

## **Council on Open Data**

November 7, 2014

12:30 - 3PM

Governor's Reception Room, 2nd Floor

The State House, Annapolis

### **Introductions (Greg Urban)**

- Welcome and thank you for coming
- Council on Open Data was established as statute and will not be impacted by the administration change
- The established charter will continue to make government transparent for general public

**Question:** Our current Governor had a real feel for data, does anyone know about the Governor Elect's perception of data and GIS specifically?

- Have not worked with him and do not know
- Think he comes from private industry, drive towards data-based decisions
- For the current Governor, GIS was one of his passions
- Cannot discount how businesses and how the new Governor will be making decisions
  - Still a lot of value to using GIS for decisions making
- Encouraged that the new Governor has experience in real estate, which would give the expectation that he is interested in maps and data
- One thing we should be looking for, the Transition Team is going to need data and hopefully we will be in a position to make data easy for them to obtain
- It will make our agencies look good if we do our jobs and continue to make data accessible and easy to get to, easy to use

### **Around-the-Room**

- There are a number of new participants at this meeting
- Conduct quick introductions of all attendees

### **Governor Office Transition (Barney Krucoff)**

(Presentation Slides Available)

- Previously we had an Executive Committee for MD iMap and a Open Data Workgroup
  - Got a lot done, for example, two new websites
- 2 separate groups by Executive Order, now we have a law in place
- We continue under the law and we keep going regardless of administration
- Work continues to help the new Governor however we can
- Hopefully data we put into these systems will prove useful and effective for getting the job done by the new Governor

### **Possible Repeal/Modification of Section 10-901 through 10-905 (Barney Krucoff)**

(Presentation Slides Available)

- Discussed last time and solicited feedback

- Got several pieces of feedback
- Got feedback from State agencies and local governments in the central region and Western MD
- Law allows for charging overhead of system in selling GIS data
  - No one objected to repealing this part
  - Can still charge for the cost of distributing, but no longer charge for collecting/creating
- No one objected to the clause to enter into a contract to use the data provided online by the State
- Concerns that the data cannot be sold at all
  - Law currently provides the opportunity to sell, but there are some local governments who want to continue to sell CDs/DVDs that include preparation and distribution costs
- Raises questions based on these comments
- Public Information Act (PIA) does allow charging after 2 hours, based on custom requests
  - Some custom request maps might be of interest to many people, how do you work this under the PIA laws?
- Are there other authorities, is PIA sufficient?
- The law would only allow for geospatial data now and why is that the case why it is geospatial only
  - Easier to leave this part in place, but this should be expanded to any type of data
- Maryland Department of Planning (MDP) gets a lot of data requests, specifically the State Data Center gets census data requests and data is put online, including some customized data that is not charged for
- For some CDs/DVDs and special maps, would like to recover some of the costs, some are large format maps, which need to be printed by MDP, recover costs of plotting, mailing and some costs for time to complete request

**Question:** Could the PIA law cover this type of request?

- If there is a walk-in request, applying the PIA law to this request would delay the response time
  - Customers are used to getting a quick turn around response

**Comment:** We do not want people to abuse staff time to manipulate datasets.

- Call out data manipulation requests or requests outside of what is available in the database
- Small businesses come in requesting boundary maps, printed out for writing on

**Question:** Can you offer an exemption for those specific agencies that have special fund fee setting authority?

- In those cases the intent was to offer charging authority for these specific service, under these specific conditions
- Would be nice to limit the law
  - Issue has been the temptation to think that a local government will get rich selling maps
  - Limiting would minimize this temptation, but downside would need to be able to understand what future changes might occur
  - Might need to come back often to add additional exemption agencies
- Some additional legislation on data requests, after 2 hours can charge a fee if you want to, but do not have to, this is related to the PIA law
- Printing maps is fairly expensive, have a request for a large quantity of maps, charge for materials, not time usually

**Question:** Is it the printing that accrues the cost or time?

- Would not charge for time unless there is a special custom program or the programmer had to take a considerable amount of time to complete this request
- Usually just try to recover the costs of the materials
- Is electronic only at no cost an option? - many maps are online already
  
- Draft language, online this data would be free, any data, CDs, DVDs, paper could be charged for
- Would this be acceptable language?
  
