

Council on Open Data

June 27, 2016

2:00PM - 4:00PM

Maryland Department of Information Technology

100 Community Place, Crownsville, MD 21032

1st Floor Conference Room

Welcoming Remarks – Luis Estrada, Deputy Secretary, Department of Information Technology

- Council on Open Data has new life with new customer service initiative from Governor
- Charged every agency with improving transparency and customer service
- Ensuring that agencies are working for the people
- Interesting opportunities for the council emerging from the new initiative
- Identify opportunities for setting goals for open data that meet the new mission
- Datasets, such as Public Information Act (PIA) requests, that can be pushed out in open data portal to show more transparency to further the mission
- Governor's Office of Performance Improvement (GOPI) is looking at measurements, anything we collect as performance of agencies
- Challenge departments to be transparent. The data belong to the people
- Welcome questions, comments, feedback to get the most out of the meeting

Attendee Introductions

Update on Barney Krucoff

- Back with D.C. Government as Chief Data Officer, running the open data and Geographic Information Systems (GIS) shops
- Changes in GIO Office
- Kenny Miller is retiring from state service in August
 - o This is Kenny's last Council on Open Data meeting
- Julia Fischer and Josh Exler remain on board and will continue to support the Council
- Going forward under strong leadership

Future Dates for Council Meetings – minimizing scheduling conflicts

- Conflicts with Mondays for a number of members
- Meet four times a year, like the idea of getting together quarterly
- Two official meetings
 - o Try to get the appointed and named members to come to these meetings
- Before annual report is submitted in January, and then other two meetings in spring and fall
- Statute indicates that this group meets at least twice per year
- Earlier in the day, traffic in the area and long day for those who do not live close

- o 10AM – 12PM is first choice and 1 – 3PM is second choice
- Factor that into the equation, not before 10AM nor too late because of travel
- Idea of November and April/May meeting for everyone is acceptable
- Decided to come up with exact days and dates and send them out and get feedback via email
- Have to put these in MD Register, put the dates into the calendar a year ahead of time

Governor Hogan’s new Statewide Service Initiative

- Launched on the June 9, 2016
- Customer Service Promise
 - o Talk about statewide 311-type systems
 - o Simple database that tracks, logs in public’s personal information, call in, email in, go to website and enter information

Improved Search on data.maryland.gov – More relevant and faster results

- Search tools have been improved over the past month
- Changes from Socrata went into effect about 2 weeks ago
- data.maryland.gov search catalog has a new interface
- Relevancy of results has been improved, more filtering and ranking the search results
- Improved search speed
- Result links would overhang off the right side of the page, this has been fixed
- Tags, keywords and date of last update showing up in the search results directly
- Default datasets are a mix of datasets from state data and some federated from other portals
- Filter on the left by keyword and type of view
- One drawback to the new design
 - o Old search catalog had an icon graphic applied to each dataset
 - o This feature has gone away
 - o Datasets all look the same, just text
 - o This is a trade-off, load speed improved in part because no longer loading graphics

Question: Do we lean on the vendor to reinstate the graphics?

- Prioritize the interests of the Council
- To know if this is a priority what is the comparable, relative to what other priorities?
- PIA tracking system brings more value to the agencies and public, then an icon on a webpage
- Sounds to be a moot point since it was an out-of-the-box feature that is no longer available
- There is another option, which would require work to build a custom front end interface
 - o Gives us complete flexibility, but a lot more time to complete and maintain by Department of Information Technology (DoIT)

Question: Is it part of the scope of work? – No

- Response from vendor: Working with five other states and local open data users, this was a feature that was not well adopted
- In the future it could be incorporated
- Some recent webinars includes the big picture
- Big enhancements coming in the next few cycles
 - In the catalog view
 - User's experience is enriched
 - See where Socrata is making investments
- Motion to not worry about icons, second, no objections

Updated links to MD iMAP layers on data.maryland.gov

- All of the MD iMAP data layers are found within the data.maryland.gov search
- Use data.maryland.gov search for both GIS and non-GIS data
- Currently migrating MD iMAP links and they will continue to be accessible through the Open Data Portal, but the links are being updated
- New links will be available at the end of processing in a few days
- For 30 days, both sets of links will persist within the system
- After 30 days the old links will be removed
- Make sure to change any links on websites, within existing documents and other places where the links are hardcoded
- This is a one-time, bulk change
- Publishing a dataset that shows the old links and the corresponding new links

Highlights from other State and City Open Data and GIS Programs

- City of Los Angeles has similar format, two data portals, new agreement with universities to allow for obtaining data quickly
- A lot of datasets updated close to real-time using 100 GBPS network speed
- Available for the ability to download data quickly, not to promote real-time data downloads
- Is there interest or a need for such high speed download capability?
- Maryland Transportation Authority (MTA) puts GPS on buses, deliver directed time to the bus in relation to the next stop
- The next step is to see where the bus is, in real-time on a hand-held device
- Police realm is active, but not public information, but not far from needing to be there

Question: What is the limiting fast connection to the universities now?

