

平成28年度 調査研究報告書の解説

「基礎自治体によるオープンデータ化と利活用の可能性に関する調査研究報告書」について ～オープンデータの生産者・流通者・消費者の適切な役割分担に向けて～

国立情報学研究所コンテンツ科学研究系 准教授 北本 朝展

1. はじめに

オープンデータについて自治体を知るべきことは何か。「基礎自治体によるオープンデータ化と利活用の可能性に関する調査研究報告書」(以下、「報告書」という。)を読めば俯瞰的な知識を得ることができるだろう。基礎知識を網羅的にカバーするため、100ページを超える報告書となっている。すべてを読み通すことは難しいかもしれないが、ざっと目次を眺めるだけでも重要なポイントを把握することはできるだろう。しかし、現場において実務を進める上で押さえておくべき考え方や、自治体がオープンデータの世界で果たすべき役割などについては、本当に重要な点が見過ごされてしまう危険性もあるのではないかと感じる。

そこで以下では、具体的な例を取り上げながら、オープンデータにおいて自治体が目すべき役割に関する私見を述べ、報告書を読み解く上での別の見方もお伝えしたいと思っている。

まず、筆者が重要だと考えるのが、データの生産者、流通者、消費者という概念である。例えば農作物の場合は、農家が作り、それが流通し、消費者が買うというルートを農作物が流れ、その逆方向にお金が行くことで全体が完結する。流通者が果たすべき役割は生産者と消費者のマッチングであり、そのためのインフラとしてマーケットや物流ネットワークなどが整備されている。一方、生産者はよい商品を生み出すこと、消費者はよい商品を求めることを役割とする。そして、財布のひもを握る消費者が決める「よい商品」の選び方が、流通者を通して生産者にまで最終的にフィードバックされることで、生産者が生み出す商品も徐々に良くなっていく。それが市場経済のメカニズムである。筆者はデータについても、基本的にこれと同じ構

造を想定すればよいのではないかと考えている。

しかし、報告書の中ではそうした構造が明示されていないため、すべての責務を自治体が負わなければならないかのように受け取ってしまうと、自治体の仕事が無限に増えていきかねないという懸念を抱いている。筆者の私見では、自治体に最も期待されているのは生産者としての役割、すなわち健全なデータ生産環境を整えて、よいデータを継続的に生み出す生産者になることである。そんな生産者になるためにどんな戦略が想定できるのか、それを本稿では考えてみたい。

2. データ生産者の役割

よい農作物とは何か。これは流通者や消費者のニーズによって異なる。例えば流通者から見れば、形が揃った野菜は運びやすく売やすいため、よい農作物である。いくら味は同じだと言っても、不揃いの野菜が流通者のニーズに合っていないければ、流通者から見た価値は低くならざるを得ない。しかし、直売所で消費者に直販するならば、むしろ不揃いの野菜の方が人気を集めることもあるだろう。また、有機農業に関心がある人、レストラン経営者などにとっても、「よい野菜」の定義は大きく異なるだろう。つまり、価値とは誰に向けて作るのかというマーケティングの要素を切り離すことができないため、万人にとってよい農作物を定義することは難しいのである。

では、自治体のオープンデータは誰に流すのだろうか。データ卸売業(流通者)向けであれば、まずは形が揃ったデータを安定的に生産することが重要である。データ卸売業としては、各地からデータを仕入れて加工して流すことがビジネスモデルなのだから、各地のデータが不

揃いであっては手間がかかって仕方がない。しかし、直売所モデルあるいは産地直送モデルならば、必ずしもデータが揃っている必要はなく、特定の顧客が喜んでくれることの方が重要となる。とはいえ、直売所モデルは地元でデータ消費者がたくさんいなければ成立しないし、産地直送モデルはまず良さを知ってもらうところに多大なコストがかかる。したがって、卸売業に流すことがまずは最も手間がかからないモデルであるとは言えるだろう。

このとき、生産者はどこまでやるべきなのだろうか。農作物であれば、ある程度きれいに整えた上で箱詰めして卸すことになるだろう。データの場合も、クリーニングを行ってそれなりのパッケージに入れて流せばよいはずである。最近はそのだけでは済まず、きれいにパッケージしたり、使い方のレシピをつけたり、生産者側でも付加価値を高める努力を行う場合が増えているとはいえ、生産者の基本的な役割は、流通者が扱いやすいものを安定的に生み出し、最低限のパッケージで送り出すことと言えるのではないか。そしてオープンデータにおいても、これが自治体の果たすべき基本的役割であるというのが筆者の考えである。

3. よいデータ生産物とは何か?

農家が土壌を耕し、作物が生育する環境を整え、プライドを持って高い品質の農作物を生み出しているのと同様に、自治体もプライドを持って高い品質のデータを生み出すことが望ましい。しかし、負担を軽減するには、生産者として責任を持つべき範囲をしっかりと限定し、その範囲に注力することも重要である。では生産者としてよいデータを生み出すためには、何に注力しなければならないのだろうか。それを次に考えてみたい。

これもいくつかの側面に分けて考える必要がある。例えば農作物の場合、味がよいという個別の農作物の品質だけでなく、形が揃っているという農作物全体の品質や、毎年安定して出荷できるという農業経営の品質など、様々なレベ

ルの品質が関係する。同様にデータの場合も、データが興味深いという個別のデータセットの品質だけでなく、フォーマットが揃っているというデータセット全体の品質や、安定して更新を継続できるという運営体制の品質も関係する。そうした全体像を把握した上で戦略を管理できないと、バラバラな動きのもとで非効率かつ低品質なデータしか生み出せない状況に陥りかねない。そこで以下では、一般論が難しい個別のデータセットの品質についてはスキップして、もっと大きい単位の品質について考えてみたい。

第一にフォーマットが揃っているという点である。この点に関しては、特定のフォーマットの優劣を論じるより、同じフォーマットを使い続けることの方が重要であり、これが安定性という観点からの品質となる。フォーマットが変わるたびに対応を迫られることほどストレスがたまることはなく、それに比べれば、フォーマットが多少奇妙であっても慣れてしまえば消費者は案外気にしないものである。細かい点にこだわって苦情を述べる人々はこの世界にもいるものだが、そういった意見を聞き入れるかどうかは生産者が自身の戦略にしたがって判断すべきであろう。

第二にデータがきちんと更新されるという点である。データをとりあえず公開することには価値があるが、そのデータの更新を続けていくことにはもっと価値がある。データは更新が命であり、更新されないデータは腐っていく。データの公開はニュースになるが、データの更新はニュースになりづらく後回しにされやすい。しかし、継続的に安定した生産を継続することにプロとしての価値が生まれることは、農家の場合と同じである。

さて、データの更新が命ということを理解した上で、さらに重要度が高い点に触れておきたい。それがIDまたは識別子の重要性である。データの基礎的な単位ごとに不変かつ唯一のIDを付与し、それを様々なデータに使いまわしていく習慣こそが、データの価値の核心であると筆者は考えている。それはなぜだろうか。