オープンデータ活用に関する研究
～　ＩＴコーディネータにとっての
オープンデータビジネスの可能性
（基本編）　～

２０１５年３月３１日

企業内 ITC・IT ガバナンス研究会

序

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」を指す。

近年公共データの活用促進、すなわち「オープンデータ」の推進により、行政の透明性・信頼性などが高まることが期待され、更には「オープンデータ」の活用による、革新的事業の創生、地方の活性化などが期待されている。

本稿では、行政機関が保有するデータの公開にフォーカスして論述する。

オープンデータは従来の情報公開と異なる点が3点あり、これは①オープンライセンス、②機械判読可能、③オープンフォーマットである。

今年度は、

- どこに、どの様なオープンデータが、どの様な形で格納されており、
- どうすれば、それらを取り出すことが出来て、
- どの様な活用が可能か…

にたどり着き、IT コーディネータとして、どのような活用局面があるだろうかを、事例を挙げて論述したい。

2015年3月

執筆者　一同

執筆メンバー　ITガバナンス研究会

久住　昭之（元 IT ｺｰﾃﾞｨﾈｰﾀ）

坂本　徳明（0064952006C）

千枝　和行（0029302004C）

古川　正紀（0005462001C）

山﨑　直和（0035252003C）

（注）本記載内容は、ITコーディネータ個人としての見解を述べたものであって、個人が所属する企業・団体としての見解を述べたもので無いことをお断りします。

また、本書において使用しているシステム名や製品名などで各メーカー等の登録商標を使用している部分があるが、文中においては TM、コピーライト表記はしておりません。

1. はじめに

オープンデータ（Open Data）とは、特定のデータが、一切の著作権、特許などの制限なしで、全ての人が望むように利用・再掲載できるような形で入手できるべきであるというアイデアである。

オープンデータの概念は新しいものではない。しかし、形式的な定義は新しい。Open Definition による形式化は、「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る。

近年、企業では内部データや SNS などの Web データを分析し、マーケティングや新サービスに活用する「ビッグデータ」の活用がビジネスの手法として広く知られるようになった。

また最近では、自らの組織が保有するデータだけでなく、外部リソースとして行政機関が公開する「オープンデータ」とを組み合わせ、新たなビジネスを立ち上げる動きが欧米を中心に活発化している。

1-1. 「オープンデータ」と言えるための条件

「オープンデータ」と言えるためには、① 機械判読に適したデータ形式で、

② 二次利用が可能な利用ルールで公開されたデータあるという条件がある。

まず、① 機械判読に適したデータ形式とは、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ことである。　つまり、画像ファイルや PDF 等の形式ではなく、Excel データ等がこれに相当する。

次に② 二次利用が可能な利用ルールで公開されたデータあるというのは、第三者がデータを一部改変して利用すること、すなわちデータの二次利用を、データ所有者が予め許諾していることを明示することが必要であると言う事である。

詳細は第 4 章にて論述する。

1-2. 政府のオープンデータ政策

政府のオープンデータ政策の目的としては、予算や調達情報を公開し、行政の透明性を高めること、白書や統計、地図などのデータを公開し、産業界での二次利用を通じてビジネスの活性化を図ることが挙げられる。　また、政府は 2015 年度末までに他の先進国と同水準のデータ公開を行うことを達成目標とし、2013 年 12 月に政府のデータカタログサイト「DATA.GO.JP」を立ち上げるとともに、データ公開の具体的な数値目標として「1 万件のデータセットの公開」を定めている。

政府のこの施策が立ち上がって約 1 年、わが国がオープンデータに本腰を入れ始めたのは、つい最近のことである。　政府の成長戦略では、2014～15 年度を「オープンデータの集中取り組み期間」と位置付けており、今年度と来年度にかけて行政

のデータ開放が進んでいくものと考えられる。

1-3. オープンデータによるビジネス展開

オープンデータが企業活動に与える影響は大きく２つある。　１つは、政府などからオープンデータとして公開されるデータを活用し、自らの事業やサービスを高度化することであり、もう１つは、企業自らがオープンデータ化を推進することで、ビジネスや事業の展開や強化を図ることである。

後者はオープンデータというキーワードで語られることは少ないが、重要である。

企業としてはデータが蓄積されていても、それを最大限に生かせるような解析技術が自社内にあるとは限らない。

そのような中、データを公開し、外部で解析技術を募れば、解決に繋がる回答を得る可能性が増える。

自社内では「使えない」データであっても、それを公開し、より金銭的な価値に変えられる企業に提供、あるいは販売する、すなわちニーズマッチングをすることでビジネスチャンスを掴める可能性もある。このため、データの流通を行う場を提供するデータマーケットプレイスというサービスも立ち上がっている。

しかし、これは現時点ではごくまれな先進的な事例であり、これに到るまでのオープンデータ利活用ノウハウをしっかりと押えた上で取り組むべき課題であろう。

今年度われわれ IT ガバナンス研究会としては、このオープンデータに関して、まずは現在公開されているデータ（多くは政府などからの公開データ）を活用し、ＩＴコーディネータが支援する事業やサービスに利するためのプロセスを追い掛けたいと考えている。

次章以降に、「日本の動向」述べた上で、「オープンデータ」の「データカタログサイト」を紹介し、最後に「IT コーディネータとオープンデータ」の関わり合いを論述し、纏めたいと考えている。

2. オープンデータに関する日本の動向

2-1.日本政府における動向

日本におけるオープンデータに対する取組は、「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 総合戦略本部）決定）に基づいて現在進められている。

本戦略においては、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。

この戦略の中では

① 政府自ら積極的に公共データを公開すること

② 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 つの基本原則が掲げられている。

上記「電子行政オープンデータ戦略」を具現化するために、IT 総合戦略本部の配下に「電子行政オープンデータ実務者会議」が設置され、直近の会合（平成 27 年 2 月 10 日開催）まで計 9 回開催されている。実務者会議においては、主に以下の 3 項目に関して基本的な事項の検討が進められている。

