

# オープンデータ活用

## 「公共データの開放」をビジネスに活かす

ビッグデータとともに注目が高まっているオープンデータ。世界各国でいま、行政の持つ膨大な量の情報を「資産」として捉え直し、誰もが自由に活用できるオープンデータとして流通させるための取り組みが進められている。本記事では、新たなビジネスの創出が期待されるオープンデータについて、国内外のこれまでの取り組みや、利用にあたって押さえておくべき法的・技術的要素について紹介するとともに、国内外の活用事例も取り上げる。

### オープンデータとは

「オープン・ナレッジ・ファウンデーション(OKF)<sup>\*1</sup>」の定義によれば、オープンデータは「誰でも自由に利用することができ、再利用や再配布も自由に行うことができるデータ」である。メディアなどでは「公共データの民間開放」と表されることも多いが、オープンデータは行政が作成したデータ(ガバメントデータ)とは限らない。政府主導で取り組んでいることから公共性の高いデータとなるケースがほとんどであるが、例えば電力使用状況など、民間企業が保有するデータであっても広義の意味でオープンデータとなる。また、ガバメントデータのすべてがオープンデータというわけではなく、個人情報や国家機密に関わるデータはオープンデータの対象外となる。

オープンデータの取り組みは、「オープンガバメント(開かれた行政)」の一部として進められており、情報公開の基本コンセプトに位置付けられている。オープンガバメントの原則として、米国のオバマ大統領は、2009年の就任時に「透明性」「市民参加」「コラボレーション」の3つを示した。オバマ大統領の目指すところは、単に

データを公開する「電子政府」ではなく、インターネット技術を使って市民が積極的に利用できるプラットフォームであり、「ガバメント(Gov) 2.0<sup>\*2</sup>」への取り組みである。そして、市民や企業によるオープンデータ活用の進展によってイノベーションや社会生活の向上といった効果が生まれることを期待し、成熟した社会の創出や経済活性化も目的としている。

オープンデータ、ガバメントデータ、オープンガバメント、オープンガバメントデータの関係性を図式化すると、図1のようになる。

なお、「ビッグデータ」は、大量データを収集・分析することにより、新たな知見を発見することを目的としており、一般に「ビジネスに役立つ知見を導出するためのデータ」との側面が強い。切り口は異なるものの、両方に交わる領域のデータもある。





