



State Representative Mike Duffey, 21st District

Sponsor Testimony – Rep. Duffey HB 321-324 (Duffey, Hagan C.) – The DataOhio Initiative

Ohio House – Finance and Appropriations Committee – November 18, 2014

Chairman Amstutz, Vice Chair Boose, Ranking Member Sykes, members of the Finance and Appropriations Committee, thank you for the opportunity to provide testimony on HB 324, one of four bills that comprise the DataOhio Initiative – today proposed to be consolidated into HB 324.

While it can be easy to get lost in the details of the “open data” movement, which is a national trend across many states, it may be helpful to think of it like this: open data does not determine what is or is not a public record; rather, it only encourages public record data to be much easier to access in a meaningful way. DataOhio, as proposed, occurs in a permissive “opt in” manner. It is not a mandate.

The value of DataOhio to public offices in Ohio, both state and local, is the ability to share information more easily. The DataOhio board, which is made up of public offices, could suggest “best practices” for how to share public data online to make it simple to find it, aggregate it with similar data and use it.

A useful example might be to imagine a number of municipalities getting together to say, “For simplicity sake, let’s put our operating budget revenues and expenditures in Excel using the same labeled columns so we can compare our data,” instead of having 900+ cities deciding to use different, incompatible programs.

This same principle led to the creation of interchangeable parts, kickstarting the industrial revolution. And it led to the computer revolution once the industry adopted standards for floppy disks, the data on them and the manufacturing of computers to use that data in programs. It all relies on standardization.

This also isn’t really that different from following a recipe for a meal. The measurements for ounces, cups and the definition of ingredients do not vary, so following a recipe is possible. Without standards for measurement or ingredients’ definitions, it would be difficult to complete a recipe or make a great meal.

And here is how we achieve our goals with DataOhio:

1. We create a collaborative board of public offices (15 members with the possibility of more non-voting members) to convene discussions about data standards and goals – authority to recommend, not mandate.
2. We create an expectation that if a public office posts data online, it should strive to do so in an open format (non-proprietary, machine-readable) if possible. And this is expressed as an “opt in”, meaning that while the expectation is implied, it is not a mandate. It is a suggestion and a means to say yes.
3. With the support of Auditor Yost, we begin development of a voluntary uniform accounting system – called “uniform charts of accounts” – for various public offices that currently lack uniform accounting. Note: public school districts in Ohio currently do use uniform accounting, but cities mostly do not.

4. For public offices that adopt their respective voluntary “chart of accounts”, it will become possible to make “apples to apples” comparisons in ways that were not previously possible.

5. To make it even easier to organize and find the data you seek, we would also create a “one stop” catalog of public data sets at data.ohio.gov. You can think of it as a Dewey decimal system for data. It does not necessarily “hold” the data, but rather “points” to it through links, etc.

6. Data.Ohio.Gov also becomes a forum for tutorials, discussion groups and support tools as well as voluntary coordination with other states’ efforts (data.ny.gov, data.texas.gov, data.hawaii.gov, etc.)

7. Finally, because we acknowledge that the lack of a mandate inherently means diminished incentive for public offices to participate, we attempt to provide a “carrot” to a subset of public offices – cities, townships, counties and libraries – to complete a checklist of “beginner level” open data and if they do so, then they are eligible to receive a \$10,000 grant. It’s not awarded by a panel, so it is virtually guaranteed.

The case for open data at its broadest is not simply about good government for citizens, journalists and researchers. It is also about job creation, elimination of waste, technology start-ups and less red tape.

The case for open data has already been made. In the modern era, we take for granted our ability to search online for stock market performance charts, housing for sale or rent, weather forecasts, GPS maps, movie show times and much more. Yet all of those are essentially made possible through open data.

Whether you are addressing applications like Google Maps, Zillow, E-Trade, the Weather Channel, Flickr or others – it requires very specific data standards to work or each of those tools will cease to function. The parts have to “fit” to allow the tools to work.