- They are currently accessing at the same speed as the general public

Question: Aside from the network connections, is there more to it? Just connect higher speed circuits? Should we install those connections? Have we seen demand, from users, to obtain the data quicker and at a higher volume? Is the speed at which the data is able to be accessed now sufficient for the types of data and size of data on the data portals? Do we need higher speed connections to the universities?

- Only a matter of time before this speed is demanded.
- Rhode Island collaborative has brought together higher education providers with state government and service data to those research institutions and facilitate better research and decision making
- Capacity vs. education engagement for groups that want and need data
- Effort into publication of data, but not as much on the side of solving problems with the available data
- Determine what datasets would be valuable for these education institutions and make this data available
- All comes down to impact of data access, ensuring that access is focused on the public and others who need to access the data

Question: Have there been problems documented? Range depending on what the data is? What are our customers? Are they universities? Are they hungry for more data, different data?

- Provide market research across the spectrum from the average citizen to an analyst, what data they want to provide and access, work to understand the community around your data
- Data experts, not marketing and outreach experts, biggest challenge is to let the public and educational institutions know what we have available now
- Need to become more successful with targeting the correct audience

Question: Anyone here interested in helping define the customer profile?

- Delegate Cullison's representative showed interest
- Governments are a big user; legislators are looking for this type of data, need to market to them
- State does not leverage the technology at the universities, representation on the Council (Morgan State University, Johns Hopkins University (JHU) and Salisbury University)
- Need to better understand customer profiles, universities and define the problem we are trying to solve, network speed, datasets are not available that would prove to be useful and respond
- Joint MD/DE Association of Librarians
 - Attended annual conference, the revelation was significant, opportunistic scenario, but not a systematic approach

Action Item: Make a plan, concrete steps, reach out to universities, cannot answer question until we know the customer base

Data USA Dashboards from MIT Media Lab

- MIT Media Lab launched Data USA for states and cities across the country, dashboards being built out on-the-fly
- At a glance, data about Maryland, similar to what DoIT and Eastern Shore Regional GIS Cooperative (ESRGC) has built out in dashboards
- Large site with a lot of data resources, a lot of this information is coming from census data (ACS data), data visualizations, charts

Center for Public Integrity Report

- Open data and open data policies were evaluated
- We took exception to some of the input in the report
- Socrata organized a follow-up call with states that were not happy with the results
- A lot of overlapping concerns to be addressed in next year's report
- Primary points of contact to evaluate each state were mostly journalists, academics, non-profits
- Dropped this on the public at the last minute, they did not reach out to governments directly, i.e., those who were trying to improve open data within a state, instead questioned third parties only
- Those asked lacked knowledge that the programs exist, and this lack of knowledge was treated as evidence that programs didn't exist
- Next year they will loop in the governments more, we are generating the products, release a draft report with commenting period, to allow review before final submission
- We will continue to have the D on our report, the next year's report should assess more fairly and accurately

Existing PIA Practices – Changes to law effective October 1, 2015

- Law change as of October 1, 2015 created Office of the Attorney General (OAG) Office ombudsman
- Open data portal will allow making good on requirement of immediacy in the legislation
- Have the data online before the data is even requested, proactive
- Not all requests are on quantitative data, a part that the open data portal can play, geographic data or non-geographic data can allow for links to the data and potentially eliminate PIA requests

Existing PIA Practices – Agency-by-agency

- Need to know what each agency is doing before we can determine where to go with this item
- Maryland Department of the Environment (MDE) receives approximately 5,000 PIA requests a year, after some are aggregated, bulk from consulting firms, law firms, NGOs and individuals
- Currently have website where requests can be made, check off what interested in

Example 1:

- \$19 to produce what was requested, materials are bound materials, so cannot make copies easily
- Database tracks hours, individuals involved and rates, account for exceeding two hours of time

Example 2:

- Plans folded up, bound reports, NGO request, no charge because collected quickly, make room available for reviewing documents, can make copies for pay or bring in copiers and scanners to make copies of the resources
- More time and dollars involved if documents have to be reviewed before they can be released to the requestor

Question: How many come to review in person versus request copies?

- Depends on the requestor
- Not practical to consider trying to digitize all of the responses and put online

Question: What happens if someone reviews in person and pages are missing?

- Sometimes happens, things get out of order
- Review room is off the lobby, monitored by receptionist, occasionally things do go missing

Question: Log request for the room?

