦略を以て実施されるべきである。

2. 迅速性

まず、現状のウェブの情報提供において、機械判読可能であることの原則を示し、個々のデータは元よりサイト全体も機械判読できるような仕組みとすることや、それに即した規約等の改定が必要である。

そして、国自身が率先してこうした行政情報の利活用を行い、自治体や関係公共機関と協力しつつ、推進されるべきである。また、その実施にあたっては、個人情報や企業情報の活用、データの標準化、データの対価の考え方などを指針として整理することも必要である。

3. 柔軟性

行政情報利活用促進のために著作権処理について明確にする必要がある。パブリックドメインとして利活用を促すか、または再配布やリミックスなどを考慮した著作権などを検討した上で、標準的な利用条件の下で、データ利用者が容易に判断できる配慮が必要である。また、国際的な枠組みの中でルール化が必要なもの(例えば、匿名情報の利用など)については、国際的な働きかけを行い、国民や我が国の産業界が不利にならないような活動が必要である。

提言の背景

以下、本提言の策定に至った背景について、以下に記述する。

1. 災害支援の観点から

2011年3月東日本大震災において、その安否情報、被害情報、生活情報などの提供にあたり、当協議会会員企業もボランティアとして多くが参加した。特に、情報提供にあたっては、地図に可視化することが、その理解を深めることになることから、国・自治体などの一次情報の利用を進めようとした。

しかしながら、多くの公開情報は、人間可読性のみ重視した PDF¹等による機械判読が困難なものであり、また、場所の名称も省略表記されているため地元の方でないと判別し難いものなどもあり、ジオコーディング²に時間を要した。(全ての避難所情報の公開に48時間を要したという報告もあった。)更に、ボランティアを含めた多くの情報サイトが立ち上がったが、国・自治体などの発表情報に関する権利(掲載許諾など)に関して現場で判断に迷うなどのケースも生じていた。

例えば、位置を表す用語(具体的には地名、施設名(公共施設、道路名称とキロポスト、河川名称と距離標等)、地番等)では、国・自治体間で標準化し、明確な著作権処理の下で、緯度経度座標との紐づけしたデータを作成・公開する仕組みが必要であろう。平時・非常時ともに、機械判読性を加味した情報が利用できる環境を整備することによって、効率的且つ、迅速な情報提供等が可能になるのではないか。

2. サービス利用の観点から

サービス事業者において、地理空間情報サービスの展開では、そのコンテンツの網羅性と鮮度が求められる。これまで、事業者は網羅性と鮮度を維持するために、多額の費用をかけて対応してきた。

他方、海外では、行政が所有する情報を機械判読可能な形で公開し、産業界に利用を促すことで、情報到達コストを圧縮し、サービス自体の高度化等に集中できる取り組みが進められている。

¹ Portable Document Format の略。Adobe Systems 社によって開発された、電子文書のためのフォーマットのこと。

² 住所などの場所の名称から地理的な座標に変換する処理・作業のこと。

特に民間での地理空間情報整備において実現することが難しいのが、現地調査などによっても得ることの出来ない「計画(将来)」等に関する情報である。具体的には、「店舗が出来る」、「空き地が駐車場になる」、「今日の午前中は工事で通行止めになる」といった情報である。これらは、国・自治体において法定図書(例えば、道路台帳や都市計画基本図など)における整備や、許認可や申請等において関係機関に予め提出されるものである。一部、ウェブ等で公開されているものもあるが、それらは 2 次利用な形式になっていない。例えば、路外駐車場の設置にあたっては、各種法令に基づく届け出を市町村窓口に提出するが、その様式には、名称、所在地は元より、収容台数(四輪車〇〇台など)、駐車場の前の道路の状況(一方通行か、中央分離帯があるかなど)、地形図、平面図などが記載(又は添付)されている。これらの情報が機械判読可能な標準的なフォーマットで提供されれば、空いている駐車場へ誘導するなどのサービスの高度化ばかりでなく、それに伴うエネルギー効率の向上など社会的な裨益も期待できる。