Does government spending data, such as is proposed in HB 175 – Rep. Dovilla’s bill to put the state’s checkbook online – become easier to access because of DataOhio? Absolutely. The two bills are complementary and do not conflict in the least. But HB 175 is, by definition, a spending transparency bill. It is not a method for organizing the non-financial outcome data that we monitor as state officials.

“Budgeting for outcomes” is increasingly important, as recognized by the Ohio Chamber and the Ohio Society of CPAs. It is an acknowledgment that knowing “where money is spent” is useful – it is a vital, critical first step. But to understand “return on investment”, we must measure what we budget against the desired outcomes – and that largely comes from data that is spread across the many statewide constitutional offices from crime data at the Attorney General’s office, to electoral data at the Secretary of State, to K-12 educational outcomes at the Dept. of Education and all the data sets DataOhio coordinates.

And it is all of this disparate data, not necessarily organized in cross-silo compatible formats, that we can begin to integrate together or at least make it easy for someone to do so if they want to do so.

If you want to compare spending on public safety to prove it reduces crime, you would need not only state spending data, but probably mostly local municipal data, as well as state-level crime rates. So you’re mashing up city data perhaps with data from the Attorney General’s office.

It is this mashup of investment together with its return in the form of desired outcomes that gives us the ability to evaluate “bang for our buck”, which I would argue is among the most key objectives we should have on the House Finance committee.

To give you a sense of the breadth of those who would benefit from open data standards, consider the following:

Every day, a variety of Ohioans make inquiries about public information. These include citizens as well as local government staff, city council members, township trustees, county commissioners, city managers and state government agencies, including LSC, the governor's staff, legislators, aides and more.

These come on top of non-governmental requests such as researchers at Ohio's public universities (OSU, Cincinnati, etc.), journalists (Columbus Dispatch, Dayton Daily News, Plain Dealer, etc.) and from think-tanks (Opportunity Ohio, the Buckeye Institute, Center for Community Solutions, etc.).

And finally, many private sector jobs exist to add value to public information. For instance, companies like LexisNexis provide computer-assisted legal research, the Cleveland Clinic uses public health data and IBM helps both government and private companies analyze their data. Companies like StreamLink in Cleveland are providing automated compliance software for federal and state grant recipients, the list goes on.

Every one is searching for accurate information without too much hassle. And yet, hassle is what we've got in many cases – either because of technology, lack of standards, or navigational issues.

An ordinary request for information in one city might return to the requestor in the form of a photocopy of a spreadsheet, or a PDF of a PowerPoint. There is no rhyme or reason why it comes to you in that format necessarily. And once you have it, the information is not always communicated with the same meaning. For instance, accounting labels among different departments or positions may be different while describing the same line item, or even labeled the same for different line items. It can get confusing.

All of this hassle results in lost time and money. It unnecessarily occupies resources: re-typing data into electronic format, translating numbers to like-to-like definitional standards, searching drop-down menus for the information you seek, calling officials on the phone to explain your request and so on.

So that's why DataOhio is important. We are taking data that is hodge-podge, disorganized and difficult to use and we are essentially fulfilling a public library function for it. We are convening the many public offices that are responsible for it, telling them to get in a room together to work out standards, to speak the same language in accounting, to share their data with each other more openly and rewarding them.

Sadly, there is little incentive for many in government to better coordinate data right now. Sharing information benefits many, yet the work may be done individually with no guarantee that if a public office shares information about itself, others will return the favor. DataOhio helps to increase the benefit.

This is an area where state government can help simply by serving as a beacon. Public offices may choose whether or not they want to follow these standards but if they do, we will all benefit from the result.

Mr. Chairman, that in a nutshell, is what we are trying to do. And it is validated by a variety of stakeholders, from university researchers to business groups, think tanks, newspapers and transparency advocates, both in Ohio and nationally. Just this month, I was interviewed by a think tank in California that heard this bill package in Ohio is actually a competitive threat to California's efforts to be ahead of other states on open data. In other words, the Ohio model for open data could be national model.

And it is important to remember that while there is a whole ocean of public information out there, we are not trying to boil the ocean. Rather, we want the data we can use and that is why this bill is so permissive.

Thank you, Mr. Chairman.

#