- The Department of Natural Resources (DNR) also struggles with these types of requests
- Can usually direct people to the online versions and that satisfies the need, but some of the requests go beyond and require some time and some printing
- Sometimes requestors are directed to local print shops
  
- D.C. would charge for certain sizes of maps, there are some requests that come in and need guidelines to understand when should give free service and when should give charge service

**Question:** Was the law's intent to charge above and beyond the costs of materials?

- Were not just charging for the CDs/DVDs and paper, were also charging for data, database and creating data, etc. time of employees fulfilling the requests
- Would still be legal for the state to charge for materials and time under PIA after 2 hours
  - Is this legal not under the PIA?
- Do state agencies have small fees built into requests for printing, etc.?
  - Under PIA, yes?
- What triggers a PIA?
- When someone walks in, can that trigger PIA and thus the option to charge?

**Question:** Is there a centralized State print shop where requests could be transferred to for printing purposes?

- Have to print out a request and send to this print shop which would delay the response time and create more work than just printing out directly
- Government print shops want to print out a few thousand times, not a custom map request, easier and cheaper to use a local plotter
- Central shops should be able to run this out, in theory

**Question:** Does a boiler plate PIA request form exist? Is there a mandatory wait time for PIA?

- Legally collect fee based on the form being submitted, but then would have to go through legal to get approved, asking customers to come back and delaying the response time
- Informal requests should be answered by State employees
- A formal letter might cost more and take longer, but can offer more rights, has to go to others for review, legal for example

**Comment:** Open Data law intent was for data, that is regularly produced, to be open by default, materials costs were not intended to be included in the law

**Action Item:** DoIT to come back next meeting with language clarification, vet with this group for additional comments

- Example where data is available, but takes time for someone to pull the data out of a system that is not accessible to the public and make it into a format that is accessible by the public, can be sent to the requestor
- Not changing PIA, only focusing on the special exemption for Geospatial data

## **Report from Public Information Act Subcommittee (Barney Krucoff and Joshua Exler)**

(Presentation Slides Available)

- Would a database of PIA requests be appropriate/desirable?
- Cannot currently answer the question of how many PIA requests are in MD
- There are some databases scattered amongst agencies of various sizes, some are customized
- Reasons it would be interesting, what data to make public based on what requests are coming in
- Inquiries from open gov advocates, interested in dates/times of the open meetings, track how government is doing with handling PIA requests, might be the group that this database would serve
  
- Difference between tracking data available and making information requested available
- Arguments on both sides of the subcommittee
- If you make the data available, not just the tracking information, make it available someone might not make the request again, but might if the data is thought to have changed
- Not data that would fall under the Open Data act, not structured data, not regularly maintained data, a lot are for emails, reports, unstructured data
- Some extenuating circumstances, can request own personnel file, a citizen cannot request someone else's personnel file, sometimes a request is made that cannot be publicly available data
- The system would have to be able to handle these exemptions
  
- A tracking database would have to come before a database that could handle the data
- Web form was discussed, where data would be entered via a web form into the database
- Number of concerns about how it would get assigned to the right person, when would the clock start, how to track down the right person in the right agency when a centralized database
- A tracking database would come before a web form, it would need to go into a tracking database
- Key groups that we need to talk to, Bar Association and Press Association to build a database
  - Solicit comments from these and other specific groups
- Subcommittee was generally cautious, did not want to change too much of tracking, but upsides to knowing trends, answer to open data community and being able to answer how many PIAs are out there and being responded to
  
- Recommend starting with a tracking database
  - There are some COTS products, the State has access to, that might do the trick

**Comment:** Want to avoid double entry for those agencies that already have these databases in place

- Many of the requests are documents thousands of pages long, people come in and review and make copies of select parts, would not be practical to digitize the entire document, in some cases

**Comment:** Thought of it as 311 system, at the State government level, there is a requestor and someone farms it out, tracking number is assigned, track the progress, using a web form or the phone

- Perhaps MDE can help guide this process
- Valuable to have a system for State government to track by departments

**Comment:** Concern at local government level, especially smaller counties, of associated costs

- Start at State level makes sense
- If State was to build out and develop, make system voluntary that locals could tap into if they so choose, open to local governments that want to build off this system and share costs

**Question:** Should we or is there concern about PIA and getting data? Do we care that requestors make money from the data they receive?

