

オープンデータの定義・目的・最新の課題

■ ■ ■

庄司昌彦(しょうじ・まさひこ)

国際大学 GLOCOM 主任研究員

■ ■ ■

▶ 自由に使えるデータ

オープンデータとは何か。端的に表現すれば「自由に使えるデータ」である*1。したがってオープンデータ活用とは、「自由に使えるデータ」の社会的流通を増やし、質を高め、活用を促して、政府活動・社会生活・経済活動にさまざまな効果を生み出そうとすることである。

ここでいう「自由」には、「法的な自由」と「技術的な自由」の二つの意味がある。そのうち「法的な自由」については誤解が多い。オープンデータであるための「法的な自由」とは、データを「公開すること」ではなく「オープンライセンスにすること」、つまり利用条件を広く開かれたものにするることなのだ。なお、中国語(台湾)ではオープンデータを「開放資料」と表記する*2。「見せる」という意味の「公開」ではなく「自由に使ってよい」という意味を持つ「開放」と訳したのは非常に適切であり、参考になる。

企業がオープンイノベーションやブランド化などビジネス面の効果を狙って「自由に使えるデータ」を社会に提供することもできるし、ウィキペディアやオープンストリートマップ*3のように個人が趣味やさまざまな動機でオープンデータを作ることでもできる。しかしオープンデータといえば、行政機関の持つデータが対象となることが多い。それは、行政機関が業務で作成しているデータはそもそも税金で作られたものであり、その中には公共財としてもっと有効活用すべきものが含まれているからである。

これまでも行政機関のデータは膨大に「公開」されてきたが、「開放」は進んでいなかった。通常、行政機関のウェブページは最下部に「All Rights Reserved」と表記し、掲載情報の著作権を主張している。これではデータや



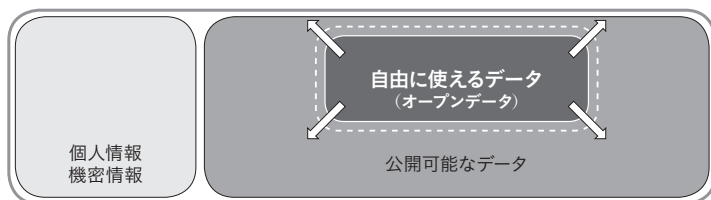
庄司昌彦 ▶ 国際大学 GLOCOM 主任研究員・講師。情報社会学、電子行政、地域情報化、社会ネットワーク、社会イノベーション等を研究。内閣官房 IT 戦略本部「電子行政に関するタスクフォース」、総務省「行政機関等が保有するパーソナルデータに関する研究会」構成員等を歴任。一般社団法人オープン・ナレッジ・ファウンデーション・ジャパン (OKFJ) 代表理事、一般社団法人インターネットユーザー協会 (MIAU) 理事も務めている。

文書等の引用はできても、改変や商用利用はできない。また利用登録や報告を求め使いにくくしている例もある。オープンデータ政策とは、このように利用しにくかったデータの利用条件を緩和し「自由に使えるデータ」を増やすことに意義がある。未公開データを新たに公開することも重要だが、それよりも自由に使えるようにすることこそが本質である。

もちろん、利用条件の緩和だけでは不十分だ。行政機関のデータはウェブ公開されていても探しにくく、形式や語彙が揃っていなかったりして使い勝手が悪い。また異なるデータ間の関連づけをしたリンクト・オープンデータ形式にしたり、プログラムの要求に応じてデータを提供する仕様 (API: Application Programming Interface) ^{★4} が提供されていたりする方が開発者には優しい。こうした技術的利便性の向上も、データ活用の促進につながる。

また、政府の外郭団体などを通じて、一部の企業や団体に対してのみデータが提供されてきた分野も数多く存在する。そうした分野で参入障壁を下げ、分野の壁などを越えてより多くの人々が創造的なデータ利用を試せるようにすることも求められる。