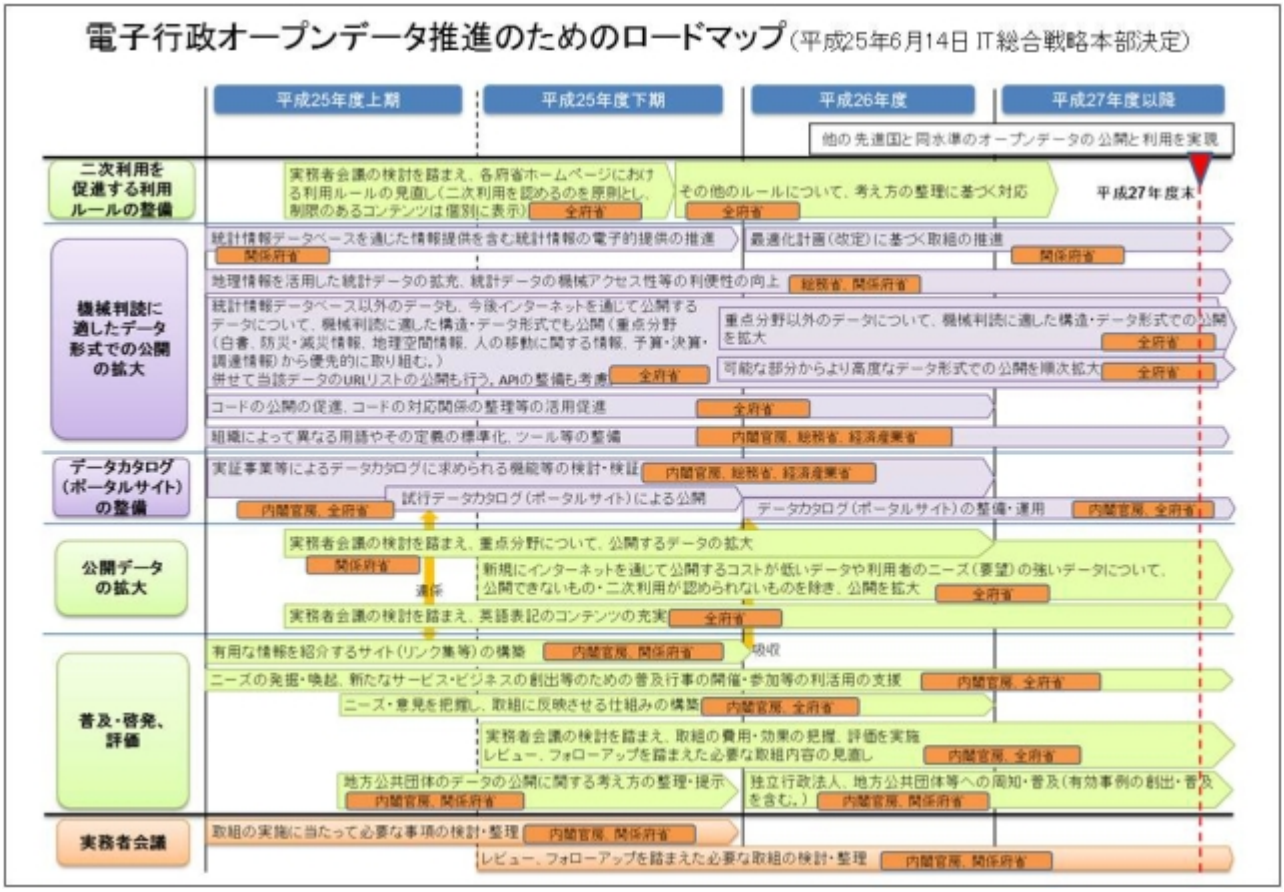
(a)公共データ活用のために必要なルール等の整備

(b)データカタログの整備

(c)データ形式・構造等の標準化の推進

また、平成 25 年 6 月 14 日の IT 総合戦略本部において「電子行政オープンデータ推進のためのロードマップ」（下図参照）が明示され、現在に至るまでそのロードマップに基づき検討が進められている。

【参考】電子行政オープンデータ推進のためのロードマップ（工程表）



平成 25 年 6 月の IT 総合戦略本部で明示された「世界最先端 IT 国家創造宣言」においては、公共データの民間開放を推進するために以下の宣言がなされており、平成 26 年度と平成 27 年度の 2 年間を集中取組期間と位置づけ、平成 27 年度末には他の先進国と同水準の公開内容を実現することを目標に掲げている。

(1) 電子行政オープンデータ戦略に基づくロードマップを策定・公表
(2) 2013 年度から公共データの自由な二次利用を認める利用ルールの見直しを行うとともに、機械判読に適した国際標準データ形式での公開を拡大
(3) 各府省が公開する公共データの横断的検索等を可能とするデータカタログサイトについて 2013 年度中に試行版を立ち上げ、2014 年度から本格運用を実施

上記(3)の成果として内閣官房情報通信技術（ＩＴ）総合戦略室が作成したものが「データカタログサイト（http://www.data.go.jp/）」である。本サイトでは、二次利

用が可能な公共データの案内・横断的検索が行えるようになっている。

2-2. 地方公共団体における動向

地方公共団体においては、前述の政府における取組に倣って裾野市、横浜市、神戸市、福岡市、鯖江市、流山市、静岡県などを始めとして、複数の団体がインターネットを活用して公共データの公開を行っている。これらのデータは二次利用として民間に活用され始めており、活用方法を巡っては各団体が民間と一緒になって活用事例を創出するなどの取り組みが見られる。

具体的には、住民が暮らしやすい街づくりや行政の「見える化」といった地域が抱える課題の解決に貢献できないか、という観点で検討されている。また地方公共団体が主体となった取り組みだけでなく、NPO 等による活動も積極的に行われている。

都道府県としては先進的な取り組みを推進している静岡県では、オープンデータを活用したポータルサイト「ふじのくにオープンデータカタログ」を平成 25 年 8 月に開設している。同年に世界文化遺産に登録された富士山をはじめ、ロケ地などに関する観光情報、気象観測点等の防災情報など、平成 26 年 2 月現在で 90 を超えるデータセットを公開している。また、同年 11 月には、静岡県裾野市が、同市の人口統計や医療機関、防災に関するデータを「ふじのくにオープンデータカタログ」に公開しているほか、民間データの掲載も始まっている。さらに、静岡県は山梨県と連携して、平成 26 年 1 月には、位置情報付きで撮影した富士山の写真を投稿してもらい、投稿された写真をオープンデータとして公開する「富岳 3776 景」を開設している。

3. データカタログサイト

オープンデータは国のみならず、地方公共団体や民間企業、地域住民が一体となってその利活用の促進に取り組むことによって、より大きな効果を発揮することができる。現在、府省庁や地方公共団体ごとに推進活動が行われているが、将来的には、日本全国を横断したオープンデータの利活用ができる環境(データカタログサイト)を整える方向で検討されており、その基盤となるデータカタログサイトとして、府省庁のデータを横断的に検索可能な「DATA.GO.JP」が、2014 年 10 月 1 日より本格稼働を開始している。

本章では、ITC がオープンデータを利活用するために必要となる基本的な知識を理解するために、この「DATA.GO.JP」について解説する。ここで得られる知識は、地方公共団体のオープンデータを利活用する際にも役に立つ。

3-1. 「オープンデータ」を利活用するための条件

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 決定)において、オープンデータの意義・目的について、以下の 3 点が挙げられている。

・透明性・信頼性の向上:

公共データが二次利用可能な形で提供されることにより、国民が自ら又は民間のサービスを通じて、政府の政策等に関して十分な分析、判断を行うことが可能となる。それにより、行政の透明性が高まり、行政への国民からの信頼を高めることができる。

・国民参加・官民協働の推進:

広範な主体による公共データの活用が進展し、官民の情報共有が図られることにより、官民の協働による公共サービスの提供、さらには行政が提供した情報による民間サービスの創出が促進される。これにより、創意工夫を活かした多様な公共サービスが迅速かつ効率的に提供され、厳しい財政状況、諸活動におけるニーズや価値観の多様化、情報通信技術の高度化等我が国を取り巻く諸状況にも適切に対応することができる。

・経済の活性化・行政の効率化:

公共データを二次利用可能な形で提供することにより、市場における編集、加工、分析等の各段階を通じて、様々な新ビジネスの創出や企業活動の効率化等が促され、我が国全体の経済活性化が図られる。また、国や地方自治体においても、政策決定等において公共データを用いて分析等を行うことで、業務の効率化、高度化が図られる。

上記のような意義・目的に合致する形での「オープンデータ」と言えるためには、

(1) 二次利用が可能な利用ルールで公開されたデータ

(2) 機械判読に適したデータ形式で、

である必要があり、それにより、人手を多くかけずにデータの二次利用が可能になる。そこ

で、これら二つについて詳しく述べる。

3-2. 二次利用が可能な利用ルール

「オープンデータ」と言えるためには、提供する公共データを情報利用者が自由に二次利用できることが重要であり、そのためには、二次利用を認める利用ルールを採用することが必要である。第三者がデータを一部改変して利用すること、すなわちデータの二次利用を、データ所有者が予め許諾していることを明示することが必要となる。例えば、著作物には著作権が発生するが、二次利用を広く認めるには、その著作権の不行使を予め宣言しておくことが求められる。他方、現在の各府省等のホームページの利用条件の中には、無断での改変を禁じているものがあり、広く二次利用を認めるものとはなっていない場合がある。また、数値データ等、著作物に該当しないデータについて、著作権の対象であるような包括的な表現となっている場合もある。オープンデータに関する著作権の考え方について説明するとともに、利用ルールに関する国際的な動向、日本政府における動向について紹介する。

3-2-1. 著作権に対する考え方

国、地方公共団体、独立行政法人、公共企業等が保有する公共データには、その表現に創作性がある場合に著作権が発生する。そして著作権が発生している場合、公共データを情報利用者が二次利用したい時には、公共データを作成した国、地方公共団体等にその都度、許諾を得なくてはならない。