- Data is supported by tax dollars which makes it public
- Concern that the data being generated from the request is used to make profits
- Such thing as open source software, a lot say you can do with the code what you want, attribute the original source, but can find ways of getting people to pay for the final product
- Not dramatically different to say you cannot sell the data, because you found someone who would be willing to pay for it
- Might also want to consider attributing the original source (metadata)
- Have language about acknowledging the source
- Data is usually reformatting and made more consumer friendly before selling
- Reasonable to accept that data will be sold as we support businesses that aggregate the data, package and use for applications, display and useful service back to the public ultimately
- Take data and provide as service, some value in encouraging these actions
- Do hear this concern from local governments
- One interest to repeal the law is to sell, so someone else cannot make money off of it instead
- It costs money to keep others from making money off your data, would take a lot of time and money to track and see if anyone is making money inappropriately from free data
- Open data is going to drive \$3 trillion dollars into the economy, Zillow at top of the list of companies built on open data
- Baltimore Sun paper requests data and then publishes it to make money

**Question:** Has the subcommittee come up with whether tracking is hard or easy to put together?

- Some have databases that can be harvested
- To what degree is it a good thing if this is not going to be easy to accomplish
- Tracking database is the lowest hanging fruit, potentially benefit, proceed with caution
- Still questioning difference between data request and PIA request, data vs. information, still confused about how PIA requests relate to transparency of open data
- Tracking is talking about a database that would say this request came in and it was handled, would not put the actual report, memo, data online, that still only goes to the requestor

- Want to collect metadata (data about the data) for the tracking database, sentence in there about the original request, not talking about the actual data
- Actual data might trigger the use of the data into something that can be published for use by the general public, for example, information that was published about restaurant health department assessments
- Do not have the ability to immediately share data out of the gate, can accomplish to some extent, but we understand the steps that are involved and discussing how to proceed

## Using [data.maryland.gov](http://data.maryland.gov) and [imap.maryland.gov](http://imap.maryland.gov) to Make Agency Websites More Interactive

### Developing and Embedding Custom Dashboards (Mike Scott, Salisbury University)

- Show how you can embed dashboards, maps and charts into your websites
- History of BayStat, Governor liked graphs on progress of maintaining the bay, DNR easement additions, many different agencies coming together to contribute data
- Originally passed by spreadsheets and put into graph in Adobe Flash format which would not work on Apple iPads
- All agencies can put data into Socrata
- Upload and the graphs for Baystat were very customized, liked the Adobe Flash site, built calling from Socrata using Javascript, allows for graph display and data download
- Talking about three different ways to get data out of different data platforms to display and show to the public
- Mike Scott from Salisbury University and Eastern Shore Regional GIS Cooperative (ESRGC), work done by the team at ESRGC
- Recreate the original Adobe Flash dataset (Causes of the Problems: <http://baystat.maryland.gov/causes-of-the-problems-map/>)
  - Click on the various basins to determine impact, can also look at individual counties
  - Data coming from MDE, MDA, DNR, some in collected from models and other sophisticated analysis
  - Want to make sure that the data is up-to-date and the dashboard is built with custom code using open source libraries, code is free and available to public, library is called D3 is the primary library, map is leaflet, ESRGC reps created the connections
- Important to have data in [data.maryland.gov](http://data.maryland.gov) to see the same data in both places and agencies responsible for putting in the data would be controlled by the data originators
- Link to pull up the list in [data.maryland.gov](http://data.maryland.gov) that are contributing to the interface
- Interested in showing the data in Socrata to build a dashboard?
- Data explorer (<http://business.maryland.gov/about/rankings-and-statistics/data-explorer>)
- Goal is for DBED to send people who are interested in doing business in the State, compare the counties, compare States and compare metro areas
- Series of charts for different kinds of data (median household income, etc.)
- Can show all counties or compare just a few counties, which can be selected by using the map
- Embed the chart into your own website, code will be provided through the interface

- Download the data, can pull into local spreadsheets and conduct rankings
- Source of where the data comes from is also noted for each data type
- Lots of statistic categories, demographics, education, transportation, etc.
- These statistics have been maintained by DBED for years, but are now being made more readily available through the use of this dashboard
- DBED wanted to be able to modify the dashboard, because sometimes it is more than the data that changes
- Administrator login provided which allows for customizing the dashboard after initial deployment
- Can add different statistics that become available, all data is stored in Socrata, change data tables in each category, can change the order of the data, custom the taglines on the page
- This allows for the agency to make the updates and does not require a specialist to conduct the updates
- Charts read from a database at ESRGC that originally comes from Socrata to speed up the response time by preloading the data
- MD Workforce Dashboard for DLLR (<http://www.esrgc.org/dashboards/dllr/>)
- Built from a mainframe database dump to ESRGC of all people who receive unemployment insurance
- Aggregate to ZIP code level to protect individual identifications
- Information about individuals drawing unemployment insurance throughout the State at one time
- Can break down by legislative districts and other geographic regions

