

# **Challenges of Public-Private Interfaces in Open Data and Big Data Partnerships**

Dennis D. McDonald, Ph.D.

---

# Challenges of Public-Private Interfaces in Open Data and Big Data Partnerships

By Dennis D. McDonald, Ph.D.<sup>1</sup>

November 7, 2014

## Partnerships

I've always been a bit wary of the term "partnership." When a large company talks about "partnership" with a smaller company there are many factors to consider especially if you're the small company, such as:

- What kind of resources will you have to spend before you see a return on your investment of time and money?
- If the partnership involves long sales cycles will you be able to sustain your efforts in light of the larger company's deeper pockets?
- Is there anything preventing the larger company from simultaneously engaging with your competitors?



After a while what you learn about partnerships is that you need to approach them with your eyes open:

1. Trust in any relationship especially a business relationship takes time to build.
2. Most companies are happy to have you *sell* their services.
3. Hoping *others* will sell your services without a formal agreement of some sort may not work.

In other words, partnering like any business activity needs to be approached with an objective view of the costs and benefits and an understanding that partnering takes work. As long as you understand this and those involved understand their mutual responsibilities you'll be OK.

## Government as partner

What happens when the partner is a government agency? What special considerations should you have about partnering?

---

<sup>1</sup> Copyright © 2014 by Dennis D. McDonald. Dennis is a project management consultant based in Alexandria, Virginia. He works with **BaleFire Global** on open data programs and with Michael Kaplan PMP on **SoftPMO** program management services. His experience includes consulting company ownership and management, database publishing and data transformation, managing the integration of large systems, corporate technology strategy, social media adoption, statistical research, and IT cost analysis. His web site is located at [www.ddmcd.com](http://www.ddmcd.com) and his email address is [ddmcd@yahoo.com](mailto:ddmcd@yahoo.com). On Twitter he is [@ddmcd](https://twitter.com/ddmcd).



I was reminded of this while reading [Frankenstein’s data](#) by Kieran Millard, Technical Director at HF Wallingford, a UK based independent consultancy in civil engineering and environmental hydraulics. Millard’s focus is on how to sustainably exploit “big geospatial data” as represented by the EU-funded [IQmulus infrastructure](#) currently under development.

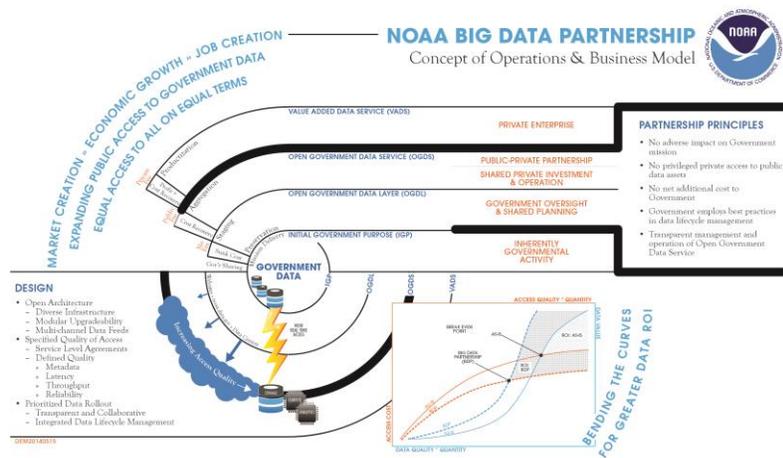
IQmulus will provide access to very large datasets and the tools for analyzing and extracting meaning from them. One of the challenges identified by Millard is the difficulty of characterizing data such as the data relevant to IQmulus, as being either explicitly public or explicitly private. Even when original data are provided by government agency satellites or remote sensing platforms, situations can arise when public data are combined with or processed by private sector resources to create something “new” with potential value that people might be willing to pay for. As pointed out by Millard, should public sector agencies — including those that originally generated the source data paid for by the public – pay “again” for products they had a hand in developing, since they supplied important data resources as inputs?

These issues are not new. As an example we have the [debate about the pricing of commercially published scholarly journals](#) where much of the data reported are based on publicly funded research.

In the data arena, as Millard discusses from the perspective of the European Union-funded IQmulus effort, the “open data” movement is forcing attention on issues of market viability and sustainability. “Who pays?” and “Who benefits?” are important questions that need to be addressed very early when such programs are being planned where one goal is to enable the private sector to add value to publicly-generated data.

## NOAA’s Big Data Partnership

In the U.S. we see similar “public/private open data interface” issues with [NOAA’s Big Data Partnership](#). In NOAA’s case much attention is already being paid during the planning process to how the interface between the public and private sectors will work when NOAA data are made available for commercial exploitation.