2014 年 3 月末現在、国や地方公共団体等のホームページを見ると、公共データの利用に際して無断での改変利用は許可されておらず、公共データを自由に利用できる環境にあるとは言えない。ホームページ上に国、地方公共団体等の公共データについて二次利用可能であることが利用ルールによって明示されると、情報利用者は自由にデータを二次利用できるようになる。「オープンデータ」には、機械判読性を高めるという技術的対応の前に、二次利用を認める利用ルールの採用が必須条件と言うことができる。

さて、国、地方公共団体、独立行政法人、公共企業等が保有する公共データについて、二次利用が可能な形で公開することを実現する手段としては、以下の 3 つの方法が考えられる。

①公共データには原則、著作権は発生しないものとする
②公共データに著作権は発生するが、これを放棄する
③公共データを二次利用可能なルールで公開する

最も望ましいのは、①の公共データには原則、著作権が発生しないものとすることである。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るが、国民、市民等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。

②の「公共データに著作権は発生するが、これを放棄する」の場合、著作権法を改正す

ることなく、国や地方公共団体等が自ら権利を放棄することで、①と同等の効果を得ることができる。ただし、著作権は、国や地方公共団体等の財産の一つであり、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可能かどうか、十分に検討する必要がある。

これらの方法に対し、③は公共データについて著作者は著作権を保持したまま自由に二次利用を認める利用ルールを採用することによって、オープンデータを進めるという考え方である。

①、②のように著作権が発生していない状態、又は、著作権が消滅した状態のことを「パブリックドメイン」と呼ぶ。データがパブリックドメインになると、著作権者によって差し止めや損害賠償等の権利が行使されず、自由に利用できることから、オープンデータにおいてはデータがパブリックドメインになることが望ましいが、法改正や国有財産法等の解釈等には時間を要することから、中長期的にはこれらの方法の検討を進めつつ、短期的に対応可能な③の方法を採ることがスモール・スタートの原則から望ましい。

3-2-2. オープンデータ利用ルールの国際的な動向

アメリカやオランダでは国等が保有する公共データには著作権はないとし、公共データをパブリックドメインの状態におくことによってオープンデータの利活用が活発化している。

それに対して、その他の諸国では政府が作成した公共データについて、著作権やデータベース権が発生するが、利用ルールにより二次利用を可能にするという方法を採っており、その利用ルールの検討に当たっては、①既存のオープンデータに関する利用ルールを採用、②独自の利用ルールを作成のどちらかを選択している。

①を選択した国としては、オーストラリアやニュージーランド、ドイツ等が挙げられる。これらの国では、著作権のある著作物の配布を許可する数種類あるパブリック・ライセンスの一つである「クリエイティブ・コモンズ・ライセンス」（以下「CC ライセンス」という。）を採用している。CC ライセンスでは4つの条件、「表示（BY）」「非営利（NC）」「改変禁止（ND）」「継承（SA）」をもとにこれらを組み合わせた 6 種類のライセンスが用意されており、この中でもオープンデータに適した「表示ライセンス（CC-BY）」を採用している。CC-BY は、基本的に出典を記載すれば自由に二次利用をすることができ、2002 年の当該利用ルール作成以降、既に多くのデータで利用されており、世界的に普及していること等が採用の理由となっている。

CC ライセンスの種類

・表示

原作者のクレジット（氏名、作品タイトルなど）を表示することを主な条件とし、改変はもちろん、営利目的での二次利用も許可される最も自由度の高い CC ライセンス。

・表示ー継承

原作者のクレジット（氏名、作品タイトルなど）を表示し、改変した場合には元の作品と同じ CC ライセンス（このライセンス）で公開することを主な条件に、営利目的での二

次利用も許可される CC ライセンス。

・表示ー改変禁止

原作者のクレジット（氏名、作品タイトルなど）を表示し、かつ元の作品を改変しないことを主な条件に、営利目的での利用（転載、コピー、共有）が行える CC ライセンス。

・表示ー非営利

原作者のクレジット（氏名、作品タイトルなど）を表示し、かつ非営利目的であることを主な条件に、改変したり再配布したりすることができる CC ライセンス。

・表示ー非営利ー継承

原作者のクレジット（氏名、作品タイトルなど）を表示し、かつ非営利目的に限り、また改変を行った際には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を主な条件に、改変したり再配布したりすることができる CC ライセンス。

・表示ー非営利ー改変禁止

原作者のクレジット（氏名、作品タイトルなど）を表示し、かつ非営利目的であり、そして元の作品を改変しないことを主な条件に、作品を自由に再配布できる CC ライセンス。

3-2-3. 日本政府におけるオープンデータ利用ルール

2014 年 6 月 19 日に各府省情報化統括責任者（CIO）連絡会議において、政府標準利用規約（第 1.0 版）として決定された。政府標準利用規約（第 1.0 版）は、各国で採用されている利用ルールと異なる利用ルールを日本政府が採用すると、海外からの情報利用者にとって利用ルール間の相違点の把握が必要なことや、国内のデータと海外のデータをマッシュアップする際の利用条件が複雑になる等、情報利用者にとって不便を強いること等から、基本的な利用条件は CC-BY と同様に出典の記載としつつも、各府省から示された意見も踏まえ、国のできるだけ多くのコンテンツに適用できるものとして検討された。その結果、「法令、条例又は公序良俗に反する利用」及び「国家・国民の安全に脅威を与える利用」を禁止する事項が盛り込まれる等、CC-BY とは別の利用ルールとなったものである。

各府省は、ホームページの利用規約として、最低限、政府標準利用規約（第 1.0 版）を採用することが求められており、これとは別の利用ルール（厳しい利用ルール）を一部コンテンツに適用する場合には、その具体的かつ合理的な根拠を説明する責任を負うこととされている。なお、CC-BY や CC0 を採用する場合には、政府標準利用規約（第 1.0 版）よりも制約の緩いルールを採用することになるため、具体的かつ合理的な根拠をホームページ上で明確に説明する必要はない。

例をあげると、「2013 年度版中小企業白書」のライセンスは「CC-BY」となっている。

3-3. 機械判読に適したデータ形式

オープンデータの編集・加工・改変等を行うのはコンピュータである。コンピュータを利

用して与えられたオープンデータを解析するのは、そのデータから新たな知見を得る作業を効率化するためである。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ことが必要となる。機械判読が容易なデータ形式には、いくつかの段階があるが、画像ファイルや PDF 等の形式は、コンピュータプログラムがその中のデータを識別することは困難となり、二次利用をするためには、人手による再入力が必要となる。また、機械判読が困難なデータ形式では、スマートフォンのアプリ等で自動処理することは非常に困難となり、民間による自発的な公共サービスの創造は期待しにくいものとなる。

「機械判読」について、「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）」では以下のように解説されている。

> ～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。～

上記ガイドラインにある「機械判読に適した度合いのいくつかの段階」について詳しく解説する。

3-3-1. 機械判読に適した度合いのいくつかの段階

機械判読可能なデータにもさまざまなレベルがある。オープンデータの評価指標の一つである「5StarOpen Data」 を基に、機械判読のレベルについて述べる。

5StarOpen Data は、Tim Berners-Lee が提唱したオープンデータの評価指標であり以下の 5 段階からなる。

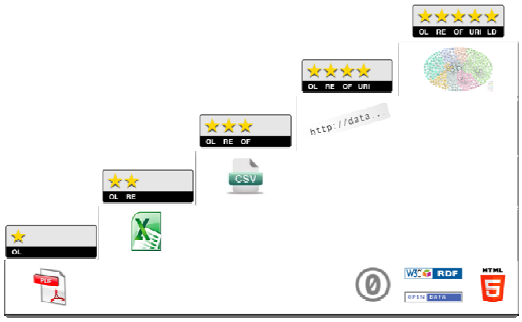


図 3-1

★1：オープンなライセンスで提供されている（データ形式は問わない／画像や PDF 等のデータでも可）

★2：構造化されたデータとして公開されている（Excel や Word 等のデータ）

★3:非独占の(標準化された)形式で公開されている(CSV 等のデータ)