平成 23 年度、当協議会の議論において、国・自治体が所有する情報の中でサービス利用等の観点から公開・利用が期待される情報として上げられたものの例を以下に示す。

情報の種類(一部)	記述されている届け出など
飲食店の開業・廃業	開店前に保健所に「営業許可申請」が提出され、工事の際に消防署に「防火対象物使用届」、開店時には個人事業主の場合は「個人事業の開廃業等の届出書」、企業の場合は「給与支払事務所等の開設届出書」が税務署に提出される等
路外駐車場の設置	一般公共の用に供する駐車場(不特定多数の人が利用できる駐車場)について、設置届や管理規定届を自治体や警察署等にする(なお、変更、休止、廃止の場合も同様)等
道路工事	道路工事(埋設管等の工事を含む)により、道路が片側通行又は通行止めとなるなど、消防隊の通行又は消火活動の支障となるような場合、道路管理者(又は工事請負者)によって、「道路工事届出書」が消防署に提出される等
診療所	所轄する保健所に「診療所開設届」が提出される等

表1. 公開・利用によってサービスの高度化などに寄与する国や自治体が所有する情報の例

以上のような観点から、それらができるだけ早く機械判読可能な形³で公開されるだけで、前記のような情報収集コストが効率化され、鮮度が高い情報を得ることができ、事業者は新サービスの創出などの投資にリソースを振り向けることができる。これによって、市場の活性化等も期待できるのではないか。

³ ここで言う機械判読可能な形とは、CSV (Comma Separated Values ;カンマ区切りしたデータ形式) 形式などによる公開はもとより、現状のウェブからの情報提供についても、マシンリーダブルであることの原則を明確にし、マ

3. 近年の情報通信技術革新の観点から

機械判読可能な情報の整備には、従来高い情報処理の技量が必要とされ、またその利用にも同様に高い技量が必要とされることから、広く社会に展開するには、整備と利用の両側面で困難が伴っていた。

これらの課題に対し、公開されるデータ形式は、迅速な情報の伝搬や、標準化の未定義によるデータ形式の乱立などを防ぐために、標準化のプロセスを考慮すべきであると考えられる。近年、HTML5⁴に代表される情報通信技術の革新により、人間が理解するための情報と、機械が判読可能な情報を同時にひとつのウェブコンテンツとして整備可能となってきている。欧米を中心として進んでいるオープンデータ(Linked Open Data⁵)ウェブサイトの構築は、これを適切に活用したものである。

例えば、HTML5 を地理空間情報に適用する(例えば、SVG⁶ Map の活用など)ことで、人間が理解可能な地図情報と、機械が判読可能な地理空間情報を一つのウェブコンテンツ(Linked Open Geo/Map Data)として提供可能となるのではないか。また、このような活動を端緒に、将来的なデータ取得に関して、人間&ヒューマンインターフェイスを介さずに、データ検索および取得が出来る仕組み(M2M⁷のアプリケーション・インターフェースの提供など)の実現が可能となるのではないか。

4. 行政と事業者の連携の観点から

「提供する地理空間情報(あるいはそれに類する台帳を含め)、情報の錯誤による責任の範囲を明確にして、提供する行政機関の負担増を避ける必要がある」ことや、「提供はしたが、その誤りによって、訴訟になる、或いは事業者が、行政が公開した最新の情報を参照せず、古い情報を原典参照もなく利用/複製

シンリーダブルなデータ形式を用いて、当該サイト全体として、機械で情報を取得できる仕組みとすることも含んでいる。

⁴ HTML (Hyper Text Markup Language) の最新バージョン。ここでは、HTML とは直接の関係を持たない JavaScript API も含めた総称として使用する。

⁵ データの「公開」と「共有」をより効率的にかつ実践的に行えるようにするための技術のこと。

⁶ Scalable Vector Graphics の略。XML ベースの 2D ベクター画像記述言語。g コンテンツ流通推進協議会では、SVG の JIS 化に協力している。

⁷ Machine to Machine の略。機械と機械が IP ネットワークを介して相互に通信し合う通信形態のこと。

した結果として誤情報が伝播する等というようなことが生じないような仕組みが必要である」ことなど、従来、行政情報の公開/利活用について、それらが懸念事項となっていた。