- People make requests to come in at a specific time, schedule appointments
- Response time includes the notification that the information is available, does not include the time to schedule an appointment

Question: What about redacted content?

- Occasionally would have to go through all of the files and review and redact information
- Get files to agency, make copy, black out parts and then make copies of the blacked out versions
- Cost approximately \$8,000

Question: Would you see more people interested in the same data or put in more requests?

- Do not think so
- Biggest universe is property transfers, history of the property, underground storage tanks, leaks, etc.

Question: Does the agency employ a records or document management system?

- No single management system used not throughout the department
- Use FORTIS for some of the programs and documents within the department
- How do you handle? Print out tracking sheet to go with the documents
- Each administration has PIA assignee, meets requestor and shows them what records are theirs
- Database, in the process of upgrading, tracks request date, what was asked for, yes or no if available documents, fee rate waiver requests

Question: Could some of the documents be published online?

- Would have to take time to scan and unbind the documents

- Do not have anyone to do that work right now
- Some of the documents are oversized or older in age and would not photocopy well

Question: Are there similarities between the requests? Would the new database capture some of the similarities that are routinely processed? How many of the requests are similar in nature?

- Similar that they are looking for information about a property
- Similar about a specific property, getting historic data online is more difficult than newer data
- Some permit search ability online now available
- Worth asking what data could be made available online to reduce requests in the future
- Recently received request for emails and other documents related to a particular project
- Continuing requests for emails within a timeframe that pertain to a specific area of interest

Question: Anything else that starts digital and then paper copies are made?

- Not much, submitted reports, newer ones could be submitted electronically

Possibility of tracking all agencies' PIA metadata as Open Data

- Office of Attorney General flag issues
- PIA includes exemptions from disclosure
- Individual asks for data about themselves they can receive protected information
- Tracking systems could include metadata and might be able to include response letter, but as long as it does not include personal materials
- Things that might not have to be redacted for one person might have to be for the public at large

Action Item – Focus on metadata for now, making the metadata, look to larger agencies to determine how they are tracking the metadata

Question: Is it possible to determine how many requests have been received by an agency?

- This would be part of the data we are trying to make public
- Also whether the agency being requested tracks those metrics
- Recommend a pilot agency to start building the metadata database.

Question: Do we know other examples, Maryland Insurance Administration (MIA) has a good example online, and should we research with them? Matrix of who is doing what?

- Part of the research effort
- If we capture metadata, should be link to response if made public already.

Draft List of Metadata Elements for a Statewide PIA Tracking Database

- What is the overall reaction to a high level, statewide database that tracks how many requests the state does a year? How many each agency does in a given time?
- Maryland State Police (MSP) tracks in an excel spreadsheet, one person shop

- o Ask that the answers not be entered into a database
- o Do not have the time to complete a request like this
- o Very few repeat requests
- o Anything to build would be great, but would not necessarily help the PIA responders
- Take a spreadsheet, regularly and map your field name to the database matching field name

Question: A PIA request itself is public information itself?

- Yes, according to state counsel
- Usually very specific requests, fairly rare that a request is duplicated
- Would it make sense to make things freely available, be proactive
- Requests sometimes require redaction before they can be made public. For example requests for criminal reports, someone has to read everything, including narratives to redact personal information

Possibility of tracking all agencies' PIA full responses as Open Data

- Received requests from E911 procurement documents, all documents had to be reviewed
- One request took year to review because it included 30,000 pages of documents
- 30 days response requirement applies unless truly cannot respond in that timeframe
- Voluminous request could not fulfill within 30 days with best intentions
- Ombudsman can call something overly burdensome
- 100% of the responses are about sending something, not people coming in to look at something
- Maryland State Department of Education (MSDE) – 25% are data requests, all responses are reviewed by Attorney General, 100 requests for the year, This is a good idea, but will add to the burden.

Question: How does MSDE track information now?

- Who is asking, request, response, making sure this information does not go into a spreadsheet
- Attorney General said not to do it, because it can be FOIAed, so MSDE's tracking system has no digital component
- Use filing system instead, because it cannot be FOIAed
- DoIT suggests as a best practice to avoid spreadsheets as a final data location: Instead make an open data portal dataset, rather than make a spreadsheet
- The dataset on the Open Data Portal looks exactly like an excel spreadsheet, can setup fields and rows and data entry is done once
- Interface created online, tracked and shared simultaneously, core of the Portal functionality
- At the core of the product is to view datasets, create new dataset, looks like blank data table, mirrors what you would see in a single excel spreadsheet, field names across the top

Question: Would the PIA Open Data tracking mimic the effort in Montgomery County?