★4:物事の識別に URI を利用している(他のデータから参照できる)

★5:他のデータにリンクしている(Linked Open Data)

各段階に該当する表形式データ、文書データ、地理空間情報のファイル形式の一覧を列挙する。ここでは、代表的な形式のデータを挙げながら、それぞれの段階の特徴を述べる。

★1に該当するファイル形式は、GIF や JPEG 等の画像ファイルや PDF ファイルである。これらのファイルからコンピュータがデータを取り出すためには、画像解析等の技術が必要でありこれは容易ではない。

★2に該当するファイル形式は、Word(.doc)や Excel(.xls)
等である。これらは構造化されているため、対応するソフトウェアを用意すればコンピュータはこの形式のファイルからデータを抽出できる。一般に「機械判読性のあるデータ」とは★2以上のデータをいう。

★3の形式のファイルは、CSV や HTML、OpenDocument(.odt、.ods)、Office Open XML(.docx、.xlsx)等である。これらのファイル形式に基づくデータの解析方法は公開されている。このため、★3 の形式のデータを解析するためのソフトウェアを構築することは★2より容易である。

RDF(Resource Description Framework)に基づいたファイルは★4 以上に該当する。このような形式のデータは相互に接続することができる。このため、これらのデータをコンピュータがマッシュアップすることが容易になる。

3-4. 日本政府のデータカタログサイト

日本政府は、公共データを広く公開することにより、国民生活の向上、企業活動の活性化等を通じ、我が国の社会経済の発展に寄与する観点から、機械判読に適したデータ形式を、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取組を推進している。この取り組みにより公開するデータが増加するにつれて、それらのデータを整理し、検索、一覧する機能に対する要求が高まってくる。このような要求に応えるものが「データカタログサイト」である。

内閣府は、各府省の保有データをオープンデータとして利用できる場をつくり、データの提供側・利用側双方にオープンデータのイメージを分かりやすく示すことを目的に「データカタログサイト(DATA.GO.JP)」を立ち上げている。この「DATA.GO.JP」は、複数の機関が保有・公開するデータの案内や横断的検索の機能を備えており、複数の機関を横断して一元的に、必要なデータを取得することが可能である。

以下に、「DATA.GO.JP」の機能や利用するに際して必要となる用語を解説する。

3-4-1. 主な機能

①省のデータ検索

府省が公開しているデータを横断検索できる機能。検索結果から、そのデータに関する説明（メタデータ）を確認でき、府省の公開URLからダウンロードできる。

②関連の取組のリンク等

利活用の参考にできるよう、政府の方針・決定、公共データ活用事例、府省等のデータベースサイトを紹介。

③利用者からの意見受付コーナー

掲載しているデータのデータ形式、現在掲載されていないデータの掲載等に関する意見を受け付ける。

④更新情報の通知機能

サイトの新着情報やデータの更新について利用者に通知する。

⑤メタデータダウンロード機能

複数のデータセットを選択し、メタデータをダウンロードできる。

図 3-2

3-4-2. 掲載しているデータ

掲載しているデータは、国以外の者が権利を有する部分を除き、基本的に、編集・加工等が可能なデータである。（国が著作権を有するデータは、クリエイティブ・コモンズ・ライセンスの CC-BY によりライセンス）。詳細な利用規約は、データカタログサイトに掲載されている内容を参照のこと。

3-4-3. 用語解説

① データセット

各府省がデータを公表する際に、取りまとめられたデータの単位で、複数のファイル（リソース）から構成されることが多い。たとえば、「○○省平成 25 年度予算の概要」、「○○白書平成 25 年版」などを、1つのデータセットとして登録している。

② リソース

データセットに含まれるファイル単位の実データ。

③ 組織

データセットを作成した府省庁で、2015 年 3 月現在つぎの 21 府省庁が登録されている。なお、括弧のなかの数字はそのグループに登録されているデータセットの件数である。

-内閣官房（63）　-内閣法制局（37）-人事院（122）
-内閣府（812）　-宮内庁（93）　-公正取引委員会（129）
-警察庁（388）　-金融庁（152）　-消費者庁（53）
-総務省（746）　-法務省（531）　-復興庁（11）

-財務省(765)　-文部科学省(1299)　-外務省(120)
-農林水産庁(536)　-経済産業省(1512)　-厚生労働省(1096)
-環境省(1025)　-防衛省(278)　-国土交通省(3202)

④ グループ

ユーザがテーマによってデータセットを閲覧しやすくするためのまとまりで、2015 年 3 月現在つぎの 17 グループが登録されている。なお、括弧のなかの数字はそのグループに登録されているデータセットの件数である。

-国土・気象(1197)　-人口・世帯(240)-労働・賃金(84)
-農林水産業(245)　-住宅・土地・建設(502)　-商業・サービス業(136)
-統計・家計・経済(531)　-行財政(3994)　-鉱工業(362)
-情報通信・科学技術(433)　-運輸・観光(1413)-エネルギー・水(359)
-司法・安全・環境(1508)　-社会保障・衛生(547)-国際(219)
-教育・文化・スポーツ・生活(1191)-その他(9)

⑤ タグ

ユーザがキーワードによってデータセットを閲覧しやすくするためのデータセットに付けられた“荷札”のことである。2015 年 3 月現在多くのタグが登録されておりすべてを記載することはできないが、主要なタグはつぎのとおりである。

・G8の重要データカテゴリ(16 項目)

-統計　-地図　-選挙結果　-予算
-企業　-犯罪と司法　-地球観測　-教育
-健康　-エネルギーと環境　-財政と契約
-国際開発　-科学と研究　-政府の行政責任と民主主義
-社会的流動性と福祉　-交通とインフラ

・電子行政アクションプランにおける業務分類(31 項目)

-国土　-気象　-人口_世帯　-労働
-企業_家計_経済　-観光　-司法　-安全
-環境　-災害　-国際
-農業_林業　-漁業　-鉱業_採石業_砂利採取業
-建設業　-製造業　-電気_ガス_熱供給_水道業
-情報通信業　-運輸業_郵便業　-卸売業_小売業　-金融業_保険業
-不動産業_物品賃貸業　-学術研究_専門_技術-サービス業
-宿泊業_飲食サービス業　-生活関連サービス業_娯楽業
-教育_学習支援業　-医療_福祉　-複合サービス事業
-サービス業_他に分類されないもの　-公務_他に分類されないものを除く
-分類不能の産業

・「電子行政オープンデータ推進のためのロードマップ」における重点分野(7項目)

-白書_年次報告　　-防災_減災関連情報　　-地理空間　　-統計
-人の移動_ツーリズムに関する情報　　-予算_決算_調達関連情報 -コード

⑥ メタデータ

メタデータとは、データについての情報を記述したデータである。 膨大なデータの山の中から目的のデータを探し出す手助けとするために作成される。 個々の情報にメタデータを付けることにより、よりデータの性質を的確に反映した検索が可能となる。

「DATA.GO.JP」では「データセット」「リソース」に以下のメタデータが付与されている。

・「データセット」

-タイトル

″データセットのタイトルを設定する。
○○白書_××年度、◆◆調査_月報_△△月といったリソースの集合体を指すタイトル名を設定する。複数の単語を組み合わせてタイトルを設定する場合は、半角アンダースコア(_)で連結することとする。
設定例:情報通信白書_平成 25 年版″

-公表組織名

″データセットの説明を 100 文字以内の文章で設定する。
説明に含まれる英数字は半角で設定する。
設定例:総務省が発行している情報通信白書の平成 25 年度版

-公表組織名

″データセットを公表している組織名を設定する。
設定する組織名は、以下から選択する。
・内閣官房 ・人事院 ・内閣法制局 ・内閣府 ・宮内庁
・公正取引委員会 ・警察庁 ・金融庁 ・消費者庁 ・復興庁
・総務省 ・法務省 ・外務省 ・財務省 ・文部科学省
・厚生労働省 ・農林水産省 ・経済産業省 ・国土交通省 ・環境省
・防衛省
設定例:総務省″

