

The State of Open Data 2011

Submitted on 25 Oct 2011 by

What is the state of the open data movement? Last week, during my opening keynote at the Open Government Data Camp (held this year in Warsaw, Poland) I sought to follow up on my talk from last year's conference. Here's my take of where we are today (I'll post/link to a video of the talk as soon as the Open Knowledge Foundation makes it available).

Successes of the Past Year: Crossing the Chasm

1. More Open Data Portals

One of the things that has been amazing to witness in 2011 is the veritable explosion of Open Data portals around the world. Today there are well over 50 government data catalogs with more and more being added. The most notable of these was probably the Kenyan Open Data catalog which shows how far, and wide, the open data movement has grown.

2. Better Understanding and More Demand

The things about all these portals is that they are the result of a larger shift. Specifically, more and more government officials are curious about what open data is. This is not to say that understanding has radically shifted, but many people in government (and in politics) now know the term, believe there is something interesting going on in this space, and want to learn more. Consequently, in a growing number of places there is less and less headwind against us. Rather than screaming from the rooftops, we are increasingly being invited in the front door.

3. More Experimentation

Finally, what's also exciting is the increased experimentation in the open data space. The number of companies and organizations trying to engage open data users is growing. ScraperWiki, the DataHub, BuzzData, Socrata, Visua.ly, are some of the products and resources that have emerged out of the open data space. And the types of research and projects that are emerging - the tracking of the Icelandic volcano eruptions, the emergence of hacks and hackers, micro projects (like my own Recollect.net) and the research showing that open data could be generating savings of £8.5 million a year to governments in the Greater Manchester area, is deeply encouraging.

The Current State: An Inflection Point

The exciting thing about open data is that increasingly we are helping people - public servants, politicians, business owners and citizens imagine a different future, one that is more open, efficient and engaging. Our impact is still limited, but the journey is still in its early days. More importantly, thanks to success (number 2 above) our role is changing. So what does this mean for the movement right now?

Externally to the movement, the work we are doing is only getting more relevant. We are in an era of institution failure. From the Tea Party to Occupy Wall St. there is a recognition that our institutions no longer sufficiently serve us. Open data can't solve this problem, but it is part of the solution. The challenge of the old order and the institutions it fostered is that its organizing principle is built around the management (control) of processes, it's been about the application of the industrial production model to government services. This means it can only move so fast, and because of its strong control orientation, can only allow for so much creativity (and adaption). Open data is about putting the free flow of information at the heart of government - both internally and externally - with

the goal of increasing government's metabolism and decentralizing societies' capacity to respond to problems. Our role is not obvious to the people in those movements, and we should make it clearer.

Internally to the movement, we have another big challenge. We are at a critical inflection point. For years we have been on the outside, yelling that open data matters. But now we are being invited inside. Some of us want to rush in, keen to make advances, others want to hold back, worried about being co-opted. To succeed, it is essential we must become more skilled at walking this difficult line: engaging with governments and helping them make the right decisions, while not being co-opted or sacrificing our principles. Choosing to not engage would, in my opinion, be to abscond from our responsibility as citizens and open data activists. This is a difficult transition, but it will be made easier if we at least acknowledge it, and support one another in it.

Our Core Challenges: What's next

Looking across the open data space, my own feeling is that there are three core challenges that are facing the open data movement that threaten to compromise all the successes we've currently enjoyed.

1. The Compliance Trap

One key risk for open data is that all our work ends up being framed as a transparency initiative and thus making data available is reduced to being a compliance issue for government departments. If this is how our universe is framed I suspect in 5-10 years governments, eager to save money and cut some services, will choose to cut open data portals as a cost saving initiative.

Our goal is not to become a compliance issue. Our goal is to make governments understand that they are data management organizations and that they need to manage their data assets with the same rigour with which they manage physical assets like roads and bridges. We are as much about data governance as we are open data. This means we need to have a vision for government, one where data becomes a layer of the government architecture. Our goal is to make data platform one that not only citizens outside of government can build on, but one that government reconstructs its policy apparatus as well as its IT systems at top of. Achieving this will ensure that open data gets hardwired right into government and so cannot be easily shut down.

2. Data Schemas

This year, in the lead up to the Open Data Camp, the Open Knowledge Foundation created a map of open data portals from around the world. This was fun to look at, and I think should be the last time we do it.

We are getting to a point where the number of data portals is becoming less and less relevant. Getting more portals isn't going to enable open data to scale more. What is going to allow us to scale is establishing common schemas for data sets that enable them to work across jurisdictions. The single most widely used open government data set is transit data, which because it has been standardized by the GTFS is available across hundreds of jurisdictions. This standardization has not only put the data into google maps (generating millions of uses everyday) but has also led to an explosion of transit apps around the world. Common standards will let us scale. We cannot forget this.

So let's stop mapping open data portals, and start mapping datasets that adhere to common schemas. Given that open data is increasingly looked upon favourably by governments, creating these schemas is, I believe, now the central challenge to the open data movement.

3. Broadening the Movement

I'm impressed by the hundreds and hundreds of people here at the Open Data Camp in Warsaw. It is fun to be able to recognize so many of the faces here, the problem is that I can recognize too many of them. We need to grow this movement. There is a risk that we will become complacent, that we'll enjoy the movement we've created and, more importantly, our roles within it. If that happens we are in trouble. Despite our successes we are far from reaching critical mass.

The simple question I have for us is: Where is the United Way, Google, Microsoft, the Salvation Army, Oxfam, and Greenpeace? We'll know we are making progress when companies - large and small - as well as non-profits - start understanding how open government data can change their world for the better and so want to help us advance the cause.

Each of us needs to go out and start engaging these types of organizations and helping them see this new world and the potential it creates for them to make money or advance their own issues. The more we can embed ourselves into other's networks, the more allies we will recruit and the stronger we will be.

(This is a [cross-posting](#) from David Eaves' personal weblog.)