

Open Data Advisory Group

August 29, 2017, 7-8pm

Attendees

Name	Role	Name	Role
David Bain	Member	Jack Belcher	Liaison
Duke Banks	Member	Alexandra Fox	Liaison
Alyssa Doom	Member	Bryna Helfer	Liaison
Nate Gould	Member	Dharmesh Jain	Liaison
Christopher Hedrick	Member	Jaime Lees	Liaison / Staff Liaison
Sarah Holland	Member	David Mota	Liaison
David Husband	Member	Katie Pegoraro	Liaison
Scott McGeary	Member	Girish Rajput	Liaison
Aaron Schroeder	Member	Jim Schwartz	Liaison
Vadim Sokolov	Member	Sanmeet Sidhu	Liaison
Lisa Sturtevant	Member		
Jason Thomas	Member / Chairperson		
Dennis Whitehead	Member		
Edith Wilson	Member		

Agenda

1. Focusing Our Initial Efforts
 - 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.
2. Open Data Portal
 - a. [Searching](#) for datasets
 - b. Exporting datasets
 - c. Site and dataset usage update
 - d. Examples how Open Data can/has been used
 - i. Arlington Restaurant Health Inspections, Event Map ([Code for NoVA](#))
 - ii. Parking Availability Predictor
 - iii. Real Estate Sale History
 - iv. What's Your Idea?
3. Updates on Data Analytics Program
 - a. Projects – completed, underway, planned
 - b. Partners – current, planned, suggestions from Group members
 - c. Metadata – expanding what we know about the open data datasets
 - d. Q&A – any questions on the data/data analytics document shared after our last meeting
4. Discussion Topic: How Good is Good Enough Data?

Much of the data requested by constituents is captured by the County. Say, for example, a constituent wants to know when a street light will be repaired. A logical step to proactively answer this type of question before it is asked is to post the work schedule online for constituents to look up any time. The challenge is that much operational work is “conditions permitting” for example it may be weather-dependent, requires certain conditions that are not met or be delayed due to reasons unrelated to the work in question. So, on the one hand, if the County doesn't put the information out there they are going to get questions and sometimes

complaints related to the lack of information. Alternatively, putting the information out runs the risk of providing some portion of constituents with the wrong information, and some of those constituents are likely to lodge complaints about data quality. What are the Group's thoughts on this dilemma and how should we advise the County Manager in this area?