-連絡先

″データの誤り等を連絡する先を設定する。
基本は課室名で記入することとし、部局名も可とする。
設定例:情報通信国際戦略局情報通信政策課情報通信経済室″

-作成者

″データセットの作成者を設定する。
基本は課室名で記入することとし、部局名も可とする。
複数設定する場合は、半角カンマ(,)でつなげて列挙する。
設定例 1:情報通信国際戦略局情報通信政策課情報通信経済室″

-タグ

″データセットの検索のための検索キーワードを設定する。
以下の(1)～(4)のそれぞれについて、該当するものがあれば設定する。それら以外に検索キーワードを設定したい場合、キーワードとなる単語を設定する。複数キーワードを設定する場合は、半角カンマ(,)でつなげて列挙する。
(1)G8の重要データカテゴリ(16 項目)
統計、地図、選挙結果、予算・決算、企業、犯罪と司法、地球観測　等
(2)「行政情報の電子的提供に関する基本的考え方(指針)」における「共通のカテゴリー」(16 項目)
組織・制度の概要、所管の法人、所管の法令、告示・通達等、国会提出法案、審議会、研究会等
(3)電子行政アクションプランにおける業務分類(31 項目)
日本標準産業分類の大項目(20 項目)と総務省統計局の統計分野から産業の観点で不足する 11 個の分類。
農業, 林業、漁業、鉱業, 採石業, 砂利採取業、建設業　等
(4)重点分野(試行版における重点 5 分野に統計とコードを加えた 7 項目)
白書(年次報告)、防災・減災関連情報、地理空間情報、人の移動・ツーリズムに関する情報、予算、決算、調達関連情報、統計情報、コード情報
設定例:情報通信業,白書(年次報告)″

-リリース日

″データセットをウェブページに公開した日付又は公開日としてウェブページ等に記載の日付を設定する。
入力形式は半角数字及び半角ハイフン(-)で、YYYY-MM-DD 形式とする 。
月や日が不明な場合は、YYYY 形式または YYYY-MM 形式で設定する。
設定例:2014″

-作成頻度

″データセットの作成頻度(作成間隔)を設定する。
以下からいずれかの項目を選択し、X には半角数字を設定する。
記入要領による記述が難しい場合(月 2 回等)は自由記述で入力する。
・X 年　・X 月　・X 週　・X 日　・X 時間　・X 分
・リアルタイム　・不定期　・更新しない
設定例:1 年″

-公開ウェブページ

データセットを公開しているウェブページの URL を設定する。

設定例:

http://www.soumu.go.jp/johotsusintokei/whitepaper/h25.html″

−対象地域

地図や対象地域が明確な調査などについて、可能な範囲で、データセットが対象としている都道府県名を設定する。

都道府県名を略さず入力し、対象となる都道府県が複数となる場合は、半角カンマ(,)でつなげて列挙する。

設定例:東京都,千葉県″

・「リソース」

−タイトル

リソースのタイトルを設定する。

データセット内に含まれるファイルのタイトルを次のような基準で設定する。

①ウェブページに表示されているハイパーリンクの表記

②ハイパーリンクがなく、ダウンロードアイコンがある場合等は、ダウンロードアイコンを表記するファイル名等。

表記やファイル名等だけでタイトルとして分かりにくい場合等、複数の単語を組み合わせて title を設定する場合は、半角アンダースコア(_)でつなげることとする。

設定例:第 2 章:ICT の活用による社会的課題の解決″

−URL

リソースのリンク先の URL を設定する。

設定例:

http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h25/html/″

−説明

リソースの説明を 100 文字以内の文章で設定する。

説明に含まれる英数字は、半角で記載する。

設定例:総務省が発行している情報通信白書の平成 25 年度版の「第 2 章:ICT の活用による社会的課題の解決」の HTML ファイルです。″

−データ形式

リソースに設定するリンク先のファイルの拡張子を設定する。

以下の拡張子から選択する。

asx、avi、csv、doc、docx、epub、exe、gif、gml、html、jpeg、jsp、jtd、kml、kmz、lzh、mp3、odp、ods、odt、pdf、php、png、ppt、pptx、rdf、sgml、shp、tiff、tsv、txt、wmv、xls、xlsx、xml、zip

設定例:html

・選択肢のない拡張子の場合は設定しない。(空欄を設定)″

-ファイルサイズ

リソースに設定するリンク先のファイルの容量を設定する。

容量はバイト単位とし、半角数字で設定する。

ウェブページのファイルサイズ等の記載を利用する場合は、キロバイトは 1000 バイト、メガバイトは 100 万バイトとしてバイト単位に換算して設定する。

設定例:5000(ウェブページで 5KB と記載されている場合)″

-最終更新日

リソースを各府省のウェブページに公開した日付、または公開日としてウェブページに記載の告知した日付を設定する。

入力形式は半角数字及び半角ハイフン(-)で、YYYY-MM-DD 形式とする 。

月や日が不明な場合は、YYYY 形式または YYYY-MM 形式で設定する。

設定例:2013″

-ライセンス

リソースのライセンスを設定する。

ライセンスは、以下のいずれかを選択する。

なお、国以外の者がリソースの著作権を有する場合には、ライセンス ID は空欄とする。

・CC-BY

・政府標準利用規約(第 1.1 版)

設定例:CC-BY

・API を利用して、rdf 形式や JSON 形式のメタデータをダウンロードした場合においては、そのライセンス条項が記載された URL が返される。″

-コピーライト

国以外の者がリソースの著作権を有する場合に以下を設定する。

・国以外の者が著作権を有する

設定例:国以外の者が著作権を有する″

-言語

リソースの主たる中身を表している言語を設定する。

ISO639-2/B に従い、以下の中から選択する。

・ara(アラビア語)・chi(中国語)・eng(英語)

・fre(フランス語) ・ger(ドイツ語)・jpn(日本語)

・kor(朝鮮語)・rus(ロシア語)・spa(スペイン語)

・zxx(地図等、言語的要素が無い場合)

設定例:jpn″

⑦ API

「DATA.GO.JP」では、組織、グループ、データセット、リソース、タグの各メタデータを取得するための API（API:Application Programming Interface）が提供されている。

3-5. 地方自治体のデータカタログサイト

地方自治体が公開するオープンデータについては、それぞれの自治体が独自にデータカタログサイトを立ち上げており、各自治体のデータを横串で検索・閲覧できる「地方自治体のデータカタログサイト」のようなサイトは今のところ存在しない。それゆえ、ここでは、静岡県が立ち上げている「ふじのくにデータカタログサイト」を例に、地方自治体のデータカタログサイトについて述べる。

3-5-1. ふじのくにデータカタログサイト

静岡県では、国の「電子行政オープンデータ戦略」を踏まえ、二次利用可能なデータの公開環境を整備し、実際に公開を進めることで、行政の透明性の向上を図るとともに、公開データを利活用したビジネスが展開される環境を整え、オープンデータによる経済活性化の促進を図っており、オープンデータを実践するサイト「ふじのくにオープンデータカタログ」（http://open-data.pref.shizuoka.jp/）を設置している。（図 3-3）

図 3-3

このサイトでは、静岡県、静岡県内の各市町、および民間や学校がもつ公共データが公開されており、メタデータとして「データ名称」「カテゴリー」「データ概要」「ライセンス」「データ提供部署」「関連ホームページ」「データ形式」が付与されている。

「カテゴリー」には、「公共施設」「観光・見どころ情報」「防災」「統計」「都市計画」「道路関連」「その他」の７つのキーワードがあり、公開されるデータはこれらのいずれかに分類されている。この中で「公共施設」が件数としては一番多く公開されている。自治体としては最も公開しやすく、また最も利用されやすいデータということであろう。