その解決のためには、まず行政側では、公開する情報の著作権利用条件、書誌情報を機械判読可能な属性情報として(可能な限りウェブ標準に準拠した形式で)付加することを前提とすることが必要と考える。これにより「公開はするが二次利用を禁ずる」、或いは「二次利用及び商用利用も可とするが表示は必要」等の行政側指定の利用条件が機械的に判断できるようになり、分野横断での迅速かつ機械的なデータ連携/二次利用が可能になるものと考えられる。また、二次利用に際しては、対象行政データの調査時期/有効期間や作成機関等の書誌情報を原典情報として、同じく機械判読可能な形式で相互に参照できるようにすること(Linked Open Data)で、派生した情報の信頼性や、その派生/利用状況のトレースによるデータ公開の効果測定も可能になるものとする。

加えて、行政側でのデータ整備(機械判読可能とするための作業を含む)ならびに公開に要するコスト(クラウド等のインフラのコストを含む)を鑑みた場合、その費用負担を全て行政側で賄うのではなく、それら情報を利用する事業者側でも、一定の費用負担を行う点も考慮することが必要である。それは例えば、個々のデータに対して DRM⁸をかけ著作権使用料を徴収する従前の非機械判読不可能な方式ではなく、システム利用料のようなサブスクリプション(期限付きの使用権)形式で包括的に徴収する等の簡易的な方法が現実的ではないかと考える(例:サブスクリプションを継続する限り最新の情報を入手可能とすることをインセンティブとし、公開データには機械判読と二次利用を可能とするため DRM はかけず、サブスクリプション解除後も既に二次利用済みのデータも継続して存続可能とする等)。なお、この点については、これまで国・自治体から公開されないために、民間投資にて整備されてきた地理空間情報も数多く存在することから、これら整備事業に対する配慮も念頭に置きつつ検討することが必要であろう。

以上のような観点から、仕組みを具体化することによって、行政と事業者が連携し、前者は行政サービスの向上、後者は事業の発展といった相乗効果が生まれることが期待できるのではないかと。

⁸ Digital Rights Management の略。デジタルデータとして表現されたコンテンツの著作権を保護し、その利用や複製を制御・制限する技術のこと。

具体化に向けたアクションアイテムの提案

欧米を中心とした先進各国では、公共データの民間利活用を促進し付加価値を生み出していくための戦略が策定され、利活用ルール・インターフェースの整備や、民間スポンサーと連携した利活用コンテストによる有用データ発掘・サービス開発促進・ベンチャー振興が積極的に行われている。これに対して、我が国では、目的外利用を前提としない、あるいは、営利利用について困難な面が未だ解決できていない。

公共データの民間利活用を促進するための環境を、まずは国が率先して整備し、自治体や民間公共機関にも拡大していくことを期待し、具体的なアクションアイテムについて、以下に記述する。これらが、透明性・簡潔性・公正さ・一貫性、並びにグローバルな視点での調和の基、具体化されることを期待する。

(1) データ形式およびその表現方法の標準化または指針・ガイドラインの作成

多くの地理空間情報は、主に位置(2次元および3次元で表現される点・線・面、または住所・施設名・地域名)・時間、およびそれに付随する追加情報で構成されている。これらの情報について、具体的なガイドライン、もしくは標準化が行われることにより、全く異なる情報種別を素早く組み合わせる可視化できるようになる。具体的には、文字コードの種類や記載、カラム種別の定義および記載、バージョン管理方法、データ精度、更新日付、CSV や XML⁹ などの業界標準形式の適用などがあり、これらを可能な限り具体的な定義を行うべきである。

なお、これらの検討にあたっては、当協議会をはじめ、それらを利用する事業者等から広く意見を求めることを希望する。なぜならば、一度決めたスキームは修正しにくく、使用時(特に災害時など)に問題がわかったとしても急な修正に対応することは困難であり、せっかく公開した情報を活用できなくなる可能性が高いためである。

⁹ Extensible Markup Language の略。

(2) データ利用条件の明確化、および標準化

公開されるデータについては、データごとに微妙に条件が違ったりすると組合せにくくなる。また、様々な情報を融合して利用する場合には、著作権などの扱いが明確になっている必要がある。国・自治体においては、パブリックドメインとして利活用を促すか、または再配布やリミックスなどを考慮した著作権などを検討した上で、標準的な利用条件の下で、データ利用者が容易に判断できる配慮が必要である。また、規約によって公開範囲を変える必要があるデータについても、同様の仕組みが必要である。

