

Council on Open Data

April 25, 2016

2:00PM - 4:00PM

Maryland Transportation Department (MDOT)
7201 Corporate Center Drive, Hanover, MD 21076
Harry Hughes Conference Room, Lower Level

Introductions (Kenny Miller, Acting Geographic Information Officer (GIO), Department of Information Technology)

- Thank you for joining us today
- Barney Krucoff is no longer in state service. Kenny Miller has assumed the role of Acting GIO and will step into the role Barney performed for the Council on Open Data
- Rest assured that the Department of Information Technology (DoIT) is still committed to assisting the Council with fulfilling its mission and the goals submitted in the 2015 Annual Report

Opening Remarks (Luis Luna, Governor's Office of Performance Improvement (GOPI) Director)

- Interested in receiving more feedback and input from the Council
- How should this Council be run? How can we best engage the Council members?

Council on Open Data Recommendations and Goals – 2016 (Josh Exler, Eastern Shore Regional GIS Cooperative (ESRGC) / DoIT)

(Slides Available)

- Recap of the recommendations for Calendar Year 2016 from the Council on Open Data's previous annual report

Goal #1 – The Department of Budget and Management (DBM), DoIT, and state agencies should collaborate with the Governor's Office of Performance Improvement (GOPI) to open and publish Managing for Results (MFR) data and other datasets identified by GOPI.

Goal #2 – GOPI should monitor agency and gubernatorial initiatives and identify opportunities to make data public.

- Working with DBM and GOPI to create data visualizations submitted as part of MFR process
- Raw data submitted related to MFR, as a starting point, publicly facing data, only public measures
- Dashboards are being developed, based on jexsubmitted data, imported into single source
- All agencies will have MFR metrics in a standard format
- Goals, objectives and metrics are counted based on total number submitted and available, more added as additional agencies are added
- Browse goals from the homepage for those agencies that have a dashboard
- Dashboard is at Stage 2 or Stage 3 of data interactions, building on top of existing data workflows
- Additional functionality includes search all goals, objectives and metrics, enter keywords and results across all agencies will be returned, find out, at a glance, how Maryland is doing
- This is a standardized solution to visualize measures coming out of spreadsheet, graphs setup the same, have different colors to help visualize the change between different agencies
- Link to Open Data Portal, ability to return results for the same keyword entered in the MFR search

- May help with limiting or diminishing the number of Public Information Act (PIA) requests coming in
- Extract, Transfer, Load (ETL) tools are setup to accommodate metric schema changes from year to year

Question: Will there be multiple years available for this data?

- Currently FY11 to FY17, FY16 and FY17 are estimates, other years are actuals
- In January, new data will be added, FY12 to FY18 will be the new date span

Question: With the MFRs going up, do other datasets that were uploaded to the Open Data Portal still need to go up also? Other health rates and other raw numbers, is this information still desired?

- MFR data is only handled with these dashboards
- The rest of the data that is submitted to the Open Data Portal is not related to the MFR data used in the dashboard
- Moving forward, performance management will be assessed and additional datasets to be submitted will be determined based on interest and need
- Continue to submit updates to existing and new datasets to the Open Data Portal

Question: What is the timeframe of a percentage change? How often do goals change?

- These are handled by each agency
- Lines establish goals and graphs show whether a goal is being met for each year that falls under that particular target

Question: How will business units of the Maryland Department of Transportation (MDOT), be reported out under MFR?

- Probably look like individual MFR dashboards for different business units
- Distinct landing page for MDOT possible
- Similar approach for other departments which also have multiple business units

Question: Have all of the MFRs been received from the agencies?

- Starting dashboard development with a batch of agencies, for which, all content has been received from DBM
- Recently met with DBM and agreed on priority listing and timeframe for dashboard development

Question: Are some of the MFRs in CY and some in FY? Will this be distinguished on the site?

- This will be indicated on each MFR, footnote language will also be included

Question: Will more ambiguous definitions be defined or described?

- Talked about building an interactive glossary, click or hover over terms to see definition pop-up
- Could also be applied to words that are used by different agencies, but have different meanings

Goal #4 – DoIT should develop and provide regular training sessions for agency staff on how to make data public.

- DoIT has already held training on how to add data to the Open Data Portal

- Entry level basics on how to use the Open Data Portal
- Training materials sent out as part of the attachments with today's meeting email
- Training materials also available for download from the [Open Data Portal website](#)

Goal #7 – DBM and DoIT should publish Maryland's Operating Budget as open data and develop an interactive dashboard to help citizens explore budget data.

