Minutes of the Open Government Advisory Group meeting on March 1, 2018

Attending:
- Tiffany Crowe, Chair - Public Member
- Barney Krucoff, Vice-Chair - the Chief Data Officer of the Office of the Chief Technology Officer, OCTO
- Karuna Seshasai - representative of the Executive Office of the Mayor
- Julie Kanzler - representative of the Chief Technology Officer, OCTO
- Shannon Turner - Public Member
- Josh Tauberer - Public Member
- Sandra Moscoso- Public Member
- Robert Becker - Public Member
- Kathryn Pettit - Public Member
- Traci Hughes - Director, Office of Open Government
- Richelle Russell, standing in for the designee of the Department of Education
- Alan Karnofsky - representative of the Office of the City Administrator

Members of the public present

- Michael, OCTO
- Peter Casey, OCTO

Not attending

- Betsy Cavendish - representative of the Executive Office of the Mayor General Counsel
- Clarence Wardell - Public Member
- Judy Greenberger - Library, Director of Strategic Planning… Not a member yet but intends to fill the vacant slot?
- Faith Leach - representative of the Deputy Mayor for Greater Economic Opportunity
- Melissa Tucker, - Mayor's Office of Legal Counsel
- Justin Grimes - Public Member
- V. David Zvenyach - designee of the Council of the District of Columbia
- Timothy White - representative of Deputy Mayor for Planning and Economic Development
- Jennifer Comey - representative of Office of the Deputy Mayor for Education
- Jay Melder - representative of Office of the Deputy Mayor for Health and Human Services

The BEGA / Office of Open Government Issue

Josh - BEGA board is appointed by mayor with advice of counsel - within the scope of OGAG. 1 vacancy on BEGA and 2 of 3 members will have expiring terms, so the mayor will take some affirmative action that it's in scope for us to advise on.

Tiffany - you have a right as a public citizen to comment, but since this group is a mix of govt/public members, not sure if it is is ok for govt employees to step out with a specific perspective. Can't say at this point if it is not in the perspective to advise on the position description. Can't say how we will be able to provide input into the selection

Traci - The work of the office is continuing. It is just a regular day today. Employment ends on 4/22. She will be held over until the board fills the position, but she is planning on not being there 4/23. It is my hope that the OGAG that we can be as forceful as we can be so the description does not infringe on the independence of the office and continues to fulfill its mission.

Bob - Essential to say operational independence of the office from BEGA was quite deliberate. The original bill. BEGA would enforce FOIA/OMA, but that was taken out by the committee.

Tiffany - No one is debating that Office of Open Government is independent, but the question is what's OGAG's role

Barney - Traci, are you on this body as ex officio as office of the government or personally as Traci Hughes? Mayor's order determines that the office's director is on the group, or his/her designee.

Traci - As the office.

Barney - I'm interested in her participation, and interested that the person to be hired will participate.

Josh - There is a seat for the director of the Office of Open Government, but they can choose or not choose to participate.

Traci - Not sure if the order says may/shall. If I am around in May, I will be here, but I'm planning for April 24.

Housekeeping: Meeting Time

Tiffany - People didn't care when the meeting started, but 80% for shorter meetings more frequently.

DC DOES Presentation

DOES people not here yet. Tiffany will send materials around if they do not come.

OCTO Presentation

Barney: Nearing completion of Enterprise Data Inventory (EDI), which has a report due Nov 1 each year, except first year to be March 11. We know statistics will change (100+ datasets from health not in). Recommendations are under review, but wants group opinion and maybe others.

[shows slide deck]

Peter Casey: Analysis of EDI metadata submitted by agencies. They are still collecting and consolidating and helping agencies complete it so subject to change. 60% participation rate across 112 agencies; 80% of mayoral agencies participated (they were required), 27% of non-mayoral agencies (they were requested but not required to participate). Many agencies are small offices and participated but do not have much data and did not think they had any data, so OCTO will help them understand what they may have.

Karuna: I am adding some datasets for them right now.

Tiffany: Non-mayoral agency participation is the bigger problem.

Traci: I am one of those agencies. It's probably not that they don't think it's important, it's more that they are not understanding that they have data to put in the repository. Their IT people are equipment-IT not data analysts.

Tiffany: I was being provocative. What else do we know about why they aren't participating?

Barney: Some of these agencies value and assert their independence.

Peter: Lots of cross-pressures - they all have limited resources and tons to do. And this was viewed as something to do, not as an important thing.

Barney: We worked harder on the mayoral agencies. When non-mayoral say no, (shrug).

Peter: There are some non-mayoral agencies that we may not want to count like DC Courts which are federal.

Bob: They're not covered by FOIA either! There is a real separation of powers issue.

Sandra - Can you give us a sense of size?

Barney: If it has a "board" that runs it, it is typically independent. DC Water is probably the biggest. The EDI covers an estimated 60 "subordinate" agencies and 40 "independent" agencies.

Traci - Agencies are either "subordinate" of the Mayor or "independent," in legal speak.

Petre - 1/2 of the agencies in the EDI contributed 14 or more datasets.

Karuna - Agencies assigned their datasets level 2 or 3 [from the District Data Policy], but also added the dataset again as level 0 (open) but with redactions.

Peter - Yes - that should count twice.

Barney - That wasn't done enough. It could be used to put pressure on in the future.

Peter - Compared the number of datasets inventoried w/budget (left out outliers of DCPS and HCFA which have very large budgets and contributed a lot). Of the data classification levels, 35% of datasets were open data; 10% public-not-released; 15% for DC use only; ⅓ confidential (HIPPA, student data, etc); restricted confidential 7%.