図1 オープンデータは公開かつ自由に使えるデータを増やすことである



出所：筆者作成

行政機関でオープンデータ化の対象となるのは、統計情報や地理空間情報（地図等）はもちろん、気象や交通、健康関連のデータや、各機関の予算や調達情報など、あらゆるものだ。ここには数値も文書も画像も、ビッグデータもスモールデータも含まれる。そして、機密情報や個人情報など、そもそも公開できない情報は対象とはならない（図1）。

▶ 目的1. 民主主義と透明性外交

政府では2012年に、内閣官房IT戦略本部が「電子行政オープンデータ戦略」を策定し、オープンデータへの取り組みを本格化させた。2013年に策定した政府全体のIT戦略「世界最先端IT国家創造宣言」（閣議決定）でも、オープンデータを柱の一つにしている。

オープンデータ政策の目的は何か。電子行政オープンデータ戦略は、①透明性・信頼性の向上、②国民参加・官民協働の推進、③経済の活性化・行政の効率化の三つを挙げている。①は政治分野で民主主義の質を高めること、②は行政分野で社会課題に社会全体で取り組むこと、③は経済分野で新たな価値を生み出したりコストを下げたりすること、と言い換えることができる。以下、三つの目的の一つずつ詳しく検討していく。

まずは民主主義や透明性の向上である。一般に、政策課題の設定、対処方針の立案、法律・条例案の作成、予算の確保、施策の実施、評価という一連のプロセスは、ほとんど行政機関が主導している。行政が関連業界や専門家等と密接に協力し、舵取りをすることが効率的で効果的なこともある。しかし情報は権力の源泉であり、政府と主権者である住民との間で情報の非対称性が大きければ、住民が政府の活動を適切に評価したりコントロールしたりすることが難しくなる。さらに、人口や税収が減少するなか、社会課題の複雑化やこれまでの制度の疲労・形骸化が進み、政府だけでは課題解決が困難になっている面もある。そのため、政府活動や社会の実態をできるだけ住民と共有していくことが求められる。行政や政治家にとっては自分たちの行動が批判される可能性があるため、透明性志向のオープンデータ化は避けられがちであるが、政府が説明責任を果たし、民主主義の質を向上させるうえで重要である。

透明性は外交課題でもある。2013年6月に英国で行われたG8サミット（主要8カ国首脳会議）では、透明性が議題の一つとなり、「オープンデータ憲章」が合意された。この憲章で日本など8カ国は、オープンデータ推進のための行動計画を作成し、毎年実施状況を報告しあっていくこととなった。また、日本は不参加だが、2010年に結成された「オープン・ガバメント・パートナーシップ」という国際組織も存在する。ここには多数の途上国を含む64カ国（2014年10月現在）が参加し、協調しながら透明性向上などに取り組んでいる。

2013年9月のオープン・ナレッジ・フェスティバル^{*5}では、英国政府担当者^{*6}が、オープンデータは「国内問題のステージからグローバルステージ」に移行したと述べた。そして英国政府は、世界銀行と組み120万ドル（約1.2億円）の予算で、途上国向けのオープンデータ支援活動を始めている。英国が透明性向上を国際連携のテーマとして進める背景には、正義の実現のような大上段の目的や個別ビジネスの創出のようなことだけではなく、世界各国の政府や社会を安定で信頼できるものとし、国際的なデータ流通を円滑化することで、中長期的に経済・社会の変化を主導したいという意図があるのだろう。

さまざまな分野のデータが標準化され、オープンデータとして公開・流通・活用される「つながった情報経済・社会」になることは、インターネットが社会に与えた影響と同様に、組織のあり方や仕事の仕方までも変えていくと考えられる。そのときに、たとえ日本の技術や制度が優れていたとしても、孤立しては「つながった情報経済・社会」の恩恵を受けることができなくなる。透明性を掲げて進む国際動向と協調し、ルール形成にも参加していく姿勢が求められる。