### 「オープン性」に重要な2つの要素

オープンデータは、「公開する」ことよりも「活用する」ことに重きが置かれており、広く活用してもらうためにはデータの「オープン性」が重要となる。

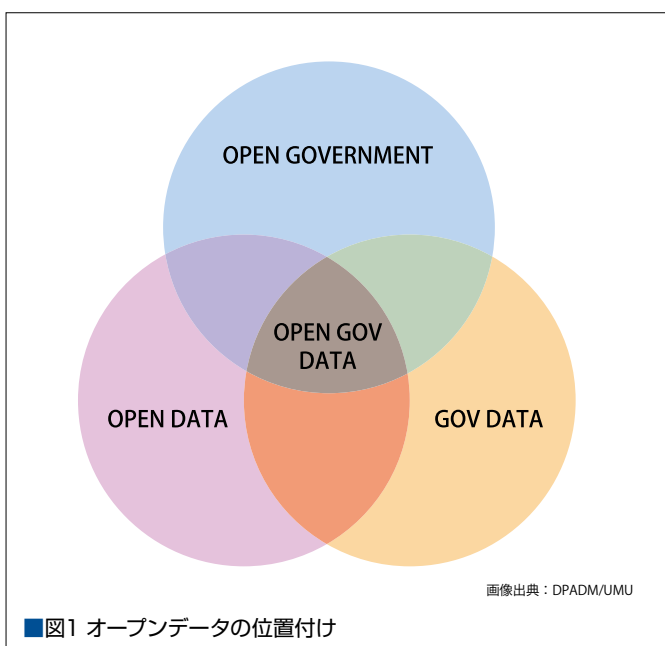
#### 1. 法的なオープン性「オープンライセンス」

公開にあたっては、データを商用で二次利用できるのかどうかのライセンス条件を明らかにする必要がある。著作権の侵害になる

表1 CCライセンスの4つの条件

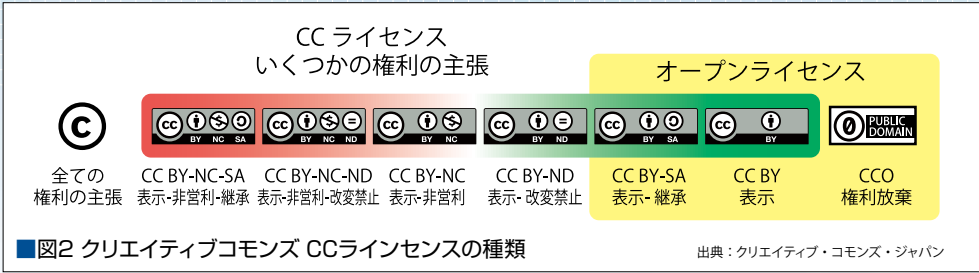
	表示(BY)	作品のクレジットを表示すること
	非営利(NC)	営利目的での利用をしないこと
	改変禁止(ND)	元の作品を改変しないこと
	継承(SA)	元の作品と同じ組み合わせのCCライセンスで公開すること

出典：クリエイティブ・コモンズ・ジャパン



\*1 オープンコンテンツやオープンデータなどのオープンナレッジを推進する非営利団体。2004年にイギリスで発足され、日本でも2012年に「オープン・ナレッジ・ファウンデーション・ジャパン(OKFJ)」が発足された。

\*2 「Web 2.0」を提唱したティム・オライリー氏が2009年に提唱。



データを結合したり分析したりといった情報活用が非常に困難になる。そのため、オープンデータはマシンリーダブルな（機械可読）フォーマットで公開することが求められる。

オープンデータのオープン性は、表2の5段階に分類される。マシンリーダブル

かもしれないと考えたり、利用申請などが煩雑であったりすると、なかなか活用が進まないからである。

採用するライセンスは、イギリスやフランスのように独自のライセンスを確立する国もあるが、多くの国で採用されているものに「クリエイティブコモンズライセンス(CCライセンス)<sup>※3</sup>」がある。CCライセンスは、「本条件を守れば自由に使用可」という意思表示をするためのツールである。CCライセンスを示すことで、データ提供者は著作権を保持したままデータを自由に流通させることができ、データ活用者はライセンス条件の範囲内で再配布や再構成などが可能になる。

CCライセンスでは、表1に示す4つの条件を用いる。これらの4つを組み合わせた基本的なCCライセンスは、全部で6種類ある。図2において、最左の「C（コピーライト：権利者の死後50年まで保護）」と最右の「PD（パブリックドメイン：保護期間終了または権利放棄）」が、これまで広く使われた著作権制度である。CCコモンズの6つの選択肢はCとPDの間に存在し、このうち日本では、最も利用範囲が広い「CC BY（表示）」ライセンスを推奨する方向で検討が進められている。すでに、国内の多くのオープンデータにCC BYライセンスが付与されており、データの出典を表示することを条件に商用利用や改変を許可している。

2. 技術的なオープン性「マシンリーダブル」

行政が印刷物やPDFフォーマットで情報を公開した場合、人間は視覚的に理解できても、機械には処理できなければ、大量の

表2 オープンデータの5つの段階

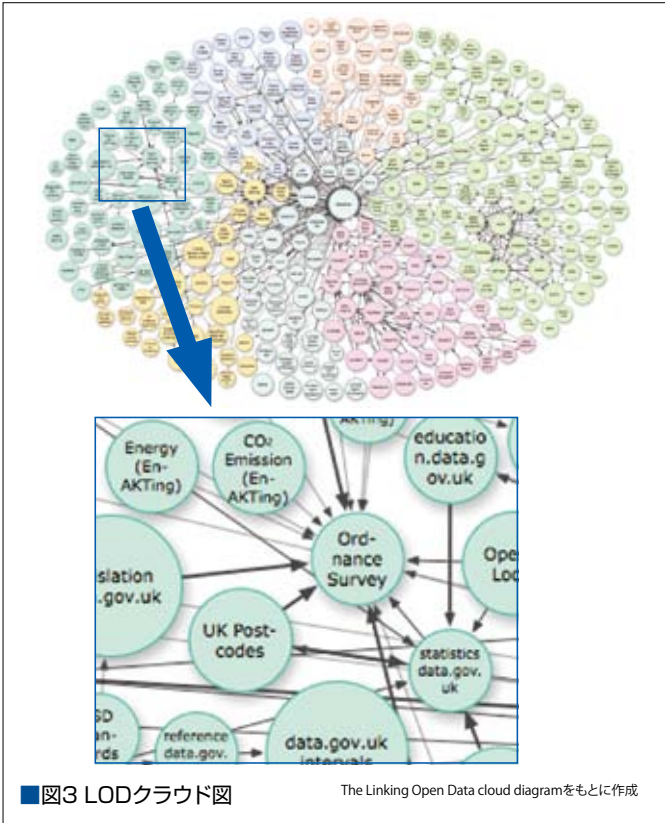
★★★★★	他へのリンクを入れたデータを公開	オープン性 ↑ 高 ↓ 低
★★★★★	RDF(と SPARQL)でデータ公開 例：RDFa、RDFストア	
★★★★	オープンに利用できるフォーマットで公開 例：CSV	
★★★	コンピュータが処理可能なフォーマットで公開 例：xls、doc	
★	どんなフォーマットでもよいからオープンライセンスでデータ公開 例：PDF、jpg	

