

第一に、データを長期的に追跡する際に、同じ項目のデータの変化を追跡しやすくなる。第二に、データを横断的に統合・比較する際に、どの項目のデータが同じなのかを判定しやすくなる。こうした利点があるにもかかわらずデータにIDが付与されている例は意外なほど少ない。これまで筆者がデータ生産者と議論してきた経験から言えば、その理由はデータ生産者自身がその重要性を理解していないからである。例えば「IDの代わりに名称を入れているから十分」という反論をよく聞く。しかし実際にデータを扱ってみると、名称はよく変わるし、表記ゆれもよく生じるため、名称でIDの代わりを果たすことは意外に難しいのである。10年ほど前の「消えた年金問題」でも、IDがない状態で名寄せを行うことの困難さが大きく報道されたが、可変の名称でなく不変のIDをきちんと使えばこうした難問もシンプルに解決できる。さらにIDは機械可読性という面でも完璧である。つまりIDを使える可能性がある場所ではすべて使うようにし、IDはデータの必須項目と考えてほしいというのが筆者からの要望である。

最近はやや様々な分野でIDが整備されてきたことも追い風である。個人にはマイナンバー、会社には法人番号など、全国的に使えるIDが普及を始めている。こうしたIDがあればそれをデータに入れる、もしなければ自治体内で定めてそれを長期的に継承していくことが、データ生産者としての責務として重要である。ただしこれは言うほど簡単ではない。例えば名称を変更した場合、変更前と変更後に同じIDを使い続けてよいのかという問題は、何を同じものとみなすのかという哲学的な問いを含み、その解は自治体内でポリシーを決め判断を重ねて見出すしかない。しかし「不変かつ唯一のID」という条件を満たすためには、同一IDの使いまわしだけは絶対に避けなければならない。そしてIDが増えても問題ないよう、ID体系には十分な余裕を持たせることが望ましい。

#### 4. データ流通者・データ消費者との繋がり

こうしてよいデータ生産物を生み出し、それをデータ卸売業に流せば、データ生産者としての役割は一通り終わる。十分に強力なデータ流通者が存在すれば、あとはデータ消費者にデータを届けてくれると期待して、また次の生産に戻る。これがデータ生産者としてのシンプルな生き方であろう。

しかし、世の中そう単純にはいかないのもまた現実である。「自分でも売ってみたい」という欲望が出たり、「もっとアピールできる取組をしてほしい」という指令が届いたりすることは多い。そうすると、自前のデータ流通プラットフォームの運営を開始し、データを生産者から消費者に直接届ける産地直送モデルや、生産者が付加価値をつけて消費者に届ける六次産業モデル<sup>[1]</sup>などに挑んでみたいという誘惑も生まれてくる。しかしそうすると、生産者としてのセンスに加え、卸売や小売のような流通者のセンスを兼ね備えたスキルが必要となり、オープンデータのハードルは自然に高くなっていく。そのあげくコストも上がって、活動が拡大するにつれて苦労ばかりが増える事態にもなりかねない。

これはある程度は自らが蒔いた種である。自らが生産したデータを他者にすぐ出荷して、生産以外には手をつけたくないという低コスト戦略もあったのに、自らそうでない道を選んだからである。しかし現在のオープンデータの潮流では、生産者が流通者を兼ねるモデルが盛んに取り上げられているため、そうしたモデルに影響されてしまうのは仕方がない面もある。

この意味では、報告書でも引用されている「5つ星オープンデータ<sup>[2]</sup>」の考え方の是非は慎重に検討する必要がある。「5つ星オープンデータ」は流通者にとって都合のよい図式を表したものであるが、この全部を生産者の責務であると捉えてしまうと、強大な流通者と消費者が一体となって生産者に際限なき要求をかける図式に陥る危険性さえある。そうならないためには、自分たちは生産者としてどのような品質を目指す

のかという認識をしっかりと持ち、それに沿った戦略を立案することが課題となる。

そして、筆者が最も好ましいと考えているのは、自治体自身が流通者と消費者にもなるという戦略である。例えばある部署が生産したデータを別の部署が活用するならば、それは自治体内部に生産者と消費者がいることを意味する。そして生産者と消費者を繋ぐ部署があれば、そこは流通者となる。このように自治体内に生産者、流通者、消費者の小さなネットワークを構築し、部署をまたいでデータを活用することこそ、自治体のオープンデータの出発点なのではないだろうか。

IT業界ではドッグフーディングという言葉がよく使われる。その意味は、自分が作ったドッグフードを自分の犬に食べさせるように、自分が使うシステムを自分で構築すること、言い換えれば自分自身が生産者であり消費者である状態を作り出すことが、品質向上に大きな役割を果たすという考え方である。同様に考えれば、自治体が生産したデータを自治体内部で使うことで、データの品質向上が自治体業務の品質向上に繋がる構造を作り上げることをまずは目指したい。そして、自治体内部という概念を少しずつ広げて、都道府県、日本、世界を相手にしていけば、オープンデータ活用の視野はより広がっていく。つまり、自分たちの業務を向上させるためにデータを活用すること、そしてその成果を外部にも「おすそ分け」していくことこそが、オープンデータ活用の本筋だと筆者は考える。

現在のように、生産者は生産する一方で消費者からフィードバックがないという分離された状況では、生産者の意気も上がらないし、その結果として品質も上がらない。こうした疎外された環境から顔の見える環境に移行するには、やはり身近に消費者を見出して「顔の見える関係」を築くことが重要である。そしてその一つの方法が、自治体内でのデータ生産・流通・消費ネットワークの構築である。そうした成功例は私の見る限りでは報告書に見当たらなかった

が、もしかしたらすでに実践している自治体があるかもしれない。これこそが、データに基づく行政そのものだからである。

#### 5. おわりに

本稿ではオープンデータにかかわる生産者・流通者・消費者という概念を提示し、それらが果たすべき役割を中心に述べた。自治体に期待されている第一の役割はデータ生産者であり、その範囲内においてプライドを持って品質が高いデータを継続的かつ安定的に生み出す仕組み作りをまずは目指していただきたい。しかし、ひたすら生産を続けるだけでは、流通者や消費者のために一方的に働かされているだけという意識を持つ心配もある。それを防ぐためにも、身近に流通者と消費者を見出し、顔の見える関係からデータ生産の重要性を体感していくことが望ましい。その一つの方法は、データに基づく行政を実現するためのインフラを構築し、部署を越えたデータ利活用を実現し、その成果を外部に「おすそ分け」していくという方法である。

長期的にはこの姿が理想だとしても、すぐに実現するには難しい面があるかもしれない。しかし、オープンデータが自治体の重要な役割であることは今後も変わらない。人々からの要請に対応するには、低コストかつ持続可能な戦略をデータ生産者として立案することが重要であることを、最後に改めて強調しておく。

そして、本稿が報告書を理解する上での一助となり、自治体におけるオープンデータ化の取組がさらに推進することを期待したい。

[1] 農林漁業者（一次産業）が、農産物などの生産物の元々持っている価値をさらに高め、食品加工（二次産業）、流通・販売（三次産業）にも取り組む試みを指す。本稿ではこの考え方を援用し、データ生産者が加工や流通などに取り組むことでデータの利用を増やす試みを「六次産業モデル」と呼ぶ。

[2] オープンデータのデータ形式を5つの星に段階（ランク）分けしたもの。機械等で二次利用しやすいデータ形式であるほど星の数が多くなる。ワールドワイドウェブの始祖であるティム・バーナーズ＝リーが提唱した。