**Question:** Underlying data is owned by the State of MD?

- Socrata facilitates the query platform upon which the data is pulled into programs or excel files
- The law requires the Portals, Socrata is a cloud provider and without this product we would have to build our own portal
- We pay for uploading, but Socrata makes downloading the data easy

**Question:** Does the State of MD contract ESRGC to build these dashboards?

- Yes, DBED built one for a reporting grant from federal government
- DoIT has entered into a contract with ESRGC to build 3, already have one in progress, but have opportunities to build 2 more
- Have room to negotiate a request for a dashboard, data must go into the Open Data Portal, but will help build the custom dashboard
- Building one for the energy projects throughout the State now, working with MEA
- Interested in testing? - ESRGC can send out once draft is available

**Question:** The components of the second and third applications are made up of what?

- Front ends are the same, D3, HTML5
- Postgres database, custom code for translation between Socrata and dashboard
- Workforce dashboard precedes the Socrata platform being purchased
- A gallery of dashboards can be seen on <http://imap.maryland.gov/Pages/dashboard-gallery.aspx>

(Presentation Slides Available)

- Available resource called ArcGIS Online that facilitates the publishing and posting of maps onto the web quickly and easily
- GIS people can help you create a map and train you to do it yourself
- Examples: [OSPREY](#), [MD Public Schools](#) map, [Enacted Capital Budget](#) map and [Maryland Made Easy](#) map
- Reviewed steps to embed a map:
  - Step 1 – Get your data. It takes 80% of your time to find, clean and organize your data. Use [data.imap.maryland.gov](#) or [data.maryland.gov](#) to search for your topic. 80% of your job has already been done if you use the data that is already available
  - Step 2 – Map development. Get an account with ArcGIS Online (<http://maryland.maps.arcgis.com>). Select a base map. Symbolize layers. Configure pop-ups.
  - Step 3 – Map deployment. Select a template. It is helpful sometimes to have a layout that is like other maps. It makes it easier for the user to navigate your site.
  - Step 4 – Embed map. Two options: First option, do not have to pick a template, just grab some basic code from GIS online. Second option, use a snippet of code that you are able to customize
- GIO/DoIT Office available to assist!

**Question:** How do maps interface with the datasets agencies are supposed to be posting on Socrata?

- Some data is easier to manage on maps vs. Socrata. In Socrata, you have industry best standards for tabular data and it can be built into a map. Agencies should post datasets in one of these 2 places.

**Question:** Have there been considerations to cross the data between the two portals?

- We have worked on some options to pass data between the systems behind the scenes
- Still working on how to make this happen
- They are two separate systems that were purchased/developed independently
- We will continue to move forward with the separate systems, but cognizant of minimizing the duplication
- There are standards in the data that are lost in the translation from one system to the other, therefore there are still incentives to maintain the separate systems

Developing and Embedding Socrata Views (Josh Exler, StateStat)

- Embed tools are available where the code is generated for you
- Additional option for those who do not want to work with HTML code directly
- Open performance module is built on top of the Socrata site, can build out without any code experience at all, the information is fed into the module and create with drag/drop interfaces
- Drag chart onto the site, select the dataset list appears and allows for dragging and dropping components
- Can split the page into different segments and then show a number of graphs on the same page
- All graphs linked to the Open Data Portal and linked as data is updated in the system, the data is updated in the graph

- Can print page as an option, print button, share options
- Can build out dashboards quickly and easily
- For something more customizable, the routes of dashboards and maps are the way to go
- To take advantage of data that is already on Socrata, we would love to see users using this functionality, it is underutilized at the moment
- Maryland.demo.socrata.gov, separate account for this site then the Socrata site, follow the same process to get a standard account, contact [Josh Exler](#) and he will create an acct for you

**Question:** How would users access the open portal site?

- Link to the individual URL or use an iframe and embed the code that is generated
- Not embedded into a department website by default, hosted at demo URL
- Hosting all of the data and content on Socrata central site and then embedded into individual agency sites using an iframe
- For a simple chart, table, use Socrata
- For a map, use Esri
- For a chance to use one of three dashboards, contact DoIT and we will engage ESRGC
- The cost is the same for all = \$0
- No new server, no change control, if you are willing to let the data go public, it is a few easy steps from there

