Open Data Advisory Group
August 31, 2017 Meeting Minutes

Introduction

A meeting of the Open Data Advisory Group (ODAG) was held on August 31, 2017 from 7-8pm at 2100 Clarendon Blvd (Conference Room 311). The following Members and County Staff Liaisons were in attendance:

Members	*County Staff Liaisons*
Duke Banks	Jack Belcher (DTS)
Alyssa Doom	Bryna Helfer (Communications)
Christopher Hedrick	Dharmesh Jain (DES)
David Husband	Jaime Lees (DTS, Staff Liaison)
Scott McGeary	Katie Pegoraro (DTS)
Vadim Sokolov	Sanmeet Sidhu (DTS)
Jason Thomas (Chair)	
Dennis Whitehead	

Charge to the Group, Focusing Efforts

The Chair opened the meeting. He briefly reviewed the 5 tasks identified by the County Manager as the charges to the group. He highlighted 3 of the 5 items as those areas where the Group will direct its initial attention:

a. Identify and prioritize datasets to be published according to standards that ensure the data is accurate, complete, timely and published in accordance with security policies and privacy laws.

b. Provide recommendations on industry standard best practices to ensure the accessibility and usability of the published datasets.

c. Recommend methods to enhance communication and collaboration between stakeholders and the County on data access and use, including the development of joint data products and services.

Open Data Portal

The Chair shared a white paper "Open Data 101: Breaking Down What You Need to Know," a copy of which is posted to the Open Data Advisory Group website. He noted that Montgomery County has a very strong open data presence, and Arlington should aim for similar. He explained that the County's open data is available to anyone in the community who has interest and ability to use it. The Open Data Policy, available on the Open Data Portal website as well as the Open Data Advisory Group website, outlines how datasets are considered for inclusion in the repository – members can use these guidelines to help shape dataset requests. He displayed a visualization of tennis court reservation data he prepared as an example how the data could be used. The interactive visualization (a copy of which is posted to the Open Data Advisory Group website) shows reservations by park, zipcode and time of day.

The Staff Liaison presented a tour of the Open Data Portal – what is in it, metadata about datasets, how to find datasets, how to request datasets and the options for downloading datasets. She mentioned a data analysis project underway regarding metered parking transaction data and the opportunity for that real-time data to be posted to the Open Data Portal and used by a community member to build an app on where one is likely to find parking. Members asked questions and submitted comments:

1. How are new dataset requests processed and updated?
   a. Requesting a dataset involved completing a simple web form, at which point the dataset is posted to an open list of requested datasets with a status of that request also posted. Members noted that some requests have been in the queue a long time. Reasons why a dataset request was declined are visible in the staff comments section of the request. The County's CIO highlighted that work may be underway or dependent on something that is longer term, and efforts are underway now to improve the status reporting of these requests as well as allocating staff to formalize a timeline for processing requests and to get quicker movement on pending requests moving forward.
2. Suggestion to add many more datasets, then reach out to local hackathons and similar.
   a. A DTS representative confirmed that Arlington open data already has been used in hackathons and that continued outreach is planned. General response was that engaging hackathon communities is beneficial, even as the County open data volume is still growing, though larger volumes of related datasets would be more compelling.
3. Updates were requested on the parking data availability. (action item)
4. Suggestion to migrate static PDFs of data on the County's website to active data in the Open Data Portal. A follow-up question was raised whether the Group should recommend based on file format or content. (action item to clarify and evaluate suggestion as a recommendation)
5. How does the Open Data Portal handle inconsistent data? The member found acreage discrepancies in the public spaces master plan posted on the County website.
   a. The Chair responded that all data has inconsistencies and one benefit of open data is it reveals these inconsistencies for reporting and reconciliation.
6. What efforts are being made to standardize data entry across all County systems for uniformity of open data? Plans for this should be incorporated in the County's data analytics strategy with the Open Data Portal as a beneficiary.
   a. The County's CIO replied that there is no Countywide effort to standardize data entry fields to a common standard. While different systems have different data entry requirements, there would be benefit to Countywide definitions of data used across systems and this could be implemented in the Open Data Portal if it existed. (action item to evaluate question as a recommendation)
7. Is there any integration of the County's data analytics initiatives with state and federal policies?
   a. The Staff Liaison answered that there is no active integration at this point and no other jurisdictions doing so. Virginia is just beginning integration at the state level while Arlington is also focusing attention initially on integrating the County's departmental data, but is keeping abreast of state and federal developments and can be opportunistic.
8. Many improvements to the Open Data Portal interface were suggested, including
   a. adding text to explain the purpose of the site,

b. having more visible links to request datasets
c. easier means to finding datasets (action item to add index)
d. updating metadata so users know what the datasets contain
e. easier way to download, and find the download options
f. creating how-tos, videos and other self-training materials
g. (action item to clarify what is requested and evaluate suggestions as a recommendation)

The Chair urged members to think not only what the County can do to improve transparency and expand the content and usability of the Open Data Portal, but also to what they as community members can/will do with the data that is there or what is needed to support their open data usage ideas. Ideas for use are something the Chair is tasking the members to do (action item).

Adjourn
The Staff Liaison reminded the Group the next meeting was planned as a Skype meeting at the end of October, and to reply to the Doodle invitation (action items to resend link and to RSVP)to indicate which date(s) Members are available to participate. Mention was made that another in-person meeting may be better. The meeting was adjourned at 8:10pm.

Action Items
1. Provide an update on the metered parking transaction data availability in the Open Data Portal (Staff Liaison, January meeting).
2. Discuss proposed recommendation to focus attention on converting data PDFs to active data in the Open Data Portal (All, next meeting).
3. Discuss proposed recommendation to standardize data definitions County wide so datasets in the Open Data Portal will be more accessible and usable to constituents (All, next meeting).
4. Discuss proposed recommendation to allocate resources to improving the Open Data Portal user interface (All, next meeting).
5. Bring ideas for using open datasets that Members are willing to pursue or encourage others to pursue (All, next meeting)
6. Create a dataset index and post to Open Data Portal (Staff Liaison, next meeting)
7. Resend Doodle meeting availability invitation (Staff Liaison, by Sept 13)
8. Reply to Doodle meeting availability invitation (All, by Sept 20)
9. Confirm meeting date and send meeting invitation (Staff Liaison, by Oct 10)