Julie - asked for "newly" open vs grandfathered.

Barney - 110 datasets from health dept are coming, so the story about having a lot of open data in the inventory may not hold. Good news - see bimodal distribution that we always thought we should see.

Peter - General counsel review helped.

Bob - If it is anonymized, can be released if PII is removed.

Barney - If they acknowledge it, citizens can ask for it to be anonymized. I may be wrong. There are vital statistics that are open.

Peter- Over 60% of open are already on data portal. A lot of them are open, but are published on the agency websites. One of the things we want to do is to get them on the portal so there's one place to search for it.

Josh - Raw number of open is not the interesting thing here. Classification levels applied correctly. It would be great to have all confidential data documented. Is there auditing/sampling to say it is correct?

Barney - Policy says general counsel decides classification. There is no appeal process in the policy, but the assumption is that if you FOIA, you'd get rejected for the reason that is stated. And then you could appeal and it would be resolved that way.

Data classified as open but not on portal is on the to-do list. What data we should start with, your opinions would matter. There were liaisons to help agencies classify - we looked at name, desc, and reason and if there was a mismatch could be a problem in the desc or in the classification (without seeing the data). Effort to ensure consistency across the board.

Karuna - Agencies are being cautious with classification. Folks are expecting FOIAs for the data, and at that time, they will think more critically and will re-evaluate it then. Agencies know the list is dynamic, and it will change, and are being cautious now knowing they can change the classification later, or when we ask them to re-evaluate it.

Tiffany: If FOIA responses use the EDI classification as a reason, you could end up in a circular loop.

Barney: With FOIA there's at least an appeal process. We want to find out datasets were released on appeal and make that a feedback loop to the EDI.

Bob: FOIA is costly and time consuming. We want the open data process to make disclosure not rely on FOIA.

Karuna: Agencies are more cautious now and are not thinking yet about derivative data that excludes e.g. PII. That's something for OGAG to think about, suggesting derivatives.

Bob: Not every disclosure of PII is unwarranted and exempt from FOIA. The issue should be how to design the database so that in disclosure it is easy to exclude fields.

Tiffany: This is foreshadowing discussion of committee work.

Peter: 80% of datasets reviewed by counsel; 60% of agencies. Working on getting 100%.

Barney: If it is not approved by general counsel, we are still including in the inventory, but it is marked as not approved.

Highest categories - general operations, education, transportation,

Barney: The EDI report has goals for the data team:
- System at OCTO to create MOU for purchasing - we want to adapt the system for data sharing agreements internally. Part of goal is to get the data flowing inside.
- Modernize data backend for analytics. Better search, haddoop
- Help agencies with the proactive component of FOIA law.
- Formalize data submission standards - other cities have this.
...and goals for the city administrator - include critical perf indicators for participation in the inventory

Sunshine week - lots of events to talk about data inventory. Wants to make future EDI report due dates the Monday of Sunshine week every year - hoping to change it in the order.

OCTO performance hearing tomorrow - one question we get is about databases. Each committee asks different questions. We want committees to agree not to ask questions if the answers are in the inventory.

For mayor and council - one problem with complying with FOIA is the time to search 294 TB of email and no retention policy. Expensive, slow to maintain and search.. Everyone wants their own email.

Karuna - Mayor Fenty issued [email?] retention policy but rescinded it 1 month after. Looking at other states for best practices. It delays FOIA productions infinitely - search 6 servers. From 10 to 60 days

Barney - wants consistency if each email had different retention policy that would be a nightmare.

Mayors orders says critical infrastructure is confidential, but doesn't define it. Grosso bill deals with it nicely.

Traci - Use the current proposed transparency legislation as a guide or pull a definition of critical infrastructure directly from the federal side.

Bob- The bill references federal definition. What would be local security that couldn't be defined as national security? If you look at it in terms of critical infrastructure, it is probably covered.

Barney - Legislation to mandate participate in EDI would cover independent agencies. There is an open data provision. We are not endorsing the bill.

Bob - The problem with referencing orders in legislation, because orders continually change.

Traci - Why reinvent the wheel - pull out the relevant part of the existing transparency bill and massage section.

Subcommittees

Tiffany: Current working groups are not really doing much. I have tried to break our actual ideas for activities down into groups. A person will lead/coordinate and be responsible for action. We all really want to be doing things and we have ideas.

Current working groups
OGAG engagement
Information access
Partnership and collaboration
Transparency and understanding

Proposed subcommittees
Open meetings
agency outreach

FOIA - most complicated - what is our role
Comms
Data policy adoption
Potential mayoral or legislative actions

Propose - 1st Thursday at every month.

Sandra - We didn't follow a robust process to form these groups, so we should talk about them.

Tiffany - Would we rather do that than the agency presentation?

Sandra - Is there a way to pull teeth to get gov memers of OGAG here?

Tiffany - Some of the invitations from EOM email or Barney - it is a problem and we are working on it.

Bob - Everything from this entity should come from dc.gov email address.

Tiffany - I can't access my dc.gov email during work hours.

Sandra - We hope we have folks with agencies that are hopefully interested in participating.

Barney - Look at our interagency data team meetings which are attende by analysts that may be more interested in OGAG than a top-down appointment. Meeting every 6 weeks on March 7. They are not the official member tho. MOTA asks agency heads for a representative and can be changed. Analysts want other people's data and see the value.

Sandra - This is about more than data.

Tiffany - We want someone who is engaged.

Alan - Proxy members can't vote on things like meeting minutes.

Next meeting = April 5 at 4pm.