- Operating budget dashboard has been developed
- Currently under review by Governor's Office, prior to public release
- Dashboard will be linked to Open Data Portal
- Dashboard shows clarification of raw data, so data can be reviewed more easily
- Shows data by different fund types (General, Special, Federal and Higher Education)
- Allows users to drill down and see different agencies, by expenditures, compare salaries vs. contracts vs. travel expenditures, compare years of data
- Kind of tool the Open Data Portal was designed to put out there, transparency of what taxpayer's dollars are being applied to

- Process for putting this all together was interesting
- Proof of concept included Financial Management Information Systems (FMIS) data, broken down by agency and spending type, block chart where you could drill down to get more granular information
- Required buy-in from management
- Concerns about public misunderstanding data or misinterpreting data, worked with GIO Office and the concerns were alleviated
- Contracted with ESRGC to generate the dashboard, do the "heavy lifting"
- DBM submitted a spreadsheet, answered questions about what the dashboard should look like, layout and functionality, draft dashboard launched, iterative process for tweaking and making corrections

- Dashboard will be launched publicly, within the next few weeks and linked from the DBM website
- Requires collaboration from within the agency, owners of the data to tell the technical folks how to interpret the data
- GIO Office has been great to work with, encourage to continue to look around for opportunities to get data out, machine readable and transparent

Question: What is the purpose of the Governor's Office review of the dashboard prior to release?

- Council on Open Data is setup to provide input and advice, recommend policies, one of these was to put this information out in a public format, the dashboard, the Governor's Office will determine if the dashboard meets the standards that were recommended by the Council on Open Data
- Vetting process to make sure the data is sound and reliable before it becomes public
- Data is already available, the dashboard is an effort to make the data transparent
- Similarly, the MFR data is available, the dashboards make it searchable and easily accessible

Question: Is the Capital Budget map also live?

- Based on the nature of the capital budget data, funding assigned to particular projects, DoIT tried to figure out, early on, how to map this budget
- An existing map is linked from the GOPI homepage (<http://gopi.maryland.gov>)
- FY17 capital budget map to be launched soon
- Drill down from a map, find out schools that are getting funding, find out roads that are getting built
- Operating budget is harder to map, because there are no points with a spatial reference
- These two resources are complimentary

Goal #6 – The Governor’s Office of Crime Control and Prevention (GOCCP) and DoIT should development a dashboard to help citizens analyze crime data by county.

- Working to make user-friendly, web-accessible crime statistics for the State of Maryland
- Recently partnered with DoIT to create state, county and municipal data on the Open Data Portal
- Next step is to create a user-friendly dropdown menu to allow for easy access to the data
- Someone coming from out of town can see these statistics, can also see how the state compares to other states
- Intent would be to have visuals to compare statistics between municipalities, counties or states
- Data coming from Maryland State Police (MSP) Uniform Crime Reports data, crimes reported to police, published annually now, in a 200 page book
- Put 1975 to 2014 data for all of the counties into the Open Data Portal, added another 10 years of data recently
- Dashboard is still preliminary, has to go up and be reviewed by Executive Branch before release

Question: Website shows crime statistics about Maryland or specific cities. Where does the data come from?

- MSP, published in annual report, manually process the data, photocopied older, print books into PDF format and then get the data entered into the system for analysis and dashboard consumption
- Can continue to add data as more data is turned into electronic format, or add other types of crime data (arrests, etc.)

Goal #9 – DoIT, with input from the Council, should report on the costs and benefits of a potential statewide Public Information Act (PIA) tracking database.

Goal #11 – The Office of the Attorney General (OAG) and DoIT should leverage new provisions of the PIA to encourage agencies to identify additional datasets for publication as open data.

- Workflows are already established; want to improve what already exists, what role does the Council on Open Data provide?
- New provisions went into effect in October, to make PIA request contacts public for state agencies
- Need to come to consensus on what exactly the Council on Open Data should and should not be doing related to this item
- Partner with OAG, OAG does not see this as “their turf”, they get requests, but have no way to track them, either the number or if the request is made more than once
- OAG has no expectations of taking the lead on this, more room for the Council to play a role
- Council will work on the fact that sometimes things work easier when there is one “champion”

Possibility of Centralizing Information Requests

Do you see an opportunity for PIA tracking to be improved? As a Council, should we avoid? Should we centralize efforts at the State level?