地方自治体のオープンデータ活用事例として取り上げられるのが、地方自治体が管理する施設を地図上で表示させるサービスがあるが、このようなサービスを可能にするために、地方自治体が公開する公共施設データにはその場所を示す「緯度」と「経度」が含まれている。これにより、施設の場所を地図上に指し示すことが容易にできる。図 3-5 は、「ふじのくにオープンデータカタログ」から参照した、裾野市が公開している防火水槽設置場所の一覧に記載されている緯度と経度のデータをもとに、地図上に防火水槽の位置をマッピングさせた例である。

図 3-4

図 3-5

4. ITCとしてのオープンデータの活用

4-1. 取り組みの基本的考え方

前章ではオープンデータを利用する上で必要な権利や所在（カタログサイト）、データを入手する上で一番重要なデータ形式等について論じた。これらの情報により利用する手がかりを理解していただけたと思う。

次に、それらを利用するにあたってＩＴＣとしてオープンデータなるものが、利用上どのようなインパクトがあり、その結果どうなるか等、直接関連することについて議論を広げてみたい。

まず、目に見える形になるものは何か。その一つの回答として、地域の課題解決が考えられる。つまり、自分の住んでいる自治体の持つ様々な情報を活用し、組み合わせ、その住民の持つ情報機器（例えばスマートフォンの様な携帯モバイル）と組み合わせることで、日常生活に密着した便利で手軽なアプリケーションを創造し、広く活用してもらう事が挙げられる。

積極的で効率的な運用が求められている地方財政や、進みつつある高齢化社会がある現実をどのように変えていくべきか。その一つの可能性として、ＩＴＣは行政コストの低減や安全・安心な社会の構築の手助け、住民サービスのレベル向上などの公共データの活用、オープンデータを住民生活に活用していくためのＩＣＴ（総合的技術）による組み合わせ提案などで、地域の課題解決を図り貢献していく事を積極的に推し進め行く推進役が期待されている。

またこれとは別に、民間企業に向けては企業自らが有する情報とオープンデータを融合させ、活用していく事で、新しいビジネスを創出するといったことを推し進めていくチャンスが訪れているので活用していきたい。

ＩＴＣとしては、自己の持つテーマに基づき、関連するオープンデータの動向をフォローし、それを利用することでどの様な事が出来るかにつて、常日頃から意識していく事が重要である。

オープンデータへのアプローチとしては、大きく２種類考えられる。ＩＴＣが持つ顧客の立場からの活用と、それに束縛されないオープンデータ側からのビジネス展開の立場からであるが、次の節以降でそれらを検討してみたい。

4-2. オープンデータを利用する上での特徴とＩＴＣが関与する意味

前章でみたとおり、オープンデータは広範囲、多項目にわたる。カタログサイトは言うなれば、官公庁が保有しているデータの巨大な電子ファイリングシステムとみなせない事もない。利用上の特徴も従来にはなかったもので、多岐にわたっている。主な特徴を列挙してみる。

・法的に自由で公開性を持つ。
・オープンライセンスで誰でもいつでも自由に使える。
・データセットを自分で自由に選択できる。
・データへの到達性、アクセス性、選択性、非拘束性を持つ。
・自己が保有しているデータや、他のデータと組み合わせて自由に加工できる。
・再利用、再配布に関しても自由にできる。
・データ提供者から自由である。
・原則としてコスト的に自由である。
・データの信頼性は公共機関の提供なので比較的高い・・・等々。

それ故に、データを自由に組み合わせて加工し、個人や民間企業で利用したり、加工データを２次的に再配布して、更に他の個人や企業で利用してもらうことが可能となる。それ故に、ビジネスチャンスとなり得るのである。

これら広範囲なデータを一つの観点から一元的に参照していく方法は存在しない。また、多様なデータに対して一つの視点のみから捉える事には意味もないと考えられる。業種別に考えてみても、企業にとって必要とされるデータは異なっており、テーマ別に見たとしても必要なデータはバラバラに存在する。個別にテーマを設定し、それに基づいてデータを参照し、利用方法を考えていく必要がある。データ検索には思考の軸が必要であるということである。ＩＴＣがデータ活用に関与する意味がここにある。利用者が求めているニーズとデータの持つ可能性をマッチングさせ、最適化を図り提案していくのである。それに、データ活用はわざわざ言うまでもなくＩＴ技術を利用して展開するものであり、得意な分野の一つとして展開していく事に意味がある。

現在設定されているデータセットからデータ検索する対象の例を挙げてみよう。耐久消費財の販売を考えているのであれば、市町村別、場合によっては地区別の耐久消費財の普及率を調べる。或いは、市町村別の給与支給額と消費財購入額を調べ、購買力と消費動向を探る。防犯装置の販売を考えているのであれば、市町村別の犯罪発生件数を調べる。介護サービスの展開を考えているのであれば、市町村の６０歳以上の人口当たりの介護・福祉施設の設置数を広範囲に調べる、等である。

4-3. ＩＴＣから見たオープンデータの可能性と顧客企業の価値向上の視点

オープンデータは新しいビジネスを創出したり、既存のビジネスに付加価値をつけたり、広範囲に活用できることが期待される。データの検索も、結合・加工も、再配信も自由である。可能性に制限は無い。データを見てからアプローチする方法もあるが、ＩＴＣの立場として、ここではまず顧客企業を通して社会に提供していくアプローチについて考察してみたい。

オープンデータを活用するには、まずその前にそれらを適用しようとする顧客企業の状況を確認しなければならない。つまり「顧客の経営課題を解決し、価値を向上する」という視点が必要である。わざわざ申し上げるまでもなく、まずデータありきではなく、問題解決と価値向上の視点から考えてみる必要がある。

経営分析を行い、問題点・課題を明確にする。市場環境の変化の問題なのか、ライバルとの力関係の問題なのか、顧客ニーズの変化の問題なのか、新技術・新製品の問題なのか、等々その原因を明らかにし共通認識とする。

次に、やりたいことに対して不足しているものは何かを探る必要がある。企業環境によって検討しなければならない項目は別々に存在するが、本報告書は情報を活用することを主眼に論じているので、論点が拡散しないように情報やデータの収集と活用に絞って論じてみる。

ここまでが情報を収集し、活用していくための準備作業である。ここから先の手順は、ＩＴＣであるならば十分理解していることであるが、念のため以下に手順の例を示す。

① シナリオに沿ったデータ収集とスクリーニング。
② 戦略目的に沿って、情報を整理・カテゴリー化。
③ 新たな問題点（阻害要因）の発見。
④ 問題点の解決策の策定。
⑤ 戦略情報マップの策定。
⑥ 実行のためのシナリオ策定。
⑦ 投資効果の算定。
⑧ ビジネスプロセスの更新提案。
⑨ 提案・実行。

4-4. 必要データの種類と所在確認

次の段階に進んだとして、ＩＴＣとしてどこから手を付ければよいか。まず、以前に類似案件があるのであれば、以前の調査手順に従って情報を検索していく。ゴールを設定し、そこから逆算するように必要な情報を過不足なくリストアップする。そして、顧客企業が保有している対象となる業務分野の現在の情報・データの内容を確認しておく必要がある。特に重点的にチェックが必要なのはデータのフォーマット（保有形式）や区切り（メッシュ）である。これは、後でオープンデータを調査する際に、社内データとマッチングするための重要な項目である。不足しているデータもしくは新たに必要とするデータは何かを明らかにし、社内・社外を問わず調査対象を順次拡張していく事が望ましい。

顧客企業が問題解決するために必要とする情報の例を、製造業を例にとり以下に列挙してみる。

① 社内技術情報（探索研究技術、開発研究技術、生産技術など）。
② 社内顧客情報（基幹システム情報、顧客情報、アンケート情報など）。
③ 外部技術情報（文献情報、特許情報、専門誌情報など）。
④ 外部顧客情報（市場情報、帝国データバンク、Ｗｅｂ、オープンデータなど）。