In the figure see the interface between “value-added services” and “open government data services” in the **Concept of Operations and Business Model** where the following high level “partnership principles” are listed:

- No adverse impact on government mission

- No privileged private access to public data assets
- No net additional cost to government
- Government employs best practices in data lifecycle management
- Transparent management and operation of open government Data Services

NOAA staff have put a lot of thought into the identification of high value data sets and the roles and rules under which these datasets will be made available to private sector “partners” for reselling and recombination with other resources. Following NOAA’s reasoning, everybody is supposed to win:

- The government continues to focus on performing services that are “inherently governmental.”
- Private sector “partners” are able to provide products and services on top of the open data made available for exploitation

## The details

NOAA has identified many datasets for potential exploitation and is currently [working through the details of making these datasets available to the public](#).

Despite the extensive planning and has gone into the NOAA program the issues raised by Millard will probably not get resolved until we get further along in the process. What if, for example, an available public dataset that is a byproduct of an ongoing NOAA program is re-analyzed or modeled in a way that makes it even more valuable than how the data were originally analyzed by NOAA? Should the government consider purchasing access to the “better” service? Will this be viewed by some as “double payment” for the data?

My belief is that such issues can be worked out if the process for doing is open and impartial — a requirement that NOAA has already established.

## Not just old wine in new bottles

Is there a reason to think that Government-provided “open data” will require a different set of criteria and practices to help distinguish between public and private uses?

That’s a key question and is why I think the openness of NOAA’s planning process will be so important. Data services and their use in development are much more accessible and advanced than in the past. Costs for developing, sharing, and providing cloud based database tools and services have dropped. New tools have made it possible to analyze and derive meaning from very large and disparate datasets.

Once an infrastructure evolves to support basic services and practices around sustainability and financing reach a basic level of maturity, the means then *should* exist for innovative new products or services to be developed that go beyond what the basic “inherently governmental” services offer. That appears to be what both NOAA and IQmulus are counting on.

## The big unknown

The big unknown here is whether a paying market exists for products and services that take advantage of the provided open data sets. From the perspective of a classical product development model, this is not how you are usually supposed to operate where you start with a *need* and *then* you develop a service to satisfy that need in a money making way. Here we appear to be starting with the data and

then hoping that partners will emerge who understand the need and can develop profitable services using that data to serve unfilled market niches.

## **Uncertainty → Risk**

The potential weakness of this approach is that the government's "basic" services may already be satisfying a significant portion of potential market demand. Potential partners are left to determine what unmet needs might be satisfied through innovative or creative use of the open data.

This uncertainty is why the model developed by NOAA is so important. By publicizing "rules of the road" regarding open data and the relationships expected between government and its "partners," and by discussing these roles in advance of any formal contracting effort, the groundwork is being laid for partners to make informed decisions about whether not it makes financial sense for them to invest in product development based on government supplied open data.

## **No more poster boys?**

Will the process work? That remains to be seen. While government-sourced weather data are the "poster boy" for showing how successful public-private partnerships are in turning data into useful and profitable services, we may now be entering uncharted territory where the financial viability and marketability of more specialized datasets now need to be considered.

## **Conclusion**

At a personal level I'm glad to see open data programs being taken so seriously. As taxpayers we pay for these data. We're justified in expecting the data to be put to good use. Government can't be expected to do "everything." Involving the private sector as a creative partner makes perfect sense if the process of doing so is fair, honest, and open.

## **One caveat**

Moving government agencies from performing basic services to becoming "data providers" may require some difficult changes for agencies and programs that have not traditionally performed the functions necessary to support open data access. Agencies and programs differ widely in how they approach data management. This may require, for example, much more attention be paid to data and metadata standards and quality control than they are accustomed to providing. I would not be surprised if government agencies find themselves experimenting with different organizational and governance structures as they begin to take open data programs more seriously.

## **Related reading**

- [\*Open Data Program Management: A compendium of quest blog posts for the BaleFire Global Open Data Blog\*](#)
- [\*Data Transparency Needs Data Strategy!\*](#)
- [\*Possible Pitfalls in Comparing Performance Measures Across Programs\*](#)
- [\*Government Performance Measurement and DATA Act Implementation Need to be Coordinated\*](#)
- [\*Open Data and Performance Measurement: Two Sides of the Same Coin\*](#)
- [\*Who Will Pay for Open Data?\*](#)

- [\*Data Standardization Scores and Changing the DATA Act\*](#)
- [\*Data Standards and Data Dictionaries Need Data Governance\*](#)
- [\*Recommendations for Collaborative Management of Government Data Standardization Projects\*](#)
- [\*Scoping Out the 'Total Cost of Standardization' in Federal Financial Reporting\*](#)
- [\*A Project Manager's Perspective on the GAO's Federal Data Transparency Report\*](#)
- [\*Data Cleanup, Big Data, Standards, and Program Transparency\*](#)
- [\*A Framework for Transparency Program Planning and Assessment\*](#)