▶ 目的2. 政府と社会のオープンマインドな協働

オープンデータ政策の背景には、「オープンガバメント」の思想がある。これは、政府の機能や情報をあらためて社会に開き、情報通信技術の特徴を活かして多くの人々や企業の参加と協力を得て社会課題に取り組むというものだ。米国オバマ大統領は、「透明性・参加・協働」の三つをオープンガバメントの原則として示してきた。政府活動や社会状況の「透明性」を高める

と人々の「参加」を呼び込み、それが成熟すると官民協力が本格的な「協働」へ進むという発展段階を描いているのだろう。

政府の意思決定や活動に対する人々の参加は、民主主義の歴史の中でも長年のテーマであり、またパソコン通信やインターネットの普及とともに長く論じられてきた。このような「参加民主主義」論では、人々が政府活動へ参加することには、社会を構成していくうえでの規範的意義や、公共的課題への関心を深める教育的意義があるとされる。

東日本大震災からの復興では、さまざまな自発的な社会活動が行政を補完してきた。この構図は、オープンデータ活用でも同様である。英、米、日本など世界各国の政府が採用しているデータポータルサイト構築ソフトのCKANも、著作物ライセンスのクリエイティブ・コモンズも、民間団体が生み出し事実上の国際標準となっている。こうした多様な民間活動と協働する行政のあり方は、政府だけではすべての課題に対処しきれない時代の参加民主主義のあり方を示唆している。

だが実際には、政府と社会組織の協力がごちない例もある。たとえば、オープンガバメントを掲げ行政機関が運営しているソーシャルメディアのアカウントが双方向のコミュニケーションを避け一方的な情報発信に終始していたり、行政機関のデータ提供において悪用を避けるという名目でさまざまな制約を設けていたりすることなどが挙げられる。

オープンガバメントにふさわしい良質な官民コミュニケーションとはどのようなものだろうか。たとえば社会課題には多様な利害関係者が存在することを踏まえ、政府、企業、消費者・国民、NPOなどの組織や個人が対等な立場で議論し、学び合い、解決策を見出していくマルチステークホルダー・プロセスという手法が参考になる^{*7}。立場の異なる人々が対話し、合意することは簡単ではなく、時間や手間が膨大にかかる。しかし、当事者たちが対等につながり情報を共有し、自発的に調整して形成した合意は、全員が自分の権利や責任を理解した良質な合意である。このような「協働」を進めるためには、当事者各自が他者を受け入れるオープンマインドと粘り強さが必要だ。

▶ 目的3. 経済効果, エコシステム

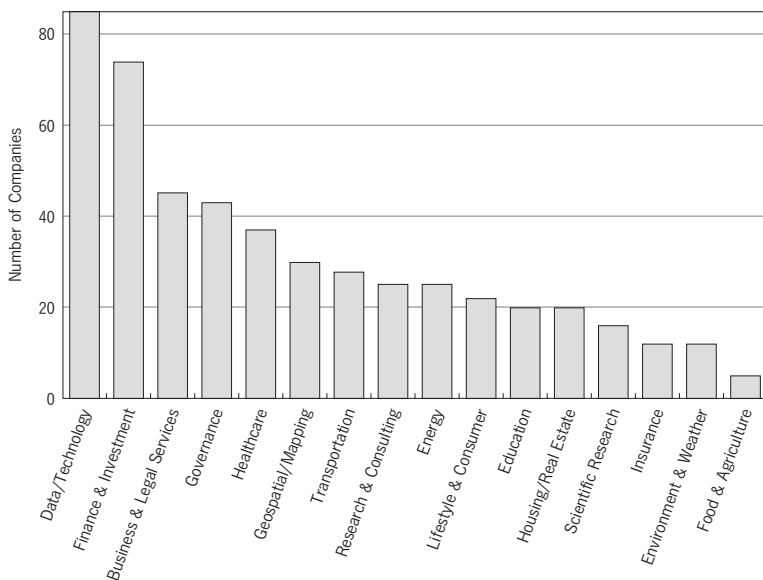
オープンデータ活用の目的の三つ目は、経済効果である。政府はオープンデータ活用が柱の一つである「世界最先端IT国家創造宣言」を成長戦略の一部に位置づけ、経済効果への期待をしだいに強調するようになってきている。