## **Creating a Maryland Data Inventory (Barney Krucoff & Josh Exler)**

(Presentation Slides Available)

- Ambition to do a data inventory
- We will never have a complete inventory in Maryland and what systems they came from
- Would like to build a web form and send to agencies, where the inventories are kept
- This is going to be confusing, is it one system with 100 data sets, is this one entry or 100?
- We have been collecting a list of what we think we want to ask about the data
  - Agency Name
  - Dataset Name
  - Contacts for the Datasets
  - Source System
    - Probably would not publish this part
  - When did you start collecting data?
  - Are you still collecting this data?
  - Does it cover the entire state?
  - At what aggregate level?
  - How often is it collected?
  - Is it published as often as it is collected?
  - What is the size?
    - Socrata contract based on number of datasets, not size
    - Do not want to have to limit by size, but might not want to post small ones either
  - What is the original format?
  - Are there existing locations where data is already being extracted and published?

- Need to learn how to consume data from an existing process
  - Are there existing publications and where are they?
  - Is this open data?
    - Want to know about data that might be open and closed for this inventory, not requesting the data itself
    - Answers would help determine if it is open data under the law or if we would just keep them in the inventory but would not request the data itself for open data publication
  - Data quality
    - This will be a judgment call, gaps, changed the way the data was collected, examples of questions to be asked
  - Caveats of the data?
    - A place to explain these circumstances
- How intrusive is the list and what might be missing?

**Comment:** Continuity is suggested, a lot of data sets got collected for awhile and then stopped and then started again, need a place to describe this information

**Question:** Is the ultimate intention to make the list publicly available?

- This is the intent, but we may get push back
- First we want to collect what is out there, some data call, through CIOs of the agencies
- DoIT already asks questions about systems, but not data in this way

**Comment:** There will be positives to this beyond open data

- Some data should be closed and remained closed, but inside State government there could be ways to leverage the data to solve problems
- Want to give assurance to CIOs that just because data is on the list there is not going to be an immediate demand to open it up or publish or work to get it ready for someone to use
- Cannot prioritize which should get funding to make them available to the public, until we know what is there
- Cannot answer many questions until we have a decent list of what is out there

**Question:** Based on the transition, what is the time it will take to get this done?

- Is there a plan for the Council on Open Data to meet with the transition team to present the initiatives that are in progress?
- We have not yet been approached, we would invite them to sit in on one of these meetings
- This would include the transition documents that will go to the team from DoIT
- Let them decide how they want to engage

**Comment:** Is this open data too broad

- Perhaps add an element to allow for saying aggregate data would be available
- Categories: Health, Public Safety, helps find redundancies and gaps
- We can debate if there should be many categories or the best category to be selected by the data entry provider

- Common in the GIS world, but not always successful
- Are there models out there that got a pretty good inventory product
- Comes down to the data call and guidance
- A lot of these inventories that happened in the early 2000s, despite making it as painless as possible, it came down to on the phone trying to understand the submissions
- Standardize the collection up front will get better responses
- Make it as easy as possible, checklist is ideal, yes/no questions
- Going to need a “hounder”, someone who is in charge of hounding people to get this done, necessary part of the process bc of the number of datasets being talked about here
- Go along with the check box, does it have personal information and the accompanying box is does the data have value without the personal information
  - Can get that during the follow up or initially, but need to get this information
- Data catalog for HSIP is a model to potentially look at

**Comment:** There is a lot of data in work, what level does the data have to be at to rise to this list?

- This is a good question, need to understand the readiness to publish

**Question:** Time frame?

- Do not see the big data call going out until after mid-January
- Wait, see about the interest of the new administration and line them up to make a data call
- Might consider working this into some other processes that are already being requested by the agencies from DoIT, some are documents which are freeform and there might be redundancies

**Question:** Is there value in a list?

- Value in a translator, sense of a brief description of high or low priority
- Start with some that are of interest and focus on collecting them first
- Name of the description is enough to know the value of the data before digging into the full list
- This could be more informal and done sooner, more quickly

**Comment:** Suggest some incentivizing

- Give credit for those who are revealing datasets that are being uncovered/unearthed
- Instead of just setting an order
- Suggest if it is already on Socrata or MD iMap we will not ask you to inventory because we already know it is out there

## **Closing Remarks (Greg Urban)**

- Thanks for coming, see you in a month and work on ideas then

## **Future Meetings**

December 18, 2014 @ 1PM

Location: Governor’s Reception Room, State House, Annapolis