(3) データの提供元責任および利用者責任の明記

公開されたデータを利用し、問題が発生した場合に、データ提供元にどのような責任が生じるのか、またデータ利用者にも責任が発生するのかについてのガイドラインの策定が必要である。これによって、データの提供・公開・活用が行いやすくなる。なお、このガイドラインは、無理に責任範囲を明確化するものではなく、提供側・利用側双方において、情報の取り扱いを行いやすくすることを目的として策定されることを期待する。

(4) データ更新条件や修正条件の明確化

データの公開にあたって、データ提供側においては、それぞれ情報種別ごとに更新条件や更新間隔、公開までにかかる時間や公開方法などが異なることが想定される。しかし、データ更新がいつ行われたかが分からなければ、その情報の信頼性が担保できない。公開にあたっては、更新の条件が明記される必要がある。これによって、データ利用者が具体的な判断を行いやすくなる。

また、公開後に誤り等があった場合に、どの時点で行われるのか(又は、データ提供元からのサポートが得られにくい場合は利用者で修正して良いなど)が明記されることも必要である。これによって、データ提供元及び利用者双方でそれらの反映を行いやすくなる。

おわりに

本提言では、国・自治体が所有する地理空間情報(例えば、ビルが建つ、道路工事がある、お店ができる等)を中心にして、まずはヒューマンユーザビリティの担保を前提に、機械判読可能な形で利活用ができる仕組みの整備が必要であることを述べた。

この仕組みの整備にあたっては、機械判読性を加味した標準的なデータ形式の採用、利用の際の窓口・契約などの明確化・平準化、利用に際しての妥当な提供価格のあり方等において考慮された仕組みであることが期待される。

これらの具体化によって、エネルギー情報、交通情報、運輸情報等の公共的なデータの機械判読可能な形で公開が進み、ネットワーク上のデータ融合によって集約されることによって、新たな利用価値(高度な人の移動、物の流通支援など)を生まれ、我が国産業界の競争力が強化されることを期待している。また、膨大な行政情報にジオコーディングを行ったり、機械判読が可能な形態に変換する作業は、それなりに工数が必要となる事が予想される。しかしその作業自体は、最適なツールの提供や、一定のトレーニングによって比較的短期間で従事できるものが多いことから、そこで緊急雇用対策事業として、被災地の雇用確保や、各地域の雇用対策とすることが可能になるのではないかと期待される。また、この行政情報は、品質・鮮度の維持のため、継続的なメンテナンスが発生するため、継続的な雇用対策にも活用できることも期待される。

なお、gコンテンツ流通推進協議会では、官民の役割分担と相互連携の促進や、関係事業者の持続的成長とそのための環境整備を図るべく、関係府省の検討等に積極的に協力し、今後とも貢献していく決意である。

gコンテンツ流通推進協議会

	氏名（敬称略）	職名
会長	柴崎 亮介	東京大学 教授
副会長	福井 弘道	中部大学 教授
副会長	寺前 秀一	加賀市市長
会員企業 （順不同）	KDDI、ゼンリン、NTTドコモ、ソフトバンク モバイル、アクリ ーグ、国際航業、インディゴ、アジア航測、セック、日本電気、明電舎、 昭文社、日本加除出版、ゼンリンデータコム、NTTデータ、パスコ、 トヨタ自動車、富士通、測位衛星技術、表示灯、野村総合研究所、電通 国際情報サービス、顧客感動システム総合研究所、NTT空間情報、マ イスター、ジオインフォシステム、デンソー、東京都ビジネスサービス、 カナエジオマチックス、エヌ・ティ・ティ・ソフトウェア、トヨタマッ プマスター、マピオン、大日本印刷、三恵エンジニアリング、日本デ イブレイク、関心空間、パナソニック、東京地区研究社、ドコモ・シ ステムズ、マルティスーフ、防災科学技術研究所、GIS 総合研究所、産業技 術総合研究所	
本提言策定に あたり御協力 頂いた企業 （順不同）	グーグル、ヤフー、マイクロソフト	

◎本提言に関する御連絡先

g コンテンツ流通推進協議会 事務局

（一般財団法人日本情報経済社会推進協会 電子情報利活用推進部内）

e-mail : dpc-gcnt@jipdec.or.jp