行政機関によるデータ開放の先駆的かつ最大の成功例は、米国政府によるGPSデータの民生利用化であろう。これは年間900億ドル(約9兆円)規模の位置情報サービス市場を生み出したとされており^{★8}、第二、第三のGPSが期待されている。

オープンデータ活用が地域経済に与える効果については、Vickery [2011] がEU 27カ国(当時)の公共データがオープン化された場合の経済波及効果を年間2,000億ユーロ(27兆円以上)と算出し、それに基づいて高木 [2012] が日本における経済波及効果を年間で最大5.5兆円と試算している。また

図2

オープンデータ500(業種別)



出所: The GovLab (CC BY-SA 4.0)

McKinsey Global Institute [2013] は世界全体で毎年最大5.4兆ドル(約540兆円)という巨大な経済効果を推計しており、期待値は高まるばかりである。

一方、米国ニューヨーク大学のGovLabは、「オープンデータ500 (Open Data 500)」と題するオープンデータ活用企業の500社リストを作成し、企業の分析を行っている(図2)。業種別の分類では「データ／技術」が最も多いが、それに「金融と投資」「ビジネス・法律サービス」「ガバナンス」が続く。つまり情報システムやアプリ等を開発するいわゆる「IT企業」以外に、データを分析し価値ある情報を生み出して提供する「情報サービス企業」が多い。

オープンデータを提供するのは、行政機関だけではない。行政とは異なり、企業にオープンデータを提供する義務はないが、社内では生まれにくい発想を取り込むオープンイノベーションや、技術者コミュニティ等との関係強化、社会的責任や社会貢献など、オープンデータの提供は経営上の合理的な戦略になり得る。たとえば、JR東日本や東京都交通局をはじめとする首都圏の公共交通事業者などは「公共交通オープンデータ研究会」を設立し、運行情報や施設情報等のオープンデータ化を検討している。

経済効果を大きく引き出していくためには、個別のシステムやアプリをどう作るかという視点だけでは足りない。官民の情報をどこまでオープン化し社会的な情報流通のどの過程に価値を加え新たな市場を生み出すのかというB2B ビジネスやエコシステム志向の視点が必要であろう。

▶ 課題1. パーソナルデータ

オープンデータの活用を進めるために、克服すべき課題も少なくない。以下では最新の課題を四つ列挙したい。

まず懸念として論じられるのが、オープンデータ化によって個人情報が多く知れ渡るのではないかという類のものだ。しかし、これはたいてい誤解である。オープンデータは「公開可能」な情報をより自由に使えるようにしようという話なのであって、第三者提供が認められないような個人情報を新たに提供し始めようという話ではない。

ただし、他のデータとの掛け合わせによって個人が特定されたり、プライバシーが明らかにされたりする可能性(モザイク効果)は存在する。たとえば、

地域の人口データやある属性の人々のデータを数百メートル四方のメッシュ形式のオープンデータとして提供すれば、地域の特性を知るうえで便利であるが、人口が少ない地域で住宅地図と重ねたりすれば、個人や世帯が特定されることがあり得る。

そこで現在、政府はパーソナルデータ活用に向けた検討を進めており、個人情報保護法の改正を予定している。ここでも、パーソナルデータを自由に使えるオープンデータにしようという話は出ていない。政府の「パーソナルデータ大綱」によると、個人を識別できるが特定が困難なデータ（個人特定性低減データ）の取り扱いについては、本人同意を取得するか、個人の再特定を禁じるなど一定のルールに従わなければ活用を認めない方針である。データ活用への不安を減らすための取り組みは進んでいる。

▶ 課題2. 透明性とビジネス

透明性の向上は、行政や政治家の監視以外に意義があることは前述のとおりであるが、ここではビジネスとしての可能性を具体的に紹介したい。欧州のアプリコンテスト等で高く評価されている「オープンコーポレイツ」とい