- To some extent, also includes links to response letters, upload the full pdf, steering away from this last part

Question: Can this tracking be customized?

- Unique tracking needs, when tracked and who it was sent to
- Do not complete the actual responses, but need to track the activity and movement of the requests

First Steps – Possibility of pilot projects

Question: How long would a modification take?

- Want to proceed with metadata tracking and not full piece, full bookshelf of documents
- DoIT to start with a pilot, MSDE has 100 annually, start with an agency with smaller number of requests
- Action item: Mock something up with Socrata and simple front end and come back to the Council with example
- Do we buy something off the shelf or create something personalized?
- Could customize through your spreadsheet or upload the portions that are requested in the central spreadsheet
- Purpose of Council looking at standard elements that would be made public, not the system you are using
- We would take the system you are using, pull data, as available and if there is a need for a statewide system, that is flexible, could take on as programming initiative to offer an enterprise solution
- Start publishing PIA metadata, need to identify the scope of the dataset, trivial in the open data portal

Question: Is there a risk of using the open data portal as the source data versus a copy? Service Level Agreement (SLA) offered, data recovery?

- Setup minimizing errors within the system, removing duplications of data entry
- Contractor can speak to redundancy and backups and fail overs

Question: Does this make sense to pursue statewide system?

- Capture 10 elements, not a laundry list of items
- Capture the basics:
 - Full name of responding state agency or office
 - Acronym of responding state agency or office
 - Contact information for responding state party (address, city, state, ZIP, phone number, email)
 - Name of request submitter (PIA requests are considered public information)
 - Date of request submission

- o Date of response deadline (date from #5 plus 30 days)
- o Date of response
- o Brief description of documents or information requested
- o Public response (approved/under review/denied)
- o Fee charged for production of response (yes/no, and dollar amount, if applicable)
- o Additional comments (optional)
- Add Link to Full Response, if available, but not make it a requirement, optional
- Response letter, standard fare, attached documents, as requested
- Request letter could include protected information, would not be asking, by policy, that it be included
- If there is a pointer to a public data file, we should link

Comment: Itemized list is included with the response letter, also includes what could not be shared because of law

Comment: Nice to make this helpful to the agencies' internal tracking also: what part of agency was sent to, names, their dates to respond back to PIA representative, do something that has added value for a particular agency

- Take data that is already available and publish it in one solution
- Make a system to track the requests is another solution
- DoIT request enterprise initiatives to generate a solution to track PIA requests from start to finish
- Invest in a statewide solution

Comment: If at least the drafted elements are available, you would have contact information to reach out to the request fulfiller, can obtain the additional information about who exactly responded

- A lot of internal departments, PIA representatives need to be able to track the movement of the request throughout the department
- Governor's Office uses IQ (COTS) - Commercial off the shelf product to track correspondence

Action Item: Create dataset for pilot agency, canvas agencies to determine how handling request, assess pain points, assess what agencies would look for in a system, deliver a solution would be for DoIT to pursue

Comment: Need to make sure that timeliness in responding is also taken into consideration in relation to the Governor's new customer service initiative

Comment: Need to take into consideration that an agency that does not have to acknowledge information because of exemption, make sure that it is clear that this does not have to be revealed

- This is about collecting information and making it available in a single place
- This is not about changing the PIA process

Question: Why are we publishing this information?

- Transparency, metadata is open data
- If the agency does not already track the data in a machine readable fashion, cannot compel you to participate in making the metadata available

Question: Will this put agencies' under additional liability?

- They are already liable; it would just be publishing information about the request.
- If we build this tool right, it can provide added value to the agencies by helping them be aware of all request timelines for which they are liable

Comment: There is a tracking system and publishing the information will increase the workload and requests numbers. General public will discover that information is accessible. The Council is outward facing, to respond to the general public, DoIT is inward facing to support the agencies. The law and customer service initiative asks us to respond to the needs of the general public and hold agencies accountable.

Comment: Open data aspect is not an ideal fit in every case, dataset by dataset examination.

- Example of Zillow that pulled in property assessment data
- Point people to transactional data and let them make of it what they will

Action Items PIA: Open Data Council will go agency by agency to determine how PIA requests are handled and provide metadata about PIA requests to Open Data Portal. If an agency does not want to participate, we will collect information. For example, if Maryland State Police decided there is a reason not to participate we will collect that information and report back at next Council meeting.

- DoIT has identified a need to centralize a tracking system, request to investigate this year
- If there is a pressing need, based on a schedule, contact Luis Estrada.
- Follow up with schedules concerning official meetings, come up with dates for upcoming meetings within the next two weeks.

Next Meeting

Date: September 26, 2016

Location: 100 Community Place, Crownsville, MD 21032