Comments from Mike Scott (ESRGC)

- There needs to be discussion about the level of tracking that goes on at each agency, thinking about finding a best practice agency and scale up from there, statewide is too heavy a lift, forwarded requests have to be considered
- Put into process some tools that would allow agencies to have an intake place to track requests, track how long to fulfill them, important process to complete

Comments from Robert Wray

- Ombudsman been hired yet? – Yes, this position would want to assess transparency as to why a dataset is too complex to post
- Two different levels: metadata (data about data) associated with the request (i.e. –name of agency request is going to, date request submitted, etc.) and a database of the data being requested so others who might be interested can go to the database and find newly available data

Comments from Andrew Gosden (MDE)

- MDE gets 2,000 PIA requests a year, PIA is not about requesting data or datasets, it is about requesting documents, not asking for database, but instead asking to produce records you already have, including emails about a particular site, background checks on environmental sites, verify environmental issues for development, there is not a way to take documents and make into PDFs and make this of use to anyone else
- Tracking who requested, who was assigned, how many hours (for bill backs), MDE is using a database to track, new version should be available soon, could get MDE IT representative to provide presentation at next meeting

Comments from Elliott Plack

- Metadata database sounds like a logical first step, is there a platform already available to create a web form or table, would a POC have to retype manually into the system?
- Try to figure out these mechanics

Comments from Josh Exler

- A handful of agencies have tracking systems already out there, we should investigate
- Montgomery County publishes requested metadata and responses to its Open Data Portal

Comments from Chunlei Ding (Howard County)

- Discussing the goals about the system would be helpful
- Request For Services requests, could be managed via CSM, for example, service requests, not PIA requests
- Informing users about the difference would be helpful
- Do we want to consolidate these different types of requests or focus only on PIA requests?

- The more we track these requests, the better the outcome, ultimately make a lot of agencies lives easier if subject matter areas were defined more clearly, PDF or data that is requested multiple times could be drawn up quickly

Comments from Scott Shaffer

- A lot of these requests are not unique. Tracking and sending out materials generated from a previous request, providing consistent answers, helps with internal processes

Question: Non-government Council members, are there similar solutions you use in your, tracking sales, contacts, capture information, processing, timing?

- Customer Relationship Management system can be useful, track response for products, WebEOC is an example, tracking workflow automation, track that all deliverables have been provided to customer
- [18F Org at GSA](#), building tools to address FOIA process, may push down for state and local organizations to use
- FOIA has a lot of tools related, managing requests and data related to those requests
- City of Oakland made a product, [RecordTrac](#), that is available for purchase
- Need to do some homework about best practices and what types of solutions ought to be included

On Agenda for Future CY2016 Meetings

Goal #3 – Departments engaged in permitting and compliance, including but not limited to Department of Labor, Licensing and Regulation (DLLR), the Department of Health and Mental Hygiene (DHMH), and the Maryland Department of the Environment (MDE) should make more information available as open data.

Goal #5 – DoIT should inform research librarians in Maryland’s public libraries to train them on finding and working with Maryland’s Open Data websites.

Goal #8 – DoIT should incorporate a statewide dataset inventory into Maryland’s Information Technology Management Plan.

Goal #10 – DoIT, with input from the Council, should evaluate open source alternatives to Maryland’s current open data portals.

Goal #12 – DoIT should make full address points and LiDAR data available to the public on open data portals.

Status of Open Data Portal (Josh Exler, ESRGC/DoIT)

(Slides Available)

- Traffic on the site and new features
- Page views vs. browser page views – 49,221 browser page views as of April 22nd

Question: Why is there a limit on the public datasets and what is the impact of this being limited?

- This limit is a result of the contract with Socrata, the resource provider for the ODP, linked to license subscription, can get more, if needed
- How many public facing datasets are hosted on our site is a consideration, consolidation is possible
- Aggregate into a single dataset, data from multiple years that is based on the same schema

Dataset Freshness Report

- Intended to round out the analytics of what is being tracked
- Live to the public, whether each dataset is up-to-date or not
- Table generated once per business day, goes through metadata and aggregates each dataset's metadata, generates additional metadata which are used for the report, including "Updated Recently Enough?" (Yes, No – or a message that more metadata is needed to determine this)
- Datasets are then scored based on date range thresholds and a color is assigned to visually indicate the freshness of the dataset (for example: green is fresh and red is stale)
- Some of the datasets cannot be calculated, because an update frequency is not assigned in the metadata of the data submission
- Also working to phase out the option of "As Needed" for update frequency, replaced with "Other" and a field to "Please describe below..."
- "As Needed" was meant to serve as a catch-all, but it is being selected too often for simple monthly, quarterly, annual etc. data updates

Question: Do updates to this new page update the metadata page of the datasets?