図3 オープンコーポレイツによる企業のグローバルな関係の可視化



出所：<<https://opencorporates.com/viz/financial/index.html#goldman/jp>>

う企業のサービスは、さまざまな企業情報を提供する信用調査会社の役割をオープンに提供するものである。法人・商業登記情報等の企業活動関連のデータを世界中から集め、2014年10月現在で世界8,468万社のデータを掲載している(図3)。オープンデータを提供するAPIも公開しており、利用者は自由にデータを使って情報サービスを作ることできる。世界中の企業の情報が自由に使えるようになることの影響は大きく、信用調査会社等のビジネスを脅かすだろう。企業情報のオープン化はG8サミット等の国際的なオープンデータ議論での重要項目であり、このようなサービスは今後さらに成長すると見込まれる。

しかし日本では法人・商業登記情報はオープンデータになっていない。一般財団法人民事法務協会が法務省から指定を受けて^{★9}ウェブサイトから提供しているが、データは有料であり、夜間や休日はサービスが停止しているため利用しにくい。

現状ではオープンコーポレイツのようなサイトに日本企業の情報を登録するのは難しいが、世界中の企業が載るデータベースに載らないことが日本企業の不利益となったり、元データを提供しない日本政府が閉鎖的であると評価されたりすることが懸念される。マイナンバーとともに導入される法人番号を活用し、企業情報の提供を進める必要があるだろう。言い換えれば、積極的に透明性を高めることが経済活動の基盤整備につながるのである。

▶ 課題3. アイディアソン・ハッカソン・マーケソン

政府や地方自治体がオープンデータの活用を進めるために、「アイディアソン」や「ハッカソン」を開催するようになってきている。アイディアソン(Ideathon)とは、人々が集まりアプリケーションやサービス開発のアイデアを出し合い、対話を通じて具体的な案にまとめ上げるイベントである。マラソンを走りきるように「アイディア出しきる」ことから二つの語を組み合わせている。ハッカソン(Hackathon)はプログラミング(ハック)を行い、限られた時間の中で試作品を「作りきる」イベントである。

これらのイベントには、データへのニーズが明らかになる、多様な参加者との対話や作業を通じて各自が発見を得る、地域社会の現状や課題などにつ

いての学習機会となる、後につながる人的ネットワーキングができる、などの効果がある。

しかし、経済効果や社会課題の解決を期待する立場からは、現実の社会課題を解決していない、いつも同じようなアプリを作っている、消費者向けの分かりやすいものになりがち、大きなビジネスが生まれそうにない、等の批判がある。つまりこの手法は、生まれたアイデアや試作品を、実用できる行政サービスやビジネスに磨き上げるところで課題を抱えている。

つまり「アイデアソン・ハッカソンの後」の展開を考えていく必要がある。そのアイデアの一つが「マーケソン」だ。これはマーケティングとマラソンを合わせた筆者の造語で、ハッカソンで生み出されたアプリケーションやサービスの有効性を検討し、使い勝手の改善をしたり、普及のための方策を考えたりするイベントである。

オープンデータを提供し、アイデアソン・ハッカソンをすれば、自動的にビジネスが生まれるわけではない。データを活用してビジネス機会や社会課題を発見・分析し、サービス等に具体化し、事業化し、さらにそこから新たなデータを生み出すことで社会に還元するという、一連のサイクルを回していくことが必要だ。マーケソンだけでそのサイクルが回るわけではないが、現在課題となっている「事業化」段階を前進させる一助となるだろう。

▶ 課題4. 準オープンデータ

オープンデータと称しているのに、正確にはオープンではないデータがある。たとえば、東京メトロは列車の位置情報等を利用した「オープンデータ活用コンテスト」^{★10}を開催し、リクルートも「オープンデータ」^{★11}と銘打って自社サービスの利用データを学術研究向けに提供している。しかし、これらのデータの利用には用途や資格等に制約があり、「自由に使える」という要件を満たしていない。そこで、オープン化が義務付けられているわけではない企業が積極的にデータを社会へ提供することは評価すべきだが、「オープンデータ」という言葉の誤用はデータ利用者の混乱を招きかねない、という指摘を受けている。