①や②の様に既に社内に蓄積している情報に加え、④のＷｅｂ上データ・オープンデータを重点的に調査し、目的に沿って活用していく事になる。これは、今までも知りたかったが、公開されていないために利用できなかった情報を利用することである。柔軟な視点が求められる。

総務省が公開している中央官庁が所有するオープンデータのカタログサイト、及びそれの参照方法は前章で述べた。統一された形式でデータが提供されることは、利用者側にとっては大変便利である。形式が分かっていれば、任意の項目を自由にダウンロードして自己の事業推進に活用できるからである。

ここで一つ問題が出てくる。それはデータ群が巨大な電子ファイリングシステムの様になっているため、必要なデータが格納されているデータセットがどこにあり、どうやってそれを入手することができるかが、直ぐには分からない事である。いずれは、それをガイドするソリューションやサービスが登場してくると思われるが、現状では必ずしも容易な作業ではない。ＩＴＣは自分が得意としている分野のデータがどこに格納されているかを知っていて自由に取り出せるだけで、データ利用のアドバイザーとしてアドバンテージを持つことができる。

「何故か？」。データ検索の難しさは、実はそのデータ構成にある。それはデータのメッシュが収集されたデータセットごとに異なっている為である。

① 時間軸に関する区切り。調査期間（Period）、間隔（Interval）など。
② 場所・空間に関する区切り（Where）。県単位、市区町村単位、地区単位など。
③ データの細かさの区切り（Mesh）。

必要なデータセットに辿りつけたとして、これらの項目が自社で利用する目的に合わないと全く使えないことに気づくだろう。例えば、時間軸に関していうなら、間隔が１年ごと・２年ごと・５年ごとの様に、調査する機関（官庁）や目的によって異なる。範囲も５年間さかのぼったり、１０年間だったり、２０年間だったりする。場所・空間も、県単位か、市町村単位か、地区単位かで異なる。データの細かさも生データ（１次データ）か、ある単位での集約データ（２次データ）かで異なる。これらを使い分けるのが難しいのである。つまり、顧客の持つ社内データとオープンデータのデータレベルがマッチングできないと、利用者から見て当然意味のないデータになってしまう。オープンデータ同士のマッチングも同様である。素人がいきなり見て検索・使いこなしができるようなものではなく、データ知識に対する高い専門性が必要となる。これがアドバイザーのビジネスとして成立する可能性を持っている理由である。

まだ検索したことがない人は、データカタログから興味のあるデータセットを検索してみることから始めてみるとよい。

データセットのメタデータに関しては前章で詳述したが、以下に総務省で公開されている主要なオープンデータの項目例を挙げる。

・人口・世帯に関する統計
・住宅・土地に関する統計
・家計に関する統計
・物価に関する統計
・労働に関する統計
・文化・科学技術に関する統計
・企業活動・経済に関する統計
・経済構造・波及効果分析や各種経済統計の基準値となる統計
・経済・金融に関する統計（IMF 公表基準掲示板）
・地域に関する総合統計

ちなみに、中央官庁以外でも、多くの自治体が個別に所有データを公開している。福井県や静岡県など先進的にデータ公開している例である。これらの多くのデータは利用の汎用性を考えてテキスト形式やＣＳＶ形式で提供されている。個々の自治体の個別に対応となっている。ゆくゆくは総務省が管理しているデータカタログサイトに準拠したように進んでいく可能性はあるが、現状では全体として統一されたフォームがあるわけでは無いので、利用者が必要に応じて個別に調査しなければならない。

また、自治体のデータ以外でも、東京大学社会科学研究所、国立社会保障・人口問題研究所、統計数理研究所などは、独自にデータの所有と提供を行っている。帝国データバンクの企業データなどは、個別に企業を評価したり、企業同士の比較を行うために以前から活用されているし、ＩＭＳなどの売り上げに関する情報なども同様である。これらのデータの多くは有料である場合が多い。

ＩＴＣは事前に行った顧客の業務改善の目的に沿い、必要に応じてこれらの情報源を組み合わせながら、データ収集を行っていく。

4-5. 情報検索とデータ解析・整理

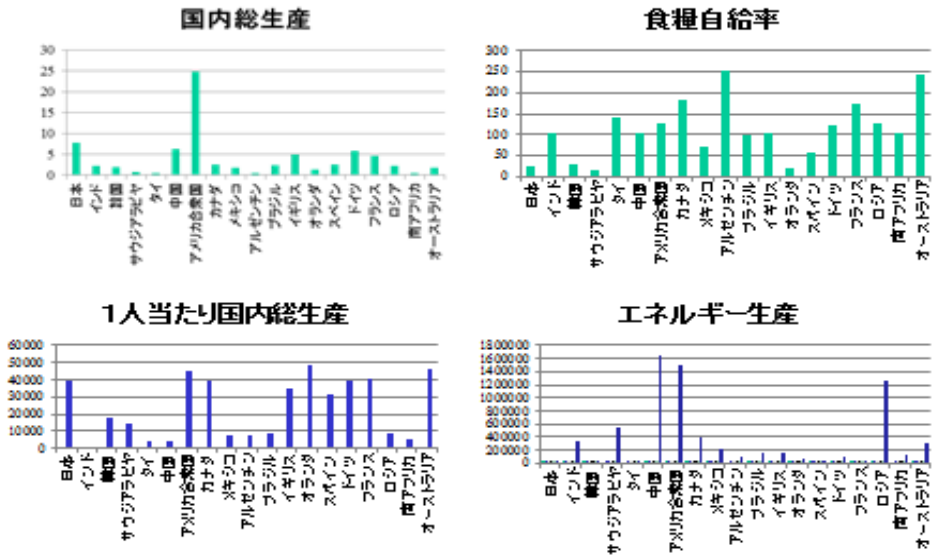
ターゲットになるデータを入手したら、データの妥当性をチェックしてみる。目的に沿っているか、データに欠損値は無いかなどをチェックし、簡単なグラフなどに描いてみて、鳥瞰的にチェックしてみる。データが大量の場合は、その一部をサンプリングし、同様にチェックを行う。

その上で、目的に沿って、それらを利用できる形に加工しなければならない。

以下にデータ加工の方法に関する例を挙げる。

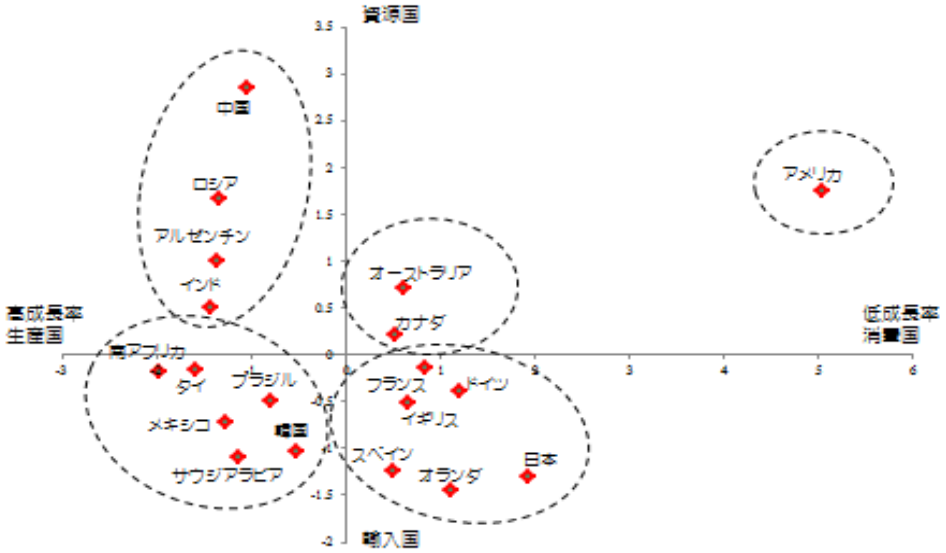
① 定量一次加工データ分析：集めたデータ、すなわち市場の売上トレンド、コスト構造、消費者嗜好などを数量的に表現できるデータを平均値などにまとめ、分かり易くするためにグラフに図示する。
総務省が公開しているオープンデータ（海外データ）のグラフ化した例を以下に示す。