- One way flow of information from the existing metadata to DoIT scripts calculating whether each dataset is "fresh"
- This new page, which is generated automatically by ESRGC/DoIT, will not be accessible for update, except by the management team.

Question: If the data has no hits on it, is the data owner informed that this data should be removed? Are there an average number of hits, when under, maybe data does not need to be updated?

- If data is available, try to measure how valuable the data is for the site, by how clean the data is, how robust, data quality is the measure, not popularity
- Leaving the dataset available is preferable, if possible with updated data

Question: Records retention, what does the Open Data Act require?

- Open Data Portal administration team, as a best practice, toggles datasets to private to "delete" them
- This allows the datasets to still be available if needed later and does not affect our contract's limit of 500 public datasets.

Question: Retention schedules for the data?

- Have not dealt with this much to date, because of the nature of making as much data available to the public
- Have not cleared out old data, except in some cases where new data is definitively no longer being generated or when new datasets deprecate or consolidate older datasets
- No set policy of when data should be cleared out, no manual backup required, have not needed to consider formal policies to date
- Council could proactively come up with expectations, if consensus agrees it should be included

Question: Where is Maryland State Archives (MSA) in this process?

- Maryland State Archives (MSA) has not been involved with this project, which has been lead by DoIT. The datasets detailed in the Dataset Freshness Report are the datasets on the Open Data Portal (data.maryland.gov) and MD iMAP only, and do not cover additional state data resources, e.g., MSA data resources.
- MSA is represented on the Council on Open Data, but its role has been similar to that of other state agencies on the Council, i.e., MSA is not yet facilitating the creation and uploading of Open Data to the Portal. This is a possible topic for future meetings: whether MSA can and should play a role in the digitalization and creation of state agencies' Open Data resources.

Question: Plan to have training, at colleges or universities? Open data allows for creativity of minds and inform on new levels, help entrepreneurs, etc.

- Outreach, training, marketing
- As a citizen of Maryland, not aware that this is out there, need a marketing strategy to make the public aware
- In May, GIO attending a MLA Conference in Ocean City, MD, partnering with Delaware to make librarians aware of these freely available resources
- Do not have a written marketing plan, takes someone experienced in marketing to help come up with a plan
- Social media links can be added to dashboards
- Could show what is possible by making a story through the resources already being used

Questions/Comments

What is the best way to conduct these meetings in the future?

- Subcommittees, PIA might be the place to start, should we continue to address PIA at each meeting to keep the ball rolling?
- Have regular PIA subcommittee meeting?
- Metadata standards, marketing, training – could be separate subcommittees?
- Really this would be a work group, narrow focus, break out of this meeting, online, here is a problem, what are solutions, bring to Council and discuss outcomes from work group meeting
- Ideas of engagement, breaking into smaller groups, discussion across the table, rather than a presentation
- Homework assignment idea? – put back in front of each member what we are trying to do, keep goals in mind, reporting out progress
- Ideas of Work Groups – Data Presentation (dashboard), Data Quality, Metadata, Data Privacy and Security
- Guided discussion, different views from different areas, breakouts will not allow for well-rounded exposure
- Remotely or take time out of this meeting to break into different work groups, need a leader for each and specific objectives, get back together and report, in addition to some of the standard format that will be followed
- Logistics of a presentation, trying to make the presentation more easily viewable
- Value when the “why” is revisited, tangible example of dataset, a successful example, dive into agency example, anchoring to the vision of why we are doing this open data

- Should the meeting be regularly scheduled, get onto calendar, next years' worth of meeting's scheduled, third Monday of every third month, afternoon at MDOT, whenever possible; get materials out as early as possible
- 4 times a year, July and August will probably be skipped, but stick with same day of the month, even if there is a longer gap between actual dates
- Like to have more interactions with the Council members between meetings

Next Meeting

Date: June 27, 2016, 2 to 4PM

Location: 100 Community Place, Crownsville, MD 21032