これは、「自由に使えるデータ」という定義が浸透していないために起き

ている。また、完全に自由に使える形式で提供しようとして組織内の調整がつかなかった等の事情もあるだろう。筆者は、このような何らかの制限によってオープンデータと呼ぶための基準を満たしていないものを「準オープンデータ」と呼ぶよう提案している。これは、基準を満たしていないものと満たしているものを名称で明確に区別し、「準」のものが完全なオープンデータになるよう促すことを意図している。準オープンデータがオープンデータになるための障害が何であるのか、その克服の方法にはどのようなものがあるのかについて、特に企業の立場に立った研究はまだ少ないと思われる。今後の課題である。

▶ 本特集にあたって

本特集は、「自由に使えるデータ」であるオープンデータの活用を通じて社会的なデータ活用の将来を展望するものである。具体的なビジネスの可能性や経済効果の展望、アジアや欧米諸国の動向、国内で活用を進めるためのビジョンや先行事例等を、最先端で活躍する方々に論じていただいた。このテーマを社会科学観点から11人もの論者が議論したものはこれまでになく、多様な知見を集めることができたと考えている。一人でも多くの読者の方にお役立ていただければ幸いである。

註

- ★1——詳細は、世界的な推進団体であるオープン・ナレッジ・ファウンデーションの「オープンの定義」を参照。要約すると、オープンデータとは「自由に使えて再利用もでき、かつ誰でも再配布できるようなデータのことであり、従うべきはせいぜい『作者のクレジットを残す』あるいは『同じ条件で配布する』程度である」。<http://opendefinition.org/od/japanese/>
- ★2——たとえば、台湾政府行政院（内閣）「政府開放資料推動策略」（政府オープンデータ推進戦略、2012年）。<http://www.ey.gov.tw/News_Content.aspx?n=40A033D0C65F76EB&s=32B9ADA064C4506C>
- ★3——世界中の人々が協力して誰でも自由に使える地図を作ろうというプロジェクト。「地図のウィキペディア」とも呼ばれる。<<http://osm.jp/>>
- ★4——総務省（統計）、気象庁（防災情報）、環境省（大気汚染物質）、復興庁（復旧・復興支援制度）によるAPI提供が行われており、政府データカタログサイト（DATA.GO.JP）も提供を予定している。
- ★5——2004年から毎年開催されているオープンデータに関する民間イベントで、世界中か

ら産官学民の人々約1,000人が参加している。

- ★6—— ポール・モルトビー (Paul Maltby) 内閣府透明性・オープンデータ担当ディレクター
- ★7—— 佐藤正弘 [2010] 「新時代のマルチステークホルダー・プロセスとソーシャル・イノベーション」『季刊 政策・経営研究』Vol.3を参照。
- ★8—— JETRO/IPA 「米国オープンデータの動向調査」, 2013年3月 <<http://www.ipa.go.jp/files/000033718.pdf>>
- ★9—— 「電気通信回線による登記情報の提供に関する法律」に基づく。
- ★10—— 東京メトロオープンデータ活用コンテスト <<http://tokyometro10th.jp/future/opendata/>>
- ★11—— リクルートオープンデータ <<http://atl.recruit-tech.co.jp/opendata/>>

参考文献

- 1) Vickery, G. [2011], “Review of Recent Studies on PSI Re-use and Related Market Developments.” <http://ec.europa.eu/information_society/newsroom/cf//document.cfm?doc_id=1093>
- 2) 高木聡一郎 [2012] 「電子行政イノベーション④ オープンデータの市場規模に迫る」『行政&情報システム』Vol.48, No.5
- 3) McKinsey Global Institute [2013], “Open data: Unlocking innovation and performance with liquid information.” <http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information>