Open Dataおよび経務省ホームページをもとに作成

ルでなくてもオープンライセンスであればよいのが第1段階の「1★（スター）」で、★が増えるほどオープン性が高くなる。

このうち、「4★」の「RDF (Resource Description Framework)」は、データの作成日時や作成者、タイトルといったメタデータを記述できるフォーマットであり、マシンによるデータセットの分類や検索などの自動化、効率化を図ることができる。さらに「5★」の「LOD (Linked Open Data)」は、RDFを使用してハイパーリンクで相互につながれた「Linked Data」のうちの公開されるデータに該当し、検索にはクエリ言語「SPARQL（スパークル）」を使用する。HTMLが「文書のWeb」であるのに対し、LODは「データのWeb」とも表現される。RDF、LOD、SPARQLは、いずれもWebに関する標準化団体「W3C」の規格である。

LODフォーマットを使っている世界のデータ公開サイトの相互関係を図3に示す。色は、データの分類(ガバメント、ジオグラフィック、メディア、ライフサイエンスなど)を表し、円の大きさと線の太さはデータ量と外部リンクを示す。



※3 国際的 nonprofit 組織「クリエイティブコモンズ (Creative Commons)」が提供している。  
 ※4 Webの発明者でありLinked Dataの創始者でもあるティム・バーナーズ＝リー氏が提案。

## 世界のオープンガバメント動向

オバマ政権のオープンガバメント政策により、米国は2009年5月、オープンデータのポータル「data.gov」を開設。約250の組織が参加し、95,000を超えるデータセット（地図情報も含む）やAPI、アプリを提供している（2014年4月時点）。同様のポータルは、世界44の国と163の自治体が開設しており（2014年4月時点）、国連やOECD、世界銀行といった国際機関もポータルを開設している。

こうした国や機関ごとの取り組みとともに、国際的な協調の枠組みにおいてもオープンデータが推進されている。2011年9月、8カ国（米国、ブラジル、インドネシア、メキシコ、ノルウェー、フィリピン、南アフリカ、イギリス）が「オープンガバメント・パートナーシップ（OGP）」を設立。OGPは、参加国に対してアクションプランの作成と実行を義務付けており、各国が設定された目標を期限までに実行することで、自国のオープンガバメント政策を推進する仕組みになっている。2014年4月時点で63カ国が参加している（日本は未参加）。

2013年6月に北アイルランドで開催されたG8サミットでは、日本を含む主要8カ国が「オープンデータ憲章」に合意。「（公表できない合理的な理由のあるものを除く）すべての政府データが原則として公表される」ことなど、5つの原則が示された。また、価値の高いデータとして、法人、犯罪と司法、地球観測、教育、エネルギーと環境、財政と契約、地理空間、世界的な開発、政府の説明責任と民主主義、健康、科学と研究、統計、社会的流動性と福祉、輸送と社会基盤、の14分野を明記。国別の行動計画などの共同アクションが明記され、オープンデータはグローバルステージへ入ってきた。

## 日本における行政の取り組み

### 1. 政府の取り組み

日本政府としての初めてのオープンデータに関する基本戦略は、内閣官房のIT総合戦略本部が2012年7月に策定した「電子行政オープンデータ戦略」である。翌年の6月に閣議決定された「世界最先端IT国家創造宣言」では、2014年度と2015年度の2年間でオープンデータ推進の集中取り組み期間と位置付けられた。前述の「オープンデータ憲章」を受けて公開された日本版data.gov「データカタログサイト試行版DATA.GO.JP」は、2013年12月にクリエイティブコモンズのCC-BYライセンスで公開。2014年度の本格版への移行を予定している。