データのグラフ　（図1）



② 定量二次加工データ分析：一次加工のデータを、更に仮説に基づいて細かく分析する。多面的に細かく分析することで、因果関係やトレンドを浮き彫りにし、より本質に近づけるようにする。
総務省が公開しているオープンデータ（海外データ）を元に解析した例を以下に示す。

主成分分析　　(図2)



データはこのようにマッピングした２次加工データとして表現すると、その意味が大変分かり易くなる場合がある。

③ 定性的データ分析：数量的に表現できないが、言葉で表現できるもの、たとえば好みやフィーリングや印象などを、分類手法を使ってまとめる。スコア付与、数量化、クラスター化など疑似的ながら定量的に表すための手法の適用も、必要に応じて行っていく。

これら一連の作業を行って情報を整理し、その意味を考察する。データを整理し解析する方法として、以下の代表的な方法があるので試してみてほしい。

① データ間の相対的位置関係を知る（主成分分析、因子分析など）。

② データ群を分類する（クラスター分析、多次元尺度構成法など）。

③ データ間の相関関係を知る(相関分析、バスケット分析、共分散構造解析など)。

④ データ間の因果関係を知る（重回帰分析など）。

手法としていろいろあるが、経験の無い手法を無理に当てはめるのではなく、知っている方法から順次拡大していくのが望ましい。解析結果を解釈するには、自由な発想と現実に根差した考察が必要になる。身近に統計解析に詳しく経験豊富な相談相手がいない場合には、類似案件の解析例や結果に対する考察が記載された論文を参考にするとよい。

統計解析を行うということは、その裏に隠れているコンテキスト（傾向や本質）を読み取るということが目的になる。間違っても、無理やり手法を当てはめて結果を導き出すことの無いようにしたい。すぐに気が付くことであるが、常識から外れた結論が導き出されることはまずあり得ない。

4-6. 顧客企業での提案と実施

次に、前節までの作業で策定した計画案を顧客企業に提案し実施を行う。ポイントとなるのは既に述べたように「顧客の経営課題を解決し、価値を向上する」という観点で考察されている点である。別の表現をするならば「このデータは貴社のこういう問題の解決に役立つと考えられます。活用してみませんか。」ということで、

① 改善策により何が良くなるのか。

② それによるコスト効果は何か。

③ それによる企業価値効果は何か。

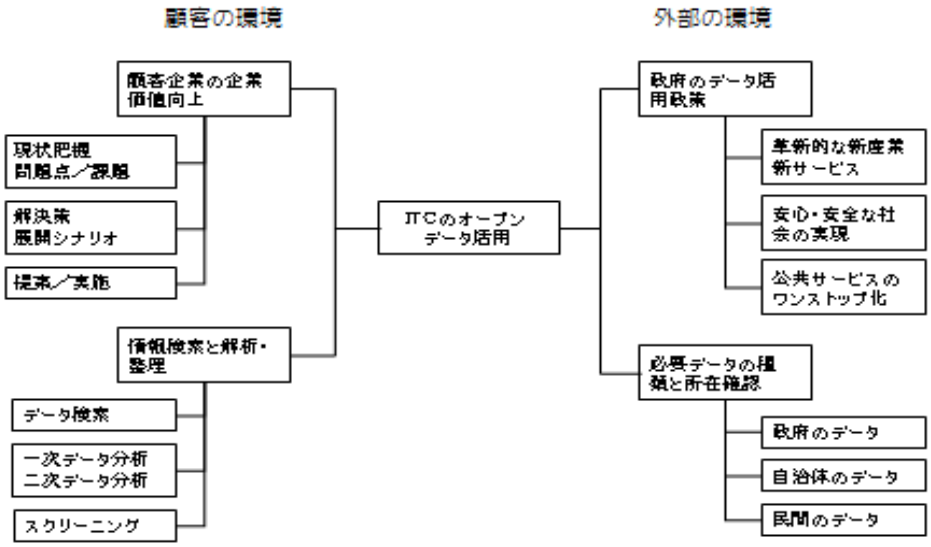
が明確に示され、納得してもらったうえで改善策を実施していくことである。

ＩＴが導入された初期の時期に、中小企業においては「ＩＴを導入したが返って効率が悪くなった」というような現象が見られた。これはなぜか。原因分析が行われた結果から分かっていることは、情報のシームレスな連携が行われないために、システム化された部分の２重入力や、システム化されていない部分の手作業によるデータ加工が増えたためである。経営者の中にはこの時のことを覚えていて、説明に対して拒否反応を示すような場合も有り得る。このようなことが発生しない様、事前に運用に沿った検証が行われ、説明資料が適切に整備される必要がある。

今までに存在しない様な新規事業・新規製品の場合、当然リスクは高くなる。マーケティングの４Ｐ（製品、価格、販路、プロモーション）に基づいた市場のサーベイは、当然必要になる。まず最初はコントロールが容易な小規模トライアルを行い、順次規模を拡大していく事が望まれる。

以下にＩＴＣがオープンデータを活用していく視点を図３に示しておく。

ＩＴＣがオープンデータを活用する視点　　（図3）



4-7. データ解析からのアプローチ

この章では、ＩＴＣとしてのミッションからずれないようにするために、主に顧客企業を通じてデータ活用するアプローチについて手順を追って論じてきた。類似案件を抱える顧客企業が多い場合には、同じ様なデータの解析で良い結果をもたらすことが想定される。しかし顧客中心のアプローチの欠点は、顧客企業の活動以外の視点が無視されてしまうことである。

もともとデータにはそのような制限は無い。従って、データを見て新しい発見をし、それを元に新しい事業を生み出しても何ら問題は無い。既存データとの融合でもっと良い結果が生み出される可能性もある。むしろ、既存の企業の活動枠にとらわれないためには、その様な自由な観点でデータを解析するスタンスが重要になってくるかもしれない。

この章の最初でもふれたが、自治体が主体となり、ＩＴベンダーが協力しながら展開しているオープンデータの活用例は、公共施設の所在を地図上にマッピングしたり、施設の利用方法に関するものが多い。公共サービスのワンストップ化を目指しているのである。防災施設や災害時の避難経路などがその代表となる。データから利用を考えている例である。

論じていることをイメージしやすくするために図１と図２を見ていただきたい。図１では個々の国の特性がバラバラで、直感でそこから一つのイメージを導き出すのは難しい。しかし、図２で示した様に、適切な統計処理を適用した場合、同じデータから特定の似通った性質を持つ特定のグループを抽出し、それに意味づけをし

て傾向を把握し、そこから価値を見出す可能性はある。潜在的な可能性を明確にできる側面で、楽しみでもある。

その様な場合でも、テーマや視点は必要である。テーマや視点を持つことは、データ解析結果の拡散を防ぎ、それを見る人達の意識を一つの方向に集中させる働きを持つ。そのことを考慮しながら、データをいろいろ分析してみると、それを応用する可能性が広がってくる。また、それが期待されていることであると考えられる。

いろいろな手段を駆使してデータ解析からアプローチすることと、その活用にチャレンジしてみていただけたらと思う。

５．おわりに

「オープンデータは活用できなければ意味がない…」、このスローガンで本年度は「先ずはどのような活用の形があるか…」までの研究を行った。

その限りにおいて、一通りの手順も織り込んだ利用形が提示出来たと判断している。

如何でしょうか？

研究会内部で意見を交わし合い、ITコーディネータとしての見識を持った成果物になっていると自負しているのであるが…。

オープンデータは大変多くの応用の切り口があり、活用次第でいろいろなアウトプットに変化し得ることが見て取れる。

そこで来年度は、更に一歩進んだ活用の形にまで踏み込んだ研究、即ち「ITコーディネータにとってのオープンデータビジネスの可能性（応用編）」を最終成果物に想定した研究を実施致したいと考えている。

以上