また総務省では、オープンデータを社会全体で円滑に活用して

いくための流通環境の整備を行っている。各種実証実験を進めるとともに、「オープンデータ流通推進コンソーシアム<sup>\*5</sup>」との連携により、技術仕様や二次利用ルールの検討、情報発信を行っている。

さらに経済産業省では、実践的にオープンデータ推進に取り組む中で、ポータル「Open DATA METI」を開設。データ公開に加えて、アプリ開発者や、ポータル開設を目指す自治体や行政機関向けの情報提供も行っている。

### 2. 地方自治体の取り組み

福井県鯖江市は、政府の「電子行政オープンデータ戦略」よりも前の2010年12月、地元のICT企業関係者の提案を同市が取り入れる形で、XMLおよびRDFフォーマットを使ってガバメントデータを公開するWebサイト「データシティ鯖江」を開始。災害時の避難所や公衆無線LANのアクセスポイント、コミュニティバスの位置情報などを公開し、APIも併せて公開している。こうした同市の取り組みは、地元の企業がアプリ開発などで支援され、市民主役、市民協働のまちづくりが行われている。

鯖江市のほか、神奈川県横浜市、千葉県の千葉市や流山市、静岡県、福井県など、20以上の自治体で具体的な取り組みが始まっている。また、近隣自治体間で公開ルールの標準化を協議する動きも見られる。

### 3. 民間の行政への参画

政府や自治体、NPO主催で、オープンデータを活用して地域課題を解決しようとするアイデアソンやハッカソン、コンテストが、各地で開催されている。アプリコンテストでは、大手ICT企業がスポンサーとなって活動を支援するケースも多く見られる。また、「コード・フォー・ジャパン<sup>\*8</sup>」や「リンクト・オープン・データ・イニシアティブ<sup>\*9</sup>」のように、ボランティアの立場でアプリ開発を行うなど、自らのスキルを活かすことで行政へ参画するエンジニアの団体も登場している。

## ビジネスにおけるオープンデータ活用

オープンデータをビジネスに活用した国内外の事例について、いくつか見ていく。

### ○農家向け保険

米国の国立気象サービス局が提供する地域ごとの気象データや、農務省が提供する過去60年間の収穫量・土壌に関するデータを解析し、地域や作物ごとの作物被害の発生確率を予測する

\*5 2012年7月設立。関係府省がオブザーバーとして参加しているほか、企業や自治体等、174の会員が参加（2014年4月現在）。

\*6 アイデアとマラソンを合わせた造語で、一定期間集中的にアプリのアイデアをまとめるイベント。ハッカソンと違い、専門知識がなくても参加できる。

\*7 ハック(hack)とマラソンを合わせた造語で、開発者が自発的に集まって一定期間集中的にアプリを開発するイベント。

\*8 米国の非営利組織「Code for America」の日本版。エンジニアを地方自治体に派遣するなどして、地域課題の解決にテクノロジーで取り組んでいる。

\*9 LODの普及促進活動を行っている非営利組織。

サービス。2011年から開始し、取扱高は3兆円。海外展開も実現している。

### ○不動産情報

米国の人口統計、公共交通機関、教育などの情報を物件情報と組み合わせ、教師一人当たりの生徒数や住民一人当たりの医師数といったエリア情報を、不動産業者だけではなく消費者へも提供するサービス。推定年間売上高は約50億円。

### ○図書館蔵書検索

日本の国立国会図書館の書誌情報、全国6,000以上の図書館の貸出状況などを組み合わせ、自分が読みたい本がどの図書館にあるかをリアルタイムで横断検索できるWebサービス。さらに、図書館システムに組み込まれるなど、図書館との連携による新たなビジネス展開も進んでいる。

### ○化粧品メーカーのキャンペーン

日本の気象庁の日照時間や降水量のデータと、自社が持つ顧客の肌データを組み合わせて分析し、都道府県別の美肌スコアをランキング。地域性を打ち出すことで、話題作りと美肌に対する関心を高めることに成功した。

## オープンデータで広がるビジネス機会

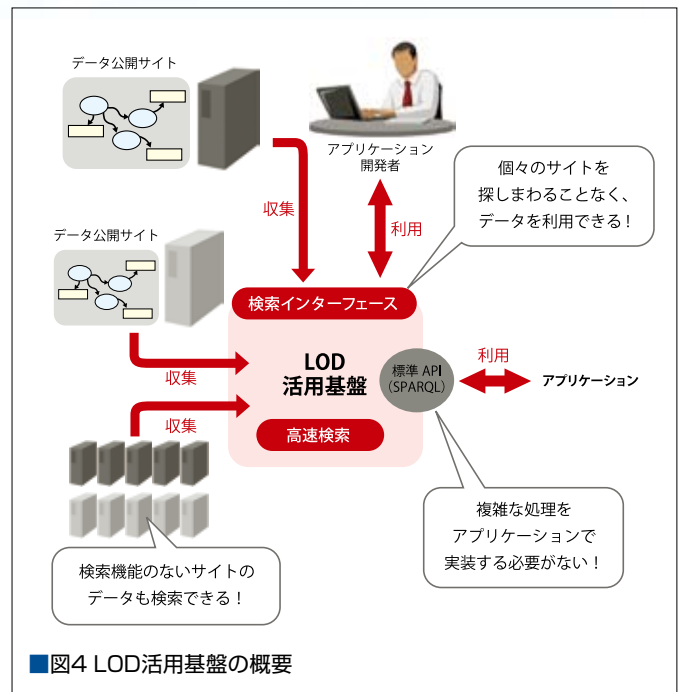
これまで、行政のホームページの多くが著作権を主張し、ほとんどの情報については、私的使用や引用など著作権法上認められた場合を除き、無断で複製・転用をすることはできない状況であった。そのため企業は、民間サービスからデータを購入したり、社内で情報収集するなどしていた。しかしオープンデータ化によって、資金をかけずにデータを入手することが可能になり、中小企業やベンチャーにも市場参入が可能になってきている。

実際、米国では気象情報の無料公開によって、ベンチャー発の新しい農業支援サービスが次々と生まれているという。日本の気象庁も2013年5月、過去の気象観測データの公開を開始し、誰でもCSVファイル形式でWebサイトから無料ダウンロードできるようになった。気象観測データのオープンデータ化は、農業分野だけに限らず、服飾や食品、流通業でも、新たなビジネスやサービスの創出機会となりうるだろう。

行政のオープンデータ推進によって、データの洪水が押し寄せてきている。これからの企業には、オープンデータを社内のデータやソーシャルメディアなどと組み合わせる(マッシュアップする)ことで、データを価値あるものとして表現できるサービスの開発が求められるだろう。

## 富士通の取り組み

富士通では、Linked Open Data (LOD)の将来性にいち早く着目してきた。世界各国でオープンデータの活用が進むに伴い、個別に公開されるLODの把握が難しくなっていることから、世界中で公開されている数百億項目にもおよぶLODを収集し、独自アルゴリズムによって従来比5～10倍の高速一括検索を実現するLOD活用基盤を開発した(図4)。



■図4 LOD活用基盤の概要

さらに、官公庁や自治体、研究機関などが保有している各種フォーマットのデータをマシンリーダブルなRDFフォーマットに短時間で変換したうえでLODとして公開するシステムの構築サービスも開始。様々な技術によって日本におけるオープンデータの取り組みを推進している。

### ●富士通関連サイト

- オープンデータの活用革新! リンクが張られた公開データ (LOD: Linked Open Data) 向け大規模データ格納・検索技術を開発  
<http://pr.fujitsu.com/jp/news/2013/04/3-1.html>
- 次世代のインターネットがすぐそこに! LODが描く"データの未来図"  
<http://journal.jp.fujitsu.com/2014/01/10/01/>

### <参考資料>

- オープン・ナレッジ・ファウンデーション・ジャパン (OKFJ)  
<http://okfn.jp/>
- 横浜市政策局政策課  
「調査季報 174号 特集:自治体の未来を切り拓くオープンデータ」  
<http://www.city.yokohama.lg.jp/seisaku/seisaku/chousa/kihou/174/>
- 独立行政法人情報処理推進機構「米国オープンデータの動向調査」  
<http://www.ipa.go.jp/about/research/20130830.html>
- 経済産業省「オープンデータのビジネス展開に向けて」  
[http://www.meti.go.jp/committee/kenkyukai/shoujo/it\\_yugo\\_forum\\_data\\_wg/pdf/h25\\_02\\_06\\_00.pdf](http://www.meti.go.jp/committee/kenkyukai/shoujo/it_yugo_forum_data_wg/pdf/h25_02_06_00.pdf)
- ITPro: 「オープンデータ」は宝の山  
<http://itpro.nikkeibp.co.jp/article/NC/20130829/501046/>

<監修>: 編集委員 井川 祥子 明治安田